Modelling and Integrating Personality as Affective Phenomenon in Agent-based Systems

vorgelegt von
Dipl.-Inf.
Sebastian Ahrndt
ORCID: 0000-0003-2716-4470

von der Fakultät IV — Elektrotechnik und Informatik
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Ingenieurwissenschaften
— Dr.-Ing. —

genehmigte Dissertation

Promotionsauschuss

Vorsitzender: Prof. Dr.-Ing. Sebastian Möller, TU Berlin
Gutachter: Prof. Dr. Dr. h.c. Sahin Albayrak, TU Berlin
Gutachterin: Prof. Dr. Catholijn M. Jonker, TU Delft
Gutachter: Michael Kaisers, PhD, Centrum Wiskunde & Informatica

Tag der wissenschaftlichen Aussprache: 16. April 2019

Berlin 2019

"Often before now have I applied my thoughts to the puzzling question – one, probably, which will puzzle me for ever – why it is that, while all Greece lies under the same sky and all the Greeks are educated alike, it has befallen us to have characters so variously constituted."

Theophrastus, The Characters of Theophrastus, 4th Century BC

Abstract

A team member's ability to predict the actions of other team members influences the efficiency of teamwork. This is named predictability and one of the challenges associated with the task of making agents team player in joint human-agent activities.

I approach this problem focusing on personality as a predictor of human behaviour and following a combination of theory and data-driven approaches. In particular, I use human-behaviour models providing insights into the human personality and related behaviour preferences. I derive these models from psychological theories and studies, which provide both, a structure describing the different facets of human personality and prediction about the relationship between these facets and preferred behaviours. These models allow agents to adapt their behaviour to other team members using learning techniques and provide the theoretical foundation for the data-driven part of my work.

In combining both, I provide different contributions. First, I present an approach to conceptualise the personality of agents and describe how this can be used to model personality effects on the agent's decision-making process. In contrast to related work, this is done in a use-case independent manner, linked to the tasks of developing adaptive systems, and generalizable w.r.t. the personality facets used. Second, I introduce two agent-models using machine-learning techniques to learn about a humans personality during interaction and validate these models within experiments with humans. My contribution is that the agent-models learn in direct interaction with the humans, whereas the related work concentrates on learning personality traits using supervised approaches. Third, I present a formalisation of the state and effects of personality within a logic of Belief, Desire, and Intention and discuss which of the characteristics are meaningful to reason about the phenomenon and its influences. The related work does not provide such a formalisation.

To make these results applicable, I finally present a development environment that guides users through the development and configuration of teamwork scenarios based on my human-behaviour models. This environment is implemented as an extension of the agent-framework developed at my institute and used within a case study that analyses the technical maturity of the approach. The case study identifies promising avenues for future work, e.g. extension of the human-personality model with further data points such as user's sex, age, background. It further identifies the linkage between the user model, holding personality information, and the behaviour model, holding personality preferences, as challenging.

Zusammenfassung

Agenten könnten Menschen einst als Teammitglieder unterstützen, wenn auch sich diese Integration in menschliche Aktivitäten herausfordernd gestaltet. Damit das Zusammenspiel erfolgreich ist, müssen geeignete Teamwork-Modelle entworfen und flexible Techniken entwickelt werden, welche die Interaktion steuern.

Die Effizienz von Teamwork wird u.a. durch die Fähigkeit der einzelnen Teammitglieder bestimmt, sich der Vorgehensweise und Intentionen der anderen Teammitglieder bewusst zu sein, um die eigene Planung, Koordination und Aktionen zu optimieren. Es entsteht die Anforderung, dass Agenten das menschliche Verhalten voraussehen müssen, um effektiv zusammenzuarbeiten.

In meiner Arbeit, bearbeite ich dieses Problem mittels einer Kombination aus theoretischen und daten-getriebenen Ansätzen. Insbesondere benutze ich menschliche Verhaltensmodelle, welche Einblicke in die menschliche Persönlichkeit und dessen Einfluss auf das Verhalten geben. Diese Modelle basieren auf psychologischen Theorien und Arbeiten, welche Strukturvorgaben für die unterschiedlichen Facetten menschlicher Persönlichkeiten liefern und gleichzeitig Aussagen über Persönlichkeitsausprägungen und Verhaltensvorlieben treffen. Während die so erstellten Modelle das theoretische Rahmenwerk liefern, welches ich in den Agenten abbilde, nutze ich sie um konkrete Verhaltensweisen zu lernen und die Unterstützung des Agenten an den individuellen Menschen anzupassen – der daten-getriebenen Teil meiner Arbeit. Hieraus ergeben sich unterschiedliche Beiträge. Als erstes, präsentiere ich einen Ansatz zur Konzeptualisierung von Persönlichkeit und beschreibe wie dieses Konzept genutzt werden kann, um den Einfluss von Persönlichkeit auf den Entscheidungsprozess der Agenten zu modellieren. Als zweites, beschreibe ich zwei Agentenmodelle welche Techniken des maschinellen Lernens nutzen um menschliche Persönlichkeiten in der direkten Interaktion zu erlernen. Beide Modelle werden in Experimenten mit Menschen evaluiert. Drittens, präsentiere ich eine Formalisierung von Struktur und Effekten von Persönlichkeit in der Entscheidungsfindung von Agenten und diskutiere die zum Ziehen von Schlussfolgerungen nützlichen Charakteristiken.

Um diese Beiträge in die Anwendung zu bringen, präsentiere ich im Abschluss eine Entwicklungsumgebung, welche es erlaubt Teamwork-Szenarien basierend auf meinen Verhaltensmodellen zu entwickeln. Diese Umgebung wurde als Erweiterung eines bestehenden Agentenframeworks umgesetzt und im Rahmen einer Fallstudie genutzt, um ihre technische Reife und Anwendbarkeit zu demonstrieren.

Acknowledgements

It is almost impossible to remember all persons involved in a project of the size and duration of my doctoral studies. Only my name and those of my doctoral committee members appear on the cover page of this document, though, a lot of people have been indirectly involved in its creation. So first, I want to apologize and thank all those that are not mentioned by name below. You know yourselves best.

I would like to provide a special thanks to my colleagues Dr.-Ing. Johannes Fähndrich and Dr.-Ing. Frank Trollmann. You guys provided insightful comments on ideas, papers, approaches, and drafts of this work. Your constant pressure to improve and challenge my own work were a real good stimulus. Your childish motivation for science inspired me over and over again. Keep it!

My colleagues at DAI-Labor and particularly those at the competence center NGS shared my way during the last years. So Marco, Andi, Nils, Stephen, Tobi, Elif, Izgh, Eva, Daniel, Kathy, Jenny: Thanks for your support, suggestions, productive criticism, and kind words. It has been a great pleasure to have worked with you over the years. It's great that I can continue to work with some of you.

I want to thank Prof. Dr. Catholijn Jonker for her comments and feedback that helped me to finish this thesis and for the open-minded Q&A we had during my defence. I also wish to thank Michael Kaisers for joining my doctoral committee when needed most and providing a review in a timely manner. I very much enjoyed to discuss my work with you.

Lastly, my deepest and most heartfelt thanks goes to my family and your constant support that pushed me to stay on track. To my mother: You laid the foundation on which I've started this journey. You have been with me on my first day at the TU Berlin and I'm happy that you also shared the last day with me. It's simply impossible to put my thankfulness into words. To Elaine: Thank you for enabling me to spend so many — if not even uncountable[1] — hours of our family time to get this work done. You, most of all, shared the sacrifices of this journey. You always give me the power to look forward and move on. To Emmi: I've finished this work they day before you arrived! Welcome!

[1] Just a joke, it must have been way less than 61.320 hours.

Contents

List of Figures

List of Tables

Part I.

Introduction

1. Introduction

In autonomous systems, it is all about decisions. An autonomous system has to choose *what* action to perform and to decide *when* this action should be performed (Wooldridge, 2009, pp. 21–28). Autonomous systems that act as team members in cooperative activities further have to decide *which* team member should or will perform an action (Wooldridge, 2009, pp. 151–156). In cooperative activities as addressed in this work—namely *joint human-agent activities* (*cf.* Klein et al., 2004)—the latter decision is identified as *function allocation* (de Winter and Dodou, 2014). Function allocation addresses the task of "*[d]eciding which functions (tasks, jobs) of a human-machine system should be allocated to the human and which to the machine (today often a computer)*" (de Winter and Dodou, 2014, p. 16:1). In the agent-community, function allocation is also known as *task sharing*, addressing the question "*...how tasks are to be allocated to individual agents*" (Wooldridge, 2009, p. 155). This task and its automation is identified as one of the challenges to make automation a team player (*cf.* Section 4.2 for an introduction to challenges and requirements). For team players, the task includes the capability to predict what the other team members will do and is one of the essential characteristics for team coordination named *predictability* (*cf.* Bradshaw et al., 2009; Hoffman and Breazeal, 2007b; Joe et al., 2014; Klein et al., 2004 or Chapter 5)

One approach to handle the above-described decisions is Artificial Intelligence (AI) planning (*cf.* Section 3.1 for an introduction). AI planning is the process of automatically searching for a sequence of actions that will achieve an objective when executed (Ghallab et al., 2004, pp. 1–16; Russell and Norvig, 2002, pp. 375–382). AI planning procedures for joint human-agent activities, *i.e.* planning processes that incorporate the humans as team member, are also named *human-aware planning* (*cf.* Cirillo, 2010 or Section 4.3). Human-aware planning is applied in situations that involve artificial and natural agents[1] in the same environment, the actions of the artificial agents being planned and those of the natural agents being predicted.

One motivation to develop such techniques is that computers and humans have different capabilities that can complement each other. An example of an early work on these differences is known as the *Fitts list* (Fitts, 1951, p. 10) and depicted in Fig. 1.1. The work lists different abilities in which humans surpass the present-day machines and in which present-day machines surpass humans. Even the summarised thoughts were

[1] Using the term 'artificial agent' I refer to software agents and robots. Indeed, robots and software agents are mostly the same; even the robot agent has a physical representation of itself (Kaisers, 2012, pp. 1–2). With the term 'natural agent' I refer to humans.

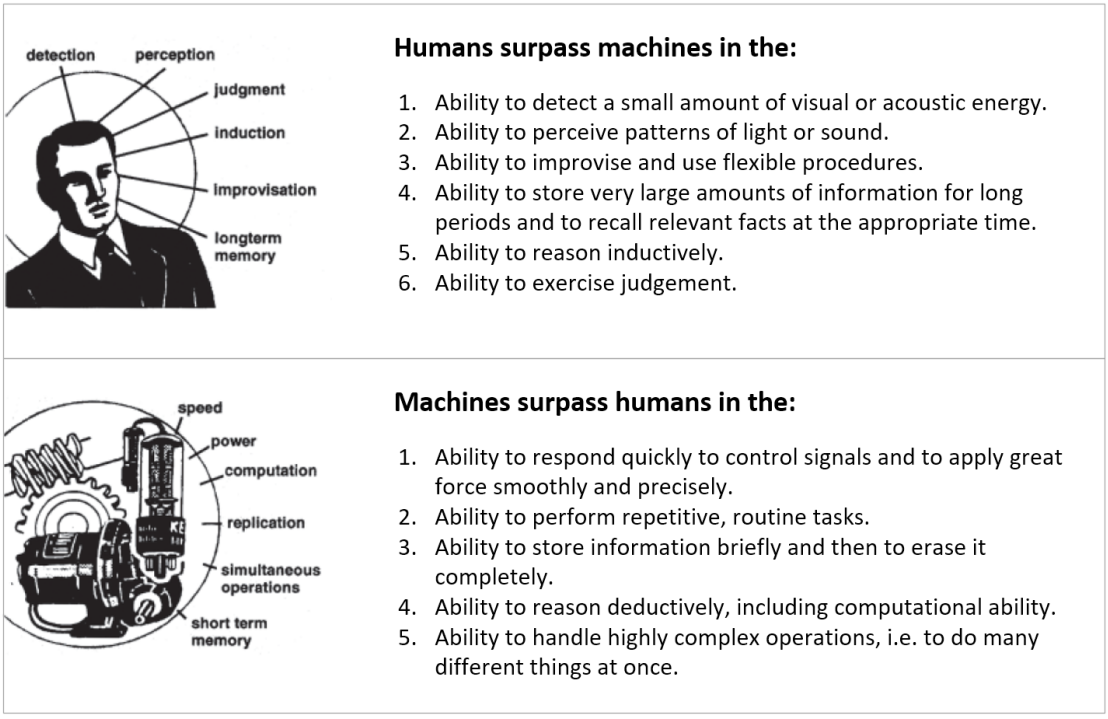


Figure 1.1.: Illustration of the Fitts list (Fitts, 1951, p. 10), which is also known as the HABA-MABA (humans-are-better-at/machines-are-better-at) or MABA-MABA (men-are-better-at/machines-are-better-at) list. This illustration is adapted from published work (Bradshaw et al., 2012, Fig. 1, p. 10).

published more than six decades ago some of them are still valid (de Winter and Hancock, 2015). In fact, the growth of computing technologies has proven that machines are getting better at several abilities (e.g., detecting, perceiving, storing very large amounts of information) and that we are now able to automate a vast amount of tasks (de Winter and Hancock, 2015, p. 5335). However, at the same time, it was shown that automating all tasks is not always the best action one can perform. This is known as the myths of autonomous systems (Bradshaw et al., 2013),[2] including the myth that "*[a]utonomous systems are autonomous*" (Bradshaw et al., 2013, p. 56). This assumption is misleading, as even humans do not have the abilities to perform autonomously in every task and situation. One consequence, among others, is that automating one task introduces other burdens to humans. Another myth is that creating "*[f]ull autonomy is not only possible, but is always desirable*" (Bradshaw et al., 2013, p. 59). That is related to the same observation, *i.e.* not reducing the workload of humans when automating tasks with machines. In fact, automation shifts the workload towards more complicated tasks. An additional reason is related to the costs associated with automating tasks, which are

[2]Prior work on this topic can be found under the terms *"Substitution Myths"* (Christoffersen and Woods, 2002) and *"Ironies of Automation"* (Bainbridge, 1983).

sometimes not viable. The authors conclude that it might be more beneficial to think about human-machine collaboration (Bradshaw et al., 2013, pp. 59–60).[3] In particular, how the capabilities of humans and machines can complement each other and what are the implications for either side (as stated by Fitts more than six decades ago).

In fact, the motivation for my thesis can be found in this area of research, particularly, at the intersection of human-computer collaboration and decision-making processes of agents. The overarching goal in this intersection is to combine the best of both types of agents—the artificial agent and the natural agent—enabling the artificial agent to support individual humans or teams of humans or to take the role of an equal team member in a human-agent team (Sycara and Sukthankar, 2006). Once artificial agents can support people and their activities (like observable in human-human teams) we[4] will see a shift from these days cooperative activities to more elaborate teamwork settings. There, agents are not only assistants but indeed become team members, which will open a broad range of application areas. Yet, such integration causes significant challenges (*cf.* Klein et al., 2004; Prada and Paiva, 2014) including the above-introduced task sharing and predictability capabilities.

After introducing the fields and the motivation to present this work, I will describe the addressed problem next (*cf.* Section 1.1). This includes the derivation and verbalisation of my research hypothesis. Afterwards, I introduce the research questions answered in this work and connect them to the individual chapters (*cf.* Section 1.2).

1.1. Problem and Research Statement

Joint human-agent activities can be considered as an extended set of actions, which are executed by a cooperating ensemble of natural and artificial agents (*cf.* Klein et al., 2004, p. 91). Both have to work in a continuous interaction and continuous interdependency. Artificial agents may increase the performance of human beings, and it is envisioned that agents may once support human beings as team members (e.g. Bradshaw et al., 2009, 2012; Prada and Paiva, 2014). Yet, as we just learned, integrating these agents into human activities is challenging (e.g. task sharing, human-aware planning, predictability). Any successful integration requires appropriate teamwork models and techniques that specify how artificial and human team members have to interact to make joint decisions and to achieve joint goals (e.g. Johnson et al., 2011; Jones et al., 2009; Klein et al., 2004; Sycara and Sukthankar, 2006).

My thesis addresses the second topic, namely the decision-making process of agents in joint human-agent activities. Henceforth I follow the argumentation of Cirillo (2010, pp. 17–30) and refer to this process as Human-Aware Planning (HAP). The ever-increasing

[3] Regarding the myths, Johnson et al. (2014a) discuss lessons learned during the development of human-robot teamwork and present related design principles.

[4] Throughout this document I use 'we' in the sense of 'the reader and the author', using it as *pluralis modestiae*, not as *pluralis maiestatis*.

use of AI in our everyday life has made HAP an evolving branch of AI planning systems (*cf.* Section 4.3 for classification, origin, and a definition of HAP). HAP can be applied when the situation involves artificial and natural agents in the same environment and when actions of artificial agents have to be planned, based on assumptions about the actions of natural agents (Cirillo, 2010, pp. 17–18). Such scenarios can be found in collaborative application areas like smart homes, where agents, robots, and humans share one environment. As an example, consider (socially assistive) robots that support the elderly (*cf.* Tapus et al., 2007).

To make autonomous agents effective team-players in such settings, planning systems have to tackle the dynamic nature of humans (e.g. Cirillo et al., 2010; Hoffman and Breazeal, 2007b; Kirsch et al., 2009; Klein et al., 2004; Kruse et al., 2013). Although humans have the ability to act better under uncertainty than computers (de Winter and Hancock, 2015, pp. 5337–5341), their behaviour features aspects of uncertainty for the agent's planning process. As an example, humans may change their goals from one moment to another without a (for computers) comprehensible reason (e.g. due to lack of information or sensing capabilities). Furthermore, humans may interrupt tasks or execute them in different ways (e.g. non-optimal way, other than expected, other than prior observed).[5] Indeed, humans "*...generally seek only good, feasible plans rather than optimal*" (Ghallab et al., 2004, p. 1) ones. These aspects of uncertainty affect the search for a feasible sequence of actions in different ways. Where a non-optimal execution might influence only the execution time of a specific task, the sudden interruption of a task endangers the whole plan and therefore the ability to reach a given goal. Hence, one has to consider that whenever a task is predicted to be fulfilled by a human, the human may perform such task and provide results either in time or delayed. Such a form of 'context-dependent' behaviour constitutes the problem space addressed in my work.

The efficiency of 'teamwork', is significantly determined by a team member's ability to predict the actions of other team members (e.g. Bradshaw et al., 2009; Bratman, 1992; Joe et al., 2014; Johnson et al., 2011; Klein et al., 2005; MacMillan et al., 2004). In other words, agents can only plan and coordinate their actions effectively, if the agent has information about the intentions of all other collaborators, including the human being(s) (Bradshaw et al., 2009, pp. 936–937). On the one hand, this requires an agent to be able to predict actions, intentions, states, goals of other team members, e.g., by observing their behaviour. Yet, it also requires an agent to act 'predictably', that is, to allow other agents to infer this knowledge. In my thesis, I approach the former problem w.r.t. human-aware planning procedures and refer to it as follows:[6]

The problem of predictability in human-aware planning is the problem of (1) determining the human team members' most likely course of actions that

[5] From the agent's point of view, this dynamic behaviour can be interpreted as the '*Quality of Service*' or '*Quality of Behaviour*' a human provides.

[6] A detailed derivation of the problem is given in Chapter 5 – Problem Analysis.

will be executed in order to reach the joint goal of the collaboration and (2) integrating this prediction into the planning process for the collaboration.

To address this problem, different authors postulate a combination of theory and data-driven approaches as beneficial (*cf.* Kirsch et al., 2010, p. 225; Prada and Paiva, 2014, pp. 6–7). In particular, the use of human-behaviour models that provide insights into the human nature from the psychological point of view is postulated as necessary (Kirsch et al., 2010, p. 225). These models are a theoretical foundation for predicting human behaviour and can provide information about, e.g. the capabilities, intentions, habits, social rules, norms, personality, emotions, and moods. Although these models are a good starting point, they must be adapted to the human's individual preferences during the actual interaction. This is because each human is different in nature and the characteristics of a human changes over time (Kirsch et al., 2010, p. 225). Thus, I aim to validate or disproof the following:

Deriving models of human behaviour from psychological studies and tailoring these models to the individual preferences and habits of a particular human being is feasible and these models can be used by artificial agents in joint human-agent activities to provide personalised assistance.

I approach this objective by collecting information about human behaviour, by adapting this information to those individuals that artificial agents communicate with, and by using this information to constrain action plans, *i.e.* influencing the action selection of the planning process. In more detail, I intend to use human-behaviour models to obtain cost-estimates for human capabilities (e.g. in terms of probabilities) and thus to establish a dynamic heuristic for the possible course of action of human users. In doing so, I focus exclusively on models that combine information about the behaviour and the personality of human beings (*cf.* Section 3.4 for an introduction to human-personality theories). Moreover, it is not my intention to develop new AI planning techniques but to use existing ones.

1.2. Research Questions

The following research questions are derived from and directly related to the problem statement. They address the task of improving the efficiency of human-aware planning components, approaching different aspects of this task in a structured manner.

The basic idea I follow is to provide information about the human capabilities and behaviour to the agents' decision-making. I refer to such information as human-behavioural models (*cf.* Section 5.3). In my work, I focus on personality as an affective phenomenon and one possible source of information about human behaviours (others being: emotions, moods, norms, laws, etc.), which I henceforth refer to as human-personality models. The all-encompassing research question leading this dissertation thus reads as follows:

Main Research Question. *How can we derive, integrate, and use human-personality models in planning procedures for joint human-agent activities?*

To answer this question, I will build a bridge from psychology to computer science on the subject of personality. The research devices employed for this are literature research, simulation, formalisation, and empirical research. The first step to derive a human-personality model is to understand how psychology defines and structures personality and its effects. By doing so, I will introduce different personality theories and provide justifications to utilise the Five-Factor Model (FFM) (McCrae and John, 1992) of personality (*cf.* Section 3.4). The FFM introduces different dimensions that are helpful to distinguish between personalities and defines a structure that is used to explain facets of human behaviour using these dimensions (*cf.* Feist et al., 2012, pp. 418 – 434). The first specific questions that I ask at this point are related to the modelling of this structure and its relevance to effects within agent-based systems. Regarding effects I refer to personality as *"...a pattern of relatively permanent traits and unique characteristics that give both consistency and individuality to a person's behavior"* (Feist et al., 2012, p. 4). In the sense of this definition, the effects (behaviour tendencies) I intend to make use of are those that contribute to the long-term behaviour preferences of humans. That is, I focus on the relatively permanent traits and their effects, neglecting other influential processes. They are, by no means, causal effects on the behaviour of a human, as such is influenced by several dynamic processes (e.g. context, role, experience, capabilities, culture, emotions) (Feist et al., 2012, pp. 424–429). However, I target on the effects solely related to the permanent traits. The questions read as follows:

Modelling State and Effect of Personality

Research Question 1. *What is the state-of-the-art for integrating personality, in particular, the FFM, as an affective phenomenon in agent-based research?*

I present an overview of the usage of personality as an affective phenomenon in agent-based systems in Chapter 7. This overview is structured according to application areas. The considered contributions approach the task of reflecting[7] psychological findings, in our case personality theories and models, within computer-processable models in different ways, ranging from rule-based approaches to utility-functions to probabilistic models. The conclusions that are drawn comprise that only a few of the contributions substantiate the decision to apply the selected representation of personality. Some of the contributions do not ground their development in psychological work. Most of the work present agent-models tailored for the specific use-case and environment, making

[7] I intentionally use the word 'reflect', referring to the trade-off between the comprehensive, complex frameworks provided by psychology and the level of complexity that can be modelled in computing systems with arguable effort. All contributions have to "*...balance between solid theoretical basis and the practicalities of parameter control*" (Bevacqua et al., 2010, p. 3).

the informational value for an adaptation of a model to other scenarios rather limited. These limitations lead to the next question:

Research Question 2. *How can we derive an agent-model representing the effects of personality with respect to the FFM?*

This question will be addressed in Chapter 7 by presenting an integration of the FFM into the Belief-Desire-Intention (BDI) model of agency (*cf.* Section 3.3). Within the presented model, I follow the idea that personality serves as a marker that contributes to consistency in the behaviour at the one hand, but which can be negated due to the influence of other affective phenomena at the other (Feist et al., 2012; Revelle and Scherer, 2010). I will discuss the influences of the different personality dimensions within the different lifecycle stages of BDI agents and thus connect two parts that are differentiated in a human-behaviour model: the *user model* holding information about the state and the *behaviour model* holding information about the effects.

Given the experiences gained answering the first two questions, the next one concentrates on the tailoring of the human-personality model to the individual preferences and habits of a particular human. The corresponding question reads as follows:

Learning Personality Information from Observation

Research Question 3. *Can agents use our model to learn the personality traits of a human during the interaction with this human?*

Chapter 8 provides an answer to this question, presenting two agent-models and experiments that use our agent-model to learn personality characteristics according to the FFM. In contrast to the related work, I do not follow an (un)supervised learning approach requiring existing data sets and thus circumvent the requirement of having (labelled) training data sets. Instead, I focus the learning task on experience gained during the interaction, *i.e.* observing the actions of humans in a controlled environment. The results indicate that we can learn the personality and that the crucial part is the linkage between expressible behaviours that encode personality information and the observable actions. Having established the prerequisites that enable to represent personality and to learn the personality of humans, the next question addresses the usage of such information:

Research Question 4. *Can we use personality information directly to make informed decisions about the potential behaviour of human?*

An answer to this question is provided in Chapter 8 using the same experiments that are used to learn about the human personality. I introduce *multi-attributed utility functions*

that make use of the personality information to predict the humans' course of action. It will be shown that the agents can outperform humans in the environment. However, the analysis of the results also reveals that these observations are not related to the personality. The discussion of the results presents several potentials to improve the approach. Those reach from the applied utility functions to the experimental design to the environment itself.

After clarifying how we can model the state and effect of personality and how such a model can be used to learn about the personality of humans using observations, I concentrate on the formalisation of personality within the agents' decision-making.[8] This is done to provide clear semantics while talking about personality in agents and leads to the following question:

Reasoning about Personality

Research Question 5. *How to represent the state and effect of personality in BDI logics?*

This question is addressed in Chapter 9 by introducing personality within one of the existing BDI logics. Personality is integrated as a new modality complementing the belief, desire, intention modalities. Therefore, I extend the BDI logic and define the necessary terms and relations that represent the state according to the FFM. I further discuss how the new modality can be used to express the effects of the personality utilising a running example and the statements that are derived thereof. Establishing these fundamentals enables us to reason about the influence of the personalities' basic tendencies on the behaviour of agents and to reason about characteristics of a personality using observations of behaviour. At this point, we can express a multitude of behaviours, but do not know which ones are meaningful/reasonable for analysing (ir)rational or personality-compliant behaviour, which leads me to the next question:

Research Question 6. *Which relations between the personality modality and the belief, desire, and intention modality are meaningful for reasoning about rational behaviour?*

Chapter 9 provides a discussion of the relations between the personality modality and the belief, desire, and intention modality. I particularly pay attention to the meaning of personality within the interplay with the other modalities. The discussion captures our intuitive understanding of the concepts in question concentrating on pairwise interactions within the possible worlds semantic in which BDI logics are grounded, *i.e.* starting with relations between worlds, detailed to the world's structures, and detailed with temporal

[8] The answer to Research Question 4 is used to substantiate design decisions in a subsequent part of this document.

relations. The conclusions that are drawn identify those relations as meaningful that express personality as an optional influence. That is to say, those relations that at the same time express the operation of enduring personality traits and enable the context-dependent behaviour.

Together, these questions address different aspects of modelling, deriving, and using human-personality models. In Chapter 5 and within the individual chapters, I connect these aspects with the different elements of adaptive systems, such as envisioned for artificial team members in joint human-agent activities. The answers provide a foundation to approach the last two questions, focusing on the integration and use of the human-personality model within joint human-agent activities:

Integration into Agent Development

Research Question 7. *What is the state-of-the-art for development environments for joint human-agent activities w.r.t. the use of human-behaviour models?*

I present an overview and analysis of the related work in Chapter 12. The analysis concentrates on the use of human-behaviour models and to which extent the approaches support the tailoring of such models during runtime. In doing so, I identify concepts and technologies that are useful for my work. The classification criteria used comprise the selected architecture, the planning and learning capabilities, and the application of social constraints, behavioural models, and intention models within the planning procedures. Comparing the contributions reveals a commitment to the same type of architecture implementing a closed loop between planning, monitoring, and executing. I further identify a focus on the integration of social constraints and intention models into teamwork development. Still, some of the authors recognise the need for a more elaborate representation of human characteristics. The conclusion is that the usage of human-behaviour models is limited. One work mentions the use of learning techniques to adapt the agent's behaviour towards the individual. This work is of conceptual nature. The gaps between the classification criteria and the available work lead to the next and last question approached in this dissertation:

Research Question 8. *How can we integrate our human-personality model into the development cycle of agents?*

Chapter 13 and 14 address this question. The former chapter introduces HPLAN, an extension of the agent-framework JIAC V (Lützenberger et al., 2013) developed at my institute[9] enabling developers to implement my human-personality model, and to use

[9] DAI-Lab, Technische Universität Berlin, `www.dai-labor.de`, last-visited: 2017-09-27

it within a lifecycle of planning, acting, and learning. The technical evaluation concentrates on verifying that HPLAN can fulfil its lifecycle and the formulated requirements. Chapter 14 describes a case study that utilises HPLAN to implement a real-world application. I show how the different abstractions and components of the environment can be calibrated and adjusted to a particular application domain and which kind of data is used to create my human-personality models. The case study is further used to evaluate the technical maturity of HPLAN and to identify possible extensions and corrections from software engineering point of view.

In the next chapter, I describe the structure of the document and explain the relations of the individual chapters to my published work.

2. Research Approach

The problem at hand is to improve human-aware planning components by providing and tailoring models of human behaviour. To approach this problem, I start with an analysis of the affected research areas, approach different aspects of the task of using personality information in agents, and introduce a development environment that enables the implementation of applications w.r.t. the problem. In the following, I will introduce the document's structure (*cf.* Section 2.1) and the relation of this work to published work of my own (*cf.* Section 2.2).

2.1. Thesis Structure

Part II – Background and Definitions The foundations part provides insights into both the touched research domains and the typically used terms and concepts. First, I give the interested reader an overview of AI planning, reinforcement learning, and agent-based systems. It is introduced how psychologists describe personality and which theory is used in this work. Afterwards, I discuss what cooperation means, how joint human-agent activities are defined and which requirements exist, and what human-aware planning is. Eventually, I present a detailed view of the problems addressed in this thesis.

Part III – Human-Personality Models – Modelling, Learning, and Reasoning The third part proceeds with a description of the individual working packages that are performed to approach the overall problem. Each chapter comes with an own problem statement, analysis of related work, approach, and evaluation/discussion. I will show how these chapters connect to each other within the conclusions of each one. The first experiments are done in a simulated environment approaching the question of how to model personality as influential characteristic within decision-making processes. The second one addresses the question if we can learn personality information during interaction with a human. The third chapter formalises personality within the decision-making process of agents, enabling us to formulate what rational behaviour means for agents with personality. The part is closed with a conclusion that wraps-up the results and integrates them into the tasks one has to satisfy to create an adaptive system.

Part IV – Human-Personality Models – Integration into Agent-Development In this part, a framework is presented that enables developers to make use of the prior presented

findings. The part follows the classical scheme of state-of-the-art analysis, conceptual and implementation work, and evaluation. I present in detail the approach and the design decisions done and finally illustrates in which way the system is implemented. The evaluation chapter provides an in-depth analysis of the approach and is twofold. It starts with a technical evaluation analysing the systems ability to fulfil a lifecycle of planning, executing and learning. Afterwards, I present a case study performed to gain insights into the usage of the presented framework. The part is closed with concluding remarks.

Part V – Thesis Summary The last part presents a wrap-up of the work described in this dissertation. The part starts with an outline of my dissertations process, comprise a discussion of the findings in relation to the research questions and emphasises the contributions. Afterwards, the limitations of the achievements are discussed and used to point out future research.

2.2. Relation to Published Work

The collected information presented within this work is based on both, the work of other authors and published work of myself. The major contributions of this thesis are based on publications at peer-reviewed conferences and workshops. Nevertheless, in-depth details about the approaches, the implementation, and the evaluations are made available for the first time within this document. The relations between the publications and the thesis are presented in the following:

Chapter 1 – Introduction motivating the work by providing insights into the domain, the addressed problem and research questions, the intended solutions, and the outcome.

- Ahrndt (2012): Introducing self-adaptive systems as one of the motivational trails to present this work, particularly focussing on self-explanation in interaction with human users.
- Ahrndt (2013): This work describes the PhDs topic, introduces a classification of the related work presented in Chapter 12, and outlines the expected contributions of the PhD project.

Chapter 4 – Human-Agent Teamwork and its Planning Procedures introducing co-operation as the umbrella term of this work building a bridge to joint human-agent activities and human-aware planning.

- Ahrndt and Albayrak (2016): There exist a bag of challenges towards making agents “team mates” – some of them are present independently from the team’s

mixture, whereas others are particularly challenging for the development of human-agent teams. This work presents an overview of the challenges and brings together knowledge from the different involved research areas, further, it provides a definition for the term joint human-agent activity.

Chapter 5 – Problem Analysis answering the questions what is meant by predictability, why it is an important requirement, and what are the relations between machine learning, human-behaviour models, and planning procedures.

- Ahrndt et al. (2016a): Within this publication, it is discussed that predictability as a challenge for joint activities is not well-defined. Furthermore, it is shown that the different involved research areas refer to the same challenge using different terms. This paper explains what is meant by predictability eventually proposing a definition for the challenge. In doing so, it discusses the relation referring to joint human-agent activities, the involved learning problem, and the inherent planning part, which is named human-aware planning.

Chapter 7 – Modelling State and Effect of Personality presenting an agent-model that combines the BDI lifecycle with the Five-Factor model of personality.

- Ahrndt et al. (2015a): Within this publication, it is demonstrated that all stages of the decision-making process are influenced by the personality and that personality-specific task assignment can alter/improve the solution's quality.
- Ahrndt et al. (2015c): Within this publication, I emphasise the impact of personality on essential elements of the behaviour of agents (e.g. decision-making processes, emotions, moods, or coping strategies). It is shown that available work on agent behaviour and contributions that investigate the nature of emotions are somewhat disconnected and that bridging this gap is able to further our efforts in conceptualising human behaviour in software agents.

Chapter 8 – Learning Personality Information from Observation approaching the task recognising the true personality of an individual using observations and the task of using such information to make an informed decision for own behaviours.

- Ahrndt et al. (2015b): Within this publication, it is shown that agents are able to learn about the personality of a human being while in interaction with humans. In order to establish this thesis, I used the *Colored Trails* game as test-bed and repeatedly played games against human beings, while at the same time adapting the human personality model. The experiment showed that some personality traits can be learned more accurately than others.

- Ahrndt and Albayrak (2017): This work extends the above-described with an extensive related work overview and a second agent-model developed to answer the question whether or not agents are able to learn the personality of a human during the interaction. The work extends the state-of-the-art in that it does not follow a supervised learning approach requiring existing data sets.

Chapter 9 – Reasoning about Personality approaching the task of formalising the concept of personality and the concept of decision-making processes jointly.

- Ahrndt et al. (2015d): Within this paper, I argue that a connection of the different cognitive characteristics requires a formalisation that specifies the concept of personality and the concept of decision-making processes jointly and present the first step towards this formalisation.

- Ahrndt and Albayrak (2016): Within this publication, we show that personality is one of the essential elements determining the behaviour of (natural) agents. It influences other cognitive mechanisms such as emotions and moods as a kind of basic heuristic, together directing an agents attention and action selection process. Motivated by the somehow disconnected research on the effects of personality and the other cognitive characteristics, the paper presents a joint formalisation of the concept of personality and the concept of decision-making processes. In doing so, one of the well-known BDI logics is extended in order to represent the state and effect of personality.

Chapter 12 – Related Work introducing a state-of-the-art analysis for development environments for joint human-agent activities.

- Ahrndt et al. (2014b): Within this publication, I analyse related work and discuss the current progress regarding the integration of human-behavioural models. This work substantiates the need for an own environment, identifying addressable gaps in existing contributions.

Chapter 13 – The HumanPlan Environment answering the question how developers can make use of personality information within the planning process for joint human-agent activities.

- Ahrndt et al. (2014a): Within this publication, I present an extension of an agent-framework enabling developers to assign capabilities and information about the influence of personality traits on these capabilities. Human beings are represented as avatars in the system, each one providing its information to the planning process in terms of cost estimates.

Chapter 14 – Case-Study: The Personality-enabled Stress Assistant approaching the task of developing a joint human-agent activity in the real world to get insights into the usage of the HPLAN environment.

- Ahrndt et al. (2016b): This demonstration paper shows the usage of relations between personalities and stress coping strategies defined as a human-behaviour model as prior knowledge to accelerate learning about user preferences in the domain of stress management.

  This work was nominated for the best demonstration award at the 15th International Conference on Autonomous Agents and Multiagent Systems.

Part II.

Background and Definitions

3. Basic Terms and Concepts

This chapter introduces concepts from (Artificial Intelligence) planning, reinforcement learning and agent-based systems that form the foundation of this dissertation. It provides insights into these fields to build a basic understanding of the terms. First, it is explained what planning means in general and in particular for the research in Artificial Intelligence. Next, I explain the core concepts of reinforcement learning and introduce reinforcement learning algorithms. This is done to provide the necessary information to comprehend the parts of the work that address the adaptation of the human-models towards the actual user. Subsequently, I explain agent-based systems and how they can be used to develop contemporary applications. Here, I also provide insights into a model of human reasoning that will be applied to derive an agent-based model that includes personality information. Anymore, I will introduce human-personality theories and provide an argumentation for a specific theory that will be used in the remainder of this work. Finally, I conclude this chapter and present insights at which points of this thesis I make use of the introduced terms and concepts.

3.1. Planning

> *"The task of coming up with a sequence of actions that will achieve a goal is called planning."* (Russell and Norvig, 2002, p. 375)

In general, planning is the process of searching for a sequence of actions that will reach a given goal (Russell and Norvig, 2002). These actions are selected from the set of all available actions, which can be extremely large, making the choice which action or which chain of actions is useful/applicable very hard (Nebel, 2000). Here, very hard means that planning is complicated in term of time and space complexity and, in consequence, that there exists a trade-off between the costs to plan and the benefits a good or maybe optimal plan will offer (Ghallab et al., 2004). This is one reason why humans rarely plan (explicitly) before they act in everyday situations. Indeed, people only search for a plan when it is strictly necessary, e.g. addressing new and/or complex tasks (Ghallab et al., 2004, p. 1). In such a case, planning can be seen as *"[...] explicit deliberation process that chooses and organizes actions by anticipating their expected outcomes."* (Ghallab et al., 2004, p. 1). In other words, one can say that planning is the process of reflecting on the available actions and organising these actions in order to achieve the goal in the

best way possible. As this reflection has to be done before actions are executed one can also say that *"[p]lanning is the reasoning side of acting."* (Ghallab et al., 2004, p. 1).

Given this general-purpose explanation of planning that holds for both natural and artificial intelligence, it can be revealed that the research in AI planning addresses the computational study of the mentioned deliberation process; answering questions such as how to find a sequence of actions, how to optimise this sequence, or how the execution of a sequence of actions transfers a system from the initial state to the targeted goal state.

In order to represent and investigate such a deliberation process in a formal way, a general model known as *state-transition system* (Dean and Wellmann, 1991) is commonly used (Ghallab et al., 2004, pp. 5–9). A state-transition system is a tuple consisting of four elements. The first element is the set of states S containing a finite or recursively enumerable quantity of states $s \in S$. Each state represents the system at a specific moment, *i.e.* it represents a set of state variables and their values. The second element is the set of actions A containing a finite or recursively enumerable quantity of actions $a \in A$. These actions are able to transfer the system from one state to another and are controlled/executed by the system itself, *i.e.* by the plan executor. Consequently, the sequence of actions we are searching for during the planning process is based on the elements of A. The third element is the set of events E containing a finite or recursively enumerable quantity of events $e \in E$. In contrast to actions, events are contingent and are not controlled/executed by the system itself. The last element is the state-transition function γ representing the influence of actions and events on the system. That is the effect of the execution of an action or the occurrence of an event that will bring the system from some state to some state.

To conclude, the 4-tuple $\Sigma = (S, A, E, \gamma)$ represents a state-transition system, where:

- $S = \{s_1, s_2, \ldots\}$ is a finite or recursively enumerable set of states;
- $A = \{a_1, a_2, \ldots\}$ is a finite or recursively enumerable set of actions;
- $E = \{e_1, e_2, \ldots\}$ is a finite or recursively enumerable set of events; and
- $\gamma = S \times A \times E \rightarrow 2^S$ is a state-transition function.

Given such a state-transition system Σ the objective of the planning process is to find a combination of actions—the plan—that transforms the initial state $s_i \in S$ into a goal state $s_g \in S_g | S_g \subseteq S$. However, finding a valid plan does not guarantee that the system actually moves from s_i to s_g during execution. That is to the presence of events; making the transition-function γ non-deterministic. Given joint human-agent activities, one could think about humans that act in the same environment as the controlled system as events. However, in Section 4.3, I will show that Humane-Aware Planning explicitly internalises humans into the system. One fact that differentiates HAP from classical AI planning techniques. Still, HAP is based on such techniques, where the most familiar

ones can be found under the umbrella terms state-space search, plan-base search, and hierarchical task networks. As it is not our intention to implement a new AI planning technique but to use existing ones, an introduction to the different approaches is out of scope. The interested reader is referred to the work of Ghallab et al. (2004) for a comprehensive introduction to the theory and practice in the field of AI planning.

3.2. Reinforcement Learning

The ability of machines to learn from past experience is another branch of research in the area of Artificial Intelligence. Mitchell (1997) defines this ability as follows: *"A computer program is said to learn from experience E with respect to some class of tasks T and performance measure P, if its performance at tasks in T, as measured by P, improves with experience E."* (Mitchell, 1997, p. 2). If we compare this definition for learning with the description of planning we can notice that the process of learning includes the reasoning about actions executed. That is based on the statement that experience can only be gained if there is some data available about actions performed for the tasks that should be fulfilled. The source where this experience originates from and the usage of this experience is one of the distinctive features of the machine learning techniques available.

Three different branches of machine learning can be distinguished: *Supervised Learning*, which is also known as predictive learning, *Unsupervised Learning*, which is also known as descriptive learning and *Reinforcement Learning* (RL) (Murphy, 2012). In Supervised Learning, the required experience is taken from a labelled set of training examples. The goal is to learn a mapping between inputs and outputs, e.g., to learn a mapping between the body height of a person (input) and the sex of this person (output). In contrast, Unsupervised Learning tries to find patterns in unlabelled input data, e.g., it tries to cluster a set of humans represented by some certain kind of features. The third type—Reinforcement Learning—receives the required experience from the interaction with the environment that provides feedback (in form of rewards or punishments) as training information for the actor. This feedback is produced by executing actions that affect the state of the environment (Sutton and Barto, 1998) and by observing whether or not or to which extent the intended effect of the executed action has been reached. The consequence of which is that each algorithm that solves *"[...] the problem of learning from interaction to achieve a goal."* (Sutton and Barto, 1998, p. 51) is a reinforcement learning algorithm. This problem is also called the *Reinforcement Learning Problem.* As we will face this problem when learning the characteristics of an individual during the interaction with this individual, I will introduce RL in more detail next.

As this is a rather vague description of RL, I will provide more detailed information about the RL cycle, the elements, and two types of RL methods next.

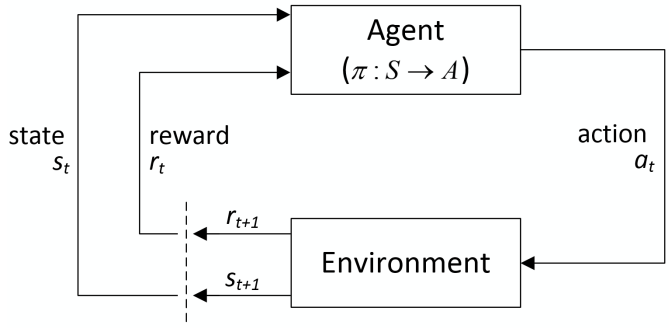


Figure 3.1.: Using RL techniques, the experience is gained during the interaction of an agent in an environment processing rewards. This illustration is adapted from published work (Sutton and Barto, 1998, p. 52).

3.2.1. Interaction Cycle and Markov Decision Processes

A schematic depiction of the reinforcement learning interaction cycle is shown in Fig. 3.1. Here, an agent acts within an environment and executes actions a at time-steps t. The environment responds to those actions by changing its state from s_t to s_{t+1} and returning a reward r_t, which is a specific numerical value (e.g., +1). As most real environments cannot produce such signals, the RL agents are required to percept the state of the environments and to judge its actions according to some metric (this component is sometimes called *Critic*). The behaviour of the agents, that means the actions that are selected, is controlled by the agent's policy π. Such a policy indicates which action is most-likely to maximise the cumulative future rewards based on the experience gained from the reward signals of previously executed actions. The agent's goals then is to find the policy π that maximises the expected rewards, which can be written as the sum of returns R with

$$R = \sum_{t=0}^{\infty} \gamma^t r_{t+1}. \tag{3.1}$$

Here the parameter γ ($0 \leq \gamma \leq 1$) is a discounting factor controlling whether the agents prefer short-term rewards and thus acts 'myopic' or long-term rewards and thus acts 'farsighted'. The problem at hand is to specify an algorithm that can find the policy with the maximum expected returns, thus leading to an agent that behaves optimally.[1] For this problem a number of algorithms have been proposed, which can be classified to be either model-free or model-based (Kaelbling et al., 1996, pp. 251–259). The difference between both is whether there is knowledge about the state-transition probability function and the reward function or not. That means, that a model-free algorithm learns a

[1] Indeed, there is a trade-off between algorithms guaranteeing the convergence to the optimum and the speed of reaching such optimum. In practice it is preferable to apply algorithms that reach a near-optimal behaviour in the short run than algorithms that guarantee the optimum in the long run but satisfy the learning speed (Kaelbling et al., 1996, p. 242).

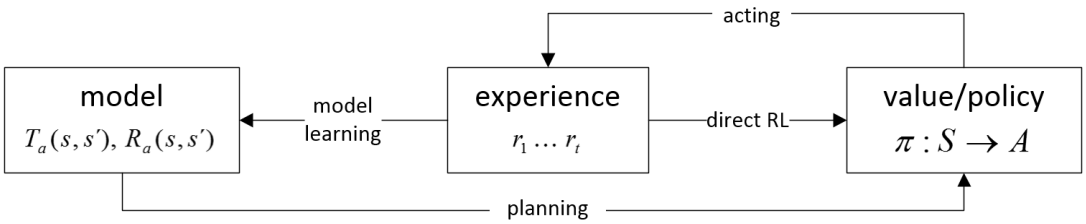


Figure 3.2.: A schematic depiction of the relation between the experience gained and the use of this experience to learn optimal behaviour: Model-free RL directly optimises the policy or value-function, model-based RL indirectly optimises those, learning a model of the environment and using such model to plan the next actions simulating future rewards. This illustration is adapted from published work (Sutton and Barto, 1998, p. 231).

policy without having or learning a model of the environment. In contrast, model-based algorithms expect a model of the environment or learn such model and derive a policy using this model.

Fig. 3.2 illustrates the differences between both types of algorithms introducing notions of a formal model to investigate agent-environment interactions named *Markov Decision Processes* (MDP) (e.g. Kaelbling et al., 1996, pp. 247–248; Sutton and Barto, 1998, pp. 66–68). A MDP is a discrete process which is at each time-step in a state s. The agent is able to select an action a that is applicable in this state. This is similar to the notion of state-transition systems introduced earlier. However, in MDP the state transitions depend on stochastic influences (e.g. effects) and the agent's actions. When an action is executed the process moves to a new state s' with some probability $T_a(s, s')$ providing some reward $R_a(s, s')$. In other words, $T_a(s, s')$ describes the transition probability of moving from s to s' when executing action a and $R_a(s, s')$ denotes the reward when successfully moving to state s'.

A MDP is represented by the 4-tuple $MDP = (S, A, T, R)$, where:

- $S = \{s_1, s_2, \dots\}$ is a set of states;
- $A = \{a_1, a_2, \dots\}$ is the set of actions;
- $T : S \times A \times S \rightarrow [0, 1]$ is the state-transition probability function, where $T_a(s, s') = T_a(s_{t+1} = s' \mid s_t = s, a_t = a)$ is the probability that the execution of an action a at time-step t in a state s leads to the state s' in time-step $t + 1$; and
- $R : S \times A \times S \rightarrow \mathbb{R}$ is the reward function, where $R_a(s, s') = E(r_{t+1} \mid s_t = s, a_t = a, s_{t+1} = s')$ is the reward received when the process successfully moves from state s to state s' applying action a.

The goal of model-based algorithms is to learn the state-transition probability function and the reward function. This knowledge is used to simulate the future expected rewards when applying actions.

Due to the inherent differences, both types of algorithms have advantages and disadvantages (*cf.* Dyan and Niv, 2008, p. 187; Kaelbling et al., 1996, pp. 251–259). Model-free algorithms "*[...] are clearly easier to use in terms of online decision-making*" (Dyan and Niv, 2008, p. 187). Their characteristic of being inflexible when facing unexpected incidents is not of relevance when learning the personality facets of humans, which are relatively stable over periods of time. However, in contrast to model-based algorithms, model-free algorithms require much trial-and-error experience to behave acceptably. Further, they are not able to respond to goal-shifts sensitively. We can emphasise that both types of RL algorithms—model-free and model-based—have in common that they are applied to estimate the value functions, *i.e.* the value of a state. The differences lie within the source of experience, which is either real experience (model-free) or a simulated one (model-based). Balancing these arguments, let me conclude to apply model-free RL. That is because personality traits are relatively stable over time and my goal is to learn about the effects independent of dynamic processes. Furthermore, I intend to learn during the actual interaction and use the made experience as cost estimate within the action planning process. The challenge that I expect is to gain enough experience within the interaction.

3.2.2. Model-free Reinforcement Learning and Policies

In model-free RL each action gets assigned an action-value $Q_t(a)$, which is used by the policy to determine the action that will be applied in the current state (Sutton and Barto, 1998). The true value of an action is denoted as $Q^*(a)$, whereas $Q_t(a)$ denotes the estimated value at time step t. Remind, that a learning agent performs an action a repeatedly for t times. At each time step t the agent receives a reward $r_t(a)$. This reward is used to iteratively improve the estimate of the action-value $Q_t(a)$ of the expected reward for executing action a. For this update process, several ways exist. A natural one is to build the average value of all rewards received when the action was selected according to the following formula known as the *Sample-Average method* (Sutton and Barto, 1998, p. 27):

$$Q_t(a) \leftarrow \frac{r_1(a) + r_2(a) + \cdots + r_{k_a}(a)}{k_a}. \tag{3.2}$$

Here, the parameter k_a denotes the amount of times the action has been chosen prior to t. Using the simple-average method it is guaranteed that $Q_t(a)$ converges to $Q^*(a)$ when $k_a \to \infty$.

Another way to update the action-values is *Q-learning* (e.g. Sutton and Barto, 1998, pp. 148–151; Watkins and Dayan, 1992), which is arguably the most influential reinforcement learning algorithm and was originally published by Watkins (1989). The

action-values are updated according to the following formula, known as the Q-learning update rule, where α ($0 \leq \alpha \leq 1$) denotes the learning rate and γ is the above-mentioned discount factor:

$$Q_a(s, t+1) \leftarrow Q_a(s, t) + \alpha \left(R_a(s, s') + \gamma \max_b Q_b(s', t) - Q_a(s, t) \right). \tag{3.3}$$

The learning rate helps to control the impact of new experiences to the currently estimated action-value. The max operator forces the update to use the best existing estimate to update the current estimate. The integration of the max operator implies an own policy for the update function (always choose the best existing estimate), which is apart of the policy of the agent determining which action to choose. For that reason, the Q-learning update rule is called off-policy.

The update rules produce Q-values that are used by the agent to select the applicable action according to the agent's policy. One strategy to do that is to always select the action with the highest Q-value, *i.e.* to select the action a^* which $Q_t(a^*) = max_a Q_t(a)$. Due to this characteristic, this action selection strategy is named greedy policy, leading to an agent that not explores the action space and thus only finds local optima (Sutton and Barto, 1998). A more advanced strategy is to select a random (non-optimal) action once a while with some probability ϵ and to select the action with the highest Q-value with probability $1 - \epsilon$, which is named the ϵ-greedy policy. This leads to an agent that explores the action space and in the long-run is guaranteed to find the global optima. Nevertheless, this strategy changes the behaviour of the agent drastically when a new action attains the highest Q-value and, further, during the exploration, it is equally likely to select the worst action available as it is to choose the one which is next-to-best. This behaviour can lead to a very slow convergence in large action spaces. Besides this strategy, other policies exist, e.g. the Boltzmann exploration scheme, which can smoothly balance between exploration and exploitation (Kaisers, 2012, p. 19).

3.2.3. Model-based Reinforcement Learning and Value-Functions

In model-based RL the experience gained during the interaction is used to build a model of the environment instead of improving the policy with a more accurate estimate of the action-values. This model, in turn, is used to determine the best policy. One approach to build such a model is to implement a lookup table, which for each state s and action a stores the resulting states s' and rewards $R_a(s, s')$ (Sutton and Barto, 1998). This lookup table can then be used to simulate the future starting in a state, selecting actions according to some policy and cumulating the expected rewards E_π. During this simulation, two different values can be calculated. The first one denotes the value of a state, indicating how good it is to be in that state and is calculated according

to the following formula known as the *state-value function* for policy π:

$$V^{\pi}(s) = E_{\pi}(R_t \mid s_t = s) = E_{\pi}\left(\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \mid s_t = s\right). \tag{3.4}$$

The second one denotes the value of taking an action in the current state, indicating how good it is to apply that action in that state and is calculated according to the following formula known as the *action-value function* for policy π:

$$Q^{\pi}(s,a) = E_{\pi}(R_t \mid s_t = s, a_t = a) = E_{\pi}\left(\sum_{k=0}^{\infty} \gamma^k r_{t+k+1} \mid s_t = s, a_t = a\right). \tag{3.5}$$

Both values can be used to decide which policy to apply by finding the optimal policy π^* with $V^*(s) = max_{\pi}\, V^{\pi}(s)$ and $Q^*(s,a) = max_{\pi}\, Q^{\pi}(s,a)$. To optimise the action-value function the different available model-free RL algorithm such as Q-learning or SARSA can be applied.

However, using such a lookup table becomes computational impractical when addressing large action and state-spaces, which especially takes effect in non-deterministic environments. The reason for this is that the 'simulated experience' (Sutton and Barto, 1998, p. 228) is applied to all available state-action pairs. A more advantage way is called *prioritized sweeping*. Here the idea is to compute only the effects for those state-actions pairs that changed their values in the most eminent way recently. That means that useless computation such as the calculation of the values for state-action pairs that were not visited are avoided, thus decreasing the computational complexity of the planning process. Planning, in this case, can be compared to the idea of the backward state-space search, starting in the goal state and working backwards using associated actions to reach predecessors.

To conclude, we can emphasise that both types of RL algorithms—model-free and model-based—have in common that they are applied to estimate the value functions. The differences lie within the source of experience, which is either real experience or a simulated one.

3.3. Agents

In the following, I will describe the meaning of the term *agent* within the scope of this work. Finding an answer to this question is not that easy, as there is an ongoing discussion on agents, multi-agent systems, and agent-based research.[2] Indeed there exist a multitude of definitions coined by different groups (*cf.* Franklin and Graesser, 1996). One of the more familiar definitions is given by Russell and Norvig (2002) stating that

[2]Such discussion can be observed during, e.g., the panel discussion AAMAS 2015. Lately, such discussions were particularly focusing on real-world applications and industrial adoptions. However, part of this panel was always a discussion on the substantial parts of agent-based research.

"[a]n agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators." (Russell and Norvig, 2002, p. 34). That is a rather vague definition, as it misses to define what sensing and acting means. Thus making it complicated to differentiate between programs and agents (Franklin and Graesser, 1996). Another definition is given by Wooldridge (2009), who states that *"[a]n agent is a computer system that is situated in some environment, and that is capable of autonomous action in this environment in order to meet its delegated objectives."* (Wooldridge, 2009, p. 21). According to this definition, an agent is an encapsulated entity that tries to fulfil its design objectives or in other words its goals in an autonomous manner. The author underpins this definition with different properties an agent should fulfil. The occurrence of such properties helps to distinguish between the weak and the strong notion of agency (Wooldridge and Jennings, 1995, p. 117). According to the weak notion of agency an agent is a computing system that presents the following four characteristics (e.g. Wooldridge and Jennings, 1995, p. 116):

- *Autonomy*, which refers to the ability to act without the intervention of humans and other agents and having control over the internal state and the actions that are applied;
- *Social ability*, which refers to the ability to act socially regarding communicating with other agents/computing-systems or even humans;
- *Reactivity*, which refers to the ability to react to changes in the environment in a timely fashion, which includes the ability to sense the environment (of course not all changes are of importance for each agent); and
- *Pro-activeness*, which refers to the ability to act in a goal-directed manner not only responding to environmental changes but taking the initiative.

These properties lead to one of the key differences between objects as the basic concept of object-oriented software engineering and agents. Despite the fact that objects also encapsulate their identity, state and behaviour, they do not act autonomous regarding deciding when to interact and with whom and which actions to choose (Weiß, 2001).[3] The strong notion of agency extends the weak one introducing additional characteristics that should be satisfied and is especially used in the field of AI (Wooldridge and Jennings, 1995, p. 117). Such additional characteristics are lent from natural agents such as human beings and characterise an agent, for example, using a mentalistic notion, namely beliefs, capabilities, intentions, and commitments (here Wooldridge and Jennings, 1995 refer to the work of Shoham, 1993).

Up to this point, the concept of an agent was explained without going into architectural considerations. Here, different types of agents can be distinguished reaching from simple

[3] To differentiate agents from web-services the interested reader is referred to the work of Dickinson and Wooldridge (2005).

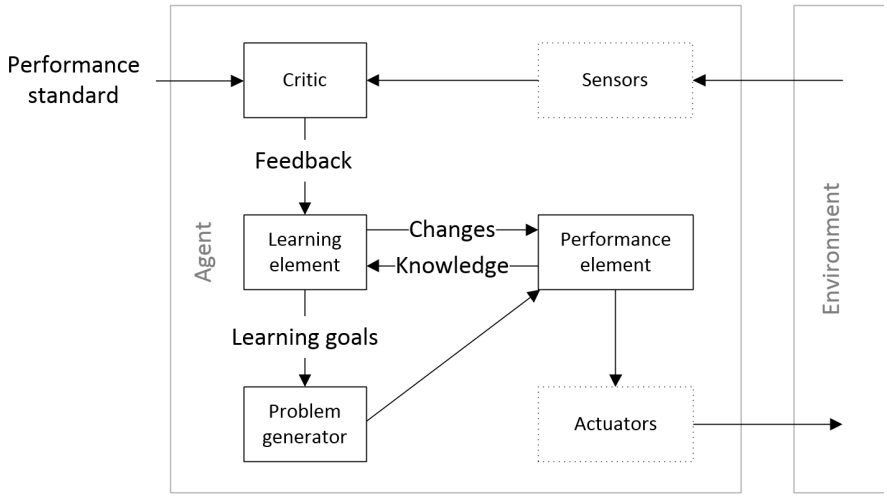


Figure 3.3.: A schematic depiction of a learning agent that improves its performance using information perceived from the environment. This illustration is adapted from published work (Russell and Norvig, 2002, p. 55).

reflex-agent to utility-based reflex and learning agents (Russell and Norvig, 2002, pp. 44–54). During this dissertation, we will concentrate on the latter, as the final system will contain multiple agents where at least one type of them can learn.

Fig. 3.3 shows a schematic depiction of a learning agent that uses information from the environment to improve its behaviour over time. It introduces a component called *critic*, that reasons about sensed information. This component evaluates the information using some performance measure. One can imagine that the agent senses some information and forwards this raw data to the critic who interprets it and creates computable feedback, e.g. in terms of the numerical value +1, which is forwarded to the learning element applying a reinforcement learning algorithm. The problem generator then uses the updated knowledge to suggest actions that are executed by the performance element. Consequently, Russell and Norvig (2002) define a learning agent as one that *"[...] improves its performance on future tasks after making observations about the world."* (Russell and Norvig, 2002, p. 693). This definition is closely related to the definition of learning introduced earlier.

Agent-based systems mainly consist of more than one agent at a time, and the weak-notion already introduces the ability to act socially regarding communicating with other agents to accomplish the application goals. We refer to such assemblies of multiple interacting agents as *multi-agent systems*.

3.3.1. Multi-Agent Systems

A multi-agent system (MAS) is a conglomeration of single-agents and is *"[...] defined as a loosely coupled network of problem solvers that interact to solve problems that are beyond the individual capabilities or knowledge of each problem solver"* (Sycara, 1998, p. 80) (this definition is based on the work of Durfee and Lesser, 1989). In the definition the word problem-solving entity refers to an agent.

One specific characteristic of a MAS is that there exists no single agent that incorporates the required resources to solve the superior problem. Here resources mean incomplete information and capabilities, leading to the following properties, which are specific for multi-agent systems:

- there exists no single agent able to solve the superior goal on its own;
- there exists no central entity that controls or observes the whole process;
- the knowledge is distributed; and
- the calculations are done asynchronously.

These properties are among others deduced from the properties of single-agents (a central instance contradicts the autonomy of an agent) and distributed system (absence of a common primary memory). One consequence is that agent-oriented software engineering is a natural approach when building complex software systems, as it enables software engineers to break down complex problems into smaller ones (Jennings, 2001, pp. 35–37). Other factors that are of interest when using MAS are extensibility, robustness, maintainability, flexibility and physical distribution (Sycara, 1998, p. 80). The latter can, for example, be utilised to avoid performance bottlenecks. The former is of interest when facing changing requirements, which can be satisfied by introducing new agents or by replacing specific agents. One disadvantage of the distributed nature is the increasing interaction effort and complexity, which can, for example, be addressed by using standardised agent communication languages and hierarchically organising the MAS.

3.3.2. The BDI Model

The Belief-Desire-Intention (Rao and Georgeff, 1995) (BDI) model of agency is a popular model for the conceptualisation of human behaviour. The computational BDI model is based on the theory of human practical reasoning introduced by Bratman (1987), which is a folk psychology theory and also named BDI. This theory was first formalised by Cohen and Levesque (1990) and later improved towards a complete computational theory by different computer scientist (main contributions have been made by Rao and Georgeff, 1995 and Wooldridge, 2000). The BDI model separates the current execution

of a plan from the activity of selecting a plan. Therefore BDI agents are equipped with three mental concepts: the current assumptions about the state of the environment (beliefs); the knowledge about intended goals, *i.e.* the set of goals the agents want to fulfil (desires); and a set of plan-elements (intentions) describing how to achieve the desires. Furthermore, each agent provides a set of basic capabilities (actions) applicable to manipulate the environment. The lifecycle of a BDI agent comprises four phases. These are the *Belief Revision*, the *Option Generation*, the *Filter Process*, and the *Actuation*.

Algorithm 3.1 The practical reasoning process of a BDI agent.

Name: bdiCycle
Input: Bel_{init}, Int_{init}
Output: -

1: Bel $\leftarrow Bel_{init}$
2: Des $\leftarrow \langle\rangle$
3: Int $\leftarrow Int_{init}$
4: **while** true **do**
5: Bel $\leftarrow$ beliefRevision(Bel, percept())
6: Des $\leftarrow$ options(Bel, Int)
7: Int $\leftarrow$ filter(Bel, Des, Int)
8: $\pi \leftarrow$ plan(Bel, Int)
9: execute(π)
10: **end while**

Algorithm 3.1 illustrates the reasoning process, which starts with the belief revision. Here given sensor values are computed by a perception function and update the beliefs about the environment (line 5). The next step is the option generation, where the agent generates its desires taking into account the updated beliefs and the available intentions (line 6). The third phase is the filter process, where the agent chooses between competing intentions and commits to achieve some of them next (line 7). Finally, the last stage is the actuation, in which the agent influences the environment performing actions executing the selected plan (line 8–9).

In summary, the BDI cycle presents a continuous *sense-decide-act* loop (Myers and Morley, 2001, p. 109). The actual planning process, *i.e.* the generation of plans is not part of the BDI model and left to developers. In earlier work on this topic, such plans were selected from plan libraries that were part of each intention. Yet, there are several works combining the reasoning process with planning techniques, both in a more theoretical consideration (e.g. de Silva et al., 2009) as well as examples for the practical realisation (e.g. Walczak et al., 2007).

3.4. Human Personality Theories

Personality can be defined as *"...a pattern of relatively permanent traits and unique characteristics that give both consistency and individuality to a person's behavior"* (Feist et al., 2012, p. 4). Such traits contribute to the individual differences in a humans behaviour on the one hand and account for the consistency of such behaviour over time and situations on the other. Human personality theories provide theoretically and frequently empirically examined tools that help to study personality as one of the affective phenomena influencing our behaviour (others are, e.g. emotions, moods, temperament). There exist a number of such theories, which can be assigned to one of the following categories (e.g. Feist et al., 2012; McCrae and Costa Jr., 2006, pp. 21–24):

- Type theories describe the personality using different existing types, whereas each individual is assigned to its personality type.
- Trait theories describe the individual difference using a number of personality traits, which are relatively stable over time.
- Psychodynamic theories describe the personality pattern based on early experiences that have been made by an individual and that is controlled by the unconscious.
- Behavioural theories describe personality using the individual interaction with its environment, measuring such behaviours.
- Humanist theories describe personality as the inherent potential and the will of each individual to self-actualise.

Whereas psychodynamic, behavioural, and humanist theories are interested in explaining how personality develops (McCrae and Costa Jr., 2006, pp. 21–24); type and trait theories focus on explaining the individual differences by identifying measurable dimensions/dichotomies that make up a human personality (McCrae and Costa Jr., 2006, pp. 24–29). As I intend to use personality to differentiate between individuals, I will concentrate on the latter.

The type and trait theories have in common that each type/trait is a characteristic feature of a human that can be used to explain the individual's behaviour, thoughts, emotions, and its motives along patterns of behaviour (Hanna, 2016, p. 35). One can identify a major theory for each theory category; both substantiated with a significant body of results (Furnham, 1996, p. 307): The *Myers-Briggs Type Indicator* (Myers and Byers, 1995) and the *Five-Factor Model* (McCrae and John, 1992). In the following, I will introduce these theories and discuss the differences between both. This discussion ends with an important decision. That is, which human personality theory to apply during this dissertation.

3.4.1. The Five-Factor Model

The Five-Factor Model of personality (McCrae and Costa, 1989; McCrae and John, 1992) (FFM) is a psychological theory that can be used to model human personality types and their influences on the decision-making process of humans. As suggested by the name, FFM introduces five dimensions characterising an individual, which briefly described are:

- ***O****penness to experience* describes a person's preference to vary its activities over keeping a strict routine and is also related to the creativity of a person (e.g., inventive, emotional and curious behaviour vs. consistent, conservative and cautious behaviour).
- ***C****onscientiousness* describes a person's preference to act duteously over spontaneous. That means the level of self-discipline when aiming for achievements (e.g., efficient, planned and organised behaviour vs. easy-going, spontaneous and careless behaviour).
- ***E****xtraversion* describes a person's preference to interact with other people and to gain energy from this interaction instead of being more independent of social interaction (e.g., outgoing, action-oriented and energetic behaviour vs. solitary, inward and reserved behaviour).
- ***A****greeableness* describes a person's preference to trust others, to act helpful and to be optimistic over an antagonistic and sceptical mind set. Directly influences the quality of relationships with other individuals (e.g., friendly, cooperative and compassionate behaviour vs. analytical, antagonistic and detached behaviour).
- ***N****euroticism* describes a person's preference to interpret external stimuli such as stress as minatory over confidence and emotional stability. Neuroticism addresses the level of emotional reaction to events (e.g., sensitive, pessimistic and nervous behaviour vs. secure, emotionally stable and confident behaviour).

These dimensions are also named the Big Five personality traits leading to acronyms such as OCEAN and NEOAC, NEO-PI and NEO-PI-R frequently used when addressing the FFM. The characteristics of each dimension are declared as a variation from the norm, whereas each dimension is an overarching container subsuming different lower-level personality traits. Table 3.1 illustrates this showing a collection of keywords and traits that are associated with each dimension. Taking this observation into account, one can argue that the FFM is a conceptual framework about human personality traits that can, for example, be used to integrate other trait inventories/personality theories into its structure (*cf.* John and Srivastava, 1999, pp. 1–3; O'Connor, 2002; Corr and Matthews, 2009, pp. 89–92).

Table 3.1.: A collection of keywords associated with the dimensions of FFM (McCrae and John, 1992, pp. 178–179). Some of these keywords are subordinated traits.

FFM trait	Associated keywords (partly subordinated traits)
Openness	Fantasy, Aesthetics, Introspective, Ideas, Values intellectual matters, Judges in unconventional terms
Conscientiousness	Competence, Order, Dutifulness, Achievement Striving, Self-Discipline, Deliberation
Extraversion	Warmth, Gregariosness, Assertiveness, Activity, Excitement Seeking, Positive Emotions
Agreeableness	Trust, Straightforwardness, Altruism, Compliance, Modesty, Tender-Mindedness
Neuroticism	Anxiety, Hostility, Depression, Self-Consciousness, Impulsiveness, Vulnerability

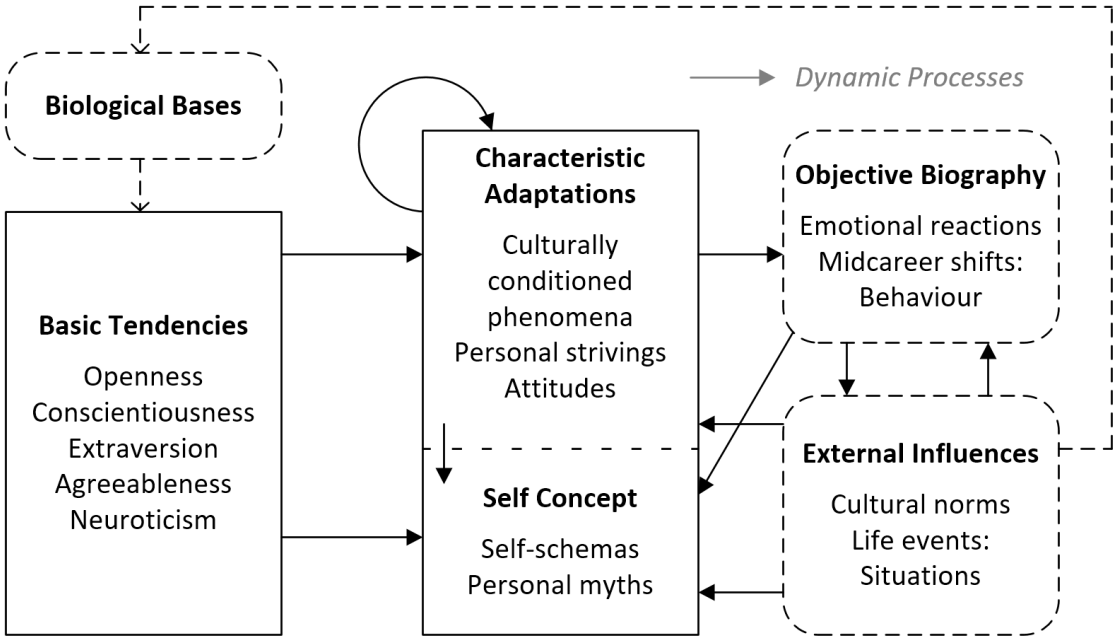


Figure 3.4.: The Five-Factor Theory includes the FFM as one of its core components named Basic Tendencies. This illustration is adapted from published work (McCrae and Costa Jr., 2010, Fig. 5.1, p. 163).

As the FFM describes the structure of personality, the Five-Factor Theory (FFT) explains the processes that lead to personality traits and the involved processes that influence the behaviour (Feist et al., 2012, p. 423). Fig. 3.4 illustrates the personality system according to FFT. It introduces three core components (Basic Tendencies, Characteristic Adaptations, and Self Concept) and three peripheral ones (Biological Bases, Objective Biography, and External Influences) that together build the personality system (Feist et al., 2012, p. 424–427). The components are interrelated. These interrelations are named dynamic processes and depicted as arrows. They constitute the direction of casual influences between the components. They are dynamic as they are subject to change over time. An introduction to FFT, its components, and related research is provided by McCrae and Costa Jr. (2010) or Feist et al. (2012, p. 423–429).

3.4.2. The Myers-Briggs Type Indicator

The *Myers-Briggs Type Indicator* (Myers and Byers, 1995) (MBTI) is based on the theory about psychological preferences introduced by Jung (1971) and provides a questionnaire to measure the psychological preferences of an individual using four pairs of dichotomies, which are:

- *Extraversion* (E) vs (I) *Introversion*, describing the dichotomy between gaining energy from interaction with other people against gaining energy from within;
- *Sensing* (S) vs (N) *Intuition*, describing the dichotomy between acting straight according to perceptions against abstractly interpreting the perceptions and acting more flexible;
- *Thinking* (T) vs (F) *Feeling*, describing the dichotomy between acting logical and detached against acting more emphatic when interacting with others; and
- *Judging* (J) vs (P) *Perception* describing the dichotomy between acting goal-directed and self-disciplined against changing commitments permanently when facing new perceptions (can be interpreted as *"[...] the choice between exploration versus exploitation."* (Salvit and Sklar, 2012, p. 151)).

Although these explanations only briefly describe the meaning of each dichotomy, they give us a hint upon the correlation of each dichotomy with the traits of FFM. However, MBTI uses these dichotomies to assign one of 16 types ($2^4 = 16$) to each human. This type determines how an individual perceives the world and makes its decisions. For example, a person that is labelled as ENTJ prefers to interact with other persons and is goal-focused, whereas a person of type ISFP prefers to work alone and is not that goal-focused (Salvit and Sklar, 2012, pp. 147–148). Consequently, MBTI tries to carve out the natural preferences of a human being according to the dichotomies and then explains how each of the resulting types typically behaves.

3.4.3. Comparing FFM and MBTI

There exist a variety of differences between both theories. To start with, the FFM emerged from empirical observations and analysis (Feist et al., 2012, pp. 420–423), whereas MBTI emerged from theoretical considerations, which were proven empirical (Feist et al., 2012, pp. 115–120). Another difference is the use of types of personalities on the one hand and personality traits on the other. The use of types presents the advantage of being distinct, but at the same time presents the disadvantage of being disjoint. That means that being classified as an extrovert (E) discerns an individual from being an introvert (I) and adds such an individual to its specific cluster, without giving any hint about the degree of extroversion. Still, this information might be valuable when this person was close to the 'artificial' border that disjoints the dichotomies or when someone wants to compare persons of the same type. At this point a continuous scale as presented by FFM delivers more information but misses the advantage of introducing standardised clusters to compare groups of people, making the implementation of FFM into agents challenging.

The completeness of a theory is another important characteristic that implies whether such theory is broad enough to understand/describe the different human personalities. Here, it was shown that MBTI misses covering different characteristics of humans (*cf.* Furnham, 1996, p. 306). In particular, the missing preference of being emotionally stable is criticised. In contrast, FFM presents a more generic structure, which is nevertheless also criticised for neglecting some domains of a human personality like honesty or religiosity (Paunonen and Jackson, 2000, pp. 827–830) (also applies to MBTI). In both cases, these criticisms are still an open discussion among psychologists and are subject to further investigation.

Besides the completeness of a theory, reliability is another characteristic one can take into account. On the one hand, reliability addresses the consistency of the results when assessing an individual using different assessment methods such as self-assessment, questionnaires and professional assessments. On the other hand, it addresses the consistency when performing the same assessment repeatedly with some timely distance, which is also named test-retest reliability. MBTI suffers in its test-retest reliability. Different experiments have shown that there is a chance of 50% to be classified as another MBTI type when repeating the test after only five weeks (Pittenger, 2005, pp. 214–215). Here, FFM delivers more accurate results for short-term intervals (1 week) (McCrae et al., 2011, p. 31, Table 2) and long-term intervals (10 years) (Terracciano et al., 2006, pp. 1000–1003).[4]

[4] This supports the finding that a developed personality is relatively stable over the lifespan of a human after it developed until early to mid-adulthood (e.g. Feist et al., 2012, p. 428; McCrae and Costa Jr., 2006, pp. 3–6, 112–115).

3.4.4. Discussion

Balancing the presented arguments and taking into account the possibility to integrate MBTI into FFM comes down to the point *"that it may be better [...] to reinterpret the MBTI in terms of the five factor model"* (McCrae and Costa, 1989, p. 37, according to Furnham, 1996). During this dissertation, I will follow this advice and apply the Five-Factor Model as the theoretical vehicle for the human-behavioural models. For the agent-community, this would imply to apply the FFM for experiments using personality traits, e.g. when modelling human-behaviour. However, one might argue that the arguments do not hold when the goal is just to produce different artificial agent traits. I argue that psychological findings should not be ignored and that knowledge transfer between both areas has a relatively long tradition, which is of vital importance. Later, the literature analysis will show that existing agent-based approaches usually use a simplified version of the MBTI or a model not based on psychology findings at all. Here, FFM enables a differentiated classification of personality types and the representation and comparison of different magnitudes of personality characteristics. This can be done, as FFM is a general model for the state of personality that subsumes other theories, if necessary. One could further argue that FFM provides a level of complexity not required for most use-cases and that we should balance between the substantiated theoretical model and the practical applicability in agents, e.g. concerning parameter control. However, FFM supports this balancing act by concentrating on its big five dimensions in the first place, and optionally allowing a fine-granular personality description using many subtraits in each of the dimensions (*cf.* Table 3.1). This is arguably an acceptable foundation that can be used to investigate different levels of personalities.

3.5. Conclusion

This chapter aimed to give a broad overview of the different research fields addressed within this thesis. I described the objective of AI planning. Later, I use this knowledge describing the requirements for human-aware planning. It will be revealed that most requirements can be satisfied out-of-the-box using contemporary approaches. Another focus of AI is related to learning. I have described what distinguished the three most important AI learning concepts from each other and provided a more detailed introduction to reinforcement learning. I introduced RL as in the setting of human-agent collaboration one can learn directly within the interaction with the human. Hence, the problem of tailoring models of human-behaviour towards the individual human is an RL problem. I introduced agents, multi-agent systems, and the BDI model as underlying concept of most of the work described next. Finally, I gave an introduction to how psychologists describe personality. This is done using personality theories. One frequently will find either the Five-Factor Model of Personality or the Myers-Briggs-Type-Indicator when reading about agent-based systems featuring personality aspects. I substantiated my

decision to use the Five-Factor Model comparing both theories according to their origin, scale, completeness, and reliability.

We will proceed with introducing cooperation, joint activities, and human-aware planning in the next chapter.

4. Human-Agent Teamwork and its Planning Procedures

This chapter introduces concepts of teamwork and cooperation. In particular, it provides insights into the motivation behind cooperation, the process of cooperating and the theoretical models underpinning cooperative activities. I will start by explaining what cooperation means in particular in agent-based systems and under which circumstance cooperation is required (*cf.* Section 4.1). Afterwards, I will introduce joint human-agent activities as one of the key concepts of this work and define the term within the work (*cf.* Section 4.2). Here I will also introduce challenges associated with joint human-agent activities, which will later be used to classify and distinct this dissertation from existing work. Next, human-aware planning, an evolving branch of AI planning systems, which accounts for joint human-agent activities, is introduced (*cf.* Section 4.3). Here I will reveal who has coined this term, in which domains human-aware planning is applied and how the term is defined. Finally, I will wrap up this chapter and remark the connection points of the presented terms to this work (*cf.* Section 4.4).

4.1. Cooperation

Normally, agents cooperate with each other to overcome some certain kind of inherent limitations. These limitations can be found either on the perception level, the cognition level and/or the execution level (e.g. Rosenthal et al., 2010, p. 915; Sycara, 1998, pp. 79–81. As examples, consider agents with a sensory malfunction (perception level), humans with a disease like dementia (cognition level), or robots that are not able to overcome obstacles like stairs (execution level). Nevertheless, cooperation would be avoided if no other stimuli exist, as it adds additional cost to an activity, e.g. in terms of a communication overhead. This can be an external stimulus like a goal that is not reachable without cooperation either caused by resource or capability constraints or an internal stimulus like an extroverted character that gains energy from cooperating with others (Jung, 1971, according to Salvit and Sklar, 2012).

In the context of multi-agent systems, cooperation is often used as one characteristic to distinguish multi-agent system from other approaches such as distributed computing, object-oriented systems, and expert systems (Doran et al., 1997). Furthermore, the characteristics of the cooperation can be used to classify different MAS and to compare them, e.g. applying the cooperation topology that is depicted in Fig. 4.1. According

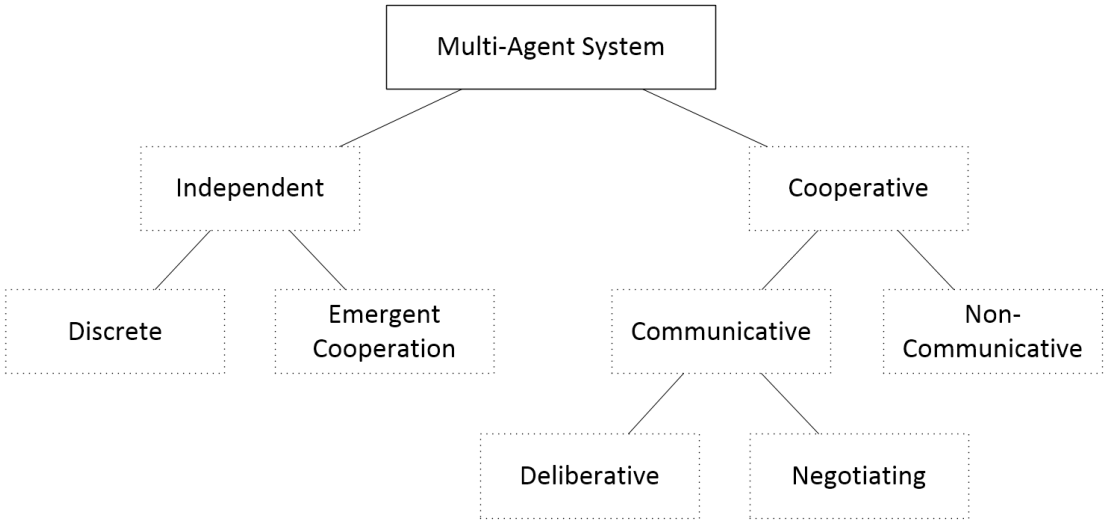


Figure 4.1.: An early version of a cooperation topology for multi-agent systems. Here MAS are divided into independent and cooperative systems. Still independent MAS can yield cooperative behaviour from the point of view of an observer. This illustration is adopted from published work (Doran et al., 1997, p. 310).

to this topology, we can classify MAS being independent or cooperative. The former means that the mode of behaviour of the agents do not depend on each other and that a cooperation only takes place from an observer's point of view. The latter addresses MAS that yield some kind of cooperation either through communication with or observation of other agents. This type of an agent-based system is more in line with our definition of MAS.

One characteristic of cooperation is not found in this topology, though, the authors discuss it. We can identify the characteristic in the statement that cooperation means *"...to act with another or others for a common purpose and for common benefit."* (Doran et al., 1997, p. 312). The statement addresses the characteristic that cooperative systems follow some common purpose, which is also named *basic compact* or *joint goal*. These joint goals are an inherent part of cooperation and present a primitive concept that can not be analysed by only taking into account the individual goals of each agent (Jennings, 1993, pp. 14–22). The presence of a joint goal requires that cooperative activities can only take place when the agents agree to act together towards that goal.[1] In other words, the agents are required to commit to working together in a joint activity towards a joint goal. Thereby they agree to act responsively to the intentions and actions of the others, which is also named mutual responsiveness (*cf.* Bratman, 1992, p. 328; Jennings, 1993,

[1]Note that the presence of an explicit joint goal conflicts with the occurrence of the emergent cooperation found in independent MAS. For that reason, a cooperative activity that is based on a joint goal, which is explicitly known by the agents, is also named shared cooperative activity (Bratman, 1992). This work only addresses such type of activities.

pp. 18–20). This agreement does not imply that the agents commit to the goal for the same reason. It only means that the agents agree to work together and to act mutually responsive and to provide mutual support (Bratman, 1992, pp. 329–331). Taking these properties into account clarifies why joint goals are a primitive concept for its own.

One specific type of a joint activity is the agreement of an ensemble of natural and artificial agents to work towards a joint goal named joint human-agent activity, which is introduced next.

4.2. Joint Human-Agent Activities and Teamwork

"ItâĂŹs not cooperation, if either you do it all or I do it all."
First Law of Cooperative Systems, David D. Woods

In general, a joint-activity is a set of behaviours that is executed by at least two people that work together towards an objective — coordinating their individual behaviours (Bratman, 1992; Clark, 1996). In joint human-agent activities, the agreement to work together is accomplished between humans and agents; building a human-agent team. The humans and agents involved coordinate with each other to reach a goal that can not be reached by a single individual, to fasten the goal achievement process, or to increase quality parameters, to name but a few reasons. In consequence, they form a symbiotic relationship in which agents fulfil tasks for humans, humans in return help agents in performing tasks, and agents and humans work together achieving joint tasks (Rosenthal et al., 2010).

The term itself evolved from two sides: The first one being agent-based teamwork settings and the second one observations of human-human teams that work together to reach a goal. Thus, synonyms and similarities for these activities can be found when reading about teamwork settings and cooperative activities in both areas, either examining human-human, human-agent, or agent-agent teams. For instance, comparing the work of Bratman (1992) on shared cooperative activities and the work of Klein et al. (2004, 2005) on joint human-agent activities reveals that they identified the same basic properties, which are sometimes renamed.

In order to build a better understanding of the nature of coordinative interactions between natural and artificial agents requirements and challenges will be discussed next. Although these challenges prevail for either teamwork setting, I highlight which parts are particularly challenging when approaching the human-agent case. After introducing the challenge, I introduce a definition of the term joint human-agent activity. The objective is to provide an overview of this area by bringing together the different terms used by different groups and pointing to related material.

4.2.1. Challenges and Requirements

There are many challenges and requirements associated with the task of developing artificial agents that act as 'team members'. Multiple authors have presented work summarising or investigating these challenges. They frequently use different terms for the same characteristics (*cf.* Bradshaw et al., 2009; Hoffman and Breazeal, 2007b; Joe et al., 2014; Johnson et al., 2012; Klein et al., 2004, 2005; Sycara and Sukthankar, 2006; Sycara and Lewis, 2004). Furthermore, the different groups sometimes distinguish ten challenges (Klein et al., 2004), sometimes six (Klein et al., 2005), or even less defining most of the requirements within the terms directability and observability (Christoffersen and Woods, 2002). As this makes it difficult to capture the entire field I provide an overview of the elements necessary for effective teams summarising the different explanations w.r.t. the ten challenges presented by Klein et al. (2004), next. Afterwards, I will show that the different authors agree on the same set of requirements while using different names and descriptions.

Elements of Joint Activities

Taking into account the overview presented by Klein et al. (2004) Table 4.1 lists ten challenges, where some of them are directly related to the above-introduced requirements. In the following, we discuss each challenge in detail.

The first challenge is named *Basic Compact* and addresses the requirement that in a joint-activity all participants must agree to work together, regardless if these activities are carried out by teams of humans, teams of agents or teams of humans and agents. This agreement is called Basic Compact. It is often tacit and a commitment of the participants to a mutual goal. Hence, *"[t]o be a team player, an intelligent agent must fulfil the requirements of a Basic Compact to engage in common-grounding activities"* (Klein et al., 2004, p. 92). In the context of MAS, we learned that this agreement is also named joint goal. However, the challenge also claims that there has to be a common ground. Meaning, that the individuals have some shared knowledge not only about the objective but also about possible actions, existing rules and norms, communication capabilities and so forth. Since coordination is a continuous process, the basic compact and the common ground are not discrete. Rather, both are subject to an everlasting communication process including negotiating, testing, updating, adapting, and repairing the mutual understanding of the joint goal and the joint knowledge (Klein et al., 2005, pp. 14–19). This process is also named *grounding* and can be found as premise/characteristic in several fields, e.g. study of conversations and negotiations or human-computer interaction (e.g. Brennan, 1998; Clark, 1996).

The second challenge is named *Adequate Models* and addresses the requirement to model the other participant's intentions and actions adequately. Furthermore, it includes the ability to reason about these models to infer knowledge about the participants. One

Table 4.1.: Ten challenges that must be addressed when developing artificial agents for joint human-agent activities (*cf.* Klein et al., 2004).

#	Challenge	Description
1.	Basic Compact	Agents must agree to work together to achieve a task and should have a common ground.
2.	Adequate Models	Agents need adequate models to reason about the others intentions and actions.
3.	Predictability	Agents should act predictable and should be able to predict the course of actions of other team members.
4.	Directability	Agents must be directable, that means they must provide the capability to adjust their level of autonomy when asked or forced by other team members.
5.	Revealing Status and Intentions	Agents must be able to communicate their intentions and their current state to other team members.
6.	Interpreting Signals	Agents must be able to infer and reason about intentions and states that are communicated by other team members.
7.	Goal Negotiation	Agents must be able to negotiate about the goals and individual sub-goals that should be achieved by the team.
8.	Collaboration	Agents should be able to cooperate with other team members in terms of understanding and reacting to the current state of the problem solving process.
9.	Attention Management	Agents should be able to highlight activities or plan changes and to manage the attention to these for the team as a whole.
10.	Cost Control	Agents should try to control the cost of the cooperation, e.g. in terms of the communication overhead during negotiations.

example for a required model is the basic compact itself. Other examples are related to the common ground that has to be established, e.g. sharing the same vocabulary, or the preferences and skills of the other participants. Solving this challenge requires contributions from different fields including knowledge representation, learning, reasoning, and planning. In fact, an adjacent research field—opponent modelling—presents several techniques for building such models and inferring knowledge (*cf.* Baarslag et al. (2015) for a comprehensive survey). Other relevant terms are (shared) mental models and team mental models. Mohammed et al. (2015) points to research on team mental models providing several further reading points. These models hold the shared knowledge about the teamwork and taskwork and are also subject to an everlasting refinement process (*cf.* Mohammed et al., 2010, pp. 879–883 for an introduction to team mental models).

The shared mental model theory states that the effort to synchronise knowledge about teamwork and taskwork is beneficial for the team performance (Smith-Jentsch et al., 2008, pp. 305–306), e.g., to anticipate needs and actions of other team members. This leads to the next challenge.

The third challenge is named *Predictability* (sometimes Interpredictability or Mutual Predicability) and addresses the requirement of building knowledge about other participants' attitudes, capabilities and course of action. Furthermore, as the participants agreed to work together it is assumed that the participants act in a way that enables the others to predict their behaviour (*cf.* Klein et al., 2005, pp. 13–14; Klein and Wright, 2016, p. 54:4). This is part of what some authors have named observability (*cf.* Christoffersen and Woods, 2002, pp. 4–6; Johnson et al., 2014b, p. 51). As predictability is the challenge that is addressed within this thesis, Chapter 5 provides a detailed discussion of the characteristics and requirements that have to be satisfied.

The fourth challenge is named *Directability* and addresses the requirement of adapting the own degree of autonomy if necessary. We can find work on this topic using different terms, e.g. sliding, flexible, adaptable, and adjustable autonomy; or levels of autonomy; or degrees of automation (Johnson et al., 2011, pp. 174–177).[2] Directability is related to the possible hierarchical structure of teams, where one team member can delegate actions, task, or sub-goals to others (*cf.* Christoffersen and Woods, 2002, pp. 7–9; Klein et al., 2004, p. 93; Klein et al., 2005, pp. 19–20). Indeed, the earlier work on agent-based systems presented two types of agents: fully-autonomous agents or teleoperated agents, which are agents that require guidance in each step (Myers and Morley, 2001, p. 108). Soon, it was recognised that in teamwork settings such as mixed-initiative approaches it is required to not only delegate tasks, which includes the acceptance of this delegation by the recipient but also to accept guidance during the decision-making process.

The fifth challenge is named *Revealing Status and Intention* and addresses the communication capabilities of agents. It focuses on the capabilities to inform other team members about the current situation and intentions. This includes information about objectives, capacities, resources to use, errors, and planned course of action. Although the spreading of this information is important, it comes with the trade-off of overwhelming others' with information. Thus, it is not only a technical issue but also a cognitive and organisational one, e.g. including the judgement whether a partner is currently interruptible or not; or which modality should be used to forward information (Klein et al., 2005, pp. 23–26).

The sixth challenge is named *Interpreting Signals* and addresses the fact that agents have to be able to receive signals and to process these signals in terms of building knowledge, e.g. models of the teammates. Indeed, it includes the possibility to interpret/reason about different types of information. This information reach from facts that

[2] In an early work, *T.B. Sheridan* presented ten degrees of automation that are widely known today (Sheridan, 1992, according to Millot and Boy, 2012).

are directly related to joint actions (like a partner has finished a task) over information about the state of the joint activity (like the fulfilment of a sub-goal or the arrival of a new team member) to information about teammates (like humans are getting tired once in a while). Thus, this challenge directly relates to the communication capabilities of agents and their capabilities to infer knowledge from observations (indirect communication) and speech-acts (direct communication). Both being long-term research areas in agent-based systems. As we address joint human-agent activities, it adds the challenge to also interpret the human's cognitive states by, e.g. learning about emotions and coping strategies, or learning about personality and preferences. Thus, this challenge is within the scope of this thesis.

The seventh challenge is named *Goal Negotiation* and addresses the active involvement of agents into the bargaining process about goals and subgoals of a joint activity. One can imagine that agreeing to a basic compact may include a step of negotiating about this compact and the intended role of the team member within the joint activity. This step involves the capability of arguing and reasoning about potential goals. Goal negotiation is also related to directability; as part of the negotiation process might be the identification of subgoals and the agreement on using specific resources or applying a specific set of norms to fulfil goals. Both restrict the team members in their level of autonomy. Several research areas work on agent-based (automated) negotiation. Lin and Kraus (2010) present a review of agent-based approaches that negotiate with humans and point out different characteristics that are notably challenging in this area. They are mainly related to the fact that agents have to handle incomplete information (e.g. not known social preferences between team members) and negotiate with an opponent that is bounded in rationality, *i.e.* that do not follow an equilibrium strategy.

The eighth challenge is named *Collaboration* and addresses collaborative approaches for the decision-making process of agents. This challenge brings together the concepts mentioned above by assuming that collaborations are a forever tentative process. One example is related to the interdependence of the actions of the individuals. It implies that one partner has to take into account the intentions, state, and course of action of the other partners during its planning process. An early approach related to agent-agent teamwork is the SharedPlan model introduced by Grosz and Sidner (1990) and Grosz and Kraus (1996). It is a theoretical vehicle for collaborative planning that takes into account that collaborative plans are not only a sum of individual plans but a *"refinement process"* (Grosz and Kraus, 1996, p. 1) of partial plans of the individuals. In fact, agents also need the ability to replan if there are changes, e.g. as one team member failed to reach an important sub-goal. Thus, to be a team player an agent should continuously monitor the overall situation as adequate as possible, should be able to inform and negotiate about changes, should be able to plan in collaboration with other team members, and so forth. As introduced in Chapter 3.1 classical AI planning techniques do not account for these requirements. However, the research areas of mixed-initiative

planning and the, soon to be discussed, human-aware planning present techniques for collaborative planning.

The ninth challenge is named *Attention Management* and addresses the necessity of spreading information in the right form using the right interaction modality at the right moment. For example, one team member should inform another team member about a resource that is running short that is required to fulfil the joint goal. Leading the attention of the other team member towards this issue. This can be done by repeatedly sending status signals, resulting in a vast amount of information that may overwhelm the partner. Hence, there exists a trade-off between not overwhelming teammates with too much information and not informing them too late. Klein et al. (2005, pp. 32–46) discuss several examples that show how bad attention management can lead to what is called 'Fundamental Common Ground Breakdown' and conclude that this challenge is an important issue for human-computer interaction research.

The tenth and last challenge is named *Cost Control* and addresses the fact that the cooperation benefits do not come for free. Rather, the advantage that is offered by joint activities is abrogated if the coordination itself is too costly. Klein et al. (2005) provides an example: *"One typical arrangement for a relay race is to use four runners to cover 400 meters. It would be inefficient to have 24 runners stationed around a 400-meter track. The coordination costs would outweigh advantages of the freshness and energy of each new runner."* (Klein et al., 2005, p. 31). Even in this example the coordinative actions include different of the above-mentioned facts, e.g. one runner has to provide signals (*i'm approaching you, you should start*), the partners have to predict each other's behaviour (*its him, the hand-over will be left-handed*), have to monitor the status of each other (*he isn't that fast today, I should wait for another few seconds before I start*), to name only a few. Thus coordinating joint activities is a continuous process and part of an effective teamwork is to handle what can be named the coordination economy.

Discussion

In the following, we elaborate the essential challenges that have been identified by the majority of the authors. However, explicitly classifying the challenges is not an easy task, as they depend on each other. For instance, predictability includes building adequate models, which includes reasoning about information, which requires that other team members reveal their status and intention. Yet, comparing the available works shows that the authors agree on three basic categories of properties necessary to make agents team player (thus can be found in work on human-human teamwork as well, *cf.* Bratman, 1992). These are:

- **Grounding** (esp. Challenge 1, 2): The grounding-process of the joint-activity that includes the (tacit) agreement to a basic compact and the continuous process of building and maintaining a common ground.

- **Mutual Engagement** (esp. Challenge 2, 3, 7, 8): The mutual predictability and mutual responsibility of each team member in the joint activity. The mutual engagement includes building models of other team members and acting cooperative regarding action planning, execution, communication, and goal management.

- **Acceptance** (esp. Challenge 4, 5, 6): The directability and observability of a team members' behaviour, which addresses among others the capability to dynamically adjust the own level of autonomy, the necessity to not act capricious, and the requirement of managing the attention of others.

In a comprehensive work, Sycara and Sukthankar (2006) discuss these categories by highlighting the importance of information exchange, fruitful communication, supporting behaviour, and team initiative. The authors state that the key factors for enabling human-agent teamwork are related to mutual predictability, to building a shared understanding, which is here named team knowledge, and the ability of the teammates to redirect and adapt to each other. Within the next section, I will present a development methodology that takes the presented requirements/challenges explicitly into account. It will serve as a guide during the next parts of this work when approaching predictability and personality in detail.

4.2.2. The Coactive System Model

One question arising is how to satisfy these challenges/requirements. Johnson et al. (2011) argue that autonomy-centred approaches are not suitable for making automation a team player. That means that research that is focusing on substituting joint actions with more complex, more autonomous agents misses taking into account the effects on the joint activity itself. This is also known as the *"Substitution Myths"* (Christoffersen and Woods, 2002), *i.e.* system designers should always take into account that autonomy in teamwork settings is no unidimensional property and must be placed within the (joint-)context in which it is operating. One example is the common motivation of offloading work from humans to machines vs. the burden of humans to perform more complex tasks, e.g. supervising the machine. Within this example, at least Challenge 5 – Revealing Status and Intention and Challenge 8 – Collaboration remain and change their characteristics. Thus system designers should remember that adding/changing automation in joint activities affects the other team members and may affect the whole cooperation. A detailed discussion on this topic is presented by Bradshaw et al. (2013). The authors work out the details of the Substitution Myths by means of what they call *"The Seven Deadly Myths of Autonomous Systems"* (Bradshaw et al., 2013, p. 54).

The *Coactive System Model* (*cf.* Johnson et al., 2011, 2012, 2014b) was developed to address these findings. The main idea is to switch the perspective from autonomy-centred system design to cooperation-centred system design when developing solutions for joint

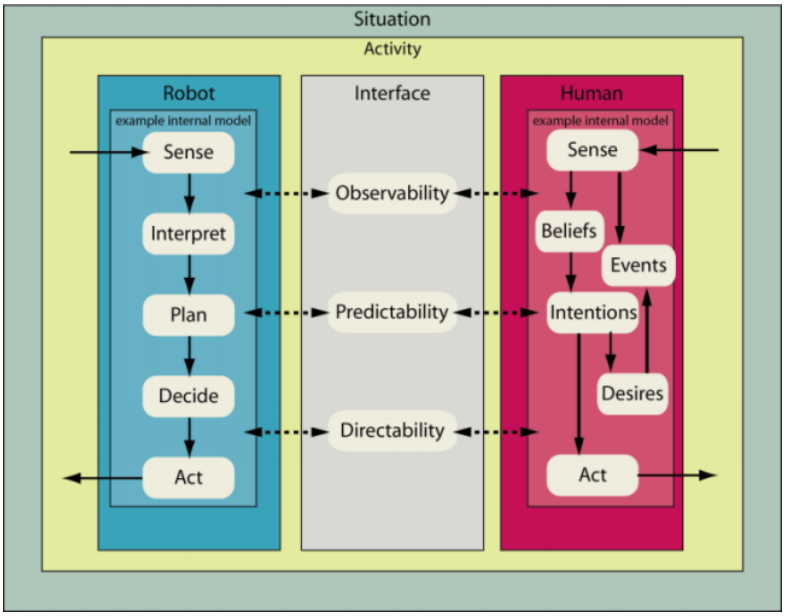


Figure 4.2.: The coactive system model that considers observability, predictability, and directability as major characteristics that lead the design process of joint activities. This illustration is taken from published work (Johnson et al., 2014b, p. 51).

human-agent activities.[3] Fig. 4.2 illustrates the coactive system model that explicitly takes into consideration three of the mentioned challenges: observability, predictability, and directability. Even the above-presented challenges do not include observability directly; it is identified as an aggregation of Challenge 5 – Revealing Status and Intention, Challenge 6 – Interpreting Signals, and Challenge 9 – Attention Management (Johnson et al., 2014b, p. 51). Notable for the coactive system model is the abstract interface between robots and humans. It is introduced to decouple the individuals and to provide an abstraction layer responsible for all elements that are required to manage the interdependencies of the teamwork. Furthermore, it abstracts the individuals in a manner that makes the internal models independent of each other. In the coactive system model interdependence is described as a twofold mutual correlation that includes *"...the set of complementary relationships that two or more parties rely on to manage required (hard) or opportunistic (soft) dependencies in joint activity."* (Johnson et al., 2014b, p. 47). Required dependencies are related to capabilities, *i.e.* the actions one participant is able (or is not able) to execute. In fact, these are the required dependencies that are necessary to reach the joint goal. Thus, one of the first steps in coactive system design is to analyse which capabilities are required. Afterwards, system designers should assess whether the team members provide such capabilities and how reliable they can perform them alone or with the support of partners. In contrast, the opportunistic relationships are related

[3] A complete description of the coactive system model can be found in the PhD work of Johnson (2014).

to soft facts that are useful to make cooperation efficient, robust, and effective. Those are supportive actions/behaviours/characteristics that facilitate team performance, e.g. an adequate communication and attention management. *Johnson et al.* argue that satisfying the soft relationships requires to *"...consider all internal cognitive processes of the parties involved"* (Johnson et al., 2014b, p. 49). Thus, the authors provide (presumably unintentionally) a justification/motivation to investigate the influence of personality on the behaviour of humans within joint activities as one of the existing cognitive processes.

In Chapter 14, the Coactive System Model will be used during the requirement analysis to identify tasks, the role of the different actors, and how the task depends on each other. Next, the term joint human-agent activity will be defined.

4.2.3. Term Definition

To clarify the understanding of the term *joint human-agent activity* within this work, I define it in the following. Therefore, different definitions provided by other authors are presented, leading to my interpretation of the meaning of the term.

We started to approach the concept of joint human-agent activities by discussing challenges and requirements for making artificial agents team members for humans. Furthermore, we have introduced properties for joint activities that originated from the human-factor analysis, *i.e.* research on human-human teamwork. One conclusion we can draw is that there are several keywords used when talking about joint human-agent activities — reaching from teamwork, human-agent teamwork, human-automation teamwork to joint activities, cooperative activities, cooperation, shared activities, to name but a few. One common aspect one can find within these terms is the concept of an action as the building block of each activity. Furthermore, it is frequently talked about the coordination of actions in this setting — leading to the concept of joint actions as the building block of joint activities. Clark (1996) defines joint actions with respect to the usage of language in communication as follows:

> *"A joint action is one that is carried out by an ensemble of people acting in coordination with each other."* (Clark, 1996, p. 3).

The first clarification we are receiving is that joint actions/coordination takes place between two or more individuals. In fact, coordination is the inherent part of each joint action. That is to say, that during a joint activity the individuals perform actions and that the coordination of these actions between each other is named joint action – note that the joint action includes the coordinated actions as well (Clark, 1996, pp. 3–4). This coordination process includes the introduced mutual characteristics; adding a context to the joint action. This context also provides the actions' goal. That means that (joint)-actions are always performed with an intention. Writing about agent teamwork, Cohen et al. (1997) claims that this intention is an essential concept for the overall teamwork as well:

> *"A team is a set of agents having a shared objective and a shared mental state – without either, there is no unified activity and hence no team"* (Cohen et al., 1997, p. 94).

In this description, the intention is named shared objective. Earlier, we have talked about it describing the concepts of a joint goal and a basic compact. Introducing the shared goal has enabled us to distinguish independent and cooperative cooperation and to focus on the latter. Indeed, the notion of a shared goal differs from the notion of a common goal in that a common goal can be established by individuals independent of each other, whereas a shared goal requires an active coordination process (Cohen et al., 1997, pp. 94–95).[4] The shared mental state is the second important concept in the statement. We have referred to this concept as the requirement to build a common ground and to build adequate models about the teammates. This allows us to infer that teams perform joint activities during the teamwork. Furthermore, we again see that coordination is a major part of the joint activity, not only at the actual activity level but also at the goal level. This is because the coordinated actions might affect the goal-directed activities of the partners, making it necessary to coordinate the individual goals. We have learned about this requirement introducing the challenges goal negotiation and collaboration. An attempt to define the term joint activity is provided by Klein et al. (2004) and reads as follows:

> *"We define joint activity as an extended set of actions that are carried out by an ensemble of people who are coordinating with each other."* (Klein et al., 2004, p. 91).

Writing about an extended set of actions this statement emphasises that joint activities are more than actions that are coordinated. Unfortunately, the authors miss to describe what is represented by this set. However, I would like to highlight that the definition is limited when talking about actions only:

- Bratman (1992) argues that each cooperative activity involves appropriated behaviour by its participants. This behaviour includes, among the actions, characteristics such as mutual responsiveness and mutual support.
- Johnson et al. (2014b) support this while writing about required interdependence relationships (capabilities, which are actions or set of actions) and opportunistic interdependence relationships (behaviours making joint work more effective).
- Joe et al. (2014) argue that not the automation of task is the main challenge anymore, but providing 'soft' skills that lead to an effective behaviour during the teamwork.

[4] Thus, the shared goal is a synonym to the prior introduced joint goal.

These comments highlight that we should talk about behaviours and not actions. Interestingly, *Klein et al.* defined joint activity in a succeeding work as follows:

> *"A joint activity is an extended set of behaviors that are carried out by an ensemble of people who are coordinating with each other."* (Klein et al., 2005, p. 8).

While arguing that a joint activity consists of coordinated behaviours, the nature of the extended set remains undefined. For this work, we follow the argumentation of Johnson et al. (2014b) and refer to interdependence relationships as what defines the extended set of behaviours. Thus the extended set of behaviours refers to all characteristics necessary for a fluent teamwork. Taking these clarifications and the explanations of the different authors together (*cf.* Bratman, 1992; Bradshaw et al., 2009; Clark, 1996; Johnson et al., 2014b; Joe et al., 2014; Klein et al., 2004; Rosenthal et al., 2010), eventually enables me to define the term joint human-agent activity:

> *A joint human-agent activity is an extended set of behaviours that is executed by an ensemble of natural and artificial agents who are coordinating with each other working in relative continuous interaction to achieve a joint goal.*

My definition builds on the idea that we should explicitly distinguish between natural and artificial agents. This idea is grounded in the observation that we are far away from achieving the fluent dynamics found in human-human teams (Joe et al., 2014, pp. 10–11). Hence, I use the differentiation between natural and artificial agents to highlight that there exists a difference in the capabilities. Additionally, I have made explicit the continuous interaction. It is necessary to enable the team partners to build a common ground and adequate models as both are subject to an everlasting communication process. Thus excluding settings, where participants negotiate a joint goal, coordinate actions once and work towards the joint goal without further interaction.

In the next section, I will approach planning procedures for joint human-agent activities as a postulated requirement to satisfy Challenge 8 – Collaboration. I will show, that these planning techniques implicitly address some of the challenges, though the main authors in the planning domain do not explicitly work on teamwork settings.

4.3. Human-Aware Planning

Human-Aware Planning (HAP) is an evolving branch of AI planning systems for collaborative settings where agents coexist with humans (Cirillo et al., 2010, pp. 15:1–15:3) (e.g. socially assistive robots in household environments). It can be considered as AI planning procedures that account for joint human-agent activities. Referring to the introduced challenges we can find the requirement of providing such technique in Challenge 8 –

Collaboration. In human-aware planning procedures, we consider the human as part of the system, not as its ruler. In particular, humans are not seen as external entities but as actors situated inside the environment, which follow (and plan) its agendas. Consequently, internalising the human weakens one of the assumptions of classical AI planning systems, where the agent planning is in control of the state of the world and such state is only influenced by the actions of the agent and events in the world (*cf.* Section 3.1).

To be more concrete, I follow the argumentation of Cirillo (2010, p. 17) stating that human-aware planning is applied in situations that involve artificial and natural agents in the same environment, the actions of the artificial agents being planned and those of the natural agents being predicted. Here, to build effective team-players, the agents are required to include the state of the human into their planning process by anticipating the actions of the human (*cf.* Hoffman and Breazeal, 2007a,b). This information then can be used to generate plans including the human as an actor and respecting a set of constraints, e.g., social, cognitive, normative or interaction constraints (*cf.* Cirillo et al., 2010; Montreuil et al., 2007). In the following, I will approach Human-Aware Planning by integrating it into the general field of AI planning.

4.3.1. Classification

Human-aware planning is based on the premise that there exists a prediction (forecast) of the humans plan or the set of possible plans. This prediction must be included during the actual search for a plan and, further, the prediction must be derived somehow. Thus a HAP-solution requires more than the actual planning component. I will use this characteristic in the following to differentiate human-aware planning from other AI planning fields:

External classification HAP distinguish itself from collaborative planning by not planning actions for humans, but using the prediction of a human's course of action for the own action selection process (*cf.* Allen and Ferguson (2002) for an introduction to collaborative planning). Furthermore, it distinguishes itself from mixed-initiative planning as it does not build a closed-loop system where planning decisions are forwarded to the human. Rather, it uses different techniques, e.g. activity recognition or plan recognition systems, to indirectly observe the human and build knowledge about its behaviour. Also, HAP differentiates itself from classical AI planning techniques that can handle external events (in fact, a human action from the planning agent's point of view is an external event) by internalising the human behaviour. Usually, this is done with respect to constraints such as human actions cannot be avoided, they can occur in every state and parallel to the agent's action execution, they have a temporal component, they can fail, and some actions might introduce new or shifting goals for the agent.

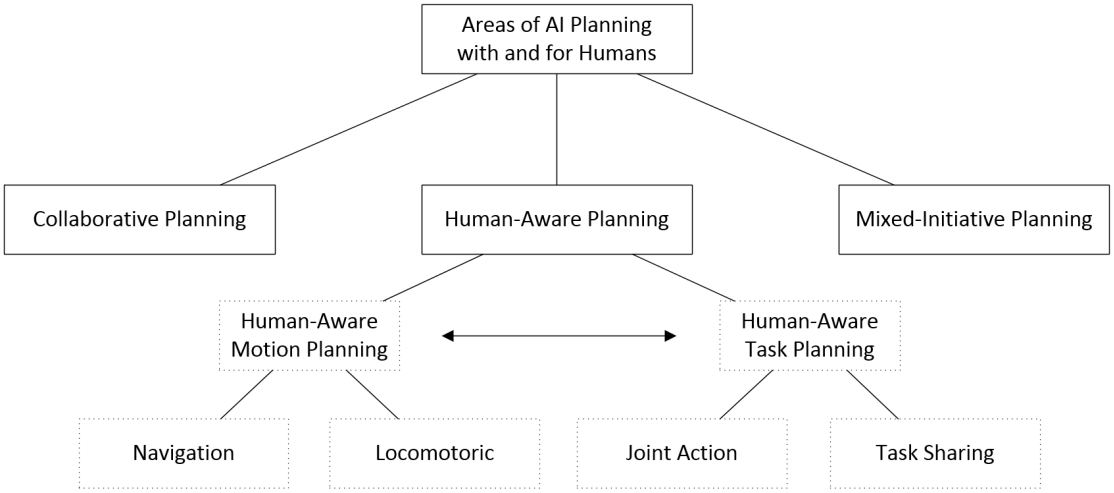


Figure 4.3.: A classification of the different subfields of human-aware planning. The envisioned socially assistive robots must support all subfields.

Internal classification As depicted in Fig. 4.3 one is also able to identify different subfields of human-aware planning. Here, we particularly distinguish motion and task planning. Motion planning includes human-aware navigation and planning for locomotory actions (e.g. fine-motion, grasp-actions). The motion planning is related to robotics, where the robot has to satisfy different requirements when approaching the humans, e.g. comfort, safety zones, and avoiding interruptions. Human-aware task planning includes the actual joint action planning and planning procedures that account for task sharing. As indicated in the figure, these fields depend on each other, especially in the case of companion robots (Kirsch et al., 2010, pp. 225–228). One example that requires all subfields are joint manipulation activities, where a robot interacts with a human to physically manipulate an object. Other use cases, like hand-over tasks in industrial environments, specifically focus on locomotory planning, whereas most (software-based) assistant functionalities belong to task planning.

4.3.2. Origin

The term Human-Aware Planning was coined by Cirillo et al. (2010) in a contribution published 2010. This work presented a conclusion of the dissertation of the main author published in the same year (Cirillo, 2010). Different other work influenced the decision to label this field with an own term. Hoffman and Breazeal (2007a,b) motivate the integration of the possible course of action of humans into the action selection using teamwork fluency. The authors argue that for a fluent action meshing (like observed in experienced human-human teams) the agents need to anticipate the future actions of team members. They describe an anticipatory action-selection mechanism and showed within experiments that the efficiency of the teamwork process indeed increases com-

pared to a reactive approach. Although, the authors neither talked about human-aware planning nor about predictability they present one approach in this area that is related to task planning.

We can also identify other work that concentrates on specific subfields. Especially, research on *human-aware navigation* was early recognised as an own research field with a growing interest during the last decade (Kruse et al., 2013, p. 1727, Fig. 1). The special challenges and requirements that are addressed are not only related to providing a natural motion and the users' comfort and safety, but also to social constraints. An in-depth analysis of this field is provided by Kruse et al. (2013). It is described as *"...intersection between research on human-robot interaction (HRI) and robot motion planning"* (Kruse et al., 2013, p. 1726). One can find a solution under the umbrella of terms like human-aware navigation, human-aware robot motion planning, human-aware manipulation planning and so forth. Kirsch et al. (2010) present a brief introduction to the different subfields by particularly focusing on planning procedures for joint human-robot activities.

In contrast, the problem of task planning in the presence of humans, namely *human-aware task planning*, is a more recent field of study (Cirillo et al., 2012, p. 544). Early work on this topic is presented by Montreuil et al. (2007) introducing the Human-Aware Task Planner (HATP). HATP was extended several times to satisfy the introduced requirements. For example, Clodic et al. (2009) combine HATP with SHARY, which is a supervision system able to recognise human activities. Lallement et al. (2014) make use of so-called social rules (policies) to specify what acceptable behaviour is and introduce reasoning as part of the HATP solution. The most recent work on HATP is presented by Devin et al. (2017), introducing a component that decides when to negotiate about a task and when to take the initiative. Lemaignan et al. (2014) present an overview of the different AI challenges for settings where agents are situated in the same environment as humans and share tasks. Also, Clodic et al. (2014) discuss what is required for human-aware planning in general. Based on observation from cognitive psychology and philosophy they propose a framework comparable to the three-layered architecture of AI planning components (*cf.* Ghallab et al., 2004, p. 8) and identify the integration of joint goals, shared knowledge, shared intentions, and models of the joint partner as future research direction for the AI planning community. Eventually leading to an exchange of findings between the agent community and the AI planning community.

4.3.3. Definition

To the best of my knowledge, there is no definition for the term human-aware planning available in the literature. Available are descriptions of the term by different authors. Cirillo et al. (2010, 2012) and Cirillo (2010) explain HAP as planning procedures for situations where a controllable agent and an uncontrollable agent (human) are situated in the same environment. Here the controllable agent plans its action using a prediction

of the actions of the uncontrollable agent. Hoffman and Breazeal (2007a,b) describe HAP as an adaptive action selection mechanism that makes anticipatory decisions based on the expectations about the future human behaviour. Comparable to this, Tomic et al. (2014) describe HAP as planning that adapts to the presence and needs of humans. One requirement to do so is a human-aware component that recognises human behaviour. Coming from the classical planning domain, Kambhampati and Talamadupula[5] identify Human-in-the-Loop Planning[6] as one of the hot-topics for AI planning.[7] They describe it as planning procedures trying to avoid the human and introduce other terms for collaborative settings. Here it becomes evident that there is no precise definition for the different subfields. However, they also describe that the major challenge is to integrate an interpretation of what the humans will do/are doing into the planning process (e.g. via plan, goal, and intent recognition). In recent work, such as presented by Sreedharan et al. (2017), the author name this the human model. In this work, we refer to such human model as human-behaviour model.

Although, there is no unified definition for HAP we can further approach the term by introducing the human-aware planning problem. Therefore, I take into consideration the commonalities of the different examined work. The first difference to classical planning problems is that HAP explicitly differentiates between the types of actors involved. That means that each HAP problem consists of a set of actions of the controllable agent and a set of actions of the uncontrollable agent. An additional, yet obvious, characteristic of any HAP application setting is partial observability, which is related to the nature of humans. That implies that the search space does not contain states but belief situations, which for itself contain different probable states. Both characteristics together imply a functionality able to plan in the presence of different alternatives of the humans' behaviour. Furthermore, the different authors introduce constraints as an essential component of each HAP component. These constraints can represent, e.g. interaction rules, social rules, norms, or other aspects like safety rules as found in the motion planning. They must be satisfied either in all or in specific sets of situations.

One attempt to define the term human-aware planning would be to concentrate on the application area, which within this work, are joint human-agent activities. This would enable a statement such as:

> *Human-aware planning are planning procedures for joint human-agent activities.*

[5] AAAI2015 Tutorial SP4 – Human-in-the-Loop Planning and Decision Support, tutorial slides can be found here: `http://rakaposhi.eas.asu.edu/HIL-Final-Combined.pdf`, last-visited: 2017-09-22

[6] Human-aware planning as introduced here is a synonym for Human-in-the-Loop Planning as both have in common that they present '...planners that inhabit the land of humans' [What's Hot: ICAPS, – `http://rakaposhi.eas.asu.edu/aaai-whats-hot-pdf.pdf`, last-visited: 2017-09-22].

[7] The same authors lately started a workshop series on Human-Aware Artificial Intelligence (HAAI) including invited talks such as *'Planning Challenges in Human-Aware AI Systems'*. HAAI 2017 can be found here: `http://researcher.watson.ibm.com/researcher/view_group.php?id=7446`, last-visited: 2017-09-22

However, this definition is not tangible as it requires the reader to have the same understanding about joint human-agent activities than me, e.g. the same interpretation of the introduced challenges. Furthermore, it would restrict HAP to cooperative and communicative cooperation (*cf.* Section 4.1). Given, e.g. human-aware navigation planning[8] as one subfield, makes it necessary to include applications that provide an independent cooperation as well. Therefore, I stay with the explanation of *Cirillo et al.* and focus on the fact that human-aware planning components include the agenda of uncontrollable agents into the own action planning. This agenda was also named the human's agenda and may include constraints that represent interaction rules, physical parameters, legal requirements, norms, or other social characteristics. Eventually leading to the following definition:

> *Human-aware planning are planning procedures that include the human's agenda, which may include physical, psychological, legal, normative, or social characteristics, into the action selection process for cooperative activities.*

A formal introduction to the human-aware planning problem, which is solved by HAP components, is presented by (Cirillo, 2010, pp. 17 – 28).

4.4. Conclusion

This chapter aimed to introduce the interested reader to the fields of cooperation particularly focusing on joint human-agent activities and planning for such activities, which is named human-aware planning. Therefore, I started with cooperation in agent-based systems in general and classified the type of cooperation this work aims for as cooperative cooperation, *i.e.* joint activities – all cooperative activities based on a joint goal that is known by the involved agents. I introduced joint activities in the shape of human and agent teamwork afterwards, highlighting the connection between the challenges postulated for the so-called joint human-agent activities and the requirements important in the planning domain. This part of the chapter further presents a methodology to develop such applications and defines the relevant terms. Last, I introduced human-aware planning and classified it within the other research areas in AI Planning approaches with and for humans.

The contributions that mainly influenced this chapter have been presented by Klein et al. (2004, 2005) (about joint activities and challenges for joint activities), by Johnson et al. (2011, 2014b) and Johnson (2014) (about interdependence as the crucial concept of joint activities), by Cirillo et al. (2010, 2012) and Cirillo (2010) (about human-aware planning), and by Kirsch et al. (2009, 2010) (about discussing requirements for planning procedures for joint activities). Based on this literature work, I presented my attempts

[8] One typical task in human-aware motion planning is to avoid injuries. Although both types of actors have this goal independent of each other, it is not necessarily a shared goal.

to define the terms joint human-agent activities and human-aware planning. First and foremost, this is done to clarify the understanding of these terms within the scope of this document. It should help the interested reader to look at the term in the same way than me. Secondly, this is done to motivate a discussion about the necessary elements that make out such definitions and their understanding to bridge the gap between the different involved communities.

In the next chapter, we will proceed with a detailed look at Challenge 3 – Predictability, which is one of the motivational factors to present this work.

5. Problem Analysis

The goal of this dissertation is to use and tailor models about human-behaviour to improve the efficiency of human-aware planning components. These models provide knowledge about the individual preferences, habits, and capabilities of humans' and should be used during the planning process in teamwork settings. Furthermore, the tailoring/shaping of these models should be done during the interaction. Up to this point, this rather vague explanation was used to emphasise the objective of this work. In this chapter, the addressed problem and the associated challenges will be clarified making use of the foundations that have been introduced in the prior chapters. I will start with highlighting the importance of a particular objective, which was already introduced briefly — namely, *Predictability* (sometimes Interpredictability or Mutual Predicability). Roughly spoken, predictability addresses the ability of an agent to build knowledge about the other participants' attitudes, capabilities and course of action, which can be distilled to the problem of learning the policy of the human team members decision-process and the transition probabilities of the team members actuation. The availability of such knowledge is one prerequisite for fluent action meshing in cooperative activities. This challenge is not well-defined, and different involved research areas refer to it using different terms. This chapter explains what is meant by predictability eventually proposing a definition for the challenge. In doing so, I discuss predictability referring to joint human-agent activities and the involved learning problem (*cf.* Section 5.1) and highlight the challenges associated with the action planning (*cf.* Section 5.2). Finally, I will connect the challenge to the envisioned use of human-behaviour models and introduce the different involved parts (*cf.* Section 5.3). Thereby, I link the development of agent's acting as team members to the steps necessary to develop adaptive systems. These steps will be used to structure the work within the next chapters.

5.1. Predictability and Teamwork

In particular, this dissertation addresses the challenge associated with one of the essential characteristics of 'good' teams: The team members' ability to act and to be mutually predictable to each other. This characteristic is required to effectively plan own actions in cooperative settings (*cf.* Bradshaw et al., 2009; Bratman, 1992; Klein et al., 2004; Klein and Wright, 2016; Sycara and Sukthankar, 2006). This explanation introduces two requirements. The first one is the ability itself. The second one is the requirement

to use this ability during the actual planning — making Challenge 3 – Predictability strongly related to Challenge 8 – Collaboration.

I will start with the first one, namely the ability of being mutually predictable, which for itself has two sides. Taking a look at the dictionary helps to approach the term. There it is defined as follows:

> *pre · dict · abil · i · ty* (noun), *pre · dict* (verb): to declare or indicate in advance; *especially*: foretell on the basis of observation, experience, or scientific reason[1]

If we map this definition to teamwork settings, being predictable means to act in a reasonable way that can be observed by the others. These observations/the gathered experiences are then used by the other team members to declare (predict) the next behaviours of one team member in advance. As a requirement, this seems to be vague. To clarify what a reasonable way is we can consider the planning process. In doing so, acting reasonably predictable means acting in a way that one's own actions are predictable enough that the team members can rely the own decision-making process on the observed information (Johnson et al., 2014b, p. 51). This implies that the agent *"...act neither capriciously nor unobservably..."* (Klein et al., 2004, p. 92), enabling other team members to observe and indeed build up the required knowledge to predict the future behaviour. In teamwork settings, we normally assume that this requirement is satisfied due to the commitment of the team members to a joint goal, which includes the agreement of the team members for mutual support and acting mutual responsive (*cf.* Section 4.1).

Sycara and Sukthankar (2006) list this ability as one of three important aspects of human-agent interaction, substantiating the importance of predictability as one essential characteristic of good teams. Psychologists examining human-human teamwork firstly observed it. Here, for example, Bratman (1992) identifies it as a part of one of the three features characterising each cooperative activity (named mutual responsiveness) stating that *"...each participating agent attempts to be responsive to the intentions and actions of the other, knowing that the other is attempting to be similarly responsive. Each seeks to guide the behaviour of the other, knowing that the other seeks to do likewise."* (Bratman, 1992, p. 328). The knowledge that the other team members also agreed to act predictably justifies that this ability is a mutual one. Klein et al. (2004) emphasise that predictability is beneficial for human-agent teams as well. The authors are pointing out that it is one basic requirement for intelligent agents when acting as a team player: *"To be a team player, an intelligent agent—like a human—must be reasonably able to predict others' actions."* (Klein et al., 2004, p. 92). For an intelligent agent that would mean to build up experience during the interaction with other team members and to

[1] The Merriam-Webster Online Dictionary — `http://www.merriam-webster.com/dictionary/predictability` (last-visited: 2017-09-22)

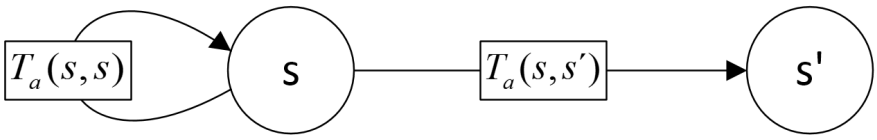


Figure 5.1.: The objective (of predictability) is to gather knowledge about (1) the policy of the team member and (2) the state-transition probability $T_a(s, s')$ for given states s, s' and actions a.

use this experience to foretell the future actions of the others. Taking into account the explanation about what makes out a reinforcement learning problem, which is described as *"...the problem of learning from interaction to achieve a goal."* (Sutton and Barto, 1998, p. 51), we can conclude that this can be partly emphasised using MDPs as formal foundations of RL (*cf.* Section 3.2 for an introduction).

To be more concrete, Fig. 5.1 shows a schematic depiction of a MDP with two states and one action (more accurately, the depiction shows part of a MDP as it misses the reward, which is not of relevance at this point). Consequently, one side of predictability means that an agent builds up knowledge about the state-transition probabilities $T_a(s, s')$ of other team members. Although this knowledge can be used to 'fine-tune' the prediction of the behaviour, the major issue is to determine the actual policy of the team member. The policy determines the behaviour of one linking the perceived state of the environment to the actions that will be selected. We refer to this as the task of anticipating the humans' agenda (*cf.* Hoffman and Breazeal, 2007a,b).

5.2. Predictability and Planning

As I have just described what predictability means in teamwork and what the link to a well-known formal model is, the next step is to approach the question why this is required to plan own actions effectively. Bradshaw et al. (2009) substantiate this requirement stating that: *"...it becomes possible to plan one's own actions (including coordination activities) only when what others will do can be accurately predicted."* (Bradshaw et al., 2009, p. 936-937). To do so, the reader should remind that each joint activity features the three characteristics of acting mutual responsive, committing to the joint activity and committing to support each other towards the joint goal mutually. These characteristics make it necessary to coordinate the team members actions, at least those where one team members' outcome directly depends on the other team members help (e.g. physical cooperation in joint-manipulation tasks). Indeed, the coordination must also take place prior performing other actions, for example, to avoid repetitions. Taking this into account, it was shown that teams are working more efficiently when coordinating actions while avoiding communication overheads (MacMillan et al., 2004, Chapter Sum-

mary). Here, the individuals anticipate the behaviour of the others before planning own actions and only explicitly communicate when it is necessary. From this observation it was followed that communication acts are to some extent costly and that there is indeed a trade-off for the team-members, deciding whether decisions should be coordinated via communications or the behaviour prediction. This observation substantiates what is claimed by Challenge 10 – Cost Control and also support the claims postulated within Challenge 5 – Revealing Status and Intentions, Challenge 6 – Interpreting Signals, and Challenge 9 – Attention Management.

The clarification and the explanations of the different authors eventually enable me to define the requirement of acting mutually predictable in human-agent teamwork settings:

> *In human-agent teamwork, the requirement of predictability is that an agent can only plan its actions—which includes coordination activities—effectively if it is assessable what the other collaborators, including the human, will do.*

Although this definition deals with both sides of the medal, planning processes for collaborative setting normally do not approach the characteristic that the individuals should act observable but consider it as a satisfied precondition. For further information on how to approach this characteristic the interested reader is referred to, e.g. Bradshaw et al. (2008, 2009) presenting a policy-based approach established within the KAoS project.

Having worked out the requirement itself enables to finally formulate it as a problem that can be approached, bringing together predictability and human-aware planning. As introduced, HAP are planning procedures that account for joint human-agent activities and address the issue of generating either collaborative plans involving humans or plans for agents acting in environments involved by humans and agents. Within HAP I define predictability as follows:

> *The problem of predictability in human-aware planning is twofold: (1) anticipating the human team members' most likely course of action that will be executed in order to reach the joint goal of the collaboration and (2) integrating this prediction into the planning process for the collaboration.*

In other words, one can say that the problem is to learn the most-likely set of agendas of a human team member. In a further step, this information must be integrated into the planning process for the collaboration. This can be done in terms of a heuristic as, for example, proposed by Doce et al. (2010); or in terms of a restriction of possible goals and acceptable actions/sequences of action as, for example, proposed by Rizzo et al. (1999); or in terms of constraints on the planning process as, for example, described by Cirillo et al. (2010).

Fig. 5.2 shows a depiction that summarises the so far provided argumentation. In fact, these arguments apply to any team mixture (human-human, human-agent, agent-agent). Concerning joint human-agent activities, we need to take into account the characteristics

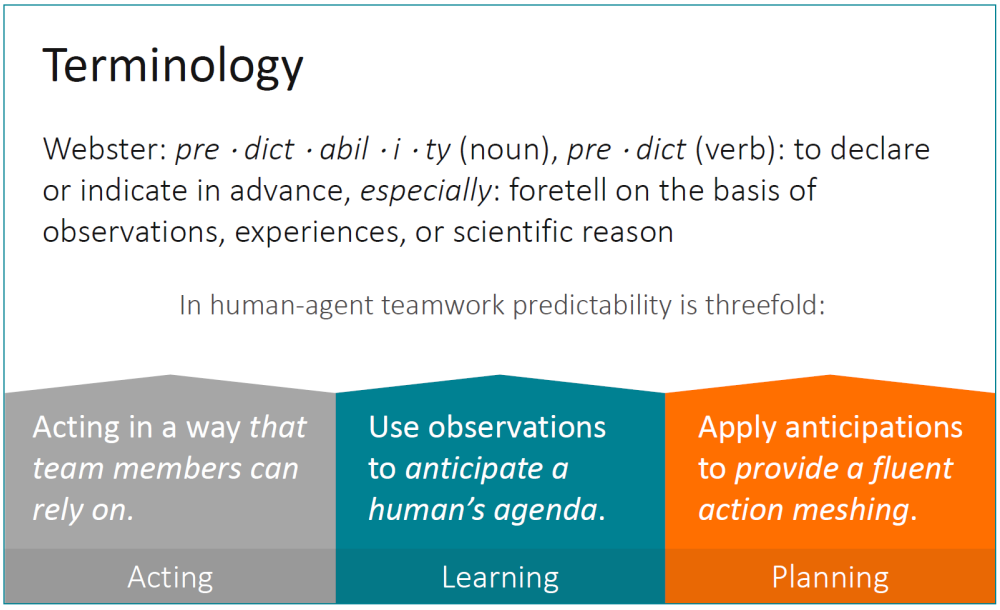


Figure 5.2.: Summarising the problem of predictability in human-agent teamwork, which for itself provides challenges in acting, learning, and planning procedures.

of human behaviours. Hence, in the following, I will introduce human-behavioural models as a prerequisite for predictability in human-agent teamwork.

5.3. Predictability and Human-Behavioural Models

As emphasised in the research statement my approach to the problem of predictability is based on the idea of using the information provided by models of human-behaviour in the planning process. These models contain information that is in some way related to the behaviour of humans, for instance, information about abilities, habits, intentions, desires, social rules, norms, disabilities, personality, and effects of emotions. Kirsch et al. (2010) argue that these models are either *"...derived from social and psychological studies and can be provided by hand-coded decision rules or constraints on plans"* or represent *"...individual preferences and abilities of a user that can change over time and should be acquired and updated constantly ... in the interaction with the human."* (Kirsch et al., 2010, p. 3).

Taking this argumentation into account, we can identify the recommendation to use the information as a heuristic. Furthermore, we can identify two parts that make out a human-behaviour model in the quote.[2] The first one being the *user model*, containing the relevant information about the individual, e.g. personal characteristics like age and

[2]Plumbaum (2015) provides a useful introduction into the domain in Chapter 2 and particularly provides a taxonomy of the involved information dimensions starting from page 18 – 24.

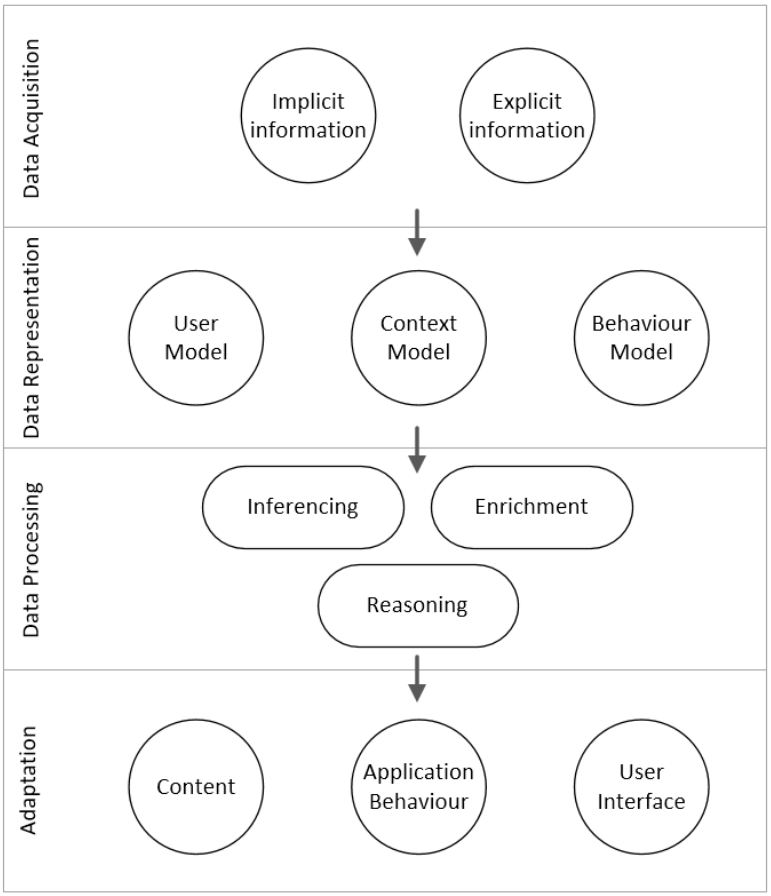


Figure 5.3.: The different task of an adaptive system reveal the components that are necessary to enable predictability. This illustration is adapted from published work (Plumbaum, 2015, p. 3).

gender, mental and physical characteristics like physical limitations or cognitive load, or individual traits like personality traits. The second one being the *behaviour model*, containing insights into behavioural preferences in an abstract manner, e.g. correlations between action-preferences and individual traits, effects of emotions on decision-making, or influences of a mental and physical state on the abilities of a human. The third part of the information is only implicitly mentioned in the quote, though, other work (*e.g.* Jameson, 2001) explicitly refer to it as the *context model*. The context model contains the information that influences the humans' or the systems' behaviour from the observer's point of view (one example could be the used interaction modality). As there is no single accepted definition for the term context, I refer to context in terms of the definition provided by Kokinov (1995): *"Context is the set of all entities that influence human (or system's) cognitive behavior on a particular occasion."* (Kokinov, 1995, p. 200).

Fig 5.3 combines all three parts in the data representation layer. I refer to this layer as the human-behavioural model. As illustrated, the representation of the human-behavioural model is one out of three necessary steps to develop an adaptive system (*cf.* Plumbaum, 2015, p. 2–4; Torre, 2009, p. 438–442). Abstractly, the steps read as follows:

- the data acquisition task — establishing modalities to collect information about users, e.g. through sensors, questionnaires, or usage of social networks;
- the data representation task — modelling the information structure that holds the users' information; and
- the data processing task – processing the information derived from the data acquisition.

To apply the derived model reasoning requires algorithms that bring together the user, the behaviour, and the context model with the application goals. Whereas generic approaches for modelling and representing user and behaviour data exist (e.g. GUMO – General User Model Ontology, *cf.* Heckmann, 2005, pp. 85–99 or UBO – User Behaviour Ontology, *cf.* Plumbaum, 2015, pp. 50–62), the reasoning algorithms are domain dependent and streamlined for the applications purpose. One major area applying these components are information retrieval systems and recommender systems (Plumbaum, 2015, p. 4). Such systems predict the user's interest into a specific type of information using search queries or try to predict the potential interest of a user in items.

Although Fig. 5.3 shows the different tasks and identifies the components that are required to adapt the behaviour of applications, it does not demonstrate the link to predictability. For our objective, the challenge is to link user model, behaviour model, context model, and application goals in a way that the inferred information can be used as a prediction of the human's agenda. Such agenda is used in a second step to make informed decisions in the joint activity. Thus, the first three layers mark the necessary steps to satisfy the first part of the problem of predictability in human-aware planning, whereas the adaptation layer refers to the second part of the problem definition.

5.4. Conclusion and Final Remarks

In this chapter, I have analysed predictability as one of the challenges associated with joint human-agent activities. In particular, predictability was identified as an important property distinguishing good and effective from flawy teamwork (e.g. Sycara and Sukthankar, 2006, pp. 2–3). The challenge was discussed in detail, substantiating the term while connecting it to different research areas such as collaborative activities, AI learning and AI planning.

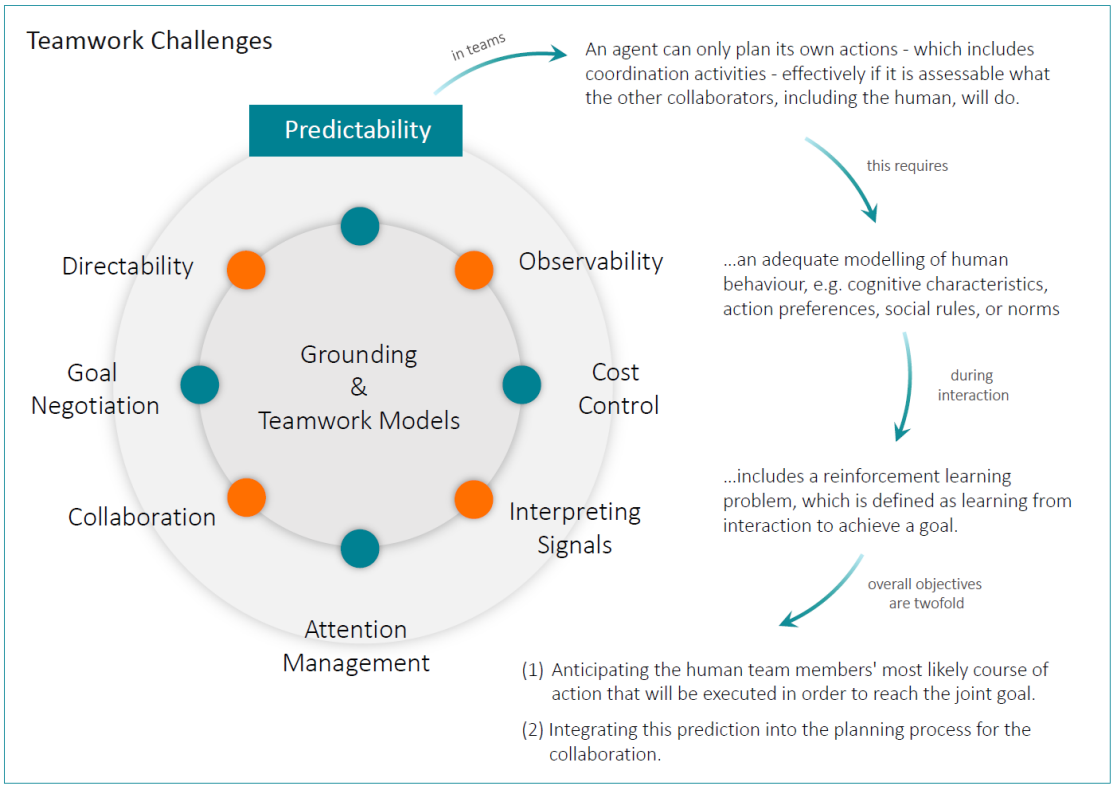


Figure 5.4.: A summary that highlights predictability as one of the challenges in joint human-agent activities.

I want to make the reader aware of the fact that predictability involves more than the knowledge about the policy and state-transition probability of other team members. Indeed, I used this formalism to explain the issues without providing details about important problems such as how to reason, recognise, or even represent, e.g. intentions or joint goals. Having a sufficient common ground and activity recognition, that enables the artificial agent to infer what needs to be done to achieve goals, is relevant for predictability as well. The provided further reading points be given as a starting point for the interested reader to take a deeper look into the individual fields. Fig. 5.4 summarises the statements that have been provided within this chapter and emphasises that predictability is only one of the challenges that are associated with teamwork.

Referring to the work on coactive systems (*cf.* Fig. 4.2) and Challenge 2 – Adequate Models we can further substantiate the requirement to have an explicit representation of the team member and to enforce an application lifecycle that supports the continuous interaction with the human. This is necessary to reason about the systems own state, the state of the environment, and the course of action of the human. Cheng et al. (2009) describe such a lifecycle from the perspective of self-adaptive systems as control loop differentiating between four phases: collecting information, analysing the informa-

tion, deciding on the next actions, and acting. In AI and particularly in agent-based systems, we know such control loops as the *sense-plan-act* or the *perceive-reason-act* cycle (*cf.* Russell and Norvig, 2002, pp. 32–55); each is introducing three steps for a continuous interaction within an environment. Researchers at IBM introduce a control loop named *MAPE-K* (Kephart and Chess, 2003) (monitor-analyse-plan-execute over a knowledge base), which received strong interest in the academic and industrial community (Kephart, 2011). The main contribution of the MAPE-K is the architectural blueprint that explicitly requires the control loop and the feedback to be part of the architecture (IBM Corporation, 2005). Joining the MAPE-K loop with control loops in agent-based systems requires to identify the agent as its autonomic manager, which is natural to the concept of agents (*cf.* Section 3.3). The managed element stays the same, in that the agent has to manage its behaviour. Even the initial work origins in the area of autonomic computing for self-adaptive/self-managing systems, we adopt MAPE-K to approach predictability as it is promising to solve the above-described problems. That is because, MAPE-K (1) specifically highlights that the agent is responsible to continually adapt its behaviour, (2) it introduces the MAPE phases, which I identify as necessary to solve human-aware planning problems, *i.e.* monitoring the humans behaviour in a joint activity, analysing the behaviour to predict the course of action, planning the own course of action taken into account the anticipated behaviour of the human, and executing own actions, and (3) it introduces the knowledge base that is necessary to learn from observations.

In the next part, I will isolate aspects of the problem by focussing on modelling personality and its influences within agents (*cf.* Chapter 7), learning personality within the interaction with humans (*cf.* Chapter 8), and formalising personality within the decision-making process of agents (*cf.* Chapter 9).

Part III.

Human-Personality Models – Modelling, Learning, and Reasoning

6. Preface

The objective of this part is to approach isolated aspects of the problems we identified. To structure this, I make use of the introduced elements of an adaptive system (*cf.* Section 5.3). The next chapters will contribute to the *Data Acquisition*, the *Data Representation*, and the *Data Processing* layer as highlighted in Fig. 6.1.

In Chapter 7 – Effects of Personality, I will focus on modelling personality within the decision-making process of agents. This is done, to examine the influence of personality on the behaviour of agents: (1) analysing the related work in agent-based research, and (2) discussing an agent-model able to simulate personality related behaviour concerning the FFM. Therefore, I will use personality information to combine it with behavioural information derived from psychological studies. I will not explicitly consider contextual information, as the personality information is part of the user model, though, I will refer to the context while interpreting the personality related behaviours.

In the Data Acquisition layer of adaptive systems, we differentiate between *explicit* and *implicit* information. Explicit information is prior knowledge, modelled by the system designer or added by the user, e.g. using onboarding or setup steps of an application. In the just-described chapter, we model the personality and behaviour information. Implicit information, on the other hand, is derived from observations of the behaviour of humans, e.g. which decisions the users make, how long a user action takes, and so forth. Thus, one main difference between these types is the way they are acquired; others are how they are modelled and what is expressed, though, both types do not necessarily present a disjoint set of information.

After modelling the information explicitly, I focus on deriving personality information in Chapter 8 – Learning Personality; approaching the task of learning about the personality of a user during the interaction with this user. Thus, focusing on deriving implicit information to establish personality characteristics in the user model. For this, I present two experiments during which agents observe the behaviour of humans while playing a scientific game against each other. Within the experiments, I apply two different agent-models that adapt their behaviours based on the estimated personality of the human opponents. Explicitly derived personality information is used to validate/falsify the personality estimates learned by the agent.

Finally, in Chapter 9 – Reasoning about Personality, I focus on the *Data Processing* layer integrating personality into BDI logics to enable reasoning about personality. The motivation to present such a formalisation is the need to provide clear semantics while talking about personality in agents. The objective of the formalisation is twofold: (1)

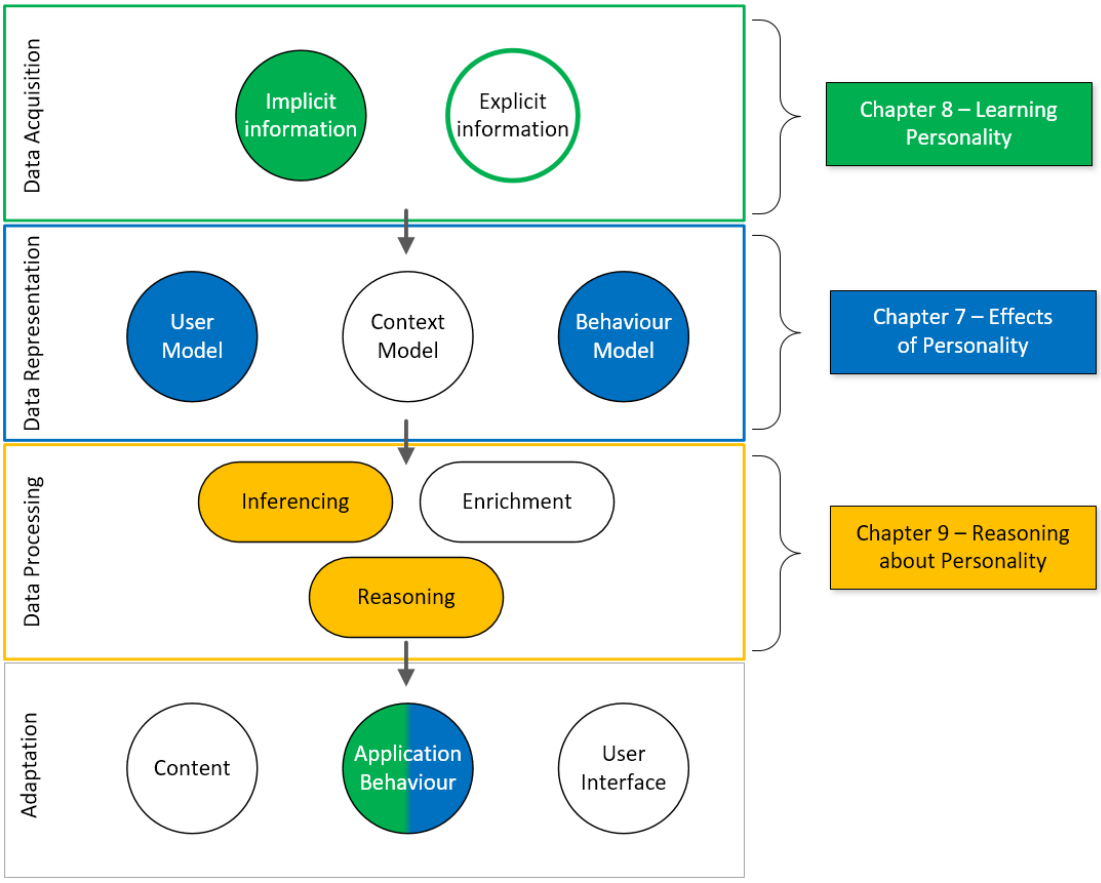


Figure 6.1.: The elements of an adaptive system that are discussed within the individual chapters of this part.

to enable reasoning about the influence of a personality on the behaviour of an agent (reasoning), and (2) to enable reasoning about characteristics of a personality using observations of behaviour (inferencing). I exclude the enrichment of information as there are several techniques in the user and data modelling community that can be used by agent researchers.

The final layer of adaptive systems handles the adaptation itself. This layer is touched upon in the first two chapters, concentrating on the application's behaviour. The adaptation of content or user interface is excluded as this thesis is about behavioural aspects of agents.

7. Modelling State and Effect of Personality

Predictability includes the ability to anticipate a team members' most likely course of action to reach the joint goal of the collaboration (Ahrndt et al., 2016a), which requires to simulate and reason about an agents' course of action. In the case of natural agents, *i.e.* human team members, this would ask for simulating and reasoning about human behaviour (Kirsch et al., 2010, p. 225). Yet, there is no accurate simulation of humans available.

Given this challenge, my goal is to investigate the concept of personality as one of the central elements determining the behaviour of natural agents. In fact, human-agent teamwork is only one area of research that can benefit from modelling and simulating cognitive characteristics in general and personality influences in particular. Other areas include the decision-making of resource-bounded agents, interface agents, crowd- and traffic-simulations, negotiation and social choice, or virtual humans.

Due to the manifold research and application areas, I will start my work on agents with personality by analysing the state-of-the-art in conceptualising personality as an affective phenomenon in agent-based systems. Therefore, I will explore available work first and discuss the identified status afterwards in Section 7.1. The objective is to provide an overview and give an answer to the question: *What is the state-of-the-art for integrating personality, in particular, the FFM, as an affective phenomenon in agent-based research?* In particular, I am interested in the use of personality theories in agent-based research and whether or not there exist work showing that personality influences the whole decision-making process of an agent independent of a specific application area or use-case.

Indeed, this part unveils that such work exist, but that it frequently applies the MBTI as theoretical vehicle or uses interpretations of personalities that are not based on a solid theoretical basis. Work that provides more generalised investigations on personality in agents is rare and the most advanced applied the MBTI or a subset of the FFM dimensions, justifying the reason to present an own complete model. Following this argumentation, the next part of the chapter (*cf.* Section 7.2) answers the question: *How can we derive an agent-model representing the effects of personality with respect to the FFM?* To achieve that, I introduce an agent-model that extends the BDI lifecycle by involving the effects of personalities according to the FFM. Referring to Fig. 7.1, which highlights the parts of an adaptive system that I will work on within this chapter,

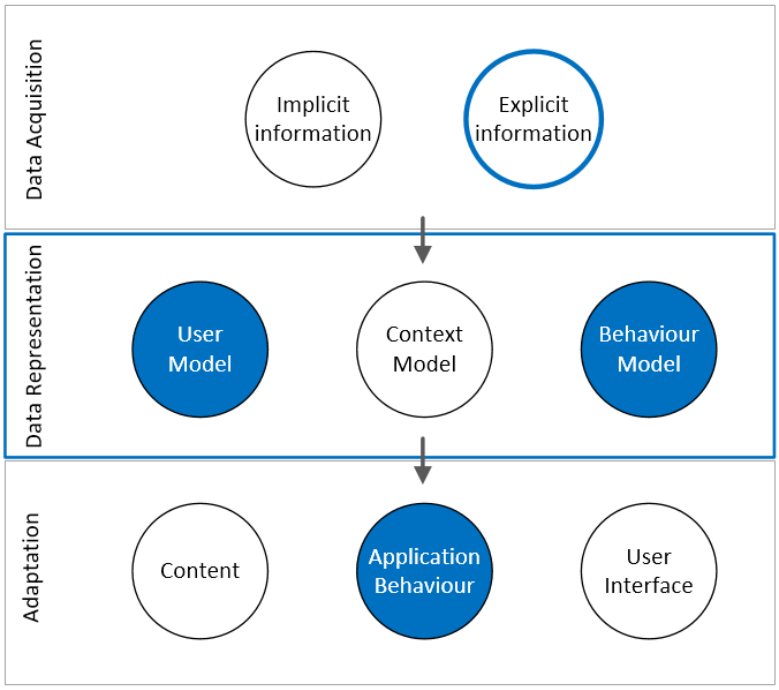


Figure 7.1.: Parts of an adaptive system that are discussed within this chapter.

the BDI paradigm represents the behaviour model whereas the dimensions of the FFM represent the user model. Afterwards, the implementation of this model for a multi-agent simulation environment is presented (*cf.* Section 7.3). Based on this implementation the effects are analysed for this specific use-case (*cf.* Section 7.4). Finally, the chapter is concluded.

7.1. Related Work

In the following, I will explore the use of personality theories in agent-based systems. Therefore, I will structure the section using the application areas and provide representative work for each identified area. The objective is to identify common approaches conceptualising personality as an affective phenomenon in agent-based systems, specifically focusing on the effects that are produced by personality traits.

Agent-Based Simulation Research on agent-based simulations frequently uses formal models of human personality for the implementation of (microscopic) traffic simulation frameworks and the agent-based simulation/visualisation of groups of people.

Canuto et al. (2005, 2006) present SimOrg, a multi-agent simulation of working environments, which aims to represent individual behaviours using aspects of personality. The authors apply personality stereotypes to receive a more natural simulation and to

analyse the influence of different personality aspects on a teams performance. The applied personality theory is MIPS: Millon Index of Personality Styles, which is typically used to assess personality disorders. MIPS introduces three key dimensions, which are named styles: (1) Motivating Style, used to assess the emotional reactions to stimuli, (2) Thinking Styles, used to examine cognitive characteristics, and (3) Behaving Styles, used to distinguish social behaviour within the interaction with others. The agent architecture incorporates the personality aspects as utility function used to assigns weights to competing actions, *i.e.* different actions that can be used to reach the same goal, and defines a set of personality profiles that are named stereotypes. During the simulation, those actions are selected that match with the stereotype of the agent.

Durupinar et al. (2008, 2011) shows how the introduction of different personalities into single agents influences the behaviour of a crowd. For this simulation, the authors applied the OCEAN model, which is another name for the FFM. The work presents an extension of the HiDAC (High-Density Autonomous Crowds) system that introduces a personality model to each simulated agent. Personality is integrated using a five-dimensional vector of weighted personality factors. The weights and dimensions are used to build different types of agents, e.g. leader agents that tend to have more confidence in themselves, which is determined by a preference to be extroverted and emotional stable. An individual's overall behaviour is derived from combining the different types. The mapping from personality traits to behaviour types is substantiated by using trait descriptive verbs. Eventually, the work presents an evaluation that setups different scenarios to validate the effects of each personality dimension on its own.

Guy et al. (2011) present a work comparable to the one of Durupinar et al. using the Eysenck 3-factor model, which is a personality trait theory introducing three factors and one of the predecessors of the FFM. The authors describe an extensive user study that applied factor analysis to identify correlations between personality types and the individual behaviour of the simulated people. The effects of different personalities are implemented by adjusting the following parameters: preferred movement speed, effective radius (distance between agents), maximum number of neighbours affecting an agent, maximum distance to other agents affecting one's own behaviour, and planning horizon, which controls the steps an agent plans ahead (foresighted vs. myopic). These parameters are commonly used in crowd simulations, making the work applicable to other frameworks of such kind, but hardly useful for other application domains.

Herpers et al. (2015) introduce the AVeSi – Agentenbasierte Verkehrssimulation mit psychologischen Persönlichkeitsprofilen (engl.: Agent-based traffic simulation with psychological personality profiles) project. The authors integrated personalities based on the FFM using a parameter named civility factor (ger.: Höflichkeitswert) that is adjusted to provide three different kinds of behaviours: Undercontrolled, Overcontrolled, and Resilient. The civility factor is used to control the extend to which an agent takes into consideration how its actions affect other traffic participants. The same factor is

used to simulate emotions in a way that it is dynamically altered during the simulations runtime.

The approach of Herpers et al. (2015) is comparable to the one presented by Guy et al. (2011) in that it simulates personalities using domain specific parameters. It differs from the work in that it restricts itself to only one parameter. However, reading work on cognitive aspects in both areas reveals comparable approaches, where the authors used elements of the simulation environment to induce personality, emotion, or mood into the agents. In fact, it was already argued that the agent-community frequently *"...focus on agent-models custom-tailored to the task at hand"* (Balke and Gilbert, 2014, line 6.1) instead of integrating cognitive aspects *"...as an integral part of an agents' architecture"* (Balke and Gilbert, 2014, line 6.1).

For a more comprehensive overview of work that incorporates cognitive aspects into agent-based simulations, the interested reader is referred to, for example, the work of Xu and Deng (2014) about crowd simulations; and the work of Lützenberger (2014) providing further reading points in agent-based traffic simulations.

Human-Agent Interaction Other work that includes personality in agents can be found in human-machine interaction, in particular, conversational agents/virtual humans and life-like characters. In an early work, Dryer (1999) empirically investigates the role of the personality of individuals in the interaction with agents. The authors eventually present guidelines for integrating personality into agents for designers that are independent of a user personality theory but focus on the design of the human-agent interaction. One of them supports the claim of that agents should mirror the personality of users (Tapus et al., 2007, pp. 37, Section III.B.): *"In general, an agent with a personality that is similar to that of a user is liked better than an agent with a personality dissimilar to a user."* (Dryer, 1999, p. 292, no. 7).

Allbeck and Badler (2002) introduce what they named Parametrized Action Representation as a foundation for planning and reasoning of agents about own actions or actions of others. They describe a system where personality and emotions are used to alter the movement of an embodied agent, thus, providing more realistic behaviour. Agents, actions, and objects are described using a domain specific language, whereas personality is represented using an integer value for each of the five dimensions of the FFM. To perform an action the agent has to satisfy the action conditions including required personality facets.

André et al. (2000) outlines three projects that apply two dimensions of FFM (extraversion and agreeableness) for life-like characters. The effects are interpreted in a rule-based or scripted manner. The user can, for instance, control the loudness of the characters by adjusting the personality traits. Although the authors describe two agent architectures, it is unclear to which extent these architectures are realised.

Concerning virtual characters focusing on (none)-verbal conversational agents, Saberi (2016) recently presented an extensive literature analysis on computational models for expressing personality. Those models reach from weighted parameters in terms of multi-attributed utility functions to Bayesian Belief Networks to neural network models. Eventually, the author concludes that there is the necessity to improve available work w.r.t. the perceived immersion as the created bodily and facial behaviour is not satisfying within the interaction with humans.

Agent-Based Modelling Another branch of research focuses on modelling and examining the effects of personalities on interactions between agents and their environments.

Campos et al. (2009) present a contribution employing the MBTI model, which is restricted to two of its dichotomies. It is integrated into the reasoning process of a BDI agent, and the work proves that different personality characteristics lead to varieties in the decision-making process. This is done in a simulation specifically designed for the paper's use-case; namely, a fire-fighter agent that has to decide whether to setup a safety-net to enable persons to jump on it or to enter the building and rescue persons this way. This decision depends on the results of the reasoning process, which is influenced by the dichotomies according to the definitions that are provided. For example, it is assumed that a thinking agent always thinks that the other agents thinks the same way, whereas a feeling agent is not applying this assumption.

In an early work, Castelfranchi et al. (1998) present a framework to investigate the effects of personalities on social interactions between agents, such as delegation and help. The agents apply opponent modelling regarding personality traits to motivate interactions. However, the work discusses personality traits as an abstract concept without relation to psychological theories.

Both contributions serve as examples for approaches that apply personality as a way for agents to derive knowledge about others and to make decisions during the cooperation. The next contribution is different from the former in that it concentrates on the task of modelling personalities in general and not on the outcome which can be achieved.

Salvit and Sklar (2011, 2012) describe a testbed, which they used to demonstrate a variety of behavioural patterns while varying the personality models. To do so, MBTI is integrated into a sense-plan-act structure, and the behaviour of each MBTI type is analysed in a simulation environment called the 'Termite World'. The authors constitute that the work concentrates on the effects of personalities on the interpretation of inputs, thus neglecting such personality traits that effect the interaction with other agents. The results underline the hypothesis of the paper that the different personality types act in quite different ways and that the observed behaviours are characteristic for the applied personalities. The presented evaluation tries to establish a baseline of single agent types, before investigating the effects of personality on team performances. One consequence is *"...that some agent personality types are better suited to particular tasks—the same*

observation that psychologists make about humans." (Salvit and Sklar, 2012, p. 147).

Preference Elicitation Compared to the above-introduced contributions, work in the domain of preference elicitation (sometimes named utility elicitation) introduce more general approaches regarding the type of information (*cf.* Chen and Pu, 2004, pp. 1-7 for an introduction and overview about elicitation methods). Preference elicitation is the process of determining a user's preferences regarding some objective, e.g. to sort a list of search results or to propose/make a good decision on behalf of a user (Chen and Pu, 2004, pp. 1-2). The goal is to infer knowledge about the preferences of a user and, given such knowledge, to derive utility functions that are used by the agent. One challenge is to find an effective elicitation strategy. In adaptive systems, we identify this as the data acquisition task (*cf.* Pommeranz et al., 2012, pp. 359–365 for background information on the design of elicitation methods).

In agent-based research, work on preference elicitation exists, e.g., in agent-based negotiation, decision making, and decision-support systems. However, publications that apply or make use of affective phenomena are rare.[1] One example is presented by Santos et al. (2011). The authors describe an argument-based group-decision support system that uses affective phenomena (personality, mood, emotion) to improve its negotiation process. Group-decision support systems are used to coordinate team efforts to solve joint tasks. The authors present an architecture that utilises the identified personality of the opponents to select the best arguments to be used in a negotiation. The personality of the users is represented according to the OCEAN model and determined using a self-reported inventory. The authors sort the set of arguments according to the categories appeal to self-interest, appeal to prevailing practice, appeal to counterexample, appeal to past reward, the promise of future reward, and threat (Santos et al., 2011, p. 62). In prior work, Santos et al. (2010, pp. 296–301) describe how they assign different personalities to the categories of arguments.

Discussion To conclude, the literature overview showed much research that includes personality in agent-based simulations, human-agent interaction, and agent-based modelling. I can summarise that the majority of contributions can be found in the area of agent-based simulation, in particular, agent-based traffic and crowd simulations. Only a couple of the considered contributions substantiate the decision to apply either the FFM, the MBTI, or a simplified representation of personality. This is problematic given the objective of adequately transferring psychological findings into computer-processable models, which is postulated by most of the contributions. One effect of the missing

[1] A possible explanation is given using the example of collaborative filtering systems approaching the *find good items* task (Cacheda et al., 2011, p. 2:1). To elicit the preferences of the user one models the observable properties of the items (in collaborative filtering the rated items and the ratings) and tries to predict the value of these properties for a user (Cacheda et al., 2011, pp. 2:2–2:13). Applying affective phenomena to this task requires to elicit the contribution of these phenomena to the observable actions and make use of it for the value prediction.

grounding in psychological work is a mixture of the effects of different cognitive characteristics, e.g., a mixture of personality and emotional influences.

Most of the contributions provide custom-tailored agent-models, where personality influences are streamlined for the environment the agents inhabit. A few contributions approach the problem more generally. In fact, influences of personality on the decision-making process of agents are discussed in the early work of Castelfranchi et al. (1998). This work provides first insights into conceptualising personality in general, though, these insights are given as abstract concept without a relation to psychological findings. The work of Salvit and Sklar (2011, 2012) discusses and evaluates these influences with respect to the MBTI demonstrating the variety of behaviours and grounding the results within the MBTI. Given our objective, it is the most advanced work that I identified.

The implication I draw from the discussion is that the question of how to incorporate personality into agents was not satisfiable approached by agent researchers. As described in Section 3.4 a multitude of human-personality theories exist, whereby type and trait theories focus on explaining the individual differences by identifying measurable dimensions/dichotomies. Discussing the major type and trait theory, I concluded that the FFM is the theoretical vehicle used to ground my work in psychology due to its origin, the used continuous scales, and the observed reliability. Furthermore, as psychologist tend to accept the FFM as personality framework and at the same time tend to refuse the MBTI, the motivation to present an own agent-model within the remainder of this chapter is to confirm the work of Salvit and Sklar (2011, 2012) with respect to the FFM.

7.2. Modulating BDI Agents with Personality

To integrate the personality of humans, we embed the FFM theory into the BDI model of agency. As introduced in Section 3.3.2, BDI agents separate the current execution of a plan from the activity of selecting a plan using the three mental concepts belief, desire and intention. The lifecycle of a BDI agent comprises four phases, namely the *Belief Revision*, the *Option Generation*, the *Filter Process*, and the *Actuation*. In our model, the phases of the BDI cycle are influenced by the characteristics of a personality in different ways. For instance, the trait conscientiousness strongly influences the goal-driven behaviour of an agent, whereas the trait extraversion influences the agent's preference to interact with others. Table 7.1 lists the influences of the different characteristics of FFM on the different phases of the BDI lifecycle. These influences address the intensity by which one personality trait influences a phase and thus (only) highlights the traits that are most influential. Indeed, this classification is discussable as it reflects an own interpretation of the FFM traits in comparison with the BDI phases. To justify this interpretation, I make use of the bag of research that substantiates the FFM. Part of personality trait theories is to explain and predict behaviour preferences using the personality traits. This knowledge can be applied to BDI agents. For instance, contributions

Table 7.1.: Our interpretation of the influences of the FFM traits on the lifecycle stages of a BDI agents. The list only highlights the traits that are most influential in each phase.

	O	**C**	**E**	**A**	**N**
Belief Revision	×			×	
Option Generation		×		×	×
Filter Process	×	×	×	×	×
Actuation	×	×	×		

that investigate the relationship between personalities and behaviour types (e.g. Du and Huhns, 2013; von der Pütten et al., 2010) describe effects on interaction preferences, trust, and even movement. Furthermore, I learned about the influences by experiments that provide findings for the relation between personalities and specific stages of the decision cycle. For instance, work that investigates the relations between personality and preferred coping strategies (e.g. Carver and Connor-Smith, 2010; Connor-Smith and Flachsbart, 2007) or work that examines the effects of varying personalities on the information processing capability/approach (e.g. Baumert and Schmitt, 2012). Other contributions describe the influence of personality on judgement and decision-making (e.g. Byrne et al., 2015; Zelenski, 2007).

In the following, I will first discuss personality influences w.r.t. the FFM and the provided references within the naive BDI lifecycle (*cf.* Section 7.2.1). This is one variant of a BDI agent following a blind-commitment strategy and being overcommitted to both, the ends (*i.e.*, the selected intentions respectively the world state the agents wants to achieve) and the means (*i.e.*, the generated plan to achieve the intended world state). Afterwards, I will extend this discussion to an advanced version of the lifecycle that balances between means and end (*cf.* Section 7.2.2.)

7.2.1. Personality and the naive BDI lifecycle

To explain the model, we will use the following syntax introduced by Wooldridge (2000, pp. 69–90) for the '$\mathcal{L}$ogic $\mathcal{O}$f $\mathcal{R}$ational $\mathcal{A}$gents' ($\mathcal{LORA}$). We represent a BDI cycle as a sequence of states of an agent. Therefore let each state of an agent be a set of variables:

- $\rho : Percepts$ is the information that the agent perceives/receives in its environment;
- $B : \mathcal{P}(Bel)$ is the set of beliefs, *i.e.* the current assumptions about the state of the environment;
- $D : \mathcal{P}(Des)$ is the set of desires, *i.e.* the set of intended goals the agent wants to fulfil;
- $I : \mathcal{P}(Int)$ is the set of intentions, *i.e.* the set of desires the agent is committed to fulfil;

Algorithm 7.1 A BDI cycle that incorporates personality into the decision-making process and presents a blindly-committed behaviour that is overcommitted to the means and ends.

Input: B_{init}, I_{init}, P; **Output:** -

```
1:  B ← B_init, I ← I_init
2:  while true do
3:     ρ ← percept(Env, Msg)
4:     B ← beliefRevision(B, ρ, P)
5:     D ← options(B, I, P)
6:     I ← filter(B, D, I, P)
7:     π ← plan(B, I, P)
8:     while not empty(π) do
9:        α ← hd(π)
10:       execute(α, P)
11:       π ← tail(π)
12:    end while
13: end while
```

- $\pi : Ac^*$ is the current sequence of actions taken from the set of plans over some set of actions Ac this agent has chosen, *i.e.* the current plan; and
- $\alpha : Ac$ is the action that is executed.

I further refer to the personality of an agent as $P : Per$, which is the collection of personality traits the agent has, *i.e.* the actual characteristics for this agent according to the dimensions of FFM. We assume that the personality does not change during the lifecycle of an agent. That is based on the finding that we as humans have a relatively stable personality over our lifespan as adults (e.g. Caspi et al., 2005; McCrae and Costa Jr., 2006; Wilks, 2009).[2]

Algorithm 7.1 shows an adapted BDI lifecycle that involves personality as influence during the different stages. All personality traits are considered during the process. The cycle starts with the perception of information. During this stage the agent receives new information from the environment (Env) using its sensors, which also comprises messages (Msg) from other agents (communication acts). The perception is not affected by the personality, as humans are not able to restrict their perception during the cognition. This is a deliberate process taking place in the next step of the cycle. Formally, the signature of the perception function *percept* is defined as:

$$percept : Env \times Msg \rightarrow Percepts.$$

The next step of the BDI lifecycle is the *Belief Revision*. That means that given per-

[2]Note, that this implies a changing personality for the whole lifespan. That means if one wants to develop an agent-based simulation of a virtual character covering birth, growing-up, being an adult and dying the personality should be adapted during the lifetime.

ceptions (ρ) are interpreted with respect to the current personality (P) to update the actual beliefs (B). The belief revision function *beliefRevision* is defined as:

$$beliefRevision : \wp(Bel) \times Percepts \times Per \rightarrow \wp(Bel).$$

After this step the set of beliefs can contain information about the environment, the state of the agent itself (e.g., energy level, injuries like sensory malfunctions) and facts that were received via communication. In our model the **O** and **A** characteristics influence this phase most frequently, as they influence the interpretation what the new measurement means for the agent and how trustful the agent is when receiving information from others. One essential reason to distinguish between perceptions/beliefs derived from the environment and perceptions/beliefs derived from other agents is the characteristic of the personality trait agreeableness, which indicates the preference to trust others.[3] That is to say, that a person with higher/lower **A** value is more likely to trust/distrust information received by other agents. Persons with higher/lower **O** values are more/less likely to adopt contrary/new information, due to the linkage of the **O** trait showing behaviour preferences for liberal (nonconformist) and conservative attitudes (McCrae and John, 1992, pp. 195–198).

The next step is the *Option Generation*, where the agent generates its desires (D) taking into account the updated beliefs, the currently selected intentions (I) and the personality. The option generation is mainly influenced by the **C**, **A** and **N** characteristics, as these traits indicate the preferences to follow picked goals, the tendency to act selfish or altruistic, and the reaction of the agent to external influences. That means, that higher/lower values within the **C** trait increase/decrease the likelihood to act planned and organised and maintain made commitments. Higher/lower **A** values indicate the likelihood that a person prefers selfish or altruistic plans and actions, given that there is a metric that can measure the own gain. The characteristic of the **N** trait provide a likelihood that a person is confident in its own commitments and that persons with higher/lower values in this trait maintain commitments longer given external influences. This deliberation process is represented by the function *options* with the following signature:

$$options : \wp(Bel) \times \wp(Int) \times Per \rightarrow \wp(Des).$$

The generated desires are a set of alternatives (goals) an agents wants to fulfil, which are often mutually exclusive. As the option generation should produce all options available to the agent the influence of the personality is restricted to the persistence of already selected intentions.

[3] In fact, it might be hard to clearly distinguish the information sources. That is because other agents are part of the environment and the observation of the behaviour of other agents might thus be both an observation of the environment and an (implicit) communication act.

The third stage is the *Filter Process* where the agent chooses between competing desires and commits to achieve some of them next. The filter process is influenced by the preferences to vary activities over keeping a strict routine and the level of self-discipline (**O, C**), the need to act in harmony with other agents (**A, N**) and even the tendency to generally interact with others (**E**). For example, variations of **C** influence an agent's preference to detach the prior selected intentions. In this sense, a person with higher/lower **C** value is more likely to stick with an intention or more likely to act spontaneous selecting other intentions. A person with higher **O** value is more likely to select intentions that it has not been committed to. Variations of **A** and **E** influence an agent's preference to commit to selfish/altruistic goals, balancing between own gain and the groups gain. A person with a higher **E** value is more likely to select intentions including group actions and being more exciting. A higher/lower value in the **N** trait controls the likelihood to which an agent is sensitive against new development, e.g. new intentions. The *filter* function is defined as:

$$filter : \wp(Bel) \times \wp(Des) \times \wp(Int) \times Per \rightarrow \wp(Int).$$

The personality helps to prioritise the different intentions and for example indicates to what extent an agent acts goal-driven, prefers interaction, varies the activities. It selects the best option from the point of view of the agent based on the current beliefs, with respect to the prior selected option.

The last stage is the *Actuation*, in which the agent creates/selects the plan (π) and influences the environment performing actions (α). This phase is mainly influenced by the creativity level of the agent (**O**), the tendency to apply actions in a decent manner (**C**) and the preference to interact with others (**E**). We refer to the influence of **O** as the likelihood to approach an intention using new plans that have not been applied before. In this sense it can be a decision support in situations in which there are multiple plans without a best solution (*cf.* Campos et al., 2009, p. 1141). A person with higher/lower value within the **C** trait is more/less likely to stay with a plan that is already generated for an intention. Whereas a higher value **E** prefers action-oriented, gregarious, and interactive actions within a plan. The actual plan is then generated for the selected intentions and executed, which is defined as:

$$plan : \wp(Bel) \times \wp(Int) \times Per \rightarrow Act^*.$$

The execution of actions as plan-elements directly influences the environment and the personality indicates how accurate an agent behaves (**C**). This is a rather vague argument for agents. To set an example, imagine a robot that should perform a motion from one point to another in a specific time frame. The level of conscientiousness then can be used to implement a noise level added to the target location or time frame boarders. Indeed, this seems to be curious when considering artificial agents but is one important

difference between humans. The actuation function *execute* is formally defined as:

$$execute : Act \times Per$$

Another example can be found referring to the extraversion trait, indicating how assertive one can act, e.g. while approaching problems. A higher value in the **N** trait indicates how sensitive a person is facing resistors or stressors during the execution.

To conclude, we can recognise that the individual traits are not independent. This is an intended interpretation as the FFM dimensions are not orthogonal to each other (van der Linden et al., 2010, pp. 316–320). Furthermore, distinctive behaviour occurs in the interplay of different traits which either intensify or diminish the agents' attitudes.

7.2.2. Balancing commitments to means and ends

The prior explained naive BDI lifecycle follows a blind-commitment strategy that is overcommitted to the ends and the means. This commitment strategy is acceptable for the simulation environment used within this work as: the domain is tick-based, the plans a rather short, and plans executed for an intention are fixed, making the time required to generate a plan negligible. However, using the provided explanation the algorithm can be adapted to produce reactive and single- or open-minded behaviour, which might be either bold or cautious. Algorithm 7.2 shows one variant of a BDI lifecycle that is not overcommitted to intentions or plans (adapted from published work, Wooldridge, 2000, pp. 31).

In order to achieve these properties the actuation stage is extended with a perception and belief revision stage. This is done as each action takes some execution time and thus the environment might change to a state where the current intention or the current selected plan is not of relevance anymore. The process outlined in Algorithm 7.1 can not recognise these facts as it strictly executes a once selected plan. Also introduced are some methods that help the agent to decide if it must *reconsider* its current intentions or if the currently selected plan is *sound*. The inner while condition is further extended with two conditions, which validate whether the intentions were *successfully* achieved or became *impossible*.

Taking the above argumentation about the influence of the personalty into account one might argue that the traits must affect the condition checks (sound, succeeded, impossible) as well. For example, the trait **C** as one of the major influences during the execution (remember the noise example) could influence the succeeded check. The trait **N** indicating the emotional stability could influence the impossible check, in terms of 'more rounds, more stressful'. However, the extension mechanism proposed affects the existing stages of the BDI cycle and we argue that these effects take place in these stages. Making the influence in the condition checks redundant. For example, if an intention was successfully achieved is recognised in the belief revision and thus will make the effects of

Algorithm 7.2 A BDI cycle that incorporates personality into the decision-making process that is not overcommitted to the means or ends.

Input: B_{init}, I_{init}, P; **Output:** -

```
1:  B ← B_init, I ← I_init
2:  while true do
3:     ρ ← percept(Env, Msg)
4:     B ← beliefRevision(B, ρ, P)
5:     D ← options(B, I, P)
6:     I ← filter(B, D, I, P)
7:     π ← plan(B, I, P)
8:     while not (empty(π) or succeeded(I, B) or impossible(I, B)) do
9:        α ← hd(π)
10:       execute(α, P)
11:       π ← tail(π)
12:       ρ ← percept(Env, Msg)
13:       B ← beliefRevision(B, ρ, P)
14:       if reconsider(I, B, P) then
15:          D ← options(B, I, P)
16:          I ← filter(B, D, I, P)
17:       end if
18:       if not sound(π, I, B) then
19:          π ← plan(B, I, P)
20:       end if
21:    end while
22: end while
```

C implicit available in the condition check.

7.3. Implementation

To evaluate the model we implemented it for the multi-agent simulation environment AntMe![4]. Major parts of the implementation were part of the bachelor thesis presented by Aria (2014) and supervised by the author. The main objective of each ant colony is to collect as much food (apples, sugar) as possible and to defend their own anthill from enemies such as other ant colonies and bugs. The possible actions are 'move'[5], 'turn'[6], 'pick-up food' and 'drop-off food', 'attack', and 'put scent-mark'. Scent-marks are used to communicate and have a lifetime. Each ant has a number of properties, comparable to the one found in other crowd-simulation frameworks, e.g. range, visibility range, maximum velocity, and the maximum energy level. The environment evolves in discrete time-steps. It is comparable to the termite world used by Salvit and Sklar (2011, 2012)

[4]For further information about the simulation environment the interested reader is referred to `http://www.antme.net/`, last-visited: 2017-09-22.

[5]Move is a category of actions comprising stay, goStraight, goAwayFromPOI, goToPOI, and goToNest.

[6]Turn is a category of actions comprising turnToPOI, turnByAngle, turnAround, and turnToGoal.

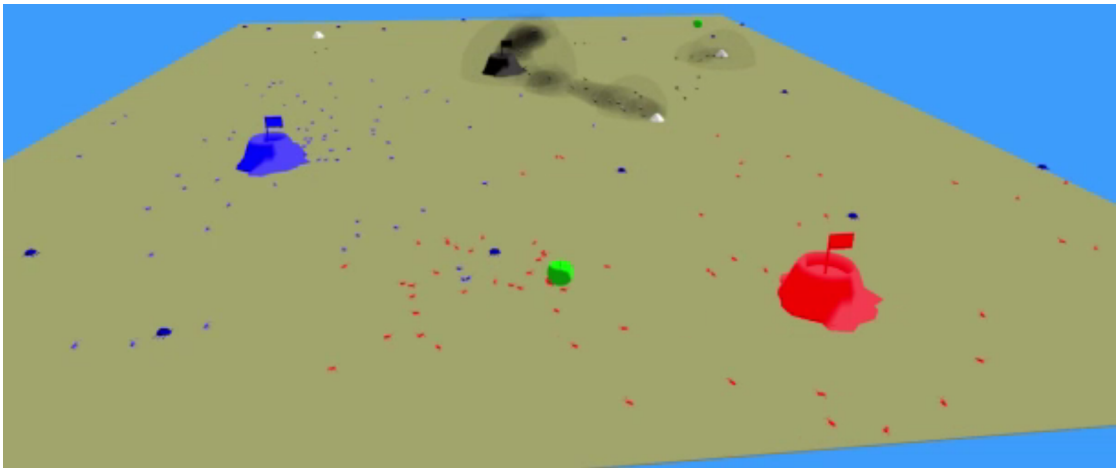


Figure 7.2.: Screenshot from an AntMe! simulation with three ant-colonies (red, blue, black). Carrying apples (green) is a teamwork task as well as defeating bugs. The black dust is the visualization of scent-marks, here used to highlight sugar. Such scent-marks disappear after a while.

in that it applies mainly the same action categories. It is different in that it introduces a communication and interaction component and provides concious move and turn actions. Fig. 7.2 shows a screenshot of the simulation environment.

In order to integrate BDI into the selected environment, it is necessary to identify what makes out the beliefs, desires, and intentions. This mapping is derived from the environment itself. Here, beliefs are given by the world model of AntMe! and comprise the current state of each ant (current energy, velocity, actual load, number of other (hostile) ants, bugs in sight, and so forth), *i.e.* everything an ant knows about itself and the world it is located in. The world itself is only partially observable for an ant. Desires are the 'motivational' factors that drive the behaviour of an agent, thus, one of the main desires for each ant is to collect food, which can be differentiated into food types. Further desires are to rest (to not starve), to be social (by communicating beliefs to others), to explore the environment, or to escape from enemies. Those are desires are individual for each ant and are generated given the ants' currents beliefs, e.g. if there is a bug insights an desire to escape from the enemy is created. Finally, intentions are used to fulfil desires by combining the basic capabilities of the agents. Those are the ones introduced above.

Fig. 7.3 shows the process of practical reasoning in a BDI agent using mainly seven components: the belief, desire, and intention attitudes and the four introduced phases belief revision, option generation, filter process, and actuation. These components form the blueprint for an intentional system, which acts motivated by internal factors and at the same time reacts to changes in the environment. To implement personality as affective phenomenon within this architecture we modified it according to our agent-

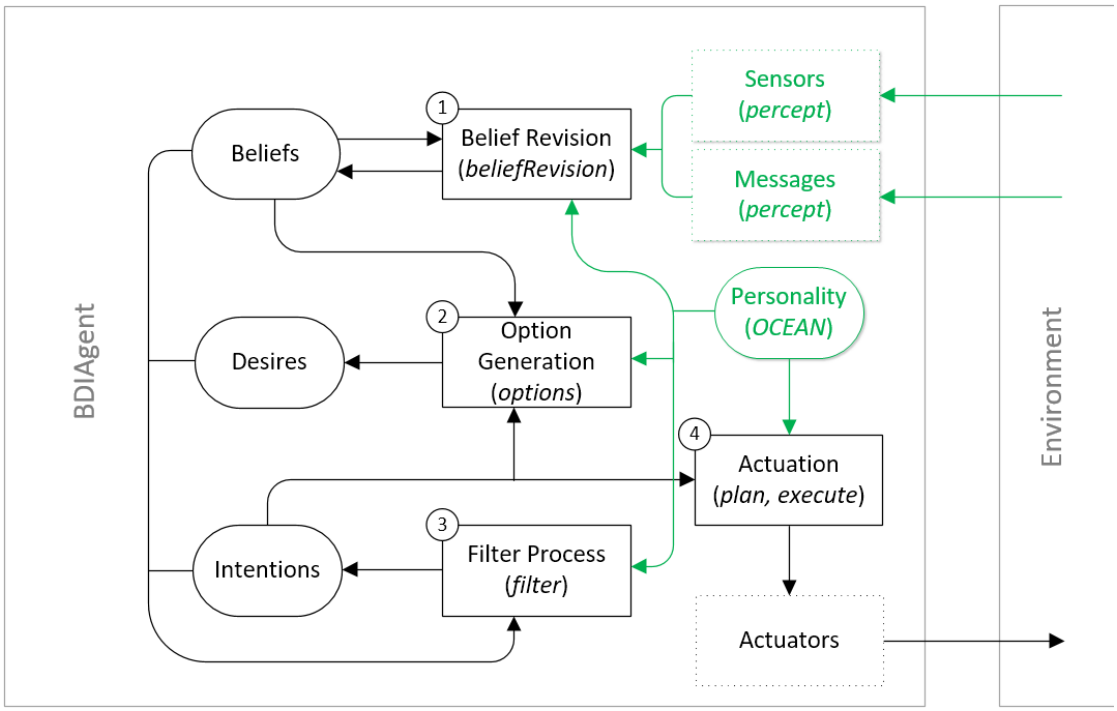


Figure 7.3.: A schematic depiction of a BDI agent architecture including personality as affective phenomenon (green components). The illustration is adapted from published work (Wooldridge, 2002, p. 31), differentiating between the lifecycle stages (1 – 4) and the influenced modalities.

model described in Section 7.2. This complements the description provided with an architectural view that is independent from the selected domain. Within the figure the modifications are shown in green extending the work of Wooldridge (2002, pp. 28–34). Personality is integrated as own modality beside the existing ones, influencing each phase of the BDI lifecycle. Section 9.4 provides a detailed discussion that substantiates this design-decision. We explicitly differentiate between perceptions derived from the environment and perceptions derived via communication acts. Although both are derived via sensors, especially in the selected environment where the ants receive messages via scent-marks, this differentiation is important for personality influences within the belief revision. That is through the observation that some personalities can ignore, e.g. sensed perception due to their openness to new facts (interested vs. conservative) or, e.g. communicated perceptions due to their agreeableness (sceptical vs. trusty). Other affective phenomena require the same distinction, as, for example, proposed by Pereira et al. (2005) and implemented by Jiang et al. (2007) in the emotional BDI (eBDI) architecture.

As described above, we implemented the influence of personality interpreting the individual personality traits as likelihood. Therefore, we represent the personality according to the main dimensions of the FFM as five-dimensional vector. Each FFM dimension is

characterised using a continuous scale with the range $[0, 1]$. Our approach is based on the idea to link personality traits to the phases by interpreting the meaning of the trait taking into consideration the individual characteristic of the phase. In fact, it applies factor analysis results (*cf.* Allbeck and Badler, 2002; Guy et al., 2011; Durupinar et al., 2011) to the environment the agents are placed in. To compare our results with the ones presented by Salvit and Sklar (2011, 2012), namely the analysis of the movement they present for the termite world, we ground our implementation in a mapping between the FFM and the EMOTE (Chi et al., 2000) model that is described by Allbeck and Badler (2002, p. 6). The authors used the following characteristics to differentiate the qualities of movement of a person:

- The Space dimension describes the attention to the surroundings as either Indirect – flexible, meandering, wandering, multi-focus; or Direct – single focus, channelled, undeviating.
- The Weight dimension describes the sense of the impact of one's own movement as either Light – agile, filigree, easily overcoming gravity, marked by decreasing pressure; or Strong – powerful, having an impact, increasing pressure into the movement.
- The Time dimension describes the lack or sense of urgency as either Sustained – lingering, leisurely, indulging in time; or Sudden – hurred, urgent.
- The Flow dimension describes the attitude towards bodily tension and control as either Free – uncontrolled, abandoned, unable to stop in the course of the movement; or Bound – controlled, restrained, able to stop.

Table 7.2 shows the mapping between the described movement components and the FFM dimensions. This mapping is based on the descriptions of both components and one possible interpretation. Such mappings can be found for a multitude of actions using the extensive body of research provided by psychologist. Within AntMe! we implement the personality influence as likelihood controlling whether or not the characteristic applies its effect for a given decision. Listing 7.1 shows some code samples taken from our implementation. For example, the influence of the agreeableness trait during the belief revision is implemented in that an agent with $A = 1.0$ always trust information received via communication acts, whereas an agent with $A = 0.0$ always rejects them. As an example for the option generation, we consider the influence of the conscientiousness trait in that an agent with $C = 1.0$ always maintains an intention as option; regardless of the current beliefs about the world. For the filter process, we implement the influence of the extraversion trait prioritising intentions that imply interaction with others using the characteristic of the trait. We add a randomisation to reflect that personality contributes to a consistent behaviour in the long run (*cf.* Section 3.4), meaning that it is optional

Table 7.2.: Correlation between movement and FFM dimensions. Table is adapted from published work (Allbeck and Badler, 2002, p. 6).

	Space	**Weight**	**Time**	**Flow**
Openness				
High	indirect	light	sustained	free
Low	direct	strong	sudden	bound
Conscientiousness				
High	direct	strong	sudden	bound
Low	indirect	light	sustained	free
Extraversion				
High	indirect	light	sustained	free
Low	direct	strong	sudden	bound
Agreeableness				
High	indirect	light	sustained	free
Low	direct	strong	sudden	bound
Neuroticism				
High	direct	strong	sudden	free
Low	indirect	light	sustained	bound

in its influences in a given situation. As major parts of the implementation of the model for AntMe! were part of the bachelor thesis presented by Aria (2014), the interested reader is referred to this publication for a complete description of the implementation.

As the influence of concrete personality always depends on the actual situation, we used a simple approach to alter the behaviour of the agent. This is comparable to the implementation described by Salvit and Sklar (2011, 2012). Providing a more advanced implementation, e.g. in terms of a reasoning engine, is not necessary as (1) we only implemented one affective phenomenon independent of the others and (2) we want to present an agent-model that confirms the work of Salvit and Sklar (2011, 2012) with respect to the FFM.

7.4. Evaluation

Each simulation run encompassed 5000 time-steps, where each ant in each time-step completes the BDI cycle of sensing its environment, updating its beliefs, desires and intentions and executing. The ants are able to sense their location and to recognise whether or not they are transporting food, the location of food, other ants, scent-marks, and enemy within their range of sight. The scent-marks are used to communicate what other ants of the own colony are targeting and to highlight the occurrence of enemies. Within one simulation configuration each single ant has the same personality and the

```
// sample from belief revision
if(A - Math.random(1.0) > 0){
  belief.update(percept.msg[i]); // trustful, belief it
}else{
  discard(percept.msg[i]); // not trustful, reject
}

// sample from option generation
if(C - Math.random(1.0) > 0){
  maintainIntention(i); // consistent behaviour
}else{
  dropIntention(i); // curious behaviour
}

// sample from filter process
if(E - Math.random(1.0) > 0){
  priorizeIfInteractionReq(option); // extroverted
}else{
  // introverted, do nothing
}

// sample from actuation process, implementation of action moveTo
if(C - Math.random(1.0) > 0){
  moveTo(poi); // decent
}else{
  moveTo(addNoise(C, poi)); // messy
}
```

Listing 7.1: Code samples illustrating how the personality influences the different phases.

measured items are averaged over 50 simulation runs. Each simulation run starts with the same point of origin of the ant hill, apples, and sugar. Occurrence of bugs is randomised and each deceased ant is instantly replaced with a new one, resulting in a constant ant population of 100.

Using the introduced model we expect that the ants' behaviours vary when adjusting the personality traits and that these variations are consistent with the personality that has been modelled. To prove this, I formulate the expected influences of the personality traits next (*cf.* Section 7.4.1). Afterwards I compare the results that can be observed in the environment with the formulated expectations (*cf.* Section 7.4.2). To substantiate this comparison, I will describe further results including the movement paths taken by the ants in Section 7.4.3. This section provides a more subjective interpretation, as it explains the effects that can be observed within the interplay of all personality traits. Finally, I will discuss the provided insights in comparison with the objective formulated in the beginning (*cf.* Section 7.4.4).

7.4.1. Expectations

We expect that an ant population with high values in the trait openness (O+) does more exploration than a population with low values (O–) as those ants are more curious about

locations they have not visit.[7] One consequence hereof is that O+ ants are expected to find sugar and apples earlier. At the same time, we expect the O– ants to harvest sugar faster as O– implies a consistent behaviour which is favourable for this task as it includes walking the same route multiple times.[8] We expect that high values in the trait conscientiousness (C+) lead to more collected food because high contentiousness implies a stronger commitment to selected goals, *i.e.* C+ maintain goals longer. The consequence is that such agents will not choose other goals, *i.e.* not drop food when facing other goals such as attacking/running away from bugs. At the same time, we expect low valued ants (C–) to have a lower chance to starve during the search for food as collecting food is the most important desire. That is to say, that C– are more likely to drop food if they desire to return to the nest and rest. Extroverted ants (E+) are expected to communicate more frequently with other ants by putting scent-marks as markers for the occurrence of sugar, apples and bugs more frequently. That is because they are more likely to favour cooperative actions, such as transporting an apple and they are more talkative. This effect correlates with the effect of the trait agreeableness, indicating whether an ant trusts information received from other ants (A+) or not (A–). We expect that high valued ants in both traits collect food more frequently as they more frequently communicate with each other and trust the information they receive, e.g. about a bugs' position and the position of sugar and apples. The neuroticism trait indicates the ants' emotional stability. We expect high valued ants (N+) to avoid dangerous situations such as bugs and hostile ants because they are more likely to avoid stressful situations, e.g. they a more sensitive to nearby killed ants (N–) instead of being confident to survive (N+) while fighting a bug – resulting in lower numbers of eaten ants and killed bugs. However, the effect of this trait correlates with the level of trust (A+ vs. A–) and the level of self-discipline (C+ vs. C–).

7.4.2. Observations

To derive observations we simulated the permutation of the minimum and maximum values for each trait, resulting in $2^5 = 32$ ant populations each consisting of 100 ants of the same personality. The features comprise the collected apples and the collected sugar, the number of eaten and starved ants, and the number of killed bugs. Table 7.3 shows the correlation matrix for all personality traits and the measurable features of an AntMe! simulation. The entries provide the linear correlation of the measurable items with the characteristic of each personality trait. An entry becomes a value between +1 and –1, where +1 marks a perfect positive correlation, *i.e.* the higher the trait the higher the measured item, 0 marks no correlation, and –1 marks a perfect negative correlation, *i.e.* the higher the trait the lower the measured item.

As indicated in the correlation matrix, the majority of effects that have been postu-

[7] The $-$, $+$ label represent a value in the interval $[0.0, 0.5)$, $[0.5, 1.0]$ respectively.
[8] In other words, openness indicates the choice between exploration and exploitation.

Table 7.3.: Correlation matrix between measured items and personality traits.

	Apple	Sugar	Eaten	Starved	Bugs
O	-0.068	-0.444	-0.043	-0.209	0.027
C	0.545	0.425	-0.454	0.893	-0.027
E	-0.150	0.072	0.002	-0.119	-0.009
A	0.261	0.501	-0.430	0.107	-0.554
N	0.305	0.114	-0.436	0.125	-0.554

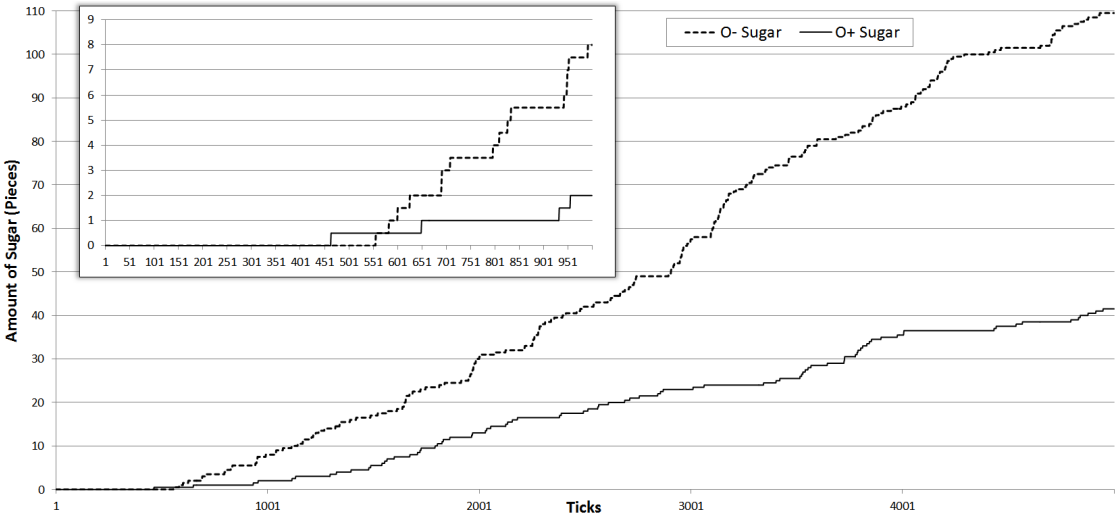


Figure 7.4.: Tick-based cumulation of all O+ and all O– ant populations and their average progress of collecting sugar. The smaller diagram highlights the segment in which the collection starts.

lated are observable in the simulation. To start with, the matrix indicates that O+ ants collect less food than O– ants and that this behaviour is most notable for the collected sugar. Still, we postulated that O+ ants will find sugar earlier. This effect is illustrated in Fig. 7.4, where the process of collecting sugar is depicted tick-wise. The figure averages all O+ and O– populations and shows that O+ ant populations start indeed approximately 2% earlier with the collection than O– populations (O+ starts after 465 ticks, O– starts after 557 ticks). At the same time it illustrates that O+ populations collect food slower than O– ant populations.

Table 7.3 further shows that C+ ant populations are more successfully in collecting apple and sugar than C– populations, though, they starve more frequently when compared to C– populations. This was expected due to the commitment to the most important desire of collecting food over the risk to starve while returning food to the nest. The frequency of communication cannot be measured in the environment, making it hard to distinguish the effect of varying extraversion. The correlations shown in the

table are weak. If we take the agreeableness trait into consideration, we can see that A+ ant populations collect more food, while they get eaten more frequently. At the same time, they kill less bugs than A– populations. While we expected the former, the latter correlation remains unclear. Within the next section, we will see that the number of defeated bugs is in general very low and that only a few personalities select the fight. The neuroticism trait shows a negative correlation for being eaten, *i.e.* N+ populations get less often eaten than N– ant populations. This effect was expected due to the preference of avoiding dangerous situations.

7.4.3. Further Results

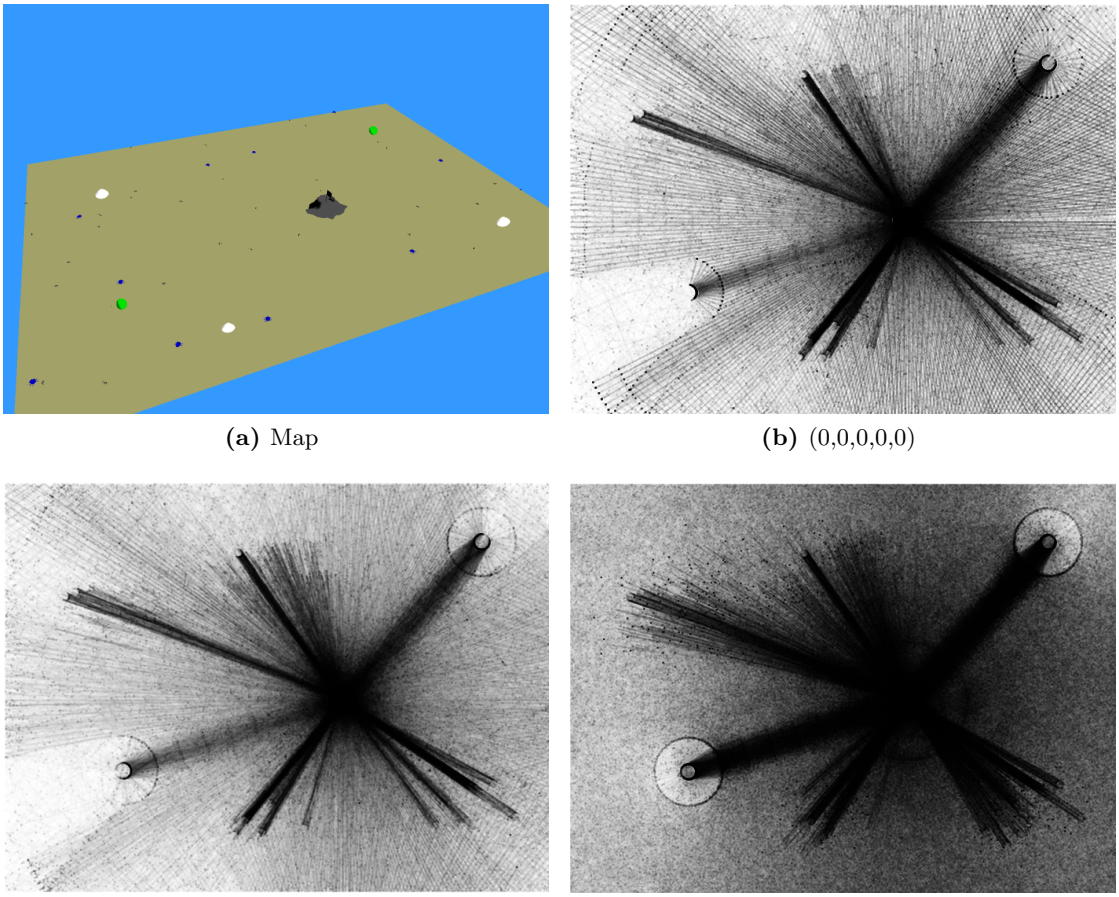
Up to this point, we used the correlation matrix to compare the expectations with the observable behaviour. Within this section, we will analyse the actual data and a visualisation of the movement of the different ant populations. Table 7.4 list the results for all permutation of the extreme values and the average ant population, emphasising the different observable behaviours we described above. For example, an ant population with maximum values (1,1,1,1,1) collects more apples and sugar, kills fewer bugs and loses fewer ants through bugs than an ant population with minimum values (0,0,0,0,0). Still, for the latter a lower number of starved ants can be observed. Here, the traits **E** and **A** influence the occurrence of scent-marks and the interpretation (trust) of the very same thing. The trait **C** implies that already picked-up food is not dropped through new percepts as collecting food is the most important goal for the ants. The trait **N** affects the fight behaviour of the ants leading to fewer/more eaten ants/killed bugs, respectively. Given the values we get insights into the variance of the different behaviours, however, the correlations between the traits makes it hard to explain the effects for a single population.

The effects of the personality traits are also visible in the paths an ant population takes. Fig. 7.5b–7.5d are showing the path heat maps for three selected populations. Here the paths taken by the ants during one simulation are cumulated, giving an indication about the movement patterns followed by one population. As mentioned above the locations of the ant-hill and the food were fixed. Fig. 7.5a shows a screenshot of the map. The spawn locations of sugar is restricted to areas. The effects of this setup appear partly as artefacts in the heatmaps.

The figures emphasise the effects of the trait **O**, which affects an ant's preference of acting explorative vs. exploitative or following a conservative vs. curious behaviour. This becomes visible, e.g., when rating the ants behaviour of staying in known areas against the behaviour of eager exploring new areas. The (1,1,1,1,1) population shows here a completely and relative uniform colourised map (Fig. 7.5d). The other two populations generate white spots and less uniform colourised map. At the same point, the figures visualise how cooperative the ants act—visible through the round artefacts highlighting the occurrence of apples. This round artefact depicts the visible range of each ant.

Table 7.4.: Collected information for single ant populations (all permutations and average). Each value is averaged over 50 simulation runs. Minimal and maximum values are highlighted in bold.

(O,C,E,A,N)	Apple	Sugar	Eaten	Starved	Bugs
(0,0,0,0,0)	8.4	18.4	281.6	6.0	2.5
(0,0,0,0,1)	19.5	81.0	84.7	130.9	0.0
(0,0,0,1,0)	**19.8**	83.2	82.8	131.5	0.0
(0,0,0,1,1)	19.4	78.2	84.7	131.6	0.0
(0,0,1,0,0)	8.2	10.6	284.8	3.5	1.8
(0,0,1,0,1)	18.6	43.1	97.2	39.8	0.0
(0,0,1,1,0)	16.7	113.1	83.0	55.1	0.0
(0,0,1,1,1)	16.1	108.3	83.4	55.3	0.0
(0,1,0,0,0)	19.0	75.6	117.4	146.9	3.3
(0,1,0,0,1)	19.4	88.3	54.5	**204.8**	0.0
(0,1,0,1,0)	19.7	90.1	54.3	203.5	0.0
(0,1,0,1,1)	19.3	86.8	55.7	203.1	0.0
(0,1,1,0,0)	19.0	52.9	98.3	162.9	2.1
(0,1,1,0,1)	19.3	58.0	**51.5**	204.7	0.0
(0,1,1,1,0)	16.5	**181.0**	65.9	174.6	0.0
(0,1,1,1,1)	16.0	175.7	64.2	175.9	0.0
(1,0,0,0,0)	8.5	8.1	**285.4**	0.0	3.0
(1,0,0,0,1)	16.2	46.3	77.1	0.0	0.0
(1,0,0,1,0)	16.8	48.1	82.7	0.0	0.0
(1,0,0,1,1)	16.2	44.5	80.1	0.0	0.0
(1,0,1,0,0)	**7.9**	**6.5**	283.9	0.1	**3.5**
(1,0,1,0,1)	15.7	28.9	74.0	0.0	0.0
(1,0,1,1,0)	16.2	40.1	74.4	0.0	0.0
(1,0,1,1,1)	15.8	39.9	75.0	0.0	0.0
(1,1,0,0,0)	19.2	70.5	90.7	167.6	1.6
(1,1,0,0,1)	19.5	65.0	54.6	193.5	0.0
(1,1,0,1,0)	19.6	69.6	54.3	193.4	0.0
(1,1,0,1,1)	19.4	66.7	52.1	197.1	0.0
(1,1,1,0,0)	18.8	47.2	99.7	153.0	2.7
(1,1,1,0,1)	19.1	50.4	52.6	193.3	0.0
(1,1,1,1,0)	19.4	78.6	55.5	184.5	0.0
(1,1,1,1,1)	19.3	75.8	54.2	188.4	0.0
$(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2})$	9.7	17.1	270.8	19.5	1.5



(a) Map

(b) (0,0,0,0,0)

(c) $(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2})$

(d) (1,1,1,1,1)

Figure 7.5.: The map used within this part of the evaluation (Fig. 7.5a) and the cumulated paths of three ant populations (Fig. 7.5b–7.5d). As the occurrence of food and the location of the ant hill are fixed a comparable structure originates. Still, the effects of exploration vs. exploitation are visible (covered area, curious behaviour, broader paths). The artefacts denote the visibility range of the ants and the points apples are spawned, giving an indication of the effects of scent-marks and the trustfulness of the ants.

Table 7.5.: Collected information for single ant populations with real personalities. Each value is averaged over 50 simulation runs. Minimal and maximum values are highlighted in bold.

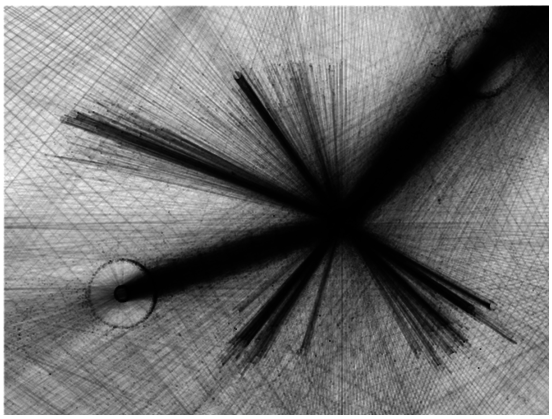
(O,C,E,A,N)	Apple	Sugar	Eaten	Starved	Bugs
(0.85, 0.70, 0.47, 0.34, 0.83)	13.4	25.8	233.6	47.1	0.7
(0.49, 0.95, 0.31, 0.70, 0.48)	17.4	49.9	166.1	104.5	1.3
(0.63, 0.63, 0.65, 0.71, 0.64)	13.10	26.68	239.18	42.58	0.70
(0.81, 0.84, 0.73, 0.60, 0.46)	13.42	27.34	234.26	46.14	0.90
(0.59, 0.59, 0.44, 0.64, 0.60)	11.64	23.14	255.48	29.60	1.24
(0.59, 0.69, 0.24, 0.43, 0.49)	**10.24**	19.26	266.32	21.64	1.40
(0.95, 0.95, 0.49, 0.83, 0.73)	**18.96**	60.66	**115.66**	**144.90**	0.26
(0.70, 0.35, 0.70, 0.66, 0.56)	10.36	**17.88**	**265.76**	**21.26**	**1.46**
(0.95, 0.69, 0.76, 0.90, 0.83)	18.20	**51.32**	138.18	123.92	**0.02**
(0.71, 0.71, 0.51, 0.64, 0.41)	11.50	21.88	256.84	28.82	0.84

As collecting apples is a cooperative task, where up to 10 ants can work together to transport one apple to the ant-hill, these artefacts are more distinct for populations that are more cooperative (E+, A+).

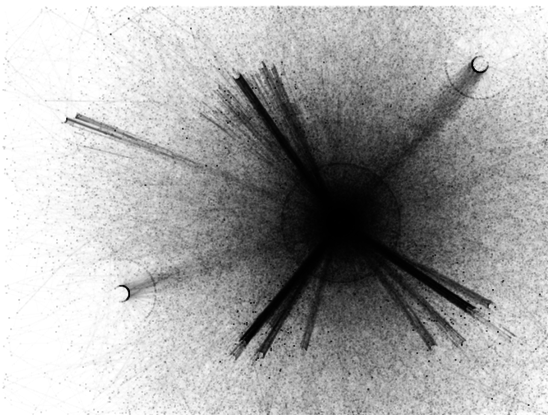
7.4.4. Discussion

Comparing the formulated expectations with the observed behaviour, we could interpret the parameters added to the BDI lifecycle as personality traits and the resulting behavioural change of the agents as personality. We have shown that some personalities are better suited for particular tasks than others. This extends the work of Durupinar et al. (2008) to the complete set of personality traits available through the Five-Factor Model. We also learned that such parameters can influence the behaviour of agents in a domain-independent manner and that one challenge is the task-dependent interpretation of the effect of a personality. Finally, the experiment confirmed the finding of Salvit and Sklar (2011, 2012) that the interpretation of the parameters as personality traits results in (personality-)consistent behaviour of agents for the Five-Factor Model instead of the MBTI.

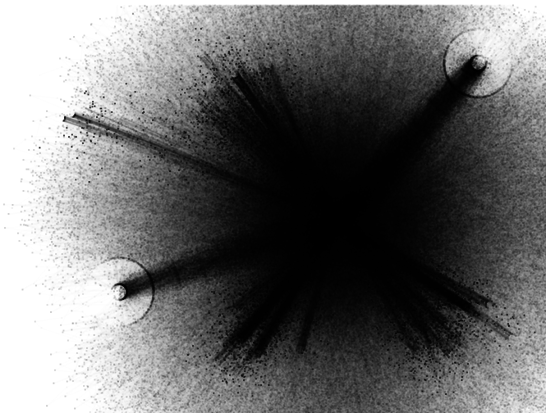
The implication is that the task performance of problem-solvers can be improved by carefully assigning personality-specific tasks. Mohammed and Angell (2003) identify this implication as personality-performance relationship and one of the factors contributing to the performance of human-human teams: "*...the role of the task and team must be carefully accounted for in designing effective teams*" (Mohammed and Angell, 2003, p. 672). To show the personality-performance relationship, we can use the derived results to determine which population performs more accurate for a specific objective. In Table 7.4 the minimal and maximal values that have been reached by the populations for the different measured categories are highlighted. Is the objective to collect as much sugar



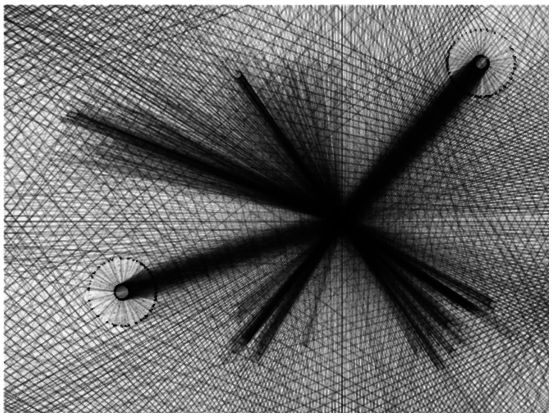
(a) (0,1,1,1,0): Max. sugar



(b) (1,0,1,0,0): Min. sugar



(c) (0,0,0,1,0): Min. starved



(d) (0,1,0,0,1): Max. starved

Figure 7.6.: The path heat maps of the ant populations that collect the most (Fig. 7.6a) and the leasts amount of sugar (Fig. 7.6b) and the paths of the populations where the individuals did not starve (Fig. 7.6c) and where the individuals starved most frequently (Fig. 7.6d).

as possible the population (0,1,1,1,0) would be the best choice, whereas the population (1,0,1,0,0) would be the worst choice. Fig. 7.6a and Fig. 7.6b show the paths walked by the ants of the population that collect the most and the least amount of sugar, respectively. Ants with character (1,0,1,0,0) collect also the least amount of apples, but at the same point attack bugs frequently leading to a high amount of eaten ants. Another example is the objective to let as least ants starve as possible (ants starve if they not rest; rest means staying in the anthill). The paths of one of the populations that are not starving is shown in Fig. 7.6c. Here the concentration on staying in near distance to the anthill becomes visible. In contrast, Fig. 7.6d shows the path heatmap of an ant population where the individuals avoid periods of rest starving when exploring the map.

Given such agent-model that covers the characteristic human behaviours in a traceable manner can help to focus development efforts towards other areas. Furthermore, it can help to identify relevant behavioural pattern in data acquisition tasks and to guide the linkage to theoretical models such as the FFM.

7.5. Conclusion

This chapter introduced a state-of-the-art analysis on the task of integrating the personality phenomenon into agent-based research. It further demonstrated how the different dimensions of the FFM can be integrated into the BDI model of agency and how this leads to variations in the interpretation of inputs and generation of outputs. This process was discussed on different levels: starting with algorithmic considerations extending two well-known variants of the BDI lifecycle, proceedings with architectural considerations integrating the personality as a new attitude within the BDI architecture, ending with a discussion on implementation details that substantiate how the extension can be adapted to a specific use case. The observations that have been made in the selected environment indicate that the decision-making process is influenced by the personality type as expected and that different personalities behave differently. That is the same observation that psychologists make about humans and which was proven for the MBTI in a related work, *i.e.* one objective of this chapter was to reproduce the results of Salvit and Sklar (2011, 2012) with respect to the FFM. The advantages of using FFM instead of MBTI are discussed in Section 3.4. Two important ones are the fuzzy classification of personality-specific behaviour and the introduction of neurotiscm as important criteria distinguishing emotional stable and emotional unstable personalities, which might be useful when connecting to emotional agents as it indicates which emotions are implied in a situation and how intense they are.

The evaluation comprises the implementation of the model into the multi-agent based simulation environment AntMe. Still, the actual implementation has some shortcomings. First of all, the effect of a personality is only based on the characteristic of the trait that decides how often such trait influences the current stages in one of two ways. In fact,

the influence of a personality is context-dependent. Here a more realistic method must be found that includes the current context of the agent, which also comprises the effects of emotions or moods. Surprisingly, agent-based research particularly emphasises the effects of emotions abandons the fact that emotions and its influences are contingent upon the personality. The evaluation is further limited to a comparison between formulated expectations and the observable behaviour in the environment. Although, I grounded the development and the formulation of the expectations in psychological literature, the agent-model should be evaluated empirical. This could be done, by applying factor analysis techniques and let humans judge the observable behaviour describing it with personality-descriptive words.

Based on the presented literature work, the introduced agent-model, and the evaluation I finally can answer the questions I formulated in the beginning of the chapter, namely:

- "What is the state-of-the-art for integrating personality, in particular, the FFM, as an affective phenomenon in agent-based research?"

  Most of the available work originate from the area of agent-based simulation, in particular, agent-based traffic and crowd simulations. Another important application area is human-agent interaction, focussing on the expressions of virtual characters (e.g., mimic and gestures). I was not able to identify contributions in the teamwork and joint human-agent activities domain. Only a couple of the contributions substantiated the decision for applying a specific personality theory. The presented agent-models are use-case specific and developed for the environments the agent inhabit. More general and more advanced work apply either the MBTI or a subset of the FFM dimensions, which justified the development of an own agent-model addressing these shortcomings.

- "How can we derive an agent-model representing the effects of personality with respect to the FFM?"

  This can be done by grounding the modelling process in psychological literature and using the structure that is given by the elements of adaptive systems. The first finding to apply is that personality influences all stages of the BDI lifecycle, though, the contributions of particular personality traits to particular stages of the lifecycle differ in intensity. The structure that is defined by the FFM can be modelled as part of the user model of an adaptive system, whereas the BDI lifecycle represents the behaviour model of an adaptive system. Having both parts, we can implement the context model to complement the data representation layer for a specific application. This can be done by using the results of studies doing factor analyses and interpreting the personality trait and it influences in a given stage of an agents lifecycle.

Giving this thesis' primary goal – predictability – it could be possible to use the presented agent-model to recognise the human team members agenda. In fact, this technique is named Agent Tracking and uses the BDI plans generated by an agent to recognise which BDI plans are executed by another agent. Thus linking the task of anticipating the humans agenda to the usage of plan-execution libraries. A recent introduction into this field is given by Vered et al. (2016). Although the available work show some achievements, the approaches have the inherent problem that all possible plans need to be modelled prior to the tasks, which is an infeasible solution for complex environments.

In the following chapter, I will describe an experiment performed w.r.t answering the questions whether or not we can learn about the personality of humans during interaction and whether or not we can use this information directly to make informed decisions about the human behaviour.

8. Learning Personality Information from Observation

Part of the task of generating human-behavioural models are user models describing the characteristics of individual users. In this domain, features that define the user as an individual and are stable over (periods of) time are named individual traits (Plumbaum, 2015, pp. 19–23). These individual traits include personality information, but also include features such as the cognitive and learning style. The collection of such information can be a challenging task and is identified as the data acquisition task of adaptive systems as illustrated in Fig. 8.1 (illustration highlights the parts of an adaptive system, which I will work on within this chapter). The information can be acquired by explicitly asking the user, e.g. using the onboarding process of an application, or by using available information sources such as the social web and aggregating the user information. Another way to acquire the information is to observe the interaction of the user with a system and use this implicit feedback to derive the necessary information for the user model. Within this chapter I will focus on the latter category to find an answer to the question: *Can agents use our model to learn the personality traits of a human during the interaction with this human?*

In doing so, I will evaluate if the prior introduced agent model can be used to learn personality from interaction by describing two experiments that use observation of the human behaviour to derive knowledge about the personality of users. Thus focusing on the acquisition of implicit information that feeds into the user model. We will use explicit information that is derived using questionnaires to evaluate the accuracy of the learning process.

Adaptive systems could utilise such information to adapt the provided content, the behaviour, or the user interface. Focusing on the agents' behaviour, I will use the implicitly derived personality information to approach predictability as a side question within this chapter. To do so, the agents utilise the estimated personality of the human user during their decision-making. This is done to find an answer to the secondary question: *Can we use personality information directly to make informed decisions about the potential behaviour of human?*

This chapter starts with analysing the related work in Section 8.1. As personality is a cognitive characteristic, I will concentrate on contributions that try to learn such characteristics and distinguish my own work from the state-of-the-art. Furthermore, a classification of the work within the broader scope of recognition techniques and Person-

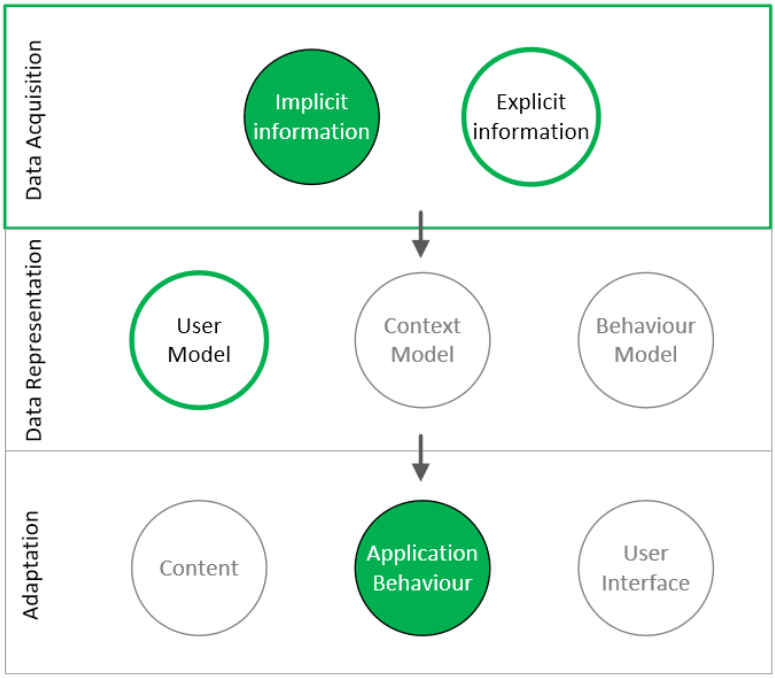


Figure 8.1.: Parts of an adaptive system that are discussed within this chapter.

ality Computing is provided. Subsequently, I will provide an introduction to the Colored Trails Game (Grosz et al., 2004) (CT) (Section 8.2) as the testbed that was used during the experiments and as an environment which was frequently used by other researchers in this domain. The agent models that are applied are described in Section 8.3. During the experiments the agents play against humans and try to estimate the personality w.r.t to the FFM. The experimental setup and the results are described in Section 8.4 indicating that some characteristics of a personality can be learned more accurately than others. Finally, I will discuss the results with respect to both questions. This is done in Section 8.5. Eventually, the chapter is concluded.

8.1. Classification and Related Work

Among others, predictability includes deliberating about other agents actions (*cf.* Chapter 5). In particular about deliberating what are the next actions the team members will execute to approach the joint goal of the collaboration. In human-agent interaction, predictability can be found in work reaching from the question whether people reason about other peoples actions (e.g. Ficici and Pfeiffer, 2008) to the question of how to recognise inter-player relationships in multi-player games (e.g. Wilson et al., 2011). The anticipation of behaviour is addressed by multiple fields, which can be described as plan, activity, goal, and intent recognition techniques. Those techniques have a long history

in AI research. Sukthankar et al. (2014) provides an introduction into the diversity of concepts and applications available.

In the following, I will concentrate on approaches that specifically focus on recognising personality as one of the phenomena that influence a humans' *"...disposition toward action, belief, and attitude formation."* (Pianesi, 2013, p. 146), *i.e.* that focus on personality information as predictive of outcomes of interest (*cf.* Vinciarelli, 2014, p. 297; Wright, 2014, pp. 292–294). I will introduce approaches where agents learn the personality of other agents (artificial and natural) in different scenarios. That is, I concentrate on one of the three main problems handling personality in computing systems, namely *Automatic Personality Recognition*. Automatic Personality Recognition is the task of recognising the true personality of an individual (Vinciarelli and Mohammadi, 2014, p. 273). The other tasks are Automatic Personality Perception, which is the task of predicting which personality is attributed to an individual by some other; and Automatic Personality Syntheses, which is the task of simulating artificial personalities in agents (as described in Chapter 7). Fig. 8.2 illustrates the cognitive processes together with these technical tasks and highlights the relations between the externalised observable behaviour on the one hand—named *distal cues*—and the perceptual process—named *proximal cues*—on the other. The agent-models that are developed in the next section are inferring the self-assessed personality traits from distal cues.[1] In the context of human-agent interaction, distal cues are all forms of a humans' behaviour an agent can observe, as most of these behaviours encode personality traces (Vinciarelli and Mohammadi, 2014, pp. 276–277) (e.g. loudness of voice, interaction patterns, appearance).

Agent-based work on this topic originated from approaches that applied social preferences to enhance predictability. Here, Hoog and Jennings (2001) present a work where agent use a weighted sum of the other agents expected outcomes as a utility function. The examined behaviour is called socially rational decision-making and is based on the idea of social welfare functions, where the individual agents have to balance their utilities and the social utilities with the intention to maximise the welfare of the group. Making it possible to have selfish or selfless personalities. The authors applied Q-learning in a way that each agent learns which interaction with other agents (personalities) is beneficial. Gal et al. (2004) present a comparable work introducing social preferences regarding the three dimensions self-interest, social welfare and inequity aversion. Agents build knowledge in these dimension about the other agents and integrate this knowledge into their decision-making process. Sen and Dutta (2002) use the dimensions equality and selfishness to learn the behaviour of other agents during a repeatedly played trading game (package delivery domain created by the same authors). In the experiments, groups of agents played the game and the composition of the group changes during the game, making it one of the early work that adds the requirement of life-long learning.

[1] The interested reader is referred to the work of Pianesi (2013) providing a more detailed discussion of the concepts of distal and proximal cues.

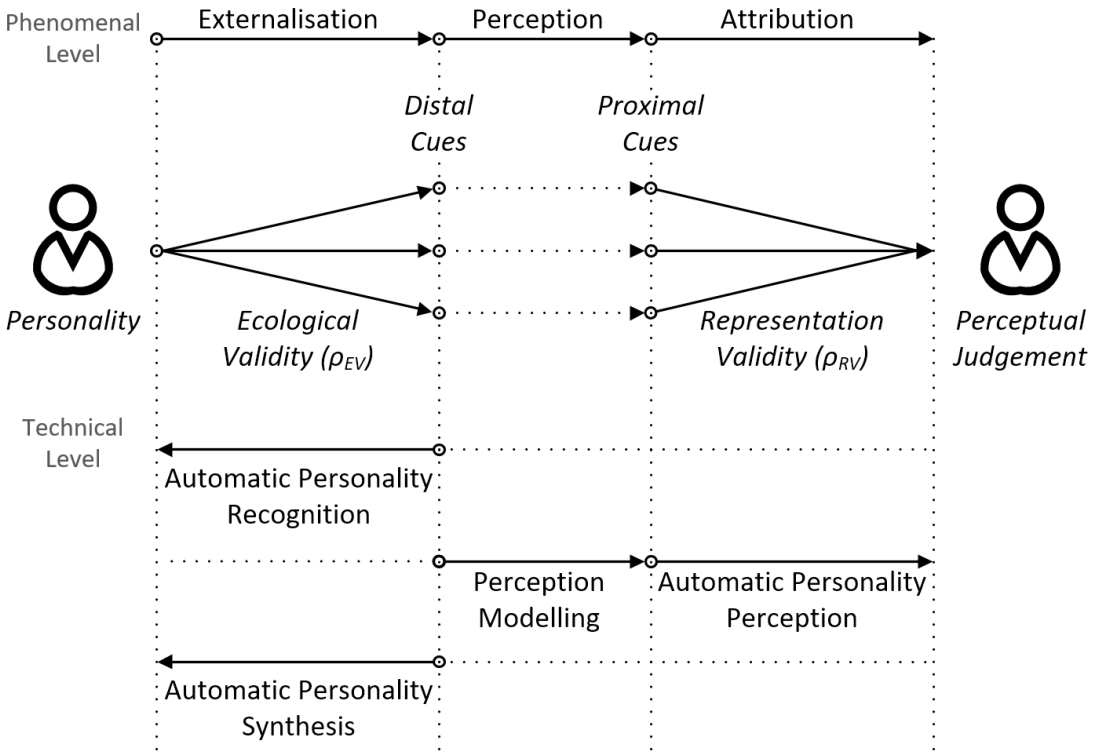


Figure 8.2.: Brunswick's Lens model and the computational tasks of automatic personality recognition, synthesis, and perception. This illustration is adapted from published work (Vinciarelli and Mohammadi, 2014, p. 276).

Talman et al. (2005) present a work that illustrates the use of a rather simple abstraction of personality types. Personalities of agents are determined through the two dimensions cooperation and reliability. The agents play the CT game and try to optimise a utility function incorporating whether or not the player reaches the goal, the distance to the goal, and the number of chips left. During repeatedly played games the agents reason about each other's helpfulness along the two dimensions. As an effect, they try to respond more effectively by customising their behaviour appropriately for different personalities. For example, an agent tries to avoid collaboration with other agents that have been identified as acting selfishly; meaning that the other agents are neither cooperative nor reliable. Otherwise, the selfish agent would always win the game and the other agent not. These decisions are modelled via a utility function. The extensive evaluation carves out that the agent that adapts its behaviour through the personalities of the opponents outperforms agents who do not adapt. Furthermore, this adaptation leads to an increased social welfare for the group in the long term.

The above-introduced approaches have in common that they learn what can be identified as sub-traits of the Big-Five Factors, *i.e.* traits which cope with specific aspects of one of the five dimensions. However, the models are usually not grounded in the

psychological literature by the authors, e.g. it is not substantiated how the authors decided what is meant by selfish or selfless behaviour. Furthermore, the agent-models applied to learn the behaviour of other artificial agents that simulate different personality aspects. In contrast, I focus on learning the personality of humans. Also, my goal is not to produce optimal group behaviour but to prove whether we can learn personality information in direct interaction with an individual human or not, and whether or not we can use information about the personality of the human directly to make informed decisions. My work is partly comparable to the discussed related work as it also applies a multi-attribute utility function for the decision-making process. In fact, it is motivated by the work of Talman et al. (2005) and transfers the presented ideas from agent-agent to human-agent cooperation.

Du and Huhns (2013) present work, which examines whether the interaction of humans with humans and agents depends on the humans' personality type according to the MBTI. The experiments were done using the cake-cutting-game[2] show that the different personalities act in different ways, but also show that there is only little clue when trying to make predictions about possible behaviour based on information about personality. In any case, the work provides a ground truth, that even in scientific games, one can identify personality characteristics using the available distal cues. The survey of Vinciarelli and Mohammadi (2014) present an overview about 55 publications on automatic personality recognition (pp. 277–282). The authors classified the approaches concerning the used datasets; classifying it as text-based, based on nonverbal communication, using data sets collected via mobile and wearable devices, social media based, and computer game behaviour based. The majority of the described approaches[3] uses supervised machine learning techniques working on different data sets and distal cues reaching from linguistic features in speech and text to physical activities to social network activities and observations of gaming behaviour. The achieved accuracy rates reach from the chance level-up to 80 percent for specific data sets and specific traits.

The majority of the here examined contributions applied the FFM using all traits. Roughly half of the described approaches aim to determine the actual personality using the theories' immanent continuous values. The other half aims to determine if a person is on one side of the traits' extremes or the other. This process is called binarization. My work distinguishes itself from the ones listed in the survey as the agents learn about the personality in direct interaction with the human. In contrast, the other approaches use existing data-sets to learn classifiers using different machine learning techniques. Furthermore, the objective of my work is to learn the actual self-assessed personality characteristics of the human.

[2] In the cake-cutting game, the players objective is to divide a cake in a fair manner, *i.e.* each player receives the amount of cake it believes to be fair (Du and Huhns, 2013, p. 240).

[3] Only one publication described an unsupervised learning technique.

8.2. The Colored Trails Game

The Colored Trails Game (Grosz et al., 2004; Gal et al., 2010) (CT) is a multi-agent computer game to investigate cooperative decision-making within a chess-like environment. The game can be played by humans, agents, or mixed groups and was designed to study the effects of different decision-making strategies in an environment that can easily be adapted regarding its complexity (e.g. board size, availability of information, chip distribution, number of colours). Within the game a players' actions influence its outcome as well as the outcome of the other players, thus, providing an environment that enforces competition. At the same time cooperation is enforced as the players have to negotiate about chips they want to share and exchange with other players.

The primal settings of the game are the following: The board is an $N \times M$ grid consisting of coloured squares with a previously defined set of available colours. Each player has a specific starting position and a set of coloured chips. The colours for the squares and chips are determined by the same palette. The player has to use chips that match the colour of an adjacent square to move to that square. The primary objective for each player is to reach one of the 'goal squares', which are marked with a 'G', as reaching it usually ends the game and provides the biggest reward for a player. Secondary objectives could be the amount of remaining chips or the number of rounds played. Having this competitive aspect, on the one hand, players have to cooperate with each other during a communication phase on the other. During the communication, they are allowed to negotiate the exchange of an arbitrary amount of chips. This phase usually consists of the proposal stage, the decision stage, and the actual exchange of the chips. Offered proposals can be accepted or refused, though, players are not enforced to hand-over the (amount, coloured) chips that have been proposed. Within each game, the players act consecutively and alternate their roles as being the proposer or the responder. In the end of each round, CT exchanges the chips and makes the best movement towards the goal square automatically using the Manhattan distance algorithm.

Fig. 8.3 shows a screenshot of the game environment that is used during the experiments that are described in the next sections. The phases display indicates the current phase of the game. The thinking phase has been inserted for human players to overlook the situation. The communication phase is used by the proposing player to create its proposal in an additional window and by the responding player to create an answer. In the exchange phase, the players transfer some amount of chips (can be adapted for that stage), which might be related to the prior accepted proposal.

CT provides various options to configure the environment, for instance, the number of goal squares, whether or not the goal squares are identical for every player, the visibility of the board, the visibility of other players chips, requirements on proposals. While some rules of the game — like the size of the board — are less important than others, especially the distribution of chips as well as the scoring evaluation are essential when

Figure 8.3.: Screenshot of the used game environment, showing the different phases (without movement), the current board, and the available chips for each player.

examining the player's behaviour. Controlling the availability of resources can be used to stimulate cooperation, as it allows to create situations where none of the players has the possibility of reaching the goal on its own, or some can while others need assistance, or all can and no cooperation is required. However, the most significance lies in the calculation of the games final score. Commonly points for reaching the goal will be granted (as the games end when one player reaches the goal(s)), but everything else can be defined to provoke or decrease cooperation.

One example for the variations of the parameters is described in the initial investigations of Grosz et al. (2004). The authors vary the interdependencies between the players by introducing a social dependency factor. This factor is used to vary the influence of individual and group utilities onto the scoring function, *i.e.* if one player's performance solely depends on its outcome or to which extent the player's performance is combined with a weighted average of the scores of all other players. Using this factor the author's thesis was that an increasing social dependency leads to more helpful behaviours, where helpful behaviours are defined as *"...reasonably balanced exchange"* (Grosz et al., 2004, p. 784).

Several researchers have used the CT environment to study different behavioural as-

pects.[4] The study objectives reach from effects of different levels of cooperative behaviour of an agent on the cooperativeness of humans (*cf.* van Wissen et al., 2009) to agent models that are able to learn social preferences in repeated games (*cf.* Gal et al., 2004) to modelling cultural sensible agents (*cf.* Haim et al., 2012). The work of Talman et al. (2005), which was introduced in Section 8.1, is one of the early work using CT and modelling explicitly different personalities. The work shows that it is beneficial for the agent to adapt its behaviour towards the personality of its opponent. Here, two dimensions of the FFM are used: conscientiousness, referred to as how reliable one behaves and agreeableness, referred to as how cooperative one behaves. Analogous, van Wissen et al. (2009) performed experiments using different shades of altruistic behaviour, which is part of the agreeableness trait. The work indicates that humans alter their cooperativeness while experience egoistic or altruistic opponents and that they reward higher cooperativeness without exploiting this; providing an indication that humans do not act completely rational. Gal et al. (2004) present an agent model developed to learn social preferences of humans. They distinguish between the humans' self-interest, their concerns about the social welfare, and their interest in fairness named inequity aversion. The results indicate that explicitly modelling these factors is beneficial while developing an agent able to play against unknown opponents, though, no data is presented that substantiate whether or not these factors were immanent in the human opponent. They indicate, that taken personality facets into account and thus allowing the agent to behave more human-like helps to predict the human's agenda.

In the following, I will present two agent-models that adapt these ideas to the task of learning the personality of human opponents according to the FFM. This is done to complement the existing results, as they are not sufficient to answer the research questions formulated for this chapter. The CT was selected as it provides a relatively simple environment, which is complex enough to learn aspects of human behaviour as shown in existing work. Last, CT is a game that was frequently used by other researchers in empirical behaviour research.

8.3. Modelling Agents that Learn Personality

Judging the personality of others given observations of behaviours is a task which is part of the daily activities of humans and which we learned about as Automatic Personality Recognition when talking about agents. Within this section, we will construct two agent-models that perform this task within the CT environment. Our approach is based on the idea to link the personality traits to the available actions by interpreting the meaning of the trait taking into consideration the effect of the action. Thus, it transfers the approach for modelling personality traits described in the prior chapter to the task of learning personality.

[4] A list of related publications can be found here: `goo.gl/skSr5o`, last visit: 2017-11-02

To achieve the goal I follow the Realistic Accuracy Model (Funder, 1995, pp. 656–665), that introduces the concepts of *Relevance* and *Accuracy*. These concepts advice one, to restrict the experiments to the relevant traits, *i.e.* the traits that can be expressed in a given environment, and the available traits, *i.e.* the traits that can be perceived by others (Wright, 2014, pp. 293–294). Given the CT environment, we have to restrict our agent-models to three of the five dimensions (conscientiousness, extraversion, agreeableness) of the FFM. The inclusion of the remaining traits (openness, neuroticism) is hard to accomplish as they cannot be expressed in the environment. That is because CT does not provide the option to reward or punish creative or conservative behaviour to perceive information about the openness trait. Furthermore, as CT is a scientific game, it is not constructed to evoke an emotional reaction in its players. One might argue that repeatedly loosing in the game will lead to an emotional reaction. However, this effect provides little indication about the emotional stability of a person and cannot be perceived observing the players' game actions. Thus, one the one hand, we do not expect users to show behaviours related to the neuroticism trait. On the other hand, we do not have the possibility ot capture emotional responses within our experiments. We include the conscientiousness trait due to the trading component and the possibility of cheating on other players, the extraversion trait due to the communication component, and the agreeableness trait due to the negotiation component of the game.

The agent-models described next have in common that they follow the same objectives:

1. Building an estimate of a humans' personality observing the humans behaviour in the CT game, which is the task of Automatic Personality Recognition that was introduced and will be used to answer the first research question asked for this chapter; and

2. Using this estimate to make informed decisions while playing the game, which is the task of predicting the possible course of action of the human and will be used to answer the second research question asked for this chapter.

To build an estimate about the personality traits of the human the agent i refers to a human k using the set $P_k = \{p_c^k, p_e^k, p_a^k\}$, where each $p \in P$ represents a personality trait (p_c – conscientiousness, p_e – extraversion, p_a – agreeableness). As the traits within the FFM are characterised using a continuous scale, the range of each p is $[0, 1]$ and the initial value is set to 0.5. This set is the main feature used by i to build the expected utility of taking action $a \in A$ in the current state $s \in S$ playing against k, which we denote as $u : S \times A \times P$. To improve the estimate of the personality the agent adapts each p during the interaction. In favour of this task, we developed two different approaches that are described next. The first approach learns about the personality of the human online, *i.e.* using the observations directly to adapt the personality estimate. The second one applies a classical Naive Bayes (Russell and Norvig, 2002, pp. 716–724) (NB) classifier

to the very same environment and is used to compare the results of the first approach with this commonly used supervised learning technique.

8.3.1. Adapting to Human Personality

The first and arguably simpler model uses different equations to build estimates from the observation of the human's actions within the game. In the following, I will describe the construction of the model and substantiate the design-decisions that have been made.

The distal cues available to both agent-models are restricted to the actions taken by the human. The action-space comprises the action to do nothing, to make proposals to exchange chips, to accept or refuse proposals made by the other, the actual exchange of chips, and the movement of the player. All of these actions, except the latter which is automatically performed by CT, can be used to reason about different facets of the human's personality, e.g. how beneficial a proposal is for oneself or the opponent or if the human sticks to reached trade agreements or frauds the agents.

To adapt the estimates for the individual personality traits we apply the *one cue–one trait process* (Funder, 1995, p. 659), in which we use observations of single behaviours of a single cue to build knowledge about a single trait as follows:

- p_c — denotes the estimate of the conscientiousness of the human and is interpreted as how reliable a player is.

  Therefore, fulfilling a trade agreement increases and not fulfilling it decreases this value. As failing to predict the reliability of a player can lead to significant score losses for the agent, this trait is of utmost importance. To update the estimate after each trading agreement, we compute the conscientiousness of a human by increasing/decreasing it with a constant factor $x_c \in [0, 1]$ using the following equation:

$$p_c \leftarrow \begin{cases} p_c + x_c & \text{if successful exchange} \\ p_c - x_c & \text{if successful exchange but fraud} \\ p_c - 2 \cdot x_c & \text{if fraud} \end{cases} .$$

  The first case applies when the proposed set of chips is equal to the one received. The second case applies when the set of proposed and received chips is not equal, but in the set of received chips exist some chips that are useful for the agent. The last case applies if the agent was fooled. This is the case when there is no exchange or when the agent only receives useless chips. Thus bailing out on an agreed trade is punished harder, as it is a greater break of trust and might critically damage the agents chance to reach the goal square.

- p_e — denotes the estimate of the extraversion of the human and is interpreted as how contact-friendly the player is.

  Therefore it is increased when the player makes a proposal of exchanging chips, which is the most extroverted actions possible in the game. It is decreased when the player acts passively not proposing anything. To update the estimate after each round, we compute the extraversion of a human by increasing/decreasing it with a constant factor $x_e \in [0, 1]$ using the following equation:

$$p_e \leftarrow \begin{cases} p_e + x_e & \text{if proposed and} \\ p_e - n \cdot x_e & \text{otherwise} \end{cases} .$$

  The first case applies when the player offers a proposal, the second case otherwise. The multiplicator n is growing until the player offers something and depicts the number of rounds played:

$$n \leftarrow \begin{cases} 0 & \text{if proposed} \\ n + 1 & \text{otherwise} \end{cases} .$$

- p_a — denotes the estimate of the agreeableness of the human and is interpreted as how friendly/altruistic a player is.

  Therefore it increases when the player accepts offers and decreases when the player declines an offer. Furthermore, the reward for an acceptance is reinforced if the proposal is favourable to the agent, *i.e.* rewarding an altruistic action twice as much. On the other side, the estimate is decreased twice if the not accepted proposal was indeed favourable for the agent. To update the estimate after each active communication phase, we compute the agreeableness of a human by increasing/decreasing it with a constant factor $x_a \in [0, 1]$ using the following equation:

$$p_a \leftarrow \begin{cases} p_a + 2 \cdot x_a & \text{if accepted and altruistic} \\ p_a + x_a & \text{if accepted} \\ p_a - x_a & \text{if not accepted} \\ p_a - 2 \cdot x_a & \text{if not accepted but favourable} \end{cases} .$$

  This equation rewards generous offers and exchanges as they might be harmful to the player's score. To analyse if the acceptance or non-acceptance was altruistic/favourable requires the CT environment to be fully observable as shown in

Fig. 8.3. At the same time, it punishes the agreeableness estimate when the exchange of important chips was declined. Thus, the level of agreeableness is a kind of measure of the selfishness of the player.

The constants x_c, x_e and x_a were adjusted and determined in test-games played prior the experiment. For reasons of readability the edge-cases when the estimates reach the minimal/maximal value of the interval are omitted within the formulas. In these cases, a positive/negative adjustment is no longer applied.

Using the personality estimates

As these parts of the agent model are used to build the estimates about the human's personality, we will next describe how the personality information is used in terms of predicting the human's most likely course of action. As mentioned above, the personality estimates are used within the decision-making by calculating a utility for each action. To do so, the estimates p_e and p_a are utilised to calculate the expectation that a proposal will be accepted, as weighted sum $e^{acc} = p_e \cdot w_e + p_a \cdot w_a$. The weights are used to adjust the influence of the traits. A second value indicates the expectation whether an agreed exchange indeed takes place and is represented as $e^{exc} = p_c$.

The second feature to build the expected utility is the score that is reachable with the current set of coloured chips (r^c), the score that is reachable after a successful trade (r^t) and the score that is reachable falling for a betrayal (r^f). Here betrayal means accepting a trade and transferring own chips without getting the promised response. All three can be easily calculated when knowing (1) that CT controls the movement phase by applying the Manhatten distance algorithm to determine the current best option to move towards the goal square and (2) the scoring function of the game, which sums the following parameters:

- 100 points for reaching the goal square and ending the round as winner;
- 5 additional points for all coloured chips left; and
- 10 penalty points for each tile between the final position and the goal square calculated using the Manhattan distance.

Both features are then used to calculate the expected value (reward) of executing action a given the current state of the game s using the following multi-attribute utility function when making a proposal:

$$u_a^i(s, P_k) = e^{acc} \cdot e^{exc} \cdot r^t + (1 - e^{acc}) \cdot r^c + \\ e^{acc} \cdot (1 - e^{exc}) \cdot r^f.$$

When the agents receives a proposal the likelihood that it will be accepted is not of relevance since the agent can choose its answer and only has to consider that the exchange

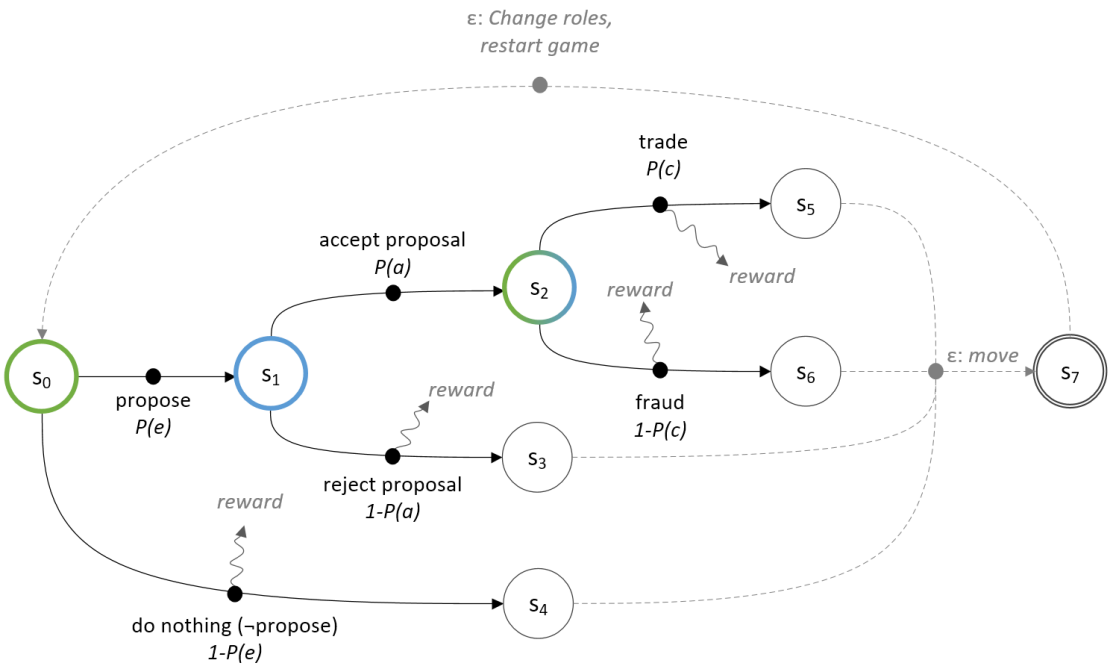


Figure 8.4.: Illustration of the CT environment using the notion of MDPs (transitions starting in green states are conducted by proposing players, transitions starting in blue states are conducted by responding players). Automatically performed actions are marked with an ϵ. The game is restarted with swapped roles after each round.

truly takes place. Therefore we remove e^{acc} when building the utility for an action in this case.

Given this function, the estimate of the personality of the human influences the policy of the agent, which tries to maximise the utility. That means that for each agent i playing against a human k an optimal action a_k^* exists, that maximises the utility in state s where $a_k^* \in \underset{a}{argmax}\, u^i(s, P_k)$. If equally valued actions exist, the one is selected that was found first. Indeed, in the implementation, the agents have no knowledge that there exist more than one action that maximises the utility. A more elaborate behaviour here would be to evaluate whether a chain of actions would lead to a higher score, leading to an agent which acts 'farsighted' instead of 'myopic' (*cf.* Section 3.2).

8.3.2. Bayesian-based Estimation of Human Personality

The second model applies Bayesian techniques to the same task and environment. To receive estimates of the humans' personality the agent is interacting with, we again apply the one cue–one trait process, following the argumentation that a humans' behaviour depends on the personality and the rewards which can be received. Fig. 8.4 depicts these assumptions using the notion of a MDP as introduced in Section 3.2. It highlights the possible state transitions for the players, the starting state being the proposer's position

to create a proposal in a one-shot round of the CT game. The move actions and the next round of the game with swapped roles are started automatically, after the process of making a proposal or not ($s_0 \rightarrow s_1 \rightarrow \ldots$ or $s_0 \rightarrow s_4$), accepting the proposal or not ($s_1 \rightarrow s_2 \rightarrow \ldots$ or $s_1 \rightarrow s_3$), and trading the agreed chips (trade) ($s_2 \rightarrow s_5$) or some other selection of chips (fraud) ($s_2 \rightarrow s_6$) are finished. Several actions lead to a reward, though, the most beneficial policy is the one that leads to a successful exchange (either trade or fraud) as the scoring function for the games played reads as follows:[5]

- each player starts with 100 points;
- 50 points are granted for reaching the goal square and ending the round as winner;
- 10 additional points are granted for all coloured chips left; and
- 20 penalty points for each tile between the final position and the goal square calculated using the Manhattan distance.

Fig. 8.4 also highlights the linkage between the personality traits and the actions. Again interpreting the meaning of the trait taking into consideration the effect of the action:

- $P(c),\ 1 - P(c)$ — denotes the likelihood that a player acts reliable, which is indicated by the conscientiousness of the human. As discussed before, we interpret this trait as how likely an agreed trade will take place $P(trade \mid c) \mathrel{\widehat{=}} P(c) \mathrel{\widehat{=}} c$ and how likely it is that the agents falls for a betrayal $P(fraud \mid c) \mathrel{\widehat{=}} 1 - P(c) \mathrel{\widehat{=}} 1 - c$.
- $P(e),\ 1 - P(e)$ — denotes the likelihood that a player is contact friendly, which is indicated by the extraversion of the human. As before, we interpret extraversion as how likely it is that a player offers a proposal $P(propose \mid e) \mathrel{\widehat{=}} P(e) \mathrel{\widehat{=}} e$ or not $P(\neg propose \mid e) \mathrel{\widehat{=}} 1 - P(e) \mathrel{\widehat{=}} 1 - e$.
- $P(a),\ 1 - P(a)$ — denotes the likelihood that a player acts friendly/altruistic, which is indicated by the agreeableness trait of the human. As before, we interpret agreeableness as how likely a player accepts an offer $P(accept \mid a) \mathrel{\widehat{=}} P(a) \mathrel{\widehat{=}} a$ and how likely an offer is declined $P(reject \mid a) \mathrel{\widehat{=}} 1 - P(a) \mathrel{\widehat{=}} 1 - a$.

Given these links, we conduct an experimental protocol where we first play 30 games against a human recording its behaviour and use the generated training data set to train a Naive Bayes classifier for each human. Thus we follow the assumption of NB, that the attributes we want to use are conditionally independent of each other, given the action space. To handle the relatively small sample size for each game, we apply a maximum-likelihood method.

[5] There are several edge cases. For example, a proposer could go without trading when he owns all chips necessary to reach the goal tile, though, proposing something and handing over nothing would be a beneficial action in the short term to increase the final score for this game.

Using the personality estimates

In the second stage, the agent uses the classifier to calculate the expected value (reward) of executing an action a in the current state. We use two utility functions for the two different positions an agent plays: proposer and responder. For the proposing case, we have to take into account the cases where the human does not accept a proposal, the human accepts a proposal and the trade takes place, and the human accepts the proposal but it dupes the agent during the trade. Again we refer to these facets with the currently reachable score (r^c) and the reachable score after a successful trade (r^t) respectively fraud (r^f). Furthermore, we introduce the weights w_0, w_1, w_2 to control the behaviour of the agent as acting more or less optimistic and more or less risky, balancing between proposals that are fair and acceptable for both sides vs. proposals that promise maximum score. For the proposing case the utility function reads as follows:

$$u_a^i(s, P_k) = \underbrace{(1 - P(a)) \cdot w_0 \cdot r^c}_{no\ trade} + \underbrace{P(a) \cdot P(c) \cdot w_1 \cdot r^t}_{accept,\ trade} + \underbrace{P(a) \cdot (1 - P(c)) \cdot w_2 \cdot r^f}_{accept,\ fraud}.$$

For the responding case the agent only needs to take the likeliness into account that human frauds it during the trade, which shortens the utility function to the following:

$$u_a^i(s, P_k) = \underbrace{P(c) \cdot w_1 \cdot r^t}_{trade} + \underbrace{(1 - P(c)) \cdot w_2 \cdot r^f}_{fraud}.$$

Giving these utilities the personality estimate influences the policy of the agent trying to maximise the utilities. This means, that the agent selects the action with the highest utility, similar to the behaviour discussed in the other agent-model.

In the following, I will introduce the experimental setup in detail and discuss the results we reached w.r.t. the questions about learning personality and using this knowledge.

8.4. Experimental Setup and Results

For the experiment, we implemented the introduced agent-models for the CT environment and invited under-graduated students to play against our agents.[6] In the beginning, the participants were asked to self-assess their personality using established assessment questionnaires (as described in the next sections). Afterwards, the game environment was explained, and each participant got a 10-minute tutorial to get familiar with the

[6] Major parts of the implementation of the first model were part of the bachelor thesis presented by Breitung (2014), supervised by the author. Major parts of the implementation of the second model were part of the bachelor thesis presented by Ly (2015), supervised by the author.

environment, the game rules, and the control elements. Here we explained the rules, the scoring function and played the game in practice with the subjects. The scoring-function consisted of reaching the goal, the distance to the goal and the chips left as described in Section 8.3.1 and Section 8.3.2, respectively. In this initial stage, the participants played against an agent that did not adapt to the opponent. Afterwards, the attendees played 30 games in a row against the adapting agent and 40 games in a row against the Bayesian agent. The latter applied the learned personality estimates during the last 20 games and used the first 20 games to build the training data set used to train the classifier. The goal of the participants was to reach the maximum score in as many games as possible. The participants were not explicitly informed about our intention to learn about their personality, but were told that we develop an AI to play CT and want to test it.

8.4.1. Adapting Agent

In the experiment with the adapting agent 22 participants took part. To assess the personality of the participants we used a 100-Item questionnaire derived from the IPIP[7]. After collecting the data, we compare the personality estimates that have been build by the agent and the estimates derived from the self-assessment. Table 8.1 lists the data. The scoring results listed in column 2 and 3 show the mean value of the points of all 30 games determined for each human player and the agent playing against the participant. It shows that the agent outperforms the human players in average, but the difference is fairly small. Setting up the CT environment, we made the single-game rounds more comparable by distributing the same amount of chips to the opponents and centralising the goal square. Taking that and the total number of 660 games played into consideration the scoring difference could be interpreted as an indication that we actually can use the personality information to make informed decisions. However, a detailed discussion of the figures is required to come up with a comprehensive conclusion. To do so, we will discuss both research questions separately next.

Can agents use our model to learn the personality traits of a human during the interaction with this human?

Table 8.1 lists the deviation between the agent's estimate of the personality traits (column 3 to 5) of its opponents and the actual self-assessed personality characteristics derived from the questionnaire including the average deviation. It shows that least variety exists with the extraversion parameter, while agreeableness and conscientiousness are drifting apart stronger (Fig. 8.5 depicts the spreading of the values in a boxplot chart). A zero value here would mean that both characterizations are perfectly equivalent, which is

[7] IPIP — International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality and Other Individual Differences — `http://ipip.ori.org/`. For the experiment, the 100-Item Set of IPIP Big-Five Factor Markers has been used. The complete questionnaire can be found in Appendix B.

Table 8.1.: Adapting agent – Listing of the average scores reached by the opponents (human and agent) within the games and the average score (μ) and deviation (σ) over all games (column 2 and 3). In addition, listing of the deviation between the agents estimate of the humans personality trait and the one derived from the questionnaire (column 4 to 6). At the bottom the Pearsons r and Spearmans ρ correlation coefficients between the reached scores and the deviation of the personality estimates.

#	**Human**	**Agent**	**Extraversion**	**Agreeableness**	**Conscientiousness**
1	111	99	0.18	0.225	0.265
2	107	154	0.09	0.09	0.12
3	98	113	0.02	0.28	0.245
4	121	127	0.075	0.025	0.215
5	118	140	0.035	0.06	0.4
6	105	113	0.05	0.19	0.175
7	132	134	0.03	0.235	0.09
8	100	107	0.14	0.335	0.24
9	88	154	0.015	0.05	0.425
10	142	102	0.045	0.225	0.11
11	104	106	0.055	0.195	0.075
12	105	112	0.07	0.295	0.37
13	99	144	0.06	0.17	0.425
14	121	120	0.065	0.19	0.12
15	126	111	0.16	0.095	0.215
16	145	137	0.04	0.05	0.22
17	86	141	0.025	0.215	0.065
18	102	107	0.015	0.06	0.13
19	138	132	0.145	0.075	0.47
20	154	110	0.05	0.21	0.275
21	101	124	0.02	0.165	0.215
22	97	138	0.125	0.285	0.23
μ	**113.64**	**123.86**	**0.07**	**0.17**	**0.23**
σ	**15.84**	**14.77**	**0.04**	**0.08**	**0.09**
The correlation coefficients between the reached scores and the deviation of the personality estimates are listed below.					
		Pearsons r	-0.23	-0.11	0.16
		Spearmans ρ	-0.32	-0.06	0.09
			$p > 0.1$	$p > 0.1$	$p > 0.1$

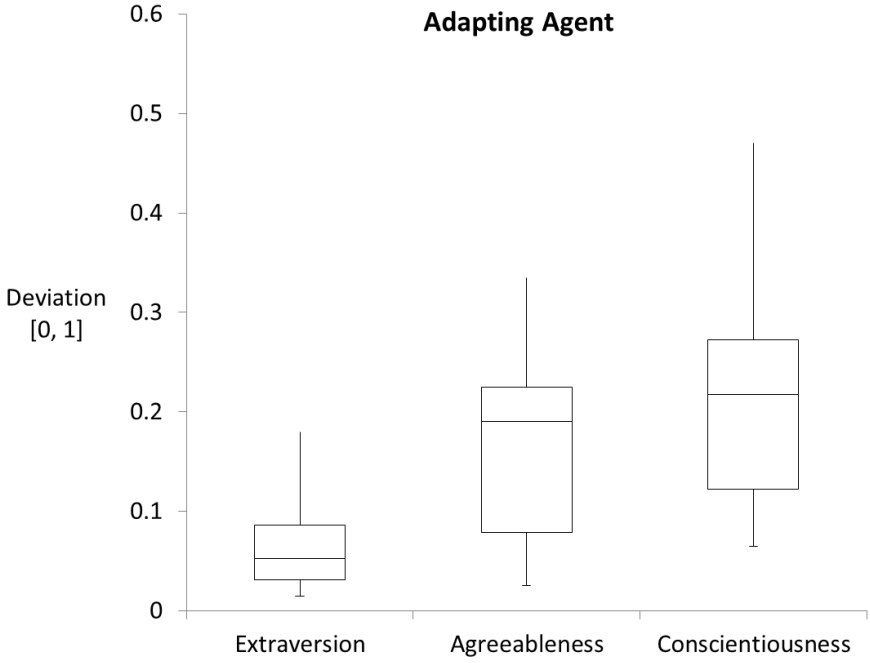


Figure 8.5.: Adapting agent – Boxplot of deviation between questionnaire and agents estimate of the player's personality.

only a theoretical option, which cannot be reached when comparing the results of different assessment instruments, e.g. self-assessment, questionnaire, professional assessment, or psychological interviews.

To get an indication of the accuracy of the reached results we have to calculate the ecological validity ρ_{EV}, which is the Spearman's ρ (McDonald, 2014, pp. 210–214) between the ranked values of the agents' personality estimate and the self-assessed personality. The ecological validity shows to which extent the findings of an experiment could be generalised to real-life settings, *i.e.* it provides a measure that answers the question whether or not we can use observations from artificial laboratory settings in more natural environments (Schmuckler, 2001, p. 419). For the deviation between the self-assessed personality and the agents' estimate of the personality the ecological validity for the traits is $\rho^e_{EV} = 0.88$, $\rho^a_{EV} = 0.36$, and $\rho^c_{EV} = 0.3$.[8] These values substantiate what is already visible in Fig. 8.5, showing that the results for extraversion are promising whereas the results for the other traits have to be improved.

One way to improve the results for agreeableness and conscientiousness would be to adapt the adaptation of the estimation function for the actions, *i.e.* our interpretation of the link between the actions and the agreeableness and conscientiousness trait is

[8]The Spearman's ρ measures the monotonic correlation of the ranked values and becomes a value between +1 and -1, where +1 marks a perfect monotone increasing correlation, 0 marks no correlation, and -1 marks a perfect monotone decreasing correlation.

not accurate. Another solution would be to use the same interpretation, and adaptation mechanism then applied for extraversion. However, several researchers showed that individual mechanisms for each distal cue and each trait are more promising (Vinciarelli and Mohammadi, 2014, pp. 283–284); substantiating our design decision to use individual mechanisms.

To conclude, the results show that we can learn about the personality of a human in direct interaction and that some characteristics of a personality can be learned more accurately/easily than others. From the experience we gained during the development and substantiated by the related work, we can conclude that one needs streamlined learning mechanisms for the distal cues that are observable in an environment.

Can we use personality information directly to make informed decisions about the potential behaviour of human?

The hypothesis to test our second research question/second objective read as follows:

- H_0 : *The agent does not perform differently, i.e. the agent's score does not increase/decrease, if the personality estimates become more accurate.*
- H_1 : *The agent performs better, i.e. the agents score increases if the personality estimates become more accurate.*

The last two rows of Table 8.1 show the Pearsons r and Spearman's ρ (McDonald, 2014, pp. 210–214). Due to the sample size, the Pearson's r is not that meaningful and can only be used to give an indication. This indication is substantiated with the Spearman's ρ, which is better suited for relatively small sample sizes. Both values show that the agent performs the worse, the more accurate the personality estimate becomes for extraversion and agreeableness (negative correlation). On the other hand, the agent performs better the more accurate the conscientiousness becomes (positive correlation). Here a zero value for either correlation score means that the scoring ranks do not correlate with the ranked personality estimates;[9] in other words, as the scoring ranks increase, the deviation of the personality estimates do not increase (or decrease). As neither r nor ρ become zero for any of the traits we cautiously reject the null hypothesis H_0 as the coefficients show a weak correlation. We cautiously reject it as the significance level of 0.1 is not reached by any of the traits. Thus, for the adapting agent, we have to conclude that the applied utility function does not use the personality information to predict the human's course of action in the way it was intended.

In the following, I will present the results reached within the experiments using the Bayesian agent.

[9]The Pearsons r measures the linear correlation of values and becomes a value between +1 and -1, where +1 marks a perfect positive correlation, 0 marks no correlation, and -1 marks a perfect negative correlation.

8.4.2. Bayesian agent

Within the experiment with the Bayesian agent, 10 other participants took part. For the personality assessment, we used a 30-Item questionnaire provided by Satow (2012). We selected another questionnaire as the applied maximum-likelihood method requires prior knowledge about the distribution of the arguments. These values are given by Satow (2012) and read as follows: $p_e = 0.65$, $p_c = 0.65$, and $p_v = 0.75$.[10] These values are not available for the IPIP questionnaire used in the first experiment, neither was the requirement known while performing the first experiment. However, the results remain comparable, as both questionnaires assess the Big Five factors with reasonable reliability coefficients. The remaining setup stayed the same. Again, we will discuss both research questions separately next.

Table 8.2.: Bayesian agent – Listing of the average scores reached by the opponents (human and agent) within the games and the average score (μ) and deviation (σ) over all games (column 2 and 3). In addition, listing of the deviation between the agents estimate of the humans personality trait and the one derived from the questionnaire (column 4 to 6). At the bottom the Pearsons r and Spearmans ρ correlation coefficients correlation coefficients between the reached scores and the deviation of the personality estimates..

#	Human	Agent	Extraversion	Agreeableness	Conscientiousness
1	112	113	0.095	0.47	0.115
2	139	110	0.06	0.225	0.03
3	113	113	0.035	0.225	0.06
4	115	123	0.175	0.155	0.195
5	116	110	0.18	0.155	0.075
6	129	129	0.07	0.03	0.175
7	115	119	0.215	0.43	0.16
8	129	127	0.39	0.12	0.3
9	125	108	0.13	0.3	0.035
10	135	111	0.05	0.05	0.175
μ	**122.6**	**116.2**	**0.14**	**0.22**	**0.13**
σ	**9.28**	**7.29**	**0.08**	**0.11**	**0.07**
The correlation coefficients between the reached scores and the deviation of the personality estimates are listed below.					
		Pearsons r	0.38	0.21	0.41
		Spearmans ρ	0.28	0.09	0.47
			$p > 0.1$	$p > 0.1$	$p > 0.1$

[10] Values are derived from a sample size of 5520 with a Cronbach's Alpha between .76 and .87 (Satow, 2012, p. 20).

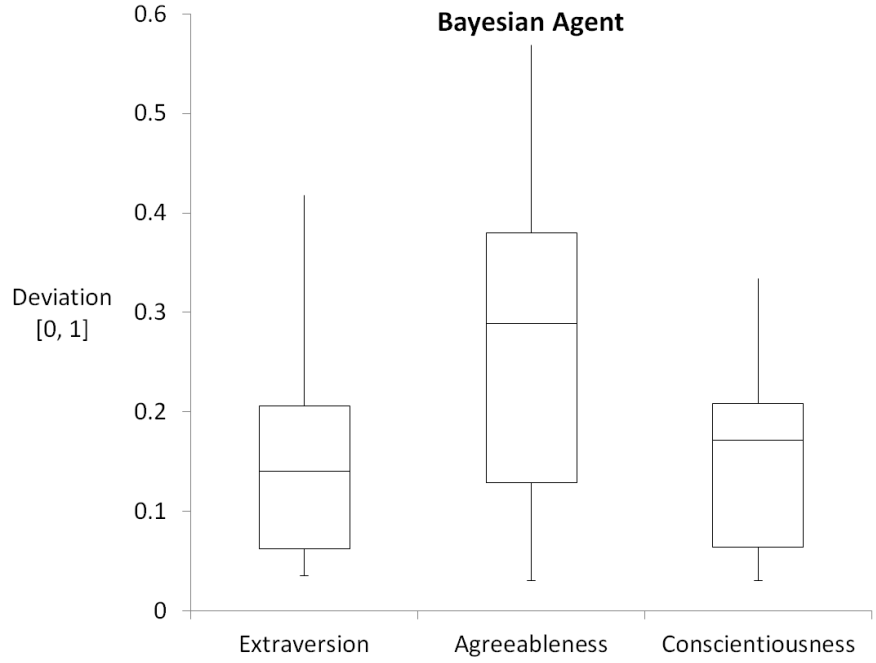


Figure 8.6.: Bayesian agent – Boxplot of deviation between questionnaire and agents estimate of the player's personality.

Can agents use our model to learn the personality traits of a human during the interaction with this human?

Table 8.2 lists the data recorded for the Bayesian agent. In column 3 to 5, the deviations between the personality assessment and the personality estimate of the agent are listed. The least variety exists with the conscientiousness trait, though, extraversion shows a closely related value. The average difference and the spreading for the agreeableness trait are much bigger. Fig. 8.6 visualises the results as a boxplot. Due to the fairly small sample size, the ecological validity has to be interpreted quite carefully. For each trait it reads as follows: $\rho^e_{EV} = 0.83$, $\rho^a_{EV} = 0.63$, and $\rho^c_{EV} = 0.84$. This indicates that the Bayesian agent might work more accurate than the adapting agent, e.g. given another population or another environment. Given much bigger sample sizes, the related work shows several approaches that have applied the NB classifier successfully to the task of automated personality recognition (Vinciarelli and Mohammadi, 2014, pp. 277–282). In contrast to the adapting agent, the Bayesian agent applied the same approach to each distal cue, eliminating one dynamic factor. Still, the figures reveal a difference between the individual traits. This substantiates the prior made statement that one should use different streamlined techniques for each distal cue that can be observed.

Can we use personality information directly to make informed decisions about the poten-

tial behaviour of human?

Column 2 and 3 of Table 8.2 list the scoring results including mean and deviation. In contrast to the adapting agent, the human outperformed the Bayesian agent in more than half of the games. Indeed the humans were successful in 224 of the 400 played games. Taking the same hypotheses as above into account, we can again reject the null hypothesis H_0 as shown by the correlation coefficients. As they are positive for each of the traits, it seems promising to repeat the experiment with a stronger null hypothesis, which could be read as follows: The agent does not perform better, i.e. the agents score does not increase, if the personality estimates become more accurate, i.e. if the deviation of self-assessed personality and the agents estimates decreases. The strongest correlations exist for the conscientiousness trait. However, neither of the correlations reach an acceptable significance level, and the number of participants is fairly small. Thus, as well as for the adapting agent we conclude for the Bayesian agent that the applied utility function does not use the personality information to predict the human's course of action in the way it was intended.

In the following, we will discuss the results in more detail before the chapter is concluded.

8.5. Discussion

The results that have been described are mixed. On the one hand, they show that we can learn about the personality in direct interaction with the human using a relatively simple approach or using only a few observations to train a classifier. On the other hand, we learned that the linkage between the observable actions and the expressible and interpretable distal cues is crucial. Furthermore, the results show that we were not able to use the learned personality estimates given the described utility functions.

Since the CT game has very limited action space available to evaluate and analyse the behaviour of the other player, it is complicated to associate these actions with factors of the FFM. Thus, the environment/action-space might have to be more complex. Another possible explanation for the estimates not depicting the assessment results is that a person might behave differently to the interpretation we had while associating actions with traits. Since the goal of the game is to reach the best possible score, it might be beneficial to use a more generalised trait just indicating how cooperative the human is (*cf.* Talman et al., 2005). Despite these hindrances, the outcome is still good, especially for the value of extraversion.

It can be concluded, that some of the traits can be estimated more accurately than others, at least in the CT environment as we modelled it. During the experiment, we followed the relevance and accuracy features. Leading to the focus on three out of the five traits of the FFM. It was shown that we were able to more accurately learn the

extraversion trait than the agreeableness and conscientiousness trait. There are multiple reasons for that. One is based on the distal cues itself, *i.e.* are those cues observable and how well does the physical cue encode the characteristics of the personality traits. This is known as the *good trait variable of accuracy* (Funder, 1995, p. 662), which describes the possibility that some traits might be easier to estimate accurately than others as some behaviours might be easier to judge than others. To approach this problem, we provided and discussed justifications how and why we made design decisions in the agent model. It is just right, to assume that we made errors during this process coming up with an inaccurate agent model and that better results can be achieved with another one. One potential problem here is an oversimplification of the effects of the personality traits on the distal cues. Another reason can be found in the perception module, *i.e.* how we used observed behaviours to generate personality estimates. This is known as the *good judge variable of accuracy* (Funder, 1995, p. 660), which describes the possibility that different individuals are differently good in judging personality. Although Funder (1995) writes about human judgement, these variables apply to machine learning as well as shown by the different approaches and their reached accuracies on same data sets in (Vinciarelli and Mohammadi, 2014, pp. 277–282). To approach this problem, we applied two different approaches within the same setting. Both of them show the same tendency. That is, in the selected environment using the identified distal cues, extraversion can be estimated more accurately than agreeableness and agreeableness can be estimated more accurately than conscientiousness.

The above-discussed points are directly related to the presented approach. Besides that, we can identify limitations w.r.t. the conducted experiments and result analysis. First of all, the sample size for both experiments is limited ($n_{adapting} = 22$, $n_{bayesian} = 10$) making it possible that the results are random, biased w.r.t. the sample, or related to other (non-human) factors (Hunter and Schmidt, 2004, pp. 3–13). Other common method biases such as the measurement and item context effects may induce further weaknesses (Podsakoff et al., 2003, pp. 881–885). Within the experimental design, I addressed this by separating the measurement of the self-assessed personality and the agents' estimation using a cover story that disconnected both parts.

Overall, although the results are mixed, we argue that the presented findings provide interesting insights into the task of learning about personality traits in interaction with a human. That is because very little is known about this task as shown in the related work section. Thus, even imperfect information and approaches offer valuable insights at this stage.

8.6. Conclusion

This chapter presented two agent-models approaching the task of Automatic Personality Recognition, *i.e.* the task of recognising the true personality of an individual. Our agent-

model approach this task by using observations of the human behaviour made while repeatedly playing rounds in the Colored Trails Game. In the beginning, I classified my work w.r.t. the state-of-the-art and elaborated that the related work concentrates on learning personality traits using supervised approaches. However, in several domains, the requirement of having labelled training data sets is not satisfied. The chapter further introduces the CT and relevant studies performed in this environment for empirical behaviour research, before describing my agent-models in detail. Both of them are based on the idea to link the personality traits to the available actions by interpreting the meaning of the trait w.r.t. the actions effect. Thus, they transfer the approach for modelling personality described in Chapter 7 to the APR task. The first agent-model adapts its personality estimates during the interaction, whereas the second one builds a data set during the interaction and learns a classifier using this relatively sparse data.

Based on the experiments performed, I finally can answer the first question that was presented in the beginning, namely: "Can agents use our model to learn the personality traits of a human during the interaction with this human?"

The results indeed show that we can learn personality information from observations using relatively sparse data, though, the results reached for the different traits are differing in accuracy. The extraversion trait reached the best accuracy in both models, while the results for agreeableness and conscientiousness are varying. My conclusion here is, that at least in the CT environment some traits are easier to observe and easier to learn than others. The related work substantiates this conclusion and recommends to use individual recognition techniques for each distal cue that can be observed (*cf.* Vinciarelli and Mohammadi, 2014). I followed these recommendations and further applied the Realistic Accuracy Model to substantiate the presented design decisions. The crucial point for the used approach is the linkage between the observable actions and the expressible and interpretable distal cues. In consequence, one action to improve the results is to fine-tune this interpretation. I discussed additional improvements in the prior section.

Based on the successful learning of personality characteristics I can provide an answer to the second question that I formulated at the beginning of the chapter, namely: "Can we use personality information directly to make informed decisions about the potential behaviour of human?"

In fact, no, at least not using the described utility functions in the CT Game. Although the results showed that the agents can outperform the humans, the statistical analysis revealed that both utility-based approaches were not working as intended, *i.e.* the agents were not able to use the learned personality characteristics to make informed decisions in terms of predicting the future behaviour of the opponents. The analysis showed that the results were not related to the personality. I provided a discussion that identifies several rationales and points-out future work directions. The results reached in this experiment are substantiated by recent literature. Du (2013); Du and Huhns (2013) describe an experiment investigating if a humans behaviour towards other human

beings and agents is related to their personality types according to MBTI and the KTS-II (Keirsey Temperament Sorter-II) test. For the experiment, the authors used the cake-cutting game where the participant one human, one simulated human and one agent played three games in a row. As participants, the study used undergraduate students that in the beginning had to fill-up the KTS-II test. The evaluation is extensive using several statistical validation instruments to prove the hypotheses whether personality and behaviour are independent or not. The authors show that the different KTS-II types act differently towards other humans and agents. However, the most interesting aspect of the work is the finding that there is no relation between personality and the actual decision-making of a human when interacting in the cake-cutting game. That is a point which falsifies the question whether we can use personality information to predict behaviour — substantiating the findings presented within this chapter. However, the authors already mentioned that the experimental setup might be too small to observe such effects and some of the effects are close to the statistical significance boarder.

In the next chapter, I concentrate on the formalisation of personality within the agents' decision-making.

9. Reasoning about Personality

We already learned about existing work on personality in agent-based systems, though, the requirement for having an underlying formalism is not satisfied yet. This requirement persists through the ongoing discussion in psychology about the existence and definition of personality traits (*cf.* Section 3.4), often leading to the use of subjective explanations of terms in the agent-based literature. As the agent-community requires clear definitions, which are underpinned with clear semantics, this requirement justifies the reason to integrate personality as mental attitude into one of the BDI logics (that is to say logics of *Belief*, *Desire*, and *Intention*, e.g. Cohen and Levesque, 1990; Herzig and Longin, 2004; Rao and Georgeff, 1998; Wooldridge, 2000) available in the literature. Hence, the goal of this chapter is to bring together our prior presented model of personality and the formal model of BDI reasoning into one formalisation. Fig. 9.1 highlights the parts of an adaptive system, which I will work on within this chapter. It highlights the focus on the Data Processing layer, particularly that the formalisations objectives contribute to the inferencing and reasoning part of the layer. In doing so, the objective of the formalisation is twofold: (1) to enable reasoning about the influence of a personality on the behaviour of an agent (reasoning), and (2) to enable reasoning about characteristics of a personality using observations of behaviour (inferencing). The main question to answer is: *How to represent the state and effect of personality in BDI logics?*

The approach is based on the '$\mathcal{L}$ogic $\mathcal{O}$f $\mathcal{R}$ational $\mathcal{A}$gents'. $\mathcal{LORA}$ is a multi-modal, branching-time logic of Belief, Desire, and Intention presented by Wooldridge (2000). This section is then structured as follows. We will start with a discussion of the related work in Section 9.1. Here, I will shortly recap the relevant work that introduced previously before concentrating on work that brings together all cognitive characteristics (personality, mood, emotions) and contributions that formalise such aspects. After-

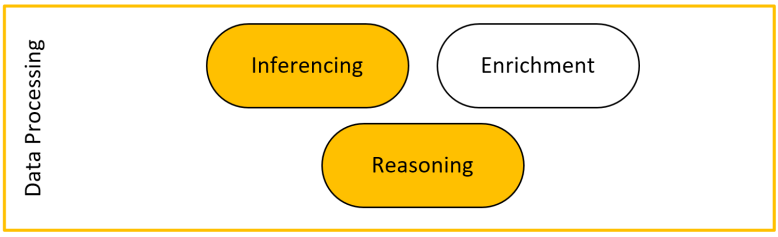


Figure 9.1.: Parts of an adaptive system that are discussed within this chapter.

wards, we will discuss the objectives of this chapter in more detail in Section 9.2. Here, I will also provide a running example, which will be used in the subsequent sections to clarify the provided arguments. In Section 9.3, I will provide an introduction to $\mathcal{LORA}$. Although this can be done in a nutshell only, it will provide all information to comprehend the other parts of the chapter. Afterwards, in Section 9.4 the existence and the influence of personality in BDI agents will be inserted into the logical framework by minimally extending the syntax and semantics of this logic with a new modal connectivity representing the personality of an agent. Having done this, we satisfied the basic prerequisite of representing personality as part of the sense-plan-act cycle of rational agents, providing the foundation for reasoning about personality influences. It follows, a detailed discussion of the properties (e.g., the pairwise interaction between the modalities, inevitability, options, weak and strong realism) that are useful for rational agents in Section 9.5. Here, I look into the question *Which relations between the personality modality and the belief, desire, and intention modality are meaningful for reasoning about rational behaviour?* I summarise the most important results of this discussion in Section 9.7 concluding the chapter.

9.1. Related Work

In the following, I discuss related work that introduces personality models in the scope of agent-based systems. As there are multiple overlaps with the already presented related work, this will be a short recap. The section concludes with a discussion on the applicability of these approaches as a formalisation of personality in the general decision-making process of agents. Here I compare the existing approaches to the work that has been presented on the formalisation of another cognitive aspect in agents, namely emotions.

9.1.1. Personality and Agents

Models of human personality are used in the implementation of (microscopic) traffic simulation frameworks (*cf.* Lützenberger and Albayrak, 2014) and the agent-based simulation/visualisation of groups of people (*cf.* Canuto et al., 2006; Durupinar et al., 2008). The work of Durupinar et al. (2008) shows how the introduction of different personalities into single agents influences the behaviour of a crowd. Other areas include human-machine interaction (*cf.* Dryer, 1999), in particular conversational agents/virtual humans (*cf.* Allbeck and Badler, 2002; Egges et al., 2004) and life-like characters (*cf.* André et al., 2000). The latter outlines three projects that apply two dimensions of FFM (extraversion and aggreeableness). The effects are interpreted in a rule-based or scripted manner.

Another branch of research focuses on modelling and examining the effects of personalities on interactions between agents and their environments. In particular, the effects

of personalities in cooperative settings. These are discussed in the remainder of this section.

Talman et al. (2005) illustrate the use of a simple abstraction of personality types. Personalities of agents are determined through the two dimensions cooperation and reliability and used to measure the helpfulness of an agent in the setting of a game. During multiple games, the agents reason about each other's helpfulness and try to customise their behaviour to the different personalities.

Campos et al. (2009) integrate a partial version of MBTI into a BDI agent. They show that personality characteristics influence the decision-making process in a simulation specifically designed for the papers' use-case.

In an early work, Castelfranchi et al. (1998) investigate the effects of personality on social interactions between agents, such as delegation and help. They apply opponent modelling via personality traits to motivate interactions. The traits are discussed as an abstract concept without relation to psychological theories.

The work presented by Salvit and Sklar (2011, 2012) experimentally validates the impact of the MBTI on the decision-making process of agents. MBTI is integrated into a sense-plan-act structure, and the behaviour of each MBTI type is analysed in a simulation environment called the 'Termite World'. The results show that different personality types act in a variety of ways.

9.1.2. Discussion

As discussed above several authors already utilise personality models in different domains. The implemented effects of personality are often specific for the considered use-case and not applicable in general. Further, most approaches define the effects/influences of personality in a rule-based or scripted manner. Unfortunately, personality traits are not inherently good or bad; their influence is context-dependent. This makes the reasoning about personality influences a problem which is hard to solve by rule-based approaches.[1] Salvit and Sklar (2011, 2012) started to investigate a more advanced model of personality based on MBTI. In Chapter 7, I showed how this could be achieved for the FFM, extending the work of Durupinar et al. (2008) to the complete set of personality traits.

Although these approaches consider personality in isolation, the overall objective is to build an agent-model that brings together all cognitive characteristics. This is discussed in some work.

The contribution presented by Gmytrasiewicz and Lisetti (2002) models the decision-making of an agent using state machines. The authors introduce emotional states and define the personality of an agent as the set of emotional states an agent can be in. Other contributions are:

[1] Main reason for that is a state explosion, *i.e.* one would have to define context-dependent rules for a multitude of personality profiles, which is – except for toy examples – not a tangible approach.

- Neto and da Silva (2010) present an architecture for artificial characters with personality, emotions, and moods based on the BDI model.
- Jones et al. (2009) propose an extension of the BDI model introducing information about the personality, emotions and the physiology of a human.
- Bevacqua et al. (2010) introduce a concept for virtual characters that include specific personalities and emotional reactions.

These approaches discuss architectural considerations from the software engineering perspective and provide first steps towards the integration of more than one cognitive characteristic.

From the literature analysis, we know that the most advanced concepts focus on bringing emotions to agents. Here, a variety of approaches formalise emotions in agents. The PhD project of Adam (2007) provides an in-depth analysis of this topic and proposes a logic of emotions in agents. The formalisation is based on the OCC model (*cf.* Ortony et al., 1988) and realised by expressing emotions based on the modalities beliefs, desires, and intentions. For instance, joy is defined as a feeling that happens when an agent is pleased about a desirable event.[2] Analogue to this, Dastani and Lorini (2012) present an approach that formalises the intensity of a set of emotions (hope, fear, joy, and distress) using the concept of graded beliefs and graded goals. The latter are further distinguished in goals that should be achieved and goals that should be avoided. The idea is to make the intensity of the emotion dependent on the strength the agent's beliefs in something and if something has actual (joy/distress) or prospective consequences (hope/fear). The authors further define associated coping strategies, *i.e.* actions that aim to handle emotions by adapting the other modalities. In contrast to the four emotions formalised in this work, the PhD project of Steunebrink (2010) describes a complete framework that formalises the emotional reactions from appraisal, *i.e.* the trigger of emotion, to the coping with this emotion. It also includes the intensity of such emotions, though, the intensity is introduced as a primitive that is not connected to the other modalities.

Comparing the work on emotion and personality in agents reveals a gap concerning theoretical and practical maturity. Given tasks, such as the creation of virtual humans, one long-term goal for the community is to bridge this gap and finally to integrate all affective phenomena into the decision-making processes (*cf.* Ahrndt et al., 2016c). This substantiates the requirement for having a formalisation of the concept of personality. In contrast, to the logic of emotions, this can not be done by using combinations of existing modalities but has to be done via introducing a new modality. Although, I will substantiate this design decision in Section 9.4 in more detail, the related work also provides approaches that create personality characteristics as combinations of the existing

[2] Using the syntax of $\mathcal{LORA}$ this could be formalised as
$(Joy\ i\ \varphi) = (Bel\ i\ \varphi) \wedge (Des\ i\ \varphi)$ (Adam, 2007, p. 100 – 101).

attitudes. This is done via introducing constraints on the interrelationship of the modalities that are known as the notions of realism (strong realism, realism, weak realism). Rao and Georgeff (1998) describe these notions as providing either over-enthusiastic, over-cautious, or balanced behaviour; each of them a characteristic described within the FFM. Based on this idea, the discussions presented by Fasli (2001a,b, 2003) introduce notions of realism that describe bold and circumspect behaviours, respectively. Common for each notion of realism is that it is streamlined to capture a specific facet of the personality phenomenon. Making the approach limited to this facet. To provide more diverse behaviours, one could argue to alter between different sets of constraints on runtime. However, it may be infeasible to find notions that model other personality facets like being extraverted/introverted or how one reacts to stress as part of the neuroticism trait, as those are other types of personality characteristics.

Within the next section, I will introduce the objectives of the formalisation and the running example, that will be used for the remainder of this chapter.

9.2. Objectives

There are different motivations to provide a formalism of personality within the decision-making process of agents. One of them is the gap we identified concerning emotional agents. Bridging this gap requires an accepted model of personality within the agent domain, to model the relationship between personality and emotions in a consistent way. This can help to establish connections substantiated in psychology (*cf.* Oatley and Jenkins, 1996; Revelle and Scherer, 2010). An established formalisation can also serve as a reference model for the integration of personality into agent-based systems. This can prevent authors from having to create their personality model and serve as a reference for the integration of personality into other BDI logics, or as a reference to integrate other personality models from psychology.

The application in the agent domain requires integrating personality into one of the BDI logics available in the literature. The general purpose of this integration is to (1) reason about the influence of personality on the behaviour of an agent and (2) derive the characteristics of personality of an agent from observations of its' behaviour. Reasoning about these two aspects requires a formalism that can represent *(a) the effect of personality on the behaviour of an agent* and *(b) the state of the personality of an agent*. The representation of this formalism and a detailed discussion of the properties are the main contributions of this chapter. In the following, the goals are illustrated via a running example. We motivate (1) and (2) based on the running example and derive the necessity for statements describing (a) and (b).

Running Example In BDI, an agent exhibits different characteristics depending on the implementation of its deliberation process. One of these characteristics is the commit-

ment strategy, *i.e.*, whether the agent follows a blind-, single-, or open-minded commitment strategy and whether or not the agent is overcommitted to the ends (*i.e.*, the selected intentions respectively the world state the agents wants to achieve) and/or the means (*i.e.*, the generated plan to achieve the intended world state). One can think about the commitment to the means as the likelihood of reconsidering the selected plan. In general, the implementation can vary between the two extremes of never reconsidering a plan once made (blind-commitment) and reconsidering it every step of the way. Blind-commitment to a goal does not allow the agent to reconsider its plan even when it becomes sub-optimal or unreachable (hence, it is sometimes referred to as *fanatical*-commitment). On the other hand, single- and open-minded commitment leads to a frequent reconsideration of means and ends which is a performance overhead.

Similar behaviour can also be attributed to humans who can stick to plans, even if they become unrealistic or change plans frequently. This can be tied to the dimension conscientiousness. Intuitively we would expect high conscientiousness to imply a high commitment to a plan and at the same time carefully reconsidering the means and ends. This relation will be used as running example. The reader should be aware, though, that this example is an over-simplification of the relation between conscientiousness and commitment.

Reasoning about Influence of Personality (1) Conscientiousness as a personality trait influences how accurate one works towards archiving a selected goal. To set a concrete example, imagine a robot that should perform a motion from one point to another in a specific time frame. The level of conscientiousness then can be used to implement a noise level added to the target location or time frame borders. Indeed, this seems to be curious when considering artificial agents, but is one important difference between humans, which most of us have experienced. Knowledge about the personality of a natural agent would enable to reason about its commitment strategy, *i.e.*, whether he/she is open-minded or single-minded, or whether he/she prefers to act boldly or cautiously. Both kinds of information are used for the action selection of an agent, as they provide preferences that can be applied to, e.g. individualise assistant functionalities.

Deriving Characteristics of Personality (2) When the personality of an agent is not available directly, it may be derived from its actions. The challenge here is to reason about the characteristics of a personality given observations of behaviour, *i.e.*, to infer the personality of an agent during interaction with that agent. An example is a navigation assistant observing the behaviour of the driver. If the driver always sticks to the calculated route, one could infer that the level of conscientiousness is rather high. If the driver reacts to all traffic incidents by changing the route the level of conscientiousness is rather low. The car assistants could use a number of such observations (e.g., noise level or stress level) to infer the personality of the user and adapt its assistance towards

that person.

Statements about Personality (a), (b) Reasoning about the influence of personality and deriving personality characteristics requires an underlying representation of personality, its effect and its state. Representing the effect of personality requires the ability to formulate that something is derived from the agent's personality. In the running example we could formulate the following informal statement:

Statement 9.2.1 *Due to its personality agent i tends towards open-minded commitment.*

This statement specifies that the cause for the open-minded commitment is the agent's personality (as opposed to other factors such as incomplete knowledge). Such statements can be used to reason about consistent behaviour that is explained by the agents' personality. The actual personality of an agent does not have to be known for this. This makes sense in situations in which the personality of an agent is not directly accessible or irrelevant. This is akin to the way humans may think about other humans regarding their behaviour. An example concerning another personality trait is the observation that a person (if given a choice) rather stays at home watching a film than going out to a club. A human that is unfamiliar with psychological personality models and the characteristics of the trait extraversion may still make this observation.

Psychological personality models are used to categorise consistent behaviour via personality traits under the assumption that persons with similar personality tend to act similarly. Thus, such models give a suitable foundation for reasoning about and comparing personalities of agents. Here statements of the following form are interesting:

Statement 9.2.2 *Agent i is conscientious.*

This statement describes an agents' personality using the concept of binarisation. In addition to such discrete classes of personality, it often makes sense to talk about the extent of personality traits, either when comparing agents evaluating the extent of a trait, e.g., to derive an increased/decreased likelihood of a person acting in a certain way. This requires to be able to formulate statements like:

Statement 9.2.3 *Agent i is more conscientious than agent k.*

Statement 9.2.4 *Agent i is very conscientious.*

Finally, both kinds of statement can be combined to formulate dependencies between the personality of an agent and its behaviour:

Statement 9.2.5 *If agent i is very conscientious then due to its personality agent i tends towards open-minded commitment.*

This statement is an implication built from statements 9.2.1 and 9.2.4. It can be used to derive the agent's behaviour when the personality state is known (1). In the example, it would be possible to derive that the agent is likely to be open-minded when it is conscientious. The statement can also be used to derive information about personality traits from the agent's behaviour (2). In the example, an agent that does not exhibit open-minded commitment is less likely to be conscientious.

In the following sections, we first introduce the logic we use within this work and afterwards describe the extension of an existing BDI logic to enable the formulation of statements 9.2.1 to 9.2.5.

9.3. Introduction to $\mathcal{LORA}$

The '$\mathcal{L}$ogic $\mathcal{O}$f $\mathcal{R}$ational $\mathcal{A}$gents' (Wooldridge, 2000) ($\mathcal{LORA}$) is a multi-modal, backwards linear, branching time logic developed to enable reasoning about the behaviour of agents that to some extent satisfies the strong notion of agency. The logic is based on the early work on this topic by Cohen and Levesque (1990) and Rao and Georgeff (1991). Multi-modal refers to the three modalities: *Belief*, *Desire*, and *Intention*.[3] These modalities—or following the strong notion of agency: mental concepts—are dictating the deliberation of decisions of the agents. To enable reasoning about such decisions $\mathcal{LORA}$ combines the BDI component with classical first-order logic, a temporal component used to express the dynamics of agent-based systems, and a component that introduces actions and the effects of actions. In the following, we will step-by-step explain the different parts of the logic, before we conclude the section given an overview of the complete structure and syntax.

Vocabulary and basic logic structure The vocabulary is based on the sorts Ag – representing agents, Ac – representing actions, Gr – representing groups of agents, and U – representing other individuals. For these sorts, constants and variables can be defined and used in first-order predicate logic formulas together with additional domain-specific formulas, e.g., denoting the actions an agent can execute. A model in the logic contains a domain description $D = \langle D_{Ag}, D_{Ac}, D_{Gr}, D_U \rangle$, specifying the available entities for each sort.

Temporal structure A temporal dimension is added by a set of time points T and a branching, temporal relation $\mathcal{R} \subseteq T \times T$; specifying which time points can follow which other time points. This relation is bounded in the past and backwards-linear. Backwards-linear means that giving a point in time 'now' there exist only one path through the past. Branching-time means that from 'now' there exist multiple paths

[3] Section 3.3.2 provides a more detailed explanation for these modalities.

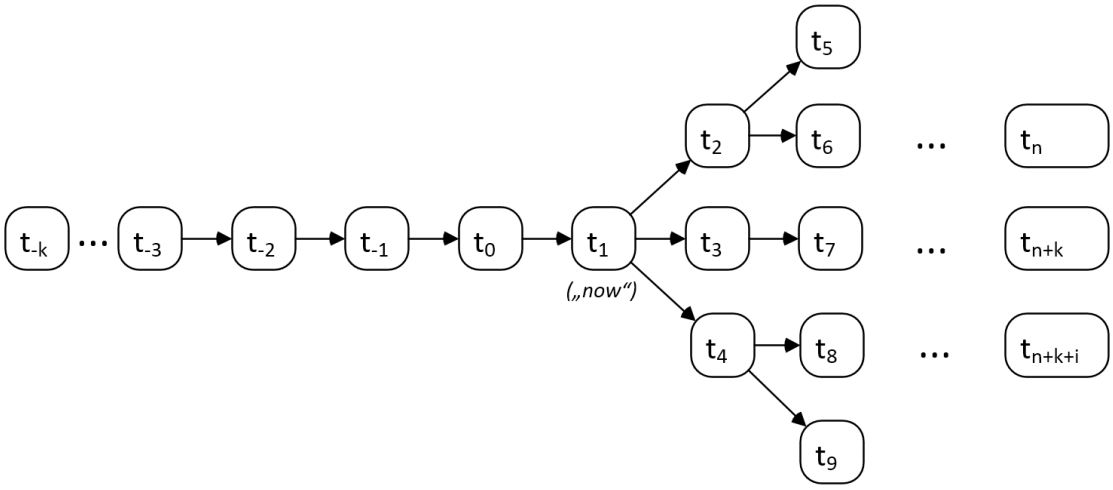


Figure 9.2.: Illustration of a possible worlds' structure of a world in $\mathcal{LORA}$.

towards the future. Fig. 9.2 illustrates the meaning depicting the structure of a possible world in $\mathcal{LORA}$. Over this structure, temporal formulas can be defined to reason about temporal properties.

Worlds and relation to temporal structure Beliefs, desires and intentions are integrated as modality operators and defined via possible worlds in Kripke Semantic Blackburn et al. (2006). For this purpose, a set of worlds W is defined. Worlds are related to temporal structures, *i.e.* each world contains a subset of $\mathcal{R}$ and the respective time points from T used in this subset.

BDI operators The operators $\mathcal{B}$, $\mathcal{D}$ and $\mathcal{I} : D_{Ag} \rightarrow \wp(W \times T \times W)$ represent the *Belief*, *Desire*, and *Intention* modalities of an agent. They map an agent to a set of triples, each of which assigns a possible world and a time point to another world. In a specific situation, *i.e.*, with a fixed agent and time point, these operators form Kripke frames, based on the set of possible worlds.

Derived modality operators To reason about belief accessible worlds from the point of view of a specific agent in a specific world at a specific point in time the shortcut function $\mathcal{B}_t^w(i) = \{w' | \langle w, t, w' \rangle \in \mathcal{B}(i)\}$ is defined. Shortcut functions for $\mathcal{D}$ and $\mathcal{I}$ are defined analogously. The elements in the sets $\mathcal{B}/\mathcal{D}/\mathcal{I}_t^w(i)$ represent i in world w at time point t. In $\mathcal{LORA}$, the operator $\mathcal{B}$ is assumed to be *serial*, *transitive* and *euclidean*; making the logic of belief correspond to KD45 (*cf.* Blackburn et al., 2006, pp. 86–138). The modalities $\mathcal{D}$ and $\mathcal{I}$ are assumed to be serial only; making the logics of desires and intentions correspond to KD (*cf.* Blackburn et al., 2006, pp. 86–138).

State formulas Based on the possible world definition of these three modalities reasoning about the fulfilment of formulas is implemented. For the $\mathcal{B}$ modality this can be done via the formula $(Bel\ i\ \varphi)$ that states that an agent i believes φ. These formulas are considered to be state formulas, meaning they are evaluated with a specific world w and a specific time point t. The evaluation of a state formula also depends on a model, describing the domain values, variables, predicates and the definitions of specific functions, e.g., $\mathcal{B}$, $\mathcal{D}$ and $\mathcal{I}$, and a variable assignment. The corresponding sets and interpretation functions are introduced shortly. The Bel formula is defined as:

$$\langle w, t\rangle \models (Bel\ i\ \varphi) \text{ iff } \forall w' \in B_t^w([\![i]\!]), \langle w', t\rangle \models \varphi, \tag{9.1}$$

where $[\![i]\!]$ is the evaluation of term i under the variable assignment. Intuitively, this statement states that an agent i believes a statement φ if this statement holds in all worlds that are accessible via its beliefs. Analogously the state formulas $(Int\ i\ \varphi)$ and $(Des\ i\ \varphi)$ can be defined from the modality operators $\mathcal{I}$ and $\mathcal{D}$ to denote that an agent intends or desires a statement to be fulfilled. Such a statement φ is any valid formula that can be built in $\mathcal{LORA}$ — enabling, for example, the nesting of the modalities.

Together these state formulas describe the state of an agent. This means that for every point in time and every of the defined modalities there exist a set of worlds that is consistent for the modality. For example, if an agent believes φ then the set of worlds that is consistent with the belief consists of such worlds where φ is true. This is also called the set of accessible worlds, which—taking into account the example—we call the belief-accessible worlds.

To conclude, a model in $\mathcal{LORA}$ is a structure $M = \langle T, \mathcal{R}, W, D, Act, Agt, \mathcal{B}, \mathcal{D}, \mathcal{I}, C, \Phi\rangle$, where:

- T is the set of time points;
- $\mathcal{R} \subseteq T \times T$ is the branching temporal structure, which is backwards-linear over T;
- W are all worlds that can be build with a non-empty set of time points T and a temporal structure $\mathcal{R}$ on T;
- $D = \langle D_{Ag}, D_{Ac}, D_{Gr}, D_U\rangle$ is a domain with D_{Ag} the non-empty set of agents, D_{Ac} the non-empty set of actions, D_{Gr} the non-empty of groups of agents as subsets of D_{Ag}, D_U the non-empty set of other objects/individuals;
- $Act : \mathcal{R} \to D_{Ac}$ is a function used to label each arc in R with actions;
- $Agt : D_{Ac} \to D_{Ag}$ is a function linking each action with an agent performing it;
- $\mathcal{B}, \mathcal{D}, \mathcal{I} : D_{Ag} \to \wp(W \times T \times W)$ are the belief-, desire-, and intention-accessibility relations, respectively;

- $C : Const \times T \rightarrow \bar{D}$ is a function that is used to interpret constants $c \in Const$ ($Const$ is a set of constants that stand for agents, action sequences, set of agents, and other individuals) for a specific time point in a domain $\bar{D} = \langle D_{Ag}, D^*_{Ac}, D_{Gr}, D_U \rangle$; and

- $\Phi : Pred \times W \times T \rightarrow \wp(\bigcup_{u \in \mathbb{N}} \bar{D}^u)$ is a function that interprets predicates $p \in Pred$ ($Pred$ is a set of predicate symbols) in a specific time point in a specific world ($\bar{D}^u$ is the set u-tuples over $\bar{D}$).

One can use this model to reason about different kinds of behaviour reaching from cautious and open-minded agents to bold agents that act blindly committed to both, the means and the ends. Also, $\mathcal{LORA}$ defines speech acts (as part of the action component), enabling the study of communication in multi-agent systems as found in joint activities, e.g., teamwork or cooperative problem-solving settings.

Table 9.1 introduces the syntax of $\mathcal{LORA}$. As $\mathcal{LORA}$ is based on first-order logic it includes the usual connectivities and quantifiers.

In the next section, I will introduce my formalisation of personality within this logic. The task at hand is to find a way to express the influence of personality on the decision-making and to answer the question how personality influences the set of accessible worlds.

9.4. Formalisation of Personality

The alphabet of $\mathcal{LORA}$ contains no symbol for the personality, which we want to represent. One approach would be to represent the effects of personality as predicates over the existing modal operators — similar to the approaches presented by different authors for the formalisation of emotions in agents. As mentioned above, we will integrate personality as an own modality besides the existing ones. This design decision is made based on the observation that the influence of emotions is frequently represented by combinations of the existing modal operators (Believes, Desires and Intentions) (cf. Adam, 2007; Dastani and Lorini, 2012; Steunebrink, 2010).

However, personality is different from emotion as it *"...is the coherent patterning of affect, behavior, cognition, and desires (goals) over time and space."* (Revelle and Scherer, 2010, p. 512). In contrast, the effects of emotions are bounded to a particular time and object (Steunebrink, 2010, p.3). In fact, emotions always occur relative to something, *i.e.* they are always triggered by an object or an expected or occurring event or action (Steunebrink, 2010, pp. 3–6). Given these observations, Revelle and Scherer (2010) conclude the difference between both affective phenomena with the quote: "A helpful analogy is to consider that personality is to emotion as climate is to weather." (Revelle and Scherer, 2010, p. 512).

Fig. 9.3 provides a classification of the existing affective phenomena concerning the duration. Psychologists consider personality to be (to some extent) a time and space

Table 9.1.: Syntax of $\mathcal{LORA}$ as introduced by Wooldridge (2000, p. 72).

$\langle$*ag-term*$\rangle$	::=	any element of $Term_{Ag}$	/* agent terms */
$\langle$*ac-term*$\rangle$	::=	any element of $Term_{Ac}$	/* action terms */
$\langle$*gr-term*$\rangle$	::=	any element of $Term_{Gr}$	/* group terms */
$\langle$*ac-exp*$\rangle$	::=	$\langle$*ac-term*$\rangle$	
	|	$\langle$*ac-exp*$\rangle$; $\langle$*ac-exp*$\rangle$	/* sequential composition */
	|	$\langle$*ac-exp*$\rangle$ | $\langle$*ac-exp*$\rangle$	/* non-deterministic choice */
	|	$\langle$*state-fmla*$\rangle$?	/* test actions */
	|	$\langle$*ac-exp*$\rangle$ $*$	/* iteration */
$\langle$*term*$\rangle$	::=	any element of $Term$	/* arbitrary terms */
$\langle$*pred-symp*$\rangle$	::=	any element of $Pred$	/* predicate symbols */
$\langle$*var*$\rangle$	::=	any element of Var	/* variables */
$\langle$*state-fmla*$\rangle$	::=	*true*	/* truth constant */
	|	$\langle$*pred-symp*$\rangle(\langle$*term*$\rangle, \ldots, \langle$*term*$\rangle)$	/* predicates */
	|	(Bel $\langle$*ag-term*$\rangle$ $\langle$*state-fmla*$\rangle$)	/* belief formula */
	|	(Des $\langle$*ag-term*$\rangle$ $\langle$*state-fmla*$\rangle$)	/* desire formula */
	|	(Int $\langle$*ag-term*$\rangle$ $\langle$*state-fmla*$\rangle$)	/* intention formula */
	|	(Agts $\langle$*ac-exp*$\rangle$ $\langle$*gr-term*$\rangle$)	/* agents of an action */
	|	($\langle$*term*$\rangle = \langle$*term*$\rangle$)	/* equality */
	|	($\langle$*ag-term*$\rangle \in \langle$*gr-term*$\rangle$)	/* group membership */
	|	A$\langle$*path-fmla*$\rangle$	/* path quantifier */
	|	$\neg\langle$*state-fmla*$\rangle$	/* negation */
	|	$\langle$*state-fmla*$\rangle \vee \langle$*state-fmla*$\rangle$	/* disjunction */
	|	$\forall\langle$*var*$\rangle \cdot \langle$*state-fmla*$\rangle$	/* quantification */
$\langle$*path-fmla*$\rangle$	::=	(Happens $\langle$*ac-exp*$\rangle$)	/* action happens */
	|	$\langle$*state-fmla*$\rangle$	/* state formula */
	|	$\langle$*path-fmla*$\rangle$ $\mathcal{U}$ $\langle$*path-fmla*$\rangle$	/* until */
	|	$\bigcirc\langle$*path-fmla*$\rangle$	/* next */
	|	$\neg\langle$*path-fmla*$\rangle$	/* negation */
	|	$\langle$*path-fmla*$\rangle \vee \langle$*path-fmla*$\rangle$	/* disjunction */
	|	$\forall\langle$*var*$\rangle \cdot \langle$*path-fmla*$\rangle$	/* quantification */

independent cognitive mechanism; that influences each stage of the decision-making process of humans (Revelle and Scherer, 2010). To substantiate this statement the interested reader is referred to work that shows that we as humans have a relatively stable personality over our lifespan as adults (*cf.* Caspi et al., 2005; Hampson and Goldberg, 2006; Wilks, 2009).

The conclusion I draw from these findings is that personality is per se independent of the *Beliefs*, *Desires*, and *Intentions* at specific times. This, by no means, says that the interpretation of the influence of personality is independent of the context. It only tells us, that we need to represent personality as dedicated modality and need to discuss the relations to the other modalities afterwards. The former will be done within this section. The latter will be done within the subsequent section.

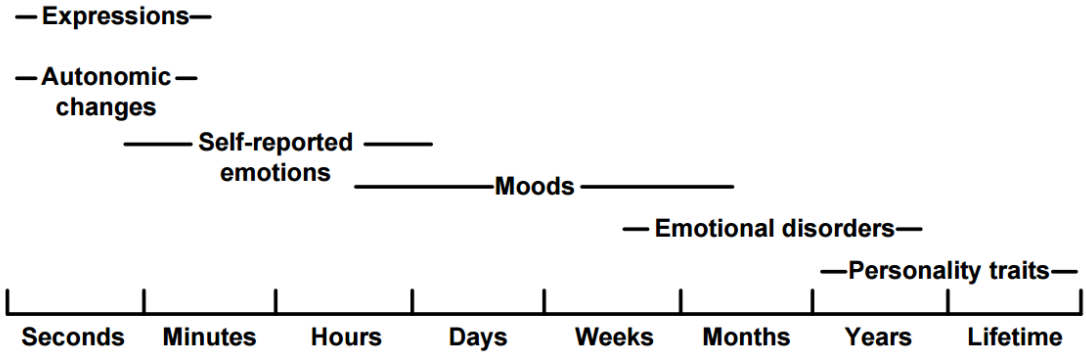


Figure 9.3.: Duration is one important factor to distinguish affective phenomena. Figure is taken from published work (Steunebrink, 2010, p. 6).

9.4.1. Representing the Effects of Personality

To define the semantics of the personality, we incorporate the definition of (Wooldridge, 2000, p. 74) introducing the *Bel*, *Des*, *Int* modalities and introduce personality as an own modal connectivity.[4] The basic idea for the definition of a modality is to use a formula with the structure (M i φ), where M represents one of the modal connectives, i is a term referring to an agent, and φ is any valid formula in $\mathcal{LORA}$. Thus, the formula (Bel i φ) verbalises the fact that agent i believes φ; (Des i φ) verbalise the fact that agent i desires φ; and (Int i φ) verbalises the fact that agent i intends φ. Consequently, the personality of each agent will be represented defining a personality modality *Per*. The associated formula (Per i φ) is then used to reason about the effects of the personality and verbalises the fact that agent i tends to φ. Here *tends to* refers to the influence of personality and not to other preferences, e.g. regarding emotions and moods. Indeed, *tends to* is a placeholder for a personality-descriptive verb that must be used in a specific situation as there is no single personality-descriptive verb such as believes, desires, or intends within the English language (*cf.* de Raad et al., 1988).[5] In a general manner, it can be interpreted as φ being aligned with the personality of agent i, as this formula holds in all worlds which are accessible with the personality of i. In the following, we will define each of the necessary elements to integrate both.

The personality of an agent (*Per*) is characterised using the function $\mathcal{P}$ that is defined as follows:

$$\mathcal{P} : D_{Ag} \rightarrow \wp(W \times T \times W). \tag{9.2}$$

The operator $\mathcal{P}$ is named personality-accessibility relation. It defines all worlds that are in line with the personality of an agent $i \in D_{Ag}$ given a specific situation $\langle w, t \rangle$, where $w \in W$ and $t \in T$. Although personality itself is described as independent of time

[4]Henceforth, I will use the terms modality, modal operator, and modal connectivity interchangeable.

[5]I will illustrate this by using personality-descriptive verbs if appropriated. I use the verb *tends to* as placeholder for the influences of the basic tendencies of the FFT.

and space its effects may be codependent on the current world and time point. This is reflected by the situation. Analogous to the other modalities a shortcut function can be used to reason about the effects of personality in a specific world and time. Formally, this is defined as:

$$\mathcal{P}_t^w(i) = \{w' | \langle w, t, w' \rangle \in \mathcal{P}(i)\} \tag{9.3}$$

We further have to define the syntax and semantics of the personality formula. Comparable to the other modalities, reasoning about the fulfilment of formulas is enabled by a state formula of the following syntax:

$$(\text{Per } \langle \textit{ag-term} \rangle \ \langle \textit{state-fmla} \rangle). \tag{9.4}$$

The semantics of this state formula is defined based on an agent i and a state formula φ. Such statements are evaluated in the context of a fixed model M and a variable assignment V. The semantic of Per is defined as follows:

$$\langle M, V, w, t \rangle \models_{\mathcal{S}} (Per\ i\ \varphi) \text{ iff } \forall w' \in \mathcal{P}_t^w([\![i]\!]), \langle M, V, w', t \rangle \models_{\mathcal{S}} \varphi. \tag{9.5}$$

Like the $\mathcal{D}$ and $\mathcal{I}$ relations, $\mathcal{P}$ is assumed to assign agents serial relations and to satisfy the world/time point compatibility property. This means, that *"if a world w' is accessible to an agent from situation $\langle w, t \rangle$, then t is required to be a time point in both w and w'."* (Wooldridge, 2000, p. 74), which formally expressed means that if $w' \in P_t^w(i)$ then $t \in w$ and $t \in w'$. Furthermore, the personality modality has a logic that corresponds to the normal modal system KD.[6]

Given the new modality the extended model is a structure $M = \langle T, \mathcal{R}, W, D, Act, Agt, \mathcal{B}, \mathcal{D}, \mathcal{I}, \mathcal{P}, C, \Phi \rangle$, which can be used to reason about the effects of a personality.

This operator can now be used to describe the first statement (Statement 9.2.2) from our running example. We can say that due to its personality agent i tends towards open-minded commitment via the state-formula:

$$(\text{Per } i\ hasOpenMindedCommitment(i)) \tag{9.6}$$

Here, $hasOpenMindedCommitment(i)$ is a predicate describing that an agent i has open-minded commitment. As the φ is a placeholder for any valid statement in $\mathcal{LORA}$ other examples could read as follows:

$$(\text{Per } sebastian\ FulfilAccordingPlan(tasks)). \tag{9.7}$$

The intended meaning of this expression is that the agent *sebastian* tends to $Fulfil\ AccordingPlan(tasks)$ or in other words that Sebastian tends to fulfil his tasks according to some prior existing plan. Here, *sebastian* is a first-order logic constant that identifies

[6] Proves are provided in Appendix D.

a particular agent. We can also generalise this statement and use the existence quantifier to verbalise that there exist some agent i who tends to fulfil its tasks according to a plan by stating:

$$\exists i \cdot (\text{Per } i \; FulfilAccordingPlan(tasks)). \tag{9.8}$$

Likewise, we can state that all agents prefer to fulfil its tasks according to a plan by stating:

$$\forall i \cdot (\text{Per } i \; FulfilAccordingPlan(tasks)). \tag{9.9}$$

The nesting of modal connectives enables statements like this one:

$$(Per \; sebastian \; (Per \; johannes \; FulfilAccordingPlan(tasks))). \tag{9.10}$$

The intended interpretation of this formula is that *sebastian* prefers that *johannes* prefers to fulfil its task according to a plan. The difference to the other modalities here is that a preference according to the personality of the agent is not necessarily affecting the deliberation process (we will discuss later why). Additionally, we are also allowed to nest different modalities:

$$(Bel \; sebastian \; (Per \; johannes \; FulfilAccordingPlan(tasks))); \tag{9.11}$$

The intention here is that *sebastian* beliefs that *johannes* tends to fulfil task according to a plan. Intuitively, this statements makes more sense than the one before. Another example for the nesting are statements such as:

$$(Per \; sebastian \; \forall \; i \; \cdot \; (Per \; i \; Hobby(extremsport))). \tag{9.12}$$

Verbalising that *sebastian* likes all other agents which prefer extreme sports as a hobby.

$$(\text{Per } i \; Places(crowded)) \Rightarrow (\text{Per } i \; Is(talkative)). \tag{9.13}$$

This formula says that for an agent that prefers crowded places it applies that this agent is talkative. In fact, this might be useful information when trying to explain the course of actions of extraverted or introverted agents. Both examples show how the reasoning about personality could be used to infer knowledge about the behaviour of the agent.

9.4.2. Representing the State of Personality

Reasoning about the state of the personality of an agent requires having a notion to represent this state. As described in Section 3.4, I consider the FFM to be most suited for the integration of personality in agents. In FFM each personality trait is represented

by a continuous scale.[7] Hence, one personality consists of a real number value for each trait. These values are interpreted with respect to a maximum and minimum (usually the continuous scale measures the personality in the interval $[0, 1]$, which we also apply), where the maximum means that the factor is fully developed, the minimum means that the factor is not developed, and the average means that the factor is balanced. For example, a value of 1 for extraversion denotes that the person is considered extroverted while a value of 0 means the person is introverted and a value of 0.5 means that neither a strong tendency towards introversion nor extraversion can be observed. To include this model into $\mathcal{LORA}$ we first need to enable handling real numbers to express and compare the extent of personality traits. For this purpose, the comparison functions $=$, $<$ and $>$ can be used. In $\mathcal{LORA}$ this is integrated as additional state formulas comparing two real number expressions:

$$\mathbb{R} = \mathbb{R} \quad | \quad \mathbb{R} < \mathbb{R} \quad | \quad \mathbb{R} > \mathbb{R} \tag{9.14}$$

For other use cases, other real-valued expressions (e.g., addition or multiplication) may be relevant. These can be integrated analogous to the statements above.

For the formalisation, we assume that the personality only depends on the agent itself and is consistent over time. Thus, the personality does not depend on the world or the time point but solely on the agent. To represent the state of personality we define one function per personality factor that maps the agent to the value representing the extent of the respective personality trait:

$$O, C, E, A, N : Ag \to \mathbb{R}. \tag{9.15}$$

The numbers derived from personality traits usually need to be interpreted in some way. On our scale, it could make sense to exclude personalities below 0.45 and above 0.55 as they may be considered to be roughly balanced. For the trait extraversion, the two extremes of the scale can be interpreted as introversion and extroversion. Here we could consider agent i to be introverted if $E(i) < 0.25$ and extroverted if $E(i) > 0.75$. This enables discrete reasoning about personality categories as it would be necessary if one wants to model the dichotomies provided by the MBTI theory. Constants and variables representing real numbers are required to express such statements. Those can be integrated into $\mathcal{LORA}$ analogously to the variables and constants of other sorts, e.g., variables denoting agents. For readability we denote constants by their actual values, e.g., 0.3 is a constant of value 0.3 whose name is "0.3".

These statements now enable expressions that refer to personality traits of agents and interpret them, either in the context of personality traits of other agents or in the context of variables or constants. They are sufficient to express Statements 9.2.2 to 9.2.4 from

[7] As before, I will use the following abbreviations: openness to experience (O), conscientiousness (C), extraversion (E), agreeableness (A) and neuroticism (N).

the running example. Statement 9.2.2 denotes that an agent i is conscientious. It could be expressed as $C(i) > 0.3$. Comparing two agents conscientiousness, e.g., to state that an agent i is more conscientious than an agent k (Statement 9.2.3) can be expressed as follows: $C(i) > C(k)$. Intervals in the continuous personality scale can be used to express more fine-grained personality trait distinctions. Here, we could consider an agent to be very conscientious when it has a higher trait than 0.8. Using this (arbitrary) line we can formulate Statement 9.2.4 as $C(i) > 0.8$.

9.4.3. Reasoning about State and Effect

The goal of the formalisation is to enable reasoning about the interdependencies of personality and behaviour of a natural agent. During the last two sections, I described how $\mathcal{LORA}$ can be extended to represent the effects and state of personality. Both extensions can be combined to express how specific personality types (*i.e.*, the state of personality) influence the agent. An example is given in Statement 9.2.5, which expresses that a very conscientious agent tends towards open-minded commitment. Given the introduced extensions, we are now able to express this statement as follows:

$$C(i) > 0.8 \rightarrow (\text{Per } i \; hasOpenMindedCommitment(i)). \tag{9.16}$$

Such statements represent the relation between state and effects of personality and can be used for reasoning, e.g., along with the implication operator. To conclude, Table 9.2 lists the syntax of $\mathcal{LORA}$ including the above-described extensions, representing the personality of an agent.

The extension of $\mathcal{LORA}$ can also be used for further discussions about the relation of the formal representation of personality to other parts of the logic. This will be done in the next section.

9.5. Characteristics of $\mathcal{P}$ in $\mathcal{BDI}$

Integrating personality as own modality provides the fundamentals for a comprehensive analysis of the properties that are useful to characterise an agent with personality. From a purely logical aspect, there is no reason to do that, as the just-presented formal system enables us to use all possible combinations of the formulas and operators available. However, that would mean to ignore the semantics of the properties, *i.e.*, to not discuss how the properties influence the behaviour of an agent and which influences are meaningful/reasonable for analysing (ir)rational behaviour or personality-compliant behaviour. In other words, this means paying little attention to the question of what the introduction of personality in the interplay with belief, desire, and intention means. Following this philosophical view, this section will discuss the extent to which the integration of

Table 9.2.: Syntax of $\mathcal{LORA}$ including the personality.

$\langle$*ag-term*$\rangle$	::=	any element of $Term_{Ag}$	/* agent terms */
$\langle$*ac-term*$\rangle$	::=	any element of $Term_{Ac}$	/* action terms */
$\langle$*gr-term*$\rangle$	::=	any element of $Term_{Gr}$	/* group terms */
$\langle$*ac-exp*$\rangle$	::=	$\langle$*ac-term*$\rangle$	
	|	$\langle$*ac-exp*$\rangle$; $\langle$*ac-exp*$\rangle$	/* sequential composition */
	|	$\langle$*ac-exp*$\rangle$ | $\langle$*ac-exp*$\rangle$	/* non-deterministic choice */
	|	$\langle$*state-fmla*$\rangle$?	/* test actions */
	|	$\langle$*ac-exp*$\rangle$ $*$	/* iteration */
$\langle$*term*$\rangle$	::=	any element of $Term$	/* arbitrary terms */
$\langle$*pred-symp*$\rangle$	::=	any element of $Pred$	/* predicate symbols */
$\langle$*var*$\rangle$	::=	any element of Var	/* variables */
$\langle$*state-fmla*$\rangle$	::=	*true*	/* truth constant */
	|	$\langle$*pred-symp*$\rangle(\langle$*term*$\rangle, \ldots, \langle$*term*$\rangle)$	/* predicates */
	|	(Bel $\langle$*ag-term*$\rangle$ $\langle$*state-fmla*$\rangle$)	/* belief formula */
	|	(Des $\langle$*ag-term*$\rangle$ $\langle$*state-fmla*$\rangle$)	/* desire formula */
	|	(Int $\langle$*ag-term*$\rangle$ $\langle$*state-fmla*$\rangle$)	/* intention formula */
	|	(Per $\langle$*ag-term*$\rangle$ $\langle$*state-fmla*$\rangle$)	/* personality formula */
	|	(O $\langle$*ag-term*$\rangle$)	/* openness formula */
	|	(C $\langle$*ag-term*$\rangle$)	/* consciousness formula */
	|	(E $\langle$*ag-term*$\rangle$)	/* extraversion formula */
	|	(A $\langle$*ag-term*$\rangle$)	/* agreeableness formula */
	|	(N $\langle$*ag-term*$\rangle$)	/* neuroticism formula */
	|	$\langle$*var*$\rangle < \langle$*var*$\rangle$	/* less than */
	|	$\langle$*var*$\rangle = \langle$*var*$\rangle$	/* equal to */
	|	$\langle$*var*$\rangle > \langle$*var*$\rangle$	/* greater than */
	|	(Agts $\langle$*ac-exp*$\rangle$ $\langle$*gr-term*$\rangle$)	/* agents of an action */
	|	($\langle$*term*$\rangle = \langle$*term*$\rangle$)	/* equality */
	|	($\langle$*ag-term*$\rangle \in \langle$*gr-term*$\rangle$)	/* group membership */
	|	A$\langle$*path-fmla*$\rangle$	/* path quantifier */
	|	$\neg\langle$*state-fmla*$\rangle$	/* negation */
	|	$\langle$*state-fmla*$\rangle \vee \langle$*state-fmla*$\rangle$	/* disjunction */
	|	$\forall\langle$*var*$\rangle \cdot \langle$*state-fmla*$\rangle$	/* quantification */
$\langle$*path-fmla*$\rangle$	::=	(Happens $\langle$*ac-exp*$\rangle$)	/* action happens */
	|	$\langle$*state-fmla*$\rangle$	/* state formula */
	|	$\langle$*path-fmla*$\rangle$ $\mathcal{U}$ $\langle$*path-fmla*$\rangle$	/* until */
	|	$\bigcirc\langle$*path-fmla*$\rangle$	/* next */
	|	$\neg\langle$*path-fmla*$\rangle$	/* negation */
	|	$\langle$*path-fmla*$\rangle \vee \langle$*path-fmla*$\rangle$	/* disjunction */
	|	$\forall\langle$*var*$\rangle \cdot \langle$*path-fmla*$\rangle$	/* quantification */

Table 9.3.: Pairwise subset interrelationship between the $\mathcal{P}$ and $\mathcal{B}$, $\mathcal{D}$, $\mathcal{I}$ modalities. The tendency column summarises the discussion w.r.t. the research question: + in general accepted as meaningful, 0 meaningful in some cases (e.g. does not hold in general, but might be useful in some contexts, e.g. useful for heuristics), − in general rejected as meaningful.

	Relationship	**Formula Schema**	**Tendency**
(1)	$\mathcal{B}_t^w([\![i]\!]) \subseteq \mathcal{P}_t^w([\![i]\!])$	$(\text{Per } i\ \varphi) \Rightarrow (\text{Bel } i\ \varphi)$	−
(2)	$\mathcal{P}_t^w([\![i]\!]) \subseteq \mathcal{B}_t^w([\![i]\!])$	$(\text{Bel } i\ \varphi) \Rightarrow (\text{Per } i\ \varphi)$	−
(3)	$\mathcal{D}_t^w([\![i]\!]) \subseteq \mathcal{P}_t^w([\![i]\!])$	$(\text{Per } i\ \varphi) \Rightarrow (\text{Des } i\ \varphi)$	−
(4)	$\mathcal{P}_t^w([\![i]\!]) \subseteq \mathcal{D}_t^w([\![i]\!])$	$(\text{Des } i\ \varphi) \Rightarrow (\text{Per } i\ \varphi)$	−
(5)	$\mathcal{I}_t^w([\![i]\!]) \subseteq \mathcal{P}_t^w([\![i]\!])$	$(\text{Per } i\ \varphi) \Rightarrow (\text{Int } i\ \varphi)$	−
(6)	$\mathcal{P}_t^w([\![i]\!]) \subseteq \mathcal{I}_t^w([\![i]\!])$	$(\text{Int } i\ \varphi) \Rightarrow (\text{Per } i\ \varphi)$	−

personality captures our intuitive understanding of the concepts in question. In doing so, I complement the discussion for the primal modalities that is provided by Wooldridge (2000, p. 91). The question to answer is: *Which relations between the $\mathcal{P}$ modality and the $\mathcal{B}$, $\mathcal{D}$, $\mathcal{I}$ modalities are meaningful for reasoning about rational behaviour?*

To answer this question, we have to discuss interrelationships between the modalities in detail. We will concentrate on the pairwise interaction between the modalities using the concept of possible worlds. First, we will discuss subset interrelationships without taking into account the structure of the worlds itself (*cf.* Section 9.5.1). Secondly, we will take into account the structure of the worlds w.r.t. inevitabilities (*cf.* Section 9.5.2) and options (*cf.* Section 9.5.3). Finally, we will take into account the temporal component of $\mathcal{LORA}$, discussing acceptable properties for reasoning about the future behaviour of agents with personality bounded rationality (*cf.* Section 9.5.4).

It shall be understood, that the discussion is limited w.r.t. the level of detail that can be captured. It presents an intuitive discussion from author's point of view and should be further funded by results from psychology. The mental concepts and states of humans are far too complicated and fuzzy to be captured entirely in a logic. However, $\mathcal{LORA}$ was not created with the intention to capture all nuances of beliefs, desires, and intentions neither is the integration of the personality phenomenon.

9.5.1. Subset interrelationships

Given the four modalities, there are a total of 12 pairwise subset relationships. We will concentrate the discussion on the ones related to personality (the remaining are discussed by Wooldridge, 2000, pp. 92–106). Table 9.3 lists the relations for the general case including the corresponding formula schema. In the general case, the structure of the world is ignored. That means that the relations hold for all w, t, i. Although all of these constructs are valid, not all of them can be considered as meaningful when reasoning about the behaviour of agents. To clarify this, let us consider each of the relations

individually. In doing so, we apply a common structure to each interpretation. Given two modalities M and M', the listed formulae follow the structure (M i φ) $\Rightarrow$ (M' i φ). The structure used for the interpretation reads as follows: *If sth. is in line with the agents M then it also is in line with the agents M'.*

- Eq. (1) says that if an agents' personality prefers a fact, then it also believes this fact.

  Although this seems reasonable for some facts (e.g., prejudice and conservative persons), it is a too strong property for rational agents. One example related to high agreeableness is the tendency to act kindly (McCrae and John, 1992, p. 178), though, there is no reason to believe that one will never argue/have a dispute with another.

  Rejecting this relation implies that an agent cannot prefer φ that are not part of the believe base. This tells us, that $\mathcal{B}_t^w([\![i]\!])$ cannot be a subset of $\mathcal{P}_t^w([\![i]\!])$.

- Eq. (2) says that if an agent believes something, then it is also in line with the agent's personality.

  A concrete example shows that this property is too strong as well. Let us suppose that I believe that I will present this work to an audience and I am an introverted person. There is no reason not to believe that there is an arbitrary situation where I have to present the work. In contrast, my personality indicates that I am willing to avoid such situations, *i.e.* that I do not like (prefer) such situations. As a concrete example, Hatemi and Verhulst (2015) present a work showing that *"...political attitudes are independent of personality traits and that changes in attitudes are primarily a function of changes in the environment or one's unique experiences"* (Hatemi and Verhulst, 2015, p. 16).

  Rejecting this relation implies that there exist φ which the agents believes, that are not preferred by the personality. Intuitively, this tells us that $\mathcal{P}_t^w([\![i]\!])$ is a proper subset of $\mathcal{B}_t^w([\![i]\!])$.

- Eq. (3) says that if an agents' personality tends to something, then it is also desired by the agent.

  In other words, the formula says that an agent desires everything that is in line with the personality. At first glance, this makes sense. Indeed it is too strong, as the personality might tend to desires that rational agents would not select. One example is the tendency of people who score high in the openness trait to gather new experiences (e.g. in extreme sports, which are potentially lethal activities) vs the actual desire to perform actions that are potentially lethal. Another example is related to the trustfulness of a person (as part of the agreeableness trait) (McCrae and John, 1992, p. 178). High trustfulness does not imply the desire to trust everyone and everything.

- Eq. (4) says that if an agent desires something, then it is also in line with the agent's personality.

  We reject this as a necessary property for rational agents in its current form, though, statements of this form might be helpful to distinguish rational from irrational behaviour; e.g. reasoning about personality disorders like Anorexia, where your otherwise vital desire of eating is affected by the characteristics of your personality. Indeed, this property would prevent desires that conflict with the agents' personality. One example of such a conflict is that even the most cooperative persons might desire to win in certain situations. Other examples can be found, when thinking about personality independent desires, e.g. the desire to not starve or the desire to eat.

- Eq. (5) says that if an agents' personality prefers something then it is also intended by the agent.

  This property represents a kind of *realism* property (Cohen and Levesque, 1990, pp. 227–228). It makes sense for rational agents as the commitment to an intention implies efforts towards the situation, and it would make little sense to commit resources to something that stays in conflict to something that is preferred by the personality. Thus, personality is some kind of constraint on the deliberation process of an agent, which was also observed by psychologists (*cf.* Connor-Smith and Flachsbart, 2007; Ortony et al., 2005; Ozer and Benet-Martínez, 2006; Revelle and Scherer, 2010). However, in its current form, it states that all situations that are in line with the personality are intended by the agent, which is a very strong property. Accepting this formula would suggest that one will never select an intention which is not preferred by the personality. All of us perform actions which we do not prefer, e.g. due to cultural norms, social contracts, or job hierarchies.

- Eq. (6) says that if an agent intends something, then it is also in line with the agent's personality.

  Let us suppose, that I intend to learn about being talkative, e.g. to be better prepared for presentations. Even this conflicts with what I prefer according to my introverted personality it makes sense as an intention in some situations. Thus, Eq. (6) seems to be too strong as a property of rational agents.

To summarise, we can say that none of these properties is meaningful to capture the rational behaviour of agents with personality. They are too strong in their presented form, even though, it seems that they might be helpful in weaker versions.

These weaker versions are provided by the possibility to analyse structural relationships using the *universal path quantifier* A for structural subsets and the *existential path quantifier* E for supersets. Structural relationships are available as the worlds in BDI logic themselves have a structure, described by a temporal relation, as introduced in

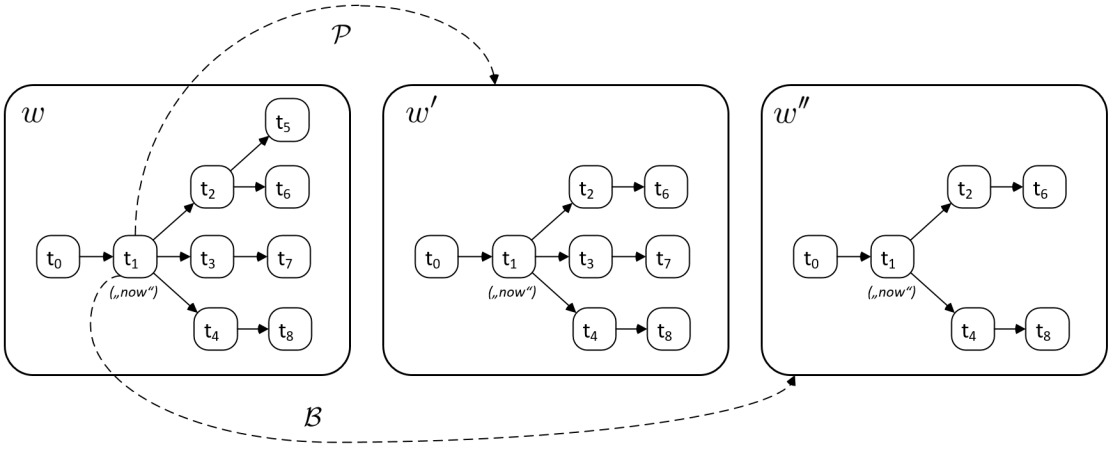


Figure 9.4.: Personality as structural superset of beliefs ($\mathcal{P} \subseteq_{sup} \mathcal{B}$). Respectively, beliefs as structural subset of personality ($\mathcal{B} \subseteq_{sub} \mathcal{P}$).

Section 9.3 and depicted in Fig. 9.2. Next, we will discuss those structural relationships, starting with inevitabilities. These relations are given by A-formula describing inevitable futures, *i.e.* expressions that are true for all possible worlds giving a situation. One restriction we can make at this point is related to equivalences. We can state that the tendency of a personality to tend to some φ, does not necessarily imply that the same personality rejects the opposite of φ. For example, people scoring high on conscientiousness tend to be hardworking (Feist et al., 2012, p. 282). This does not imply that these people do not prefer to be lazy the other day.

9.5.2. Inevitabilities

Fig. 9.4 illustrate how the structure of the world is taken into consideration depicting a relationship between the belief and personality modality. Here, the personality-accessibility relation is considered as structural superset of the belief accessibility relation.[8] The figure contains the three worlds w, w', and w'', such that w'' is a subworld of w' ($w'' \sqsubseteq w'$) and w' is a subworld of w ($w' \sqsubseteq w$) and additionally shows that $w' \in \mathcal{P}^{w}_{t_1}([\![i]\!])$ and $w'' \in \mathcal{B}^{w'}_{t_1}([\![i]\!])$. This means, that for a belief-accessible world w'' there exist a personality-accessible world w' such that $w'' \sqsubseteq w'$. Giving this example, inevitability means that some φ is true on all paths in w, w', and w''.

This case introduces the first pairwise interrelationship between personality and the other modalities. All structural subset relations capturing an agent's attitudes to inevitabilities are listed in Table 9.4. In general, these relations say that all φ that are inevitable true on the left-hand side are also inevitable on the right-hand side. For a worlds' structure, this implies that φ is true in all possible futures, which we refer to as

[8] A formal definition for the structural subset/superset relation is provided elsewhere (Wooldridge, 2000, p. 96). There, Theorem 5.3. also defines what is meant with inevitabilities.

paths. Following the same argumentation as above, most of the relations are too strong for determining the behaviour of agents, which possess the concept of personality. Indeed, the influence of a concrete personality always depends on the actual situation, which has recently lead to agent-models such as the Social Context based Personality model (*cf.* Kochanowicz et al., 2015). Being aware of this fact lets us conclude that personality induced inevitabilities are – in general – not a tangible concept for explaining the behaviour of rational agents. To clarify this, let us again consider each of the relations individually. Given two modalities M and M', the listed formulae follow the structure (M i A(φ)) $\Rightarrow$ (M' i A(φ)). The structure used for the interpretation reads as follows: *If sth. is inevitable according to the agents M it is also inevitable according to the agents M'.*

- Eq. (7) says that all φ that are inevitable according to the personality are also inevitable believed.

  The relation says that all facts that are inevitable true according to the personality of an agent are in consequence believed. Although, personality influences the opinions of a person, by, e.g., contributing to factors such as how fast we accept new knowledge and how interested we are in other arguments enhancing the own belief base, this relation is far too strong to be acceptable.

- Eq. (8) says that an agents' beliefs that are inevitably true are also inevitable according to the agent's personality.

  This relation is a variant of Eq. (2) and is still too strong as property for explaining the effects of personality. It would intend that the entire belief base of an agent stays in relation to the personality. It is easy to find facts that are independent of the personality like norms or laws or my name, for example. Thus, this property is too strong, but, in fact, psychologists found that a weaker version of this property may be applicable (*cf.* Baumert and Schmitt, 2012). We will come back to this kind of relation later.

Table 9.4.: Pairwise **structural subset** interrelationship that capture an agents attitudes to inevitabilities. The tendency column summarises the discussion w.r.t. the research question: + in general accepted, 0 meaningful in some cases, − in general rejected.

	Relationship	**Formula Schema**	**Tendency**
(7)	$\mathcal{B}(\llbracket i \rrbracket) \subseteq_{sub} \mathcal{P}(\llbracket i \rrbracket)$	(Per i A(φ)) $\Rightarrow$ (Bel i A(φ))	−
(8)	$\mathcal{P}(\llbracket i \rrbracket) \subseteq_{sub} \mathcal{B}(\llbracket i \rrbracket)$	(Bel i A(φ)) $\Rightarrow$ (Per i A(φ))	−
(9)	$\mathcal{D}(\llbracket i \rrbracket) \subseteq_{sub} \mathcal{P}(\llbracket i \rrbracket)$	(Per i A(φ)) $\Rightarrow$ (Des i A(φ))	−
(10)	$\mathcal{P}(\llbracket i \rrbracket) \subseteq_{sub} \mathcal{D}(\llbracket i \rrbracket)$	(Des i A(φ)) $\Rightarrow$ (Per i A(φ))	−
(11)	$\mathcal{I}(\llbracket i \rrbracket) \subseteq_{sub} \mathcal{P}(\llbracket i \rrbracket)$	(Per i A(φ)) $\Rightarrow$ (Int i A(φ))	−
(12)	$\mathcal{P}(\llbracket i \rrbracket) \subseteq_{sub} \mathcal{I}(\llbracket i \rrbracket)$	(Int i A(φ)) $\Rightarrow$ (Per i A(φ))	−

- Eq. (9) says that all φ that are inevitable according to the personality are also desired to be inevitable.

  This is another version of Eq. (3) and must be rejected for the same reason.

- Eq. (10) says that all desires that are inevitably true cannot stay in conflict with the ones the agents' personality tends to.

  There are desires independent of the personality. Indeed, there always exist essential desires like survival that are independent of the personality of an individual. Still, it seems to be too strong. Suppose that I am introverted and have the desire to be extraverted. Then, one could argue it makes no sense for me to desire to be a more extraverted person, as there are no actions I can process to fulfil this desires (remember: personality is stable during adulthood). So this schema is too strong.

- Eq. (11) says that all φ that are inevitable according to the personality are inevitably intended by the agent.

  This is a variant of Eq. (5) and is helpful for the intention selection process of an agent. However, it indicates that all intentions are related to the personality of an agent, which is not an interesting property for personality-bounded agents.

- Eq. (12) says that if an agent intends something as inevitably true, it cannot stay in conflict to the agents' personality.

  We can reject this property as there are intentions motivated by stronger attitudes than the personality of an agent. Examples are the satisfaction of basic requirements of an individual, e.g. the physiological needs described in Maslow's hierarchy of needs (Maslow, 1943).

As already mentioned, personality is a kind of heuristic for the deliberation process of agents. So it seems that it is mainly related to the option generation, the filter process and the actual action selection. In the following, we will discuss an agent's attitudes to options and will find situations where personality indeed influences the facts an agent believes.

9.5.3. Options

An agent attitudes to options say that for if φ is possible on the left-hand side it is also possible on the right-hand side. For a worlds' structure, this implies that φ is true on at least one path. Fig. 9.5 provides an example for such a relation adapting the prior used relationship between the belief and personality modality.[9] Giving this example, optional means that some φ is true on at least one path in w, w', and w''. In our example, this applies for r on path (t_0, t_1, t_4, t_8).

[9] Theorem 5.4. (Wooldridge, 2000, p. 97) defines what is meant with options.

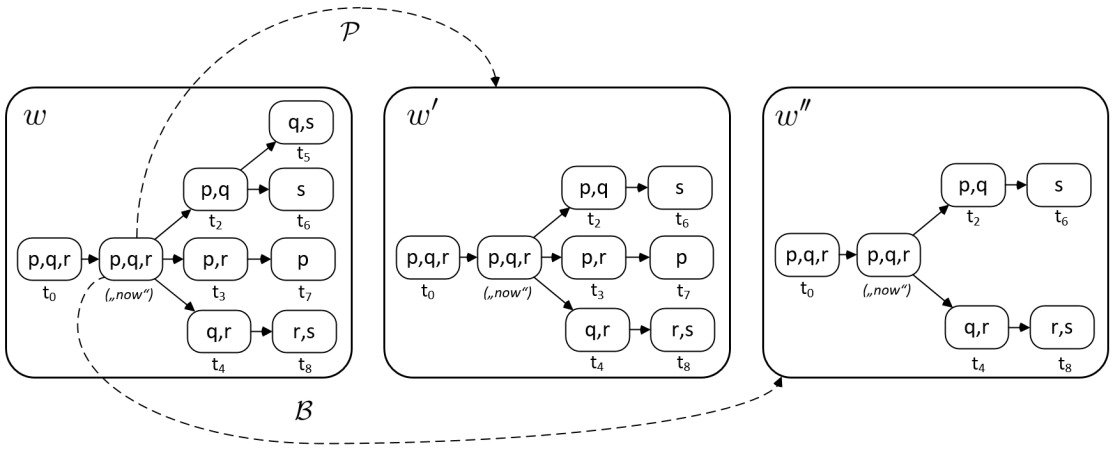


Figure 9.5.: Illustration of an agent's attitudes to options. Here, r is true on both sides on path (t_0, t_1, t_4, t_8). This could be some r the agent is interested in on the one hand and beliefs in in some future on the other.

Table 9.5.: Pairwise **structural superset** interrelationship that capture an agents attitudes towards options. The tendency column summarises the discussion w.r.t. the research question: + in general accepted, 0 meaningful in some cases, − in general rejected.

	Relationship	**Formula Schema**	**Tendency**
(13)	$\mathcal{B}(\llbracket i \rrbracket) \subseteq_{sup} \mathcal{P}(\llbracket i \rrbracket)$	(Per i E(φ)) $\Rightarrow$ (Bel i E(φ))	0
(14)	$\mathcal{P}(\llbracket i \rrbracket) \subseteq_{sup} \mathcal{B}(\llbracket i \rrbracket)$	(Bel i E(φ)) $\Rightarrow$ (Per i E(φ))	0
(15)	$\mathcal{D}(\llbracket i \rrbracket) \subseteq_{sup} \mathcal{P}(\llbracket i \rrbracket)$	(Per i E(φ)) $\Rightarrow$ (Des i E(φ))	0
(16)	$\mathcal{P}(\llbracket i \rrbracket) \subseteq_{sup} \mathcal{D}(\llbracket i \rrbracket)$	(Des i E(φ)) $\Rightarrow$ (Per i E(φ))	0
(17)	$\mathcal{I}(\llbracket i \rrbracket) \subseteq_{sup} \mathcal{P}(\llbracket i \rrbracket)$	(Per i E(φ)) $\Rightarrow$ (Int i E(φ))	0
(18)	$\mathcal{P}(\llbracket i \rrbracket) \subseteq_{sup} \mathcal{I}(\llbracket i \rrbracket)$	(Int i E(φ)) $\Rightarrow$ (Per i E(φ))	0

The structural superset relations capturing an agent attitudes to options are listed in Table 9.5. Intuitively, these relations seem to be more acceptable than inevitabilities as they capture behaviours that are optional. However, they restrict the agent's behaviour to those that keep the optional path open for all possible future. To clarify this behaviour, let us again consider each of the relations individually. Given two modalities M and M', the listed formulae follow the structure (M i E(φ)) $\Rightarrow$ (M' i E(φ)). The structure used for the interpretation reads as follows: *If sth. is possible to fulfil according to the agents M it is also possible to fulfil according to the agents M'.*

- Eq. (13) says that if sth. is possible to fulfil according to the agent's personality it is also possible to fulfil according to the agent's beliefs.

  In other words, one could say that Eq. (13) expresses that if an agents' personality is interested into φ to be true in at least one future, then the agent optionally believes that it is true in at least one future. This schema intends that the personality

of an agent prefers particular facts to be true. We can easily find examples where such statements make sense, e.g. thinking about conservative people that have specific opinions regardless of the argumentation; making Eq. (13) an interesting property for personality-bounded agents.

- Eq. (14) says that if sth. is possible to be believed it is also possible to be preferred by the personality.

  We can find many facts which are independent of the personality (e.g. one's tax burden) on the one hand, and facts that are not preferable by (healthy) personalities, on the other. For example, suppose I believe there is some future in which I am not, which is a fact which applies to all humans in some future. There is no reason always to keep an option where I prefer such a situation. However, Eq. (14) is a weaker variant of Eq. (2) and Eq. (8) and might be meaningful in some situations.

- Eq. (15) says that if sth. is possible to fulfil according to the agent's personality it is also possible to fulfil according to the agent's desires.

  In general, it says that if an agents' personality tends to φ to be true in at least one future, then the agent desires that φ is true in at least one future. In other words, one tries to follow its preferences, which makes more sense than being restricted solely to the own preferences as expressed by Eq. (9). It is a reasonable property for personality-bounded agents, as it provides a prerequisite for acting consistently over time and situations, which is one of the properties of personality traits, as argued in Section 3.4: Personality traits contribute to the consistency of a humans behaviour over time and situations. That is, the agent keeps the option to select its personal preference.

- Eq. (16) says that if sth. is possible to be desired it is also possible to be preferred by the personality.

  The schema describes a behaviour where the agent tries to prefer all desires. It seems that such behaviour does not hurt anything, though, it would restrict the potential objectives to the set that are not contrary to the personality and that are related to my personality. This might make sense in some situations. However, after generating all options, neither all of them are necessarily related to one's personality, nor it is assured that none of them conflicts with the personality. For example, becoming a Dr.-Ing. includes a lot of unavoidable tasks that can stay in conflict with one's personality (introverted doing presentations, careless persons that have to work conscientiously).

- Eq. (17) says that if an agents' personality prefers φ to be true in at least one future, then it intends it be optionally true.

This characterises a behaviour where the agent tries to prefer all options committed to. In other words, an agent that never drops intentions that are preferred by the personality. Although this is useful in some situations, we already learned that the human's decisions are strongly related to the actual context/situation neglecting the personality influences in some cases.

- Eq. (18) says that if sth. is possible to be intended it is also possible to be preferred by the personality.

  Again we decline this kind of relationship as an interesting property, as there exist intentions which are not related to the personality.

We can see that these properties are weaker versions of the ones' we have already discussed. To set an example, about the possible influence of personality on the belief base of agents — particularly the impact of personality on the belief update function — we refer to the dimension agreeableness. In the FFM, each dimension serves as an overarching container subsuming different lower-level personality traits. Thus, agreeableness is among others associated with the trustfulness of an agent, *i.e.* how trustful the agent is when receiving information from others (McCrae and John, 1992, pp. 178–179). Thus, the personality determines to some extent whether new facts are dropped or added to the belief base. The dimension openness to new experiences at the same time influences what the new measurements mean for the agent, e.g., influencing whether new facts are of interest and overwrite old beliefs or are dropped as the agents only slowly changes its knowledge/motives. When reasoning about the behaviour of agents with personality this can be helpful to explain curious or conservative behaviour.

From the above-discussed formulae, we know that the agent's attitudes to E-formulae capture exactly such behaviour in that it allows the agent to behave with some degrees of freedom. In contrast, the A-formulae are too restrictive, qualifying personality as the major influence on the agent's behaviour that has to be taken into account in every situation. However, for the desires and intentions of the agent, the personality controls to what extent it tries to *"keep its options open"* (Wooldridge, 2000, p. 104) and contributes to a consistent behaviour in the long run. At the same time, these formulae — the E-formulae — allow the agent to decide situation depending. We can imagine co-workers, friends, relatives that provide consistent behaviour either preferring to keep options open or to committing to steady behaviour.

9.5.4. Weak Realism

Until now, we have discussed realism properties, which are mainly too strong to be used to characterise the behaviour of agents that represent personality as mental attitude. Giving the available connectivities and both quantifiers, we can also identify relations

Table 9.6.: Pairwise **weak realism** interrelationships solely, weak realism and inevitabilities, and weak realism and options. The tendency column summarises the discussion w.r.t. the research question: + in general accepted, 0 meaningful in some cases, − in general rejected.

	Relationship	**Formula Schema**	**Tendency**
		Weak realism	
(19)	$\mathcal{B}_t^w([\![i]\!]) \cap \mathcal{P}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Per } i\ \varphi) \Rightarrow \neg(\text{Bel } i\ \neg\varphi)$	0
(20)	$\mathcal{P}_t^w([\![i]\!]) \cap \mathcal{B}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Bel } i\ \varphi) \Rightarrow \neg(\text{Per } i\ \neg\varphi)$	0
(21)	$\mathcal{D}_t^w([\![i]\!]) \cap \mathcal{P}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Per } i\ \varphi) \Rightarrow \neg(\text{Des } i\ \neg\varphi)$	0
(22)	$\mathcal{P}_t^w([\![i]\!]) \cap \mathcal{D}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Des } i\ \varphi) \Rightarrow \neg(\text{Per } i\ \neg\varphi)$	0
(23)	$\mathcal{I}_t^w([\![i]\!]) \cap \mathcal{P}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Per } i\ \varphi) \Rightarrow \neg(\text{Int } i\ \neg\varphi)$	0
(24)	$\mathcal{P}_t^w([\![i]\!]) \cap \mathcal{I}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Int } i\ \varphi) \Rightarrow \neg(\text{Per } i\ \neg\varphi)$	0
		Weak realism and inevitabilities	
(25)	$\mathcal{B}_t^w([\![i]\!]) \cap_{sup} \mathcal{P}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Per } i\ \text{A}(\varphi)) \Rightarrow \neg(\text{Bel } i\ \neg\text{A}(\varphi))$	−
(26)	$\mathcal{P}_t^w([\![i]\!]) \cap_{sup} \mathcal{B}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Bel } i\ \text{A}(\varphi)) \Rightarrow \neg(\text{Per } i\ \neg\text{A}(\varphi))$	−
(27)	$\mathcal{D}_t^w([\![i]\!]) \cap_{sup} \mathcal{P}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Per } i\ \text{A}(\varphi)) \Rightarrow \neg(\text{Des } i\ \neg\text{A}(\varphi))$	0
(28)	$\mathcal{P}_t^w([\![i]\!]) \cap_{sup} \mathcal{D}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Des } i\ \text{A}(\varphi)) \Rightarrow \neg(\text{Per } i\ \neg\text{A}(\varphi))$	+
(29)	$\mathcal{I}_t^w([\![i]\!]) \cap_{sup} \mathcal{P}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Per } i\ \text{A}(\varphi)) \Rightarrow \neg(\text{Int } i\ \neg\text{A}(\varphi))$	+
(30)	$\mathcal{P}_t^w([\![i]\!]) \cap_{sup} \mathcal{I}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Int } i\ \text{A}(\varphi)) \Rightarrow \neg(\text{Per } i\ \neg\text{A}(\varphi))$	0
		Weak realism and options	
(31)	$\mathcal{B}_t^w([\![i]\!]) \cap_{sub} \mathcal{P}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Per } i\ \text{E}(\varphi)) \Rightarrow \neg(\text{Bel } i\ \neg\text{E}(\varphi))$	0
(32)	$\mathcal{P}_t^w([\![i]\!]) \cap_{sub} \mathcal{B}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Bel } i\ \text{E}(\varphi)) \Rightarrow \neg(\text{Per } i\ \neg\text{E}(\varphi))$	0
(33)	$\mathcal{D}_t^w([\![i]\!]) \cap_{sub} \mathcal{P}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Per } i\ \text{E}(\varphi)) \Rightarrow \neg(\text{Des } i\ \neg\text{E}(\varphi))$	0
(34)	$\mathcal{P}_t^w([\![i]\!]) \cap_{sub} \mathcal{D}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Des } i\ \text{E}(\varphi)) \Rightarrow \neg(\text{Per } i\ \neg\text{E}(\varphi))$	0
(35)	$\mathcal{I}_t^w([\![i]\!]) \cap_{sub} \mathcal{P}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Per } i\ \text{E}(\varphi)) \Rightarrow \neg(\text{Int } i\ \neg\text{E}(\varphi))$	0
(36)	$\mathcal{P}_t^w([\![i]\!]) \cap_{sub} \mathcal{I}_t^w([\![i]\!]) \neq \emptyset$	$(\text{Int } i\ \text{E}(\varphi)) \Rightarrow \neg(\text{Per } i\ \neg\text{E}(\varphi))$	0

which we refer to as weaker realism properties.[10] Tables 9.6 list the relevant relations for the general case and the associated attitudes to inevitabilities and options.

Considering the weak realism properties pairwise helps to explain the consistency of behaviour, *i.e.* we can capture the *personality-belief consistency* taking into account Eq. (19) and Eq. (20), the *personality-desire consistency* taking into account Eq. (21) and Eq. (22), and the *personality-intention consistency* taking into account Eq. (23) and Eq. (24). Consistency here refers to expressions that avoid conflicts in behaviour, *i.e.* given two modalities M and M' the interpretation reads as follows:

- (M i (φ)) $\Rightarrow \neg$(M' i $\neg(\varphi)$) – *If sth. is inline with the agents M then its negation is not in line with the agents M'.*

[10] For an introduction into strong realism, realism, and weak realism the interested reader is referred to Rao and Georgeff (1998) or the work of Fasli (2001a,b) on further realism notions.

- (M i A(φ)) $\Rightarrow \neg$(M' i $\neg$A(φ)) – *If sth. is inevitable according the the agents M its negation is not inevitable according to the agents M'.*
- (M i E(φ)) $\Rightarrow \neg$(M' i $\neg$E(φ)) – *If sth. is possible to fulfil according to the agents M its negation is not possible to fulfil according to the agents M'.*

Note that the accessibility relation of each modality gives the set of possible worlds that are relevant from the agents perspective. The consistency, on the other hand, restricts the possible worlds (or the agents perspective) to those worlds that are persistent — or in other words not in conflict/accessible — for two modalities at the same time. In particular, the schemas listed in Table 9.6 read as follows:

- Eq. (19) says that if an agents' personality tends to φ to be true, then the agent does not believe that φ is false.
- Eq. (20) says that if an agent believes that φ is true, then the agents' personality should not tend to φ as false.
- Eq. (21) says that if an agents' personality tends to φ to be true, then the agent should not desire φ to be false.
- Eq. (22) says that if an agent desires that φ is true, then the agents' personality should not tend to φ as false.
- Eq. (23) says that if an agents' personality tends to φ to be true, then the agent should not intend φ to be false.
- Eq. (24) says that if an agent intends that φ be true, then the agents' personality should not tend to φ as false.

To provide a more fine-granular analysis of the consistency between personality, beliefs, intentions and desires, we can again apply the universal path quantifier expressing an agents attitudes towards inevitabilities for weak realism. These schemas read as follows:

- Eq. (25) says that if an agents' personality tends to φ to be true in all possible futures, then the agent should not believe that φ can be avoided.
- Eq. (26) says that if an agent believes that φ is true in all possible futures, then the agents' personality should not tend to avoid φ.
- Eq. (27) says that if an agents' personality tends to φ to be inevitably true in all possible futures, then the agent should not desire to avoid φ.
- Eq. (28) says that if an agent desires that φ is true in all possible future, then the agents' personality should not tend to avoid φ.

- Eq. (29) says that if an agents' personality tends to φ to be inevitably true in all futures, then the agent should not intend to avoid φ.
- Eq. (30) says that if an agent intends that φ be true in all possible future, then the agents' personality should not tend to avoid φ.

Finally, the existing relations concerning options are listed read as follows:

- Eq. (31) says that if an agents' personality tends to φ to be possible on at least one future, then the agent should not believe that φ is not possible.
- Eq. (32) says that if an agent believes that φ is true in at least one future, then the agents' personality does not tend to φ as not optional.
- Eq. (33) says that if an agents' personality tends to φ to be possible in at least one future, then the agent should not desire that φ is not possible.
- Eq. (34) says that if an agent desire that φ as optional, then the agents' personality does not tend to φ as not optional.
- Eq. (35) says that if an agents' personality tends to φ to be possible in at least one future, then the agent should not intend φ as not possible.
- Eq. (36) says that if an agent intends that φ be optional true, then the agents' personality does not tend to φ as not optional.

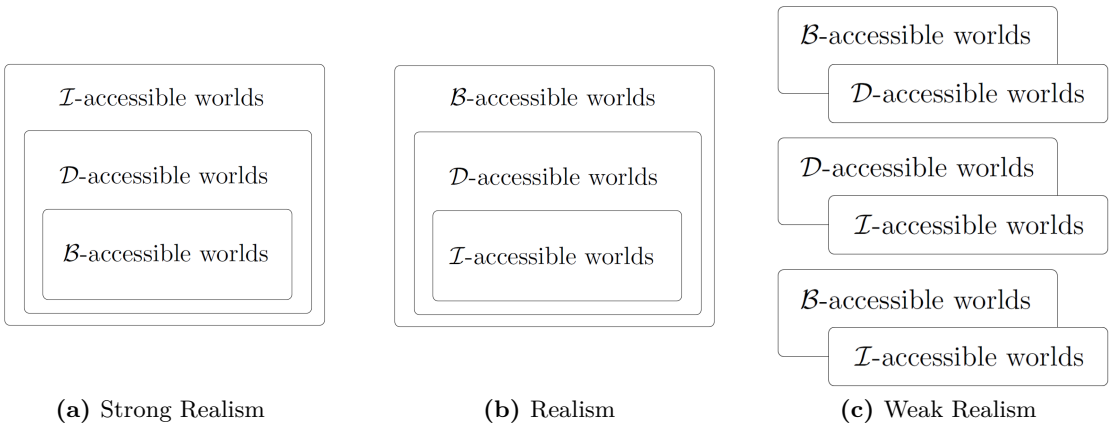
All relations that we have introduced here are variants of the already discussed schemas in that they provide weaker versions of the same relations. This implies that we can accept those schemes that we identified as meaningful in the above discussion. Again the formulae that require that something holds in every future provide an unacceptably strong relation for most agents, in particular for the relationships between personality and beliefs. In contrast, the formulae describing an optional relation are more acceptable. We would expect such behaviour, e.g. for virtual humans. In such real-world cases, the optional relations enable us to reason about opposed characters provided by the same agent. For instance, an agent that behaves selfishly in most cases and generous in the interaction with a subset of other agents. We can observe such demeanour in humans interacting with strangers or colleagues and their relatives or in distinct movie/game characters such as the godfather from the identically named movies.

To conclude, it should be clarified that the introduction of personality leads to the possibility to reason about behaviour that is not rational in the first place. Consider, that rationality regarding our actions means that the actions we apply must be conducive to obtain the selected intentions (Manktelow, 2004, pp. 170–175). Personality here can be seen as a heuristic influencing the action selection and the action execution process even in a way that leads to irrational behaviour, e.g. due to the effects of conscientiousness

which among others indicates how accurately an agent behaves (planned and organised vs easy-going and careless). This can lead to belief-intention inconsistencies, which we now can reason about by taking into account the personality of the agent.

9.6. System of $\mathcal{PBDI}$

Given the four modalities, we can also discuss ternary and quaternary relationships step-by-step. However, given the results from the prior section, it is not very likely that we will find new significant relationships. Furthermore, switching the focus from binary to more extensive relationships makes it more difficult to understand and argue about the meaning of the relation — in particular, with our focus on providing and discussing the intuitive understanding of the concepts, on the one hand; and the ongoing discussion about the definition of personality as an phenomenon and the existence and definition of personality traits, on the other. Thus, we should find another way to integrate the provided discussion into one $\mathcal{PBDI}$ system.



(a) Strong Realism **(b)** Realism **(c)** Weak Realism

Figure 9.6.: Notions of realism as set constraints between the accessible worlds. These illustrations are adapted from published work (Rao and Georgeff, 1998, pp. 318).

To achieve this, we can adopt the view that is provided by the notion of realism an agent follows. Given the modalities and the quantifiers, one can construct arbitrary types of realism, though, the well-known types have been mentioned already: strong-realism, realism, and weak-realism (Rao and Georgeff, 1998). Each of these notions provides a set of constraints, that describe the permitted interrelationships between the modalities. Fig. 9.6 provides an illustration of the notions of realism. Strong-realism, as depicted in Fig. 9.6a, considers a structure in which the set of belief-accessible worlds is a subset of the set of desire-accessible worlds which for itself is a subset of the set of intentions-accessible worlds. This leads to a behaviour where an agent only desires options that it beliefs in and only intends something which it desires and believes. This behaviour could be seen as very cautious. Realism is depicted in Fig. 9.6b and describes

an *'over-enthusiastic'* (Rao and Georgeff, 1998, p. 325) behaviour, in which the set of intention-accessible worlds is a subset of the set of desire-accessible worlds which is a subset of the set of belief-accessible worlds. Thus encoding a behaviour in which the agents believes in all desires and intentions. Finally, Fig. 9.6c depicts weak realism characterising an agent that shows a more balanced behaviour and is enabled to intend something if it does not have a desire for the opposite and does not believe the opposite, whereas it can desire something when it at least does not not believe in it.

The conclusion we can draw from the prior section is that personality induced inevitabilities are not a tangible concept for agents showing personality and that the same applies for relations that ignore the worlds structure. In contrast, the E-formulae are clearly more acceptable as they enable the relations to be optional. However, we also learned that they are not helpful in every situation and that an agents behaviour is influenced by several different aspects. This leads to situation in which a proposition φ, which is preferred by the personality, needs to be dropped for the possible paths an agent can take. Sticking with the personality definition that personality provides *"...both consistency and individuality to a person's behavior"* (Feist et al., 2012, p. 4), we have to find a system that provides personality as such an optional influential factor, which can be negated due to, for instance, the influence of other affective phenomena. This can be done by proposing a set of constraints about the notions of realism.

To do so, we adopt the weak realism as it describes a balanced behaviour between an agent showing a too cautious or too enthusiastic behaviour. Both characteristics, cautious and enthusiastic, are sub traits of the dimensions provided by the FFM (being cautious to an extent is a sub trait of Openness, being enthusiastic to an extent is a sub trait of Extraversion) and would otherwise bias the personality influences towards such behaviours. We interpret agents with personality in the BDI framework as agents that:

1. may adopt a belief even if the personality does not appreciate this belief and may adopt beliefs that are not related to the personality; and

2. may desire a proposition even if it is not preferred by the agents personality and may have desires that are not related to the personality; and

3. may intend something even if the personality does not tend to it and may select and satisfy intentions that are not related to the personality.

In terms of the accessibility relations the set of personality-accessible worlds intersects with the set of belief-accessible, desire-accessible, and intention-accessible worlds. An agent based on these constraints provides a balanced behaviour which is optionally influenced by the personality of the agent. This enables us to apply the E formula discussed in Section 9.5.4 and to narrow these implications to be optional.

9.7. Conclusion

Within this chapter I presented a formalisation that enables reasoning about the influence of personality and enables inferring characteristics of personality using observations of behaviour. By doing so, this chapter contributes to the data processing layer of adaptive systems. The described formalisation extends the approach presented in Chapter 7 in that it formally defines the state and effect of personalities within one of the available BDI logics.

The chapter starts with a state-of-the-art analysis that reveals a gap w.r.t. theoretical considerations of personality influences in agent-based systems. The analysis further shows point of contacts that can be used in future work to integrate the personality formalisation with other affective phenomena. Before defining the necessary elements to integrate the state and effects of personality into the BDI logic, I formulated objectives in terms of statements that should be expressible using my formalisation. After introducing the reader to the selected BDI logic, $\mathcal{LORA}$ – '$\mathcal{L}$ogic $\mathcal{O}$f $\mathcal{R}$ational $\mathcal{A}$gents', I introduced personality into the logic. This enables me to answer the first question that was formulated in the beginning of the chapter, namely: "How to represent the state and effect of personality in BDI logics?"

I define personality as own modality (beside Belief, Desires, and Intentions), as it is a time and space independent phenomenon, which changes only very slowly during lifetime. Furthermore, it is defined as a characteristic that influences the occurrence, intensity, and duration of other affective phenomena such as emotions and moods on the one hand; and influences the decision-making in the BDI lifecycle on the other. While defining the new modality I introduced the necessary elements (additional state formulas to express the personality traits) and operators (additional state formulas to compare personality facets) required to express the prior formulated statements and carefully extended the syntax and semantic of the logic.

Having established the fundamentals, theoretically enables us to express different personalities, to reason about the influence of the phenomenon, and to infer knowledge about personality using observations of behaviour. To elaborate these options, the chapter further provides an analysis of the newly introduced personality modality in the interplay with belief, desire, and intention. This analysis discuss our intuitive understanding of the concepts, and finally enables me to answer the second question presented in the introduction of this chapter, namely: "Which relations between the existing modalities are meaningful for reasoning about behaviour?"

The meaningful relations are those ones that express an optional influence. The rationale for this is that personality contributes to a consistent behaviour of a human in the long run. This means that the personality influence is always present, but there exist the requirement that the influence can be dropped due to other, stronger concerns in every situation. Strong relations, *i.e.* those ones that require personality to be the main

concern in every situation, are unacceptable strong for such behaviours.

It is important to mention, that I presented a basic formalisation and a discussion of helpful relations between the available mental attitudes rather than a comprehensive logical system that represents personality in BDI agents. Thus, there are several aspects that are promising avenues for future research and possible extension/adaptation points:

- **Ternary and quaternary relationships** – Given the four modalities, we can also discuss ternary and quaternary relationships step-by-step. However, given the results from the prior section, it is not very likely that we will find new significant relationships. Furthermore, switching the focus from binary to more extensive relationships makes it more difficult to understand and argue about the meaning of the relation — in particular, with our focus on providing and discussing the intuitive understanding of the concepts, on the one hand; and the ongoing discussion about the definition of personality as an phenomenon and the existence and definition of personality traits, on the other. To achieve this, it maybe feasible to adopt the view that is provided by the notion of realism an agent follows.

- **Personality and Emotions** – One relation that could be beneficial to observe is the relation between formalisations of emotions and our formalisation of personality.[11] Several authors represent emotions via formulas over the believe, desire and intention modalities. Personality also influences the way in which we react to situations emotionally, *i.e.*, the occurrence, intensity, and duration of emotions. The representation of personality presented in this chapter provides a foundation for considering such relations between personality and emotions. To do so, one could integrate the presented findings with the work of Adam (2007) that presents a formalisation of emotions within the decision-making process of BDI agents.

- **Other Logics** – The formalisation of personality in $\mathcal{LORA}$ can be seen as a guideline for integrating personality into other logics of belief, desire and intention. For effects of personality, I propose the formulation as modality operator using the possible world semantic. Other logic systems, that formulate beliefs, desires and intentions in a similar way can transfer the formalisation. This is for example possible when using Kripke models. Our integration of FFM only requires the availability of a representation of agents, which can be expected of an agent-based logic, and of double-valued expressions, for which standard operations exist and can be integrated analogously to our integration in $\mathcal{LORA}$.

- **Other Personality Theories** – One benefit of our model is that it is modular, *i.e.*, the representation of states of personality is independent of the representation of effects of personality. This enables the easy integration of alternate personality

[11] Zelenski (2007) discusses the influence of personality on emotional reactions and provides several reading points for the interested reader.

models. For example, if an approach intends to make use of the MBTI model, only the representation of state of personality has to be adapted.

- **Reasoning** – Although our extension enables the integration of personality into the reasoning process in general, it does not enable to derive relation between personality and behaviour directly. Doing so in a general way requires formalising findings from psychology in the form of statements that can be used for reasoning among multiple approaches. Our formalisation provides a vocabulary to express those statements.

In the next chapter, I will conclude this part of the PhD highlighting the contributions and how they influence the work presented in the next part.

10. Concluding Remarks

Within this part, I worked on isolated aspects of the thesis objective focusing on the integration of personality as an affective phenomenon into agent-based systems. These aspects are related to the elements of an adaptive system that have been introduced in Section 5.3 and individually contribute to the *Data Acquisition*, the *Data Representation*, and the *Data Processing* layer. Fig. 10.1 highlights these relations visualising the main contributions that have been distilled. In the following, I will summarise this part of the PhD chapter-by-chapter and will lead over to the next.

In Chapter 7 – Effects of Personality, I focused on the Data Representation layer and in doing so on the usage of personality models and theories in agent-based systems. The chapter contributes to the Data Representation layer by modelling personality information as user model, using the BDI lifecycle as behaviour model, and discussing how both can be integrated to adapt the application's behaviour to show personality consistent behaviour. Presenting an analysis of the related work, I can summarise that the majority of contributions can be found in the area of agent-based simulation, in particular, agent-based traffic and crowd simulations. Only a couple of the considered contributions substantiate the decision to apply either the FFM, the MBTI, or a simplified representation of personality, which is problematic given the task of adequately transferring psychological findings into computer-processable models. The introduced agent-models are tailored for the specific use-case and environment the agents are living. Work that provides more generalised investigations on personality in agents is rare, and the most advanced applied the MBTI or a subset of the FFM dimensions, justifying the reason to present an own complete model. To do so, I presented an approach that integrates the dimensions of the FFM, which represents the user model, into each stage of an agents' lifecycle according to the BDI paradigm, which represents the behaviour model. Although, personality influences all stages of the BDI lifecycle I substantiate that some traits are more influential in a specific stage and how they can be interpreted and implemented. The implementation grounded the agent-model in a context, completing all elements of the Data Representation layer to a human-behavioural model that was tested in an agent-based simulation. The evaluation showed that the agent-model can simulate personality-related behaviour according to the FFM, lifting the existing results to a state-of-the-art personality theory and all its dimensions. Furthermore, the evaluation provides evidence that a personality-specific task-assignment is beneficial when facing different kinds of tasks as it improves the overall performance reached in the simulated environment that was used – the same observation that is made by psychologist

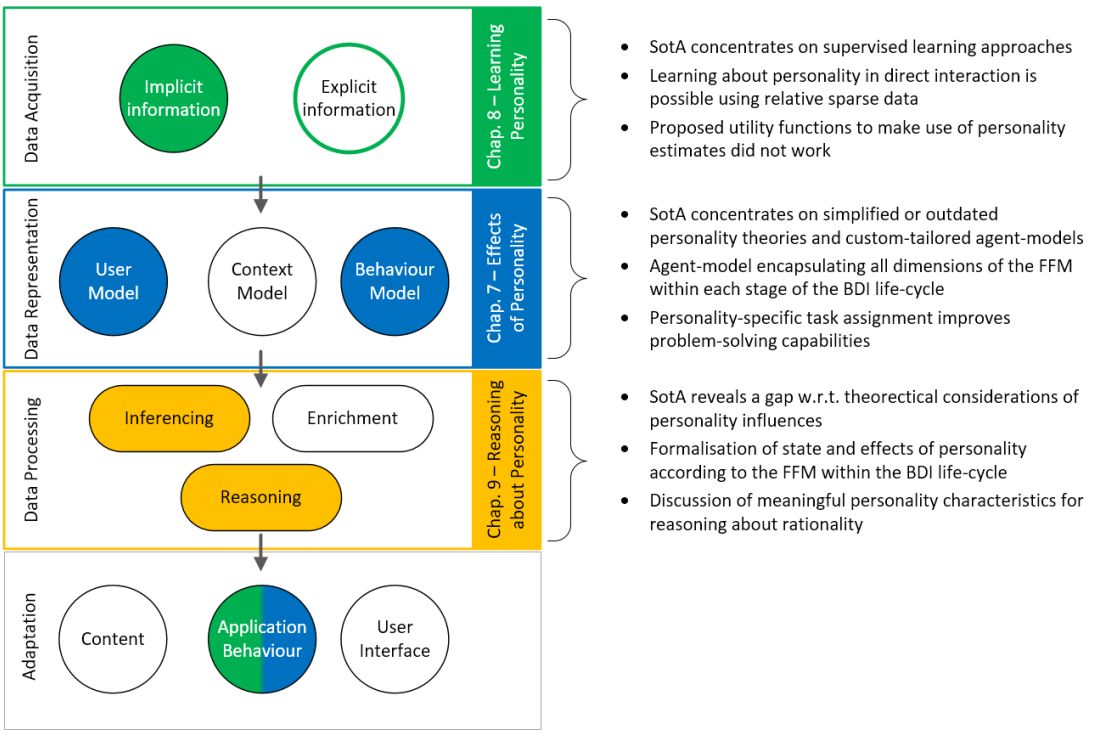


Figure 10.1.: The elements of an adaptive system that have been discussed within this part including the contributions of the individual chapters.

analysing human behaviour.

In Chapter 8 – Learning Personality, I focused the Data Acquisition layer differentiating between explicit and implicit information. While a system designer explicitly models the former or added by the user, the latter is derived indirectly observing the behaviour of users. The chapter contributes to the Data Acquisition layer approaching the task of observing the behaviour of humans and using these observations to build knowledge about the personality of the human according to the FFM. This task is also named Automatic Personality Recognition and utilises implicit information (the distal cues in Brunswick's Lens are implicit information within adaptive systems) to recognise the true personality of an individual. The related work shows many approaches for this topic, though, the existing strand of research focuses on supervised learning methods using (huge) available data sets. Agent-based research on this topic focuses on learning aspects of personalities simulated by other agents, e.g. to improve negotiation outcomes or to study effects on the social welfare of groups of agents. I presented two agent-models w.r.t. this task, designed to learn three out of the five dimensions of the FFM within the Colored Trails Game environment. For this, I used the same approach as before, linking the personality traits to the available actions by interpreting the meaning of the trait taking into consideration the effect of the action. The evaluation shows that both mod-

els can learn about the personality of the humans, though, the results reached for the different traits are differing. The best accuracy was reached for the extraversion trait, while the results for the agreeableness and the conscientiousness trait are varying. One conclusion is that, at least in the selected environment, some traits are easier to learn than others. This conclusion is substantiated by the related work proposing individual solutions for each trait and each observable distal cue. I finally provide a discussion on factors that affect the presented results. As a side question, the agents used the personality estimates to improve their outcome in the game. Therefore, I used utility functions to determine the possible course of action of the human. Using this utility functions, the agents outperformed the humans in average. However, the analysis of the results showed that the results were not related to the personality.

In Chapter 9 – Reasoning about Personality, I focused on the Data Processing layer presenting a formalisation of the concept of personality in BDI logic. The chapter contributes to the Data Processing layer in that it presents a formalisation that enables reasoning about the influence of personality and enables inferring characteristics of personality using observations of behaviour. Personality is defined as own modality (beside Beliefs, Desires, and Intentions), as it is a time and space independent phenomenon, which changes only very slowly during lifetime and influences the occurrence, intensity, and duration of other affective phenomena such as emotions and moods. I defined the necessary elements to integrate the personality modality, thus, providing the foundation for reasoning about the influence of personality on the action selection and inferencing the personality of other agents from their actions. Also, I provide a comprehensive analysis of the newly introduced personality modality in the interplay with beliefs, desires, and intentions. This is done to identify how the properties influence the behaviour of an agent and which characteristics are meaningful/reasonable for analysing (ir)rational behaviour. After discussing the intuitive understanding of the relations, I can argue that personality is an influential factor that contributes to a consistent long-term behaviour on the one hand, and which can be overwritten by more pressing concerns (e.g., goal fulfilment) in every situation on the other. To capture this behaviour, I introduce a notion of realism based on weak-realism that includes personality as world accessibility relation.

In the next part, I will use these findings from the software engineering point of view, introducing an extension to the agent-framework JIAC V that facilitates the development of human-agent teamwork applications. From the discussed work, I know how to model personality information in agents, how to learn about the personality of humans using behaviour observations, and that the personality of human beings accounts for a consistent behaviour which, however, does not mean that the preferred behaviour will be observed in a situation. These points will contribute to the design-decisions that have to be made to provide a programming environment for the monitor, analyse, plan, and execute lifecycle necessary for adaptive systems as seen in human-agent teamwork.

Part IV.

Human-Personality Models – Integration into Agent-Development

11. Preface

This part focuses on the engineering side of joint human-agent activities. It describes how the so far presented findings are used to create a development environment supporting the actual implementation of agents that serve as team members in human-agent teams on the one hand and make use of personality information to predict the behaviour of humans on the other. The outcome of the part is an extension for the agent-framework JIAC V Lützenberger et al. (2013). The extension is named HPLAN (HumanPlan) and brings together different technologies to enable the lifecycle that is common for adaptive systems and is necessary to solve human-aware planning problems.

In Chapter 12 – Related Work, I analyse the related work in agent platforms supporting the development of joint human-agent activities. This is done to investigate to which extent contemporary frameworks support the integration of affective phenomena (e.g. regarding human behavioural models), particularly concentrating on human-aware planning approaches. For this, I introduce a set of categories of interest that is used to classify the related work and highlight the current frontier. Such categories are derived from the elements of joint activities, which are discussed in Section 4.2.1, though they focus on the utilisation of human-behaviour models and the support of learning techniques as those are essential for the thesis objective.

After reviewing development tools for joint activities, I introduce the concept, implementation, and evaluation of HPLAN in Chapter 13 – The HumanPlan Environment. The chapter shows how our personality model can be integrated into the development cycle of agents that plan and learn in teamwork settings. Within the conceptual work, I bring together the design decisions of how to represent the personality, how to learn about the personality, and how to reason about personality with the technical requirements and necessities developers face while implementing agents. For this, I ground the development of HPLAN in practical and theoretical work from the planning domain, applying the conceptual model for planning components (Ghallab et al., 2004, p. 8) and Planning4J[1], which is a library providing a Java API to connect to various available AI planning components; the learning domain, applying reinforcement learning with a look-ahead advice method (Wiewiora et al., 2003, pp. 794–796) named action biasing (Knox and Stone, 2012, pp. 476–478); and the agent engineering community, integrating the extension into the JIAC V framework. I evaluate the technical feasibility of HPLAN

[1] Cerny, M. 2012. Planning4J - Java API for AI planning. Available at `http://code.google.com/p/planning4j`, last-visited: 2017-09-25

using an adapted version of a classic planning domain and testing step-by-step the requirements that I use to distinguish the related work.

At this point, I achieved to present a functional development environment concerning the thesis objective. Developers are enabled to integrate personality information explicitly derived from psychological studies, and the agents are enabled to tailor the interpretation of this information to the individual preferences and habits of a particular human while interacting with this human. Such interpretation is used as a cost estimate for the action selection process during the planning stage, to more efficiently plan in the interaction with humans. As this is a programming environment, it provides the necessary components, interfaces, and communication infrastructure that needs to be calibrated and adjusted to a particular application domain. I do not provide a methodology to do so. However, the *Coactive Design Model* as presented in Section 4.2.2 describes such a methodology and is applied in the final chapter of this part.

Finally, in Chapter 14 – The Personality-enabled Stress Assistant, I present *PeSA*. PeSA is a case study that uses Hplan to implement an agent-based application that accounts for the individualism in the stress-coping of humans by utilising the personality of users to personalise its assistance. The chapter aims to demonstrate how Hplan is used to develop a real-world application based on our human-personality model, psychological findings of a specific domain, and experience gained through the observation of the human's behaviour.

12. Related Work

As many fields can benefit from representing and modelling their systems using the agent-oriented perspective, many platforms and tools exist to develop agent applications. Recent surveys on this topic (*cf.* Kravari and Bassiliades, 2015 focussing on agent-oriented software engineering or Nikolai and Madey, 2009 focussing on agent-based modelling), lists and analyses a vast amount of agent-platforms, at which most of them are still on active deployment. However, the available surveys concentrate on top-level requirements such as the used programming language, license model, user support, performance. In contrast, this chapter outlines the related work in agent platforms supporting the development of joint human-agent activities. In particular, I am interested in the usage of human-behaviour models and to which extent the approaches support the tailoring of such models during runtime. An overview with this focus directly contributes to my hypothesis—formulated in Section 1.1—in that it identifies the authors and groups working in my field, reveals the approaches that cope with the very same idea, and determines concepts and technologies that are useful for my work. Having clarified this, the purpose of this chapter is to answer the question: *What is the state-of-the-art for development environments for joint human-agent activities w.r.t. the use of human-behaviour models?* In doing so, I will carve out if human-behaviour models are a feature in contemporary approaches and how lifecycles, such as the MAPE-K loop introduced for adaptive systems, are supported w.r.t. the task of learning about the human behaviour.

We will start by discussing requirements that should be satisfied w.r.t. human-aware planning problems. To some degree, these requirements have been introduced in Section 4.2.1 as the challenges associated with the task of developing artificial agents that act as team members. From this discussion, we know that many aspects have to be taken into consideration; reaching from the basic compact as the substantial requirement of a joint activity to cost control, which introduces what is named the coordination economy. In conjunction with human-aware planning, these challenges introduce a lot of technical tasks that lead to abilities an agent needs to provide. Within the core of these abilities lies the requirement that an agent needs to be able to understand and solve (human-aware planning) problems in an incremental fashion (e.g. using plan supervision and plan revision techniques). This claim is substantiated by the statement that every joint human-agent activity is a continuous process and subject to an everlasting process. This includes negotiating, testing, updating, adapting, and repairing the mutual understanding of the joint goal, the joint knowledge, and – in consequence – the course of action

(*cf.* Section 4.2.3).

Given the identified challenges and the definition for joint human-agent activities and human-aware planning, I formulate the following categories of interest that are used to classify the related work in the remainder of this chapter:

The first categories of interest are related to the task of solving HAP problems (*cf.* Section 4.3) and restrict the architectures taken into consideration to the ones that can plan and monitor the environment. This restriction is necessary due to the nature of HAP problems which internalise the human behaviour, e.g. using activity and plan recognition techniques, and in consequence, have to act in dynamic and partially observable environments.

Manage State Identifies the ability to monitor and analyse the execution of the current plan. This is required to react to changes in the environment (e.g., effects that take place, appearing/disappearing actors) and to detect possible failures during the plan execution. It includes the ability to reveal the current status of the overall plan and to communicate the intention of the artificial agents.

(Re)Planning Identifies the capacity to generate plans given an objective and an environment, *i.e.* the ability to search for a feasible set of actions that solves a given problem. This capacity includes the reaction to identified failures or changing conditions by replanning or retasking, e.g. when an individual agent's capabilities are outperformed. Retasking is favourable to replanning as it does not require as many resources, though, the ability to retask comes with the requirement to evaluate the viability of the current plan. This requirement is directly related to the demand to solve HAP problems.

I further evaluate if the approaches take the representation of humans into account. Here, I distinguish between social constraints and the introduced human-behavioural models and whether these models are static or introduce some certain kind of learning:

Human-Behaviour Models We have to consider the context-dependent behaviour of humans, that means, that whenever a human is assumed to fulfil a task, the human may fulfil, postpone, fail, or ignore the task and provide results either in time or delayed (Ahrndt, 2013). This ability relaxes one of the assumptions of currently available human-aware planning components assuming that whenever a human is predicted to fulfil a task, the human will provide results in a timely fashion (*cf.* Section 1.1). The human-behaviour models account for the ability of solutions to utilise knowledge about human behaviour, characteristics, and intentions while acting in cooperation with humans; particularly concentrating on the influence of affective phenomena (personality, emotions, moods).

Social Constraints The behaviour of humans is not only influenced by cognitive aspects such as personality, emotions, and moods but also by social constructs like laws, norms, and social rules. With this category of interest, I subsume such types of information, which become necessary, e.g. in multi-user scenarios or while working in teams which provide a hierarchy.

Learning Identifies the ability to learn from experience and to use this knowledge within the interaction in the joint human-agent activity. Learning is necessary, e.g. to produce plans more efficiently or to adjust the communication modality that is used to the addressed agents' capabilities (human agents require UIs where software agents expect procedural calls).

Given the thesis objective, the possibility to integrate human-behaviour models and to learn from the interaction with humans is of utmost importance for me. The ability to monitor and analyse the environment and the actors is a prerequisite that has to be fulfilled just as the ability to plan and replan.

In the following, I will describe the search strategy that was applied to examine the State-of-the-Art in Section 12.1. Afterwards, in Section 12.2, I will introduce the approaches that have been included and describe them in detail. In Section 12.3, I will discuss the results with respect to the former introduced classification. This is done, to distinguish the work on HPLAN from the available strand of research and to highlight the focus I will apply during the next chapters. At last, I will provide final remarks and further reading points by concluding this chapter.

12.1. Search Strategy

In order to identify relevant work, I carry out a literature research in public available sources. That includes querying the databases ACM Digital Library[1] and IEEE Xplore Digital Library[2] and using the research search engine Google Scholar[3]. Table 12.1 lists the used keywords for the search. I extend the search results with the approaches that I already know from the prior parts of the PhD and further scan the paper database of the AAMAS conference series for the last 10 years. From the identified papers, I start a transitive search using the references that are included in the publications to identify further work on the same topic by the authors, e.g. identifying the PhD projects some publications are based on, and by other authors.

Criteria for the inclusion of work are derived from the research question of this chapter. That means I include work from the area of human-aware planning, joint activities, and human-agent interaction (or human-robot interaction) that proposes architectural

[1] ACM Digital Library – `http://dl.acm.org/`, last-visited: 2017-09-25
[2] IEEE Xplore Digital Library – `http://ieeexplore.ieee.org`, last-visited: 2017-09-25
[3] Google Scholar search engine – `https://scholar.google.de`, last-visited: 2017-09-25

Table 12.1.: Keywords used during the search for relevant related work.

Category	Keywords
Planning	Human-Aware Planning, Human-in-the-Loop Planning, Cooperative Planning
Learning	Reinforcement Learning, Learning in Interaction, Human Behaviour Observation
Engineering	Agent-Platform, Agent-Framework, Agent-Runtime
Others	Constructing, Engineering, Implementing, Developing, Modelling, Building, Programming, Concept, Design, Model, Framework, Architecture

concepts for implementing such type of applications. As discussed in Chapter 4, these areas are cannot be sharply differentiated from each other. Thus, as a rule of thumb, I include work in which an autonomous system (agent or robot or a mix of both) plans its interactions with a human being. The work considered is limited to those that provide a clear element of social interaction. I exclude contributions that describe general purpose agent-platforms, as it is not my objective to provide a review of the existing solutions.

12.2. Existing Solutions

Several of the presented requirements can be satisfied using available approaches either from the field of human-aware planning or adjacent research fields (Talamadupula et al., 2010, p. 14:22); namely the primary architecture of such solutions and the actual planning techniques.[4] From the architectural view, the core concepts of managing the state, being able to plan and replan, revealing the current status of the overall plan, and to detect possible failures during the plan execution are supported by the 3-layer architecture of current dynamic planning components (Ghallab et al., 2004, p. 9). Here, a closed loop between the planning level, the monitoring level and the execution/controller level enables interleaved planning and execution, which can be utilised for plan supervision, plan revision and replanning. Fig. 12.1 depicts the three components that are distinguished in such an architecture. Given the initial state, the objectives, and some description of the environment, the planner generates plans that are executed by the controller to achieve the objectives. The controller performs the actions according to the plan within the environment and derives observations about the effects or events of the actions that happen in the environment. The system itself, which represents the environment accessible through some effectors and sensors, evolves caused by the actions and events. This is a more architectural view on what has been described in Section 3.1

[4]The latter are frequently available as open-source libraries, e.g. Planning4J, and are made available after participating at venues such as the ICAPs conference series.

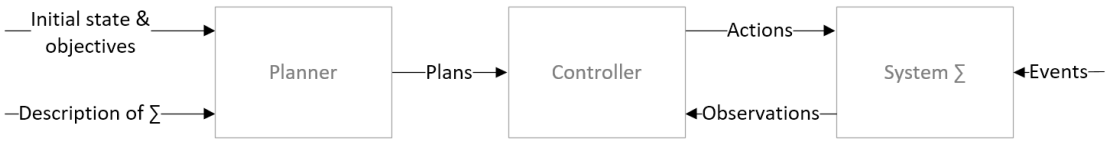


Figure 12.1.: The conceptual model for planning components. This figure is adapted from published work (Ghallab et al., 2004, p. 8).

as the state-transition system. In the remainder of this section, we will see that the different approaches use architectures that a grounded onto this 3-layer structure.

12.2.1. HAP Framework

Cirillo et al. (2008, 2010) and Cirillo (2010) present the HAP framework, which applies the multi-layered structure for dynamic planning components within a human-robot interaction use case. The approach describes the use of intention models to decide whether an agent is allowed to perform its task or if the agent would disturb the human user and should not carry out the task now. This information is used to postpone agent tasks to a more acceptable time frame. Fig. 12.2 provides an overview of the proposed HAP framework. Monitoring of the environment is realised via a human plan recognition module, able to analyse available sensor input. The activity monitoring and recognition engine used for this task is part of the PhD project of Cirillo (2010) and was developed for the use case (Cirillo, 2010, pp. 65–89). The interpretations are forwarded to a monitoring module which triggers the (re)planning process or informs the executor about completed human actions. The controller in this work is named the executor and is an automatic vacuum cleaner. The whole concept was implemented for the described use case, though, it was not realised as a framework that can be used by others. The development of the HAP framework seems to be inactive.

To conclude, the HAP framework can manage the state and (re)plan. In principle, the architecture allows failure detection, though, this is not described in the associated publications. Social constraints are part of the approach, in that it can plan with multiple human actors and in that the task is not to distract humans. Behavioural models are supported in a limited way (as intention models), in that the activity recognition engine predicts the human's behaviour based on typical daily activities (e.g. after waking up in the morning, humans leave the bedroom and go to the kitchen). The HAP framework does not include learning techniques.

12.2.2. Decisional Framework for HRI

Clodic et al. (2005, 2009) present a decisional framework for HRI, introducing different modules together handling a joint human-agent activity. The authors use the example

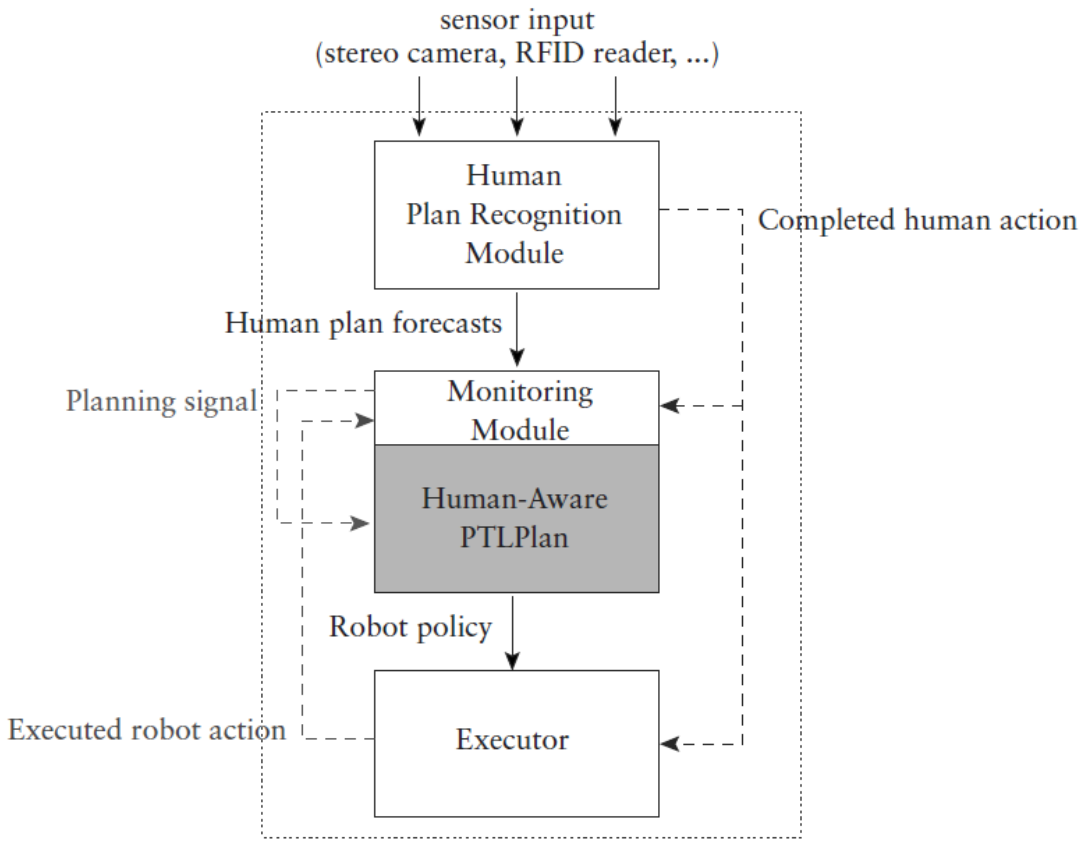


Figure 12.2.: The HAP framework for human-aware planning. This illustration is taken from published work (Cirillo, 2010, p. 94).

of a museum guide robot, to illustrate the purpose of the various modules, whereas the main parts are described in different publications by different authors of the same group:

- **Planning** – The Human Aware Task Planner (HATP) is presented by Montreuil et al. (2007). HATP can estimate the viability of a plan according to several social constraints. The work introduces six different kinds of social rules reaching from undesirable states and sequences of actions to effort balancing and abstraction legibility. The latter one addresses plans which are too difficult to understand for human partners. HATP is, in general, addressed to agents that have to act in the presence and in collaboration with the human users available. The authors introduce a set of human-centred constraints, which are used to calculate a social score to produce socially acceptable plans. However, the agent model for a single agent is rather simple and consists of a set of actions and a set of context-dependent costs for these actions. Consequently, the plan evaluation consists of the context-dependent costs and the social costs. Although the authors mark that a task which is assigned to a human user who can fulfil it can be declined, the introduced approach does not handle this situation.

- **Managing State** – An extension of the HATP is presented by Alili et al. (2009). The work introduces a two-layer structure utilising HATP as task planner and a component called SHARY (Clodic et al., 2009) handling the higher level goal management. SHARY enables the system to refine tasks based on the current context in an incremental fashion and monitors the human behaviour, for example, whether a human user commits to an assigned task or not. The latter ability allows the system to recognise in which context a human user accepts or declines a specific task. Here, the work lacks details of the usage of such information.

- **Human** – Alami et al. (2005, 2006) present a way to adjust the planning procedures to different types of humans using so-called InterActionAgents (IAA). Each IAA represents a human as a kind of proxy or avatar. The authors further discuss a concept for a framework using the information provided by the IAA to produce legible behaviour.[5] Each IAA is responsible for the interaction with a specific human and can provide information about the humans' abilities and preferences. As the authors are working in the area of human-aware motion planning, they use the IAAs to track geometric attributes such as the position, posture, or line of sight.

The architecture of the decisional framework for HRI is depicted in Fig. 12.3 and provides two different components monitoring the environment. On the one hand, the IAA which are responsible for monitoring and analysing the state and actions of the human. On the other, the robot itself, which monitors its actions. The robot supervision kernel encapsulates the controller component and assigns and updates tasks to the IAA using task delegates. Those tasks are received from the agenda, which is responsible for the goal selection and planning and for ensuring that there are no conflicting goals.

Taken into account all the contributions to the decisional framework for HRI, we can conclude that it is capable of managing the state and to (re)plan and that it further provides the most advanced concepts w.r.t. social constraints due to the usage of the HATP. Human-behaviour models are supported in that HATP models humans using a set of attributes including physical characteristics and the mood of the human. However, this representation is limited to the estimate to which degree a human wants to be involved in a particular task (its mood towards a particular task). Although the used architecture is a reasonable approach to learn about the behaviour of humans, in particular using the concept of IAA, learning is not part of the framework.

12.2.3. Plan-based Adaptive Control

Kirsch et al. (2009) present a concept for an integrated planning and learning framework for HRI. This framework accounts for a plan-based adaptive control of HRI applications

[5] Kirsch et al. (2010) give an introduction to the term 'legible behaviour' in the context of joint human-agent activities.

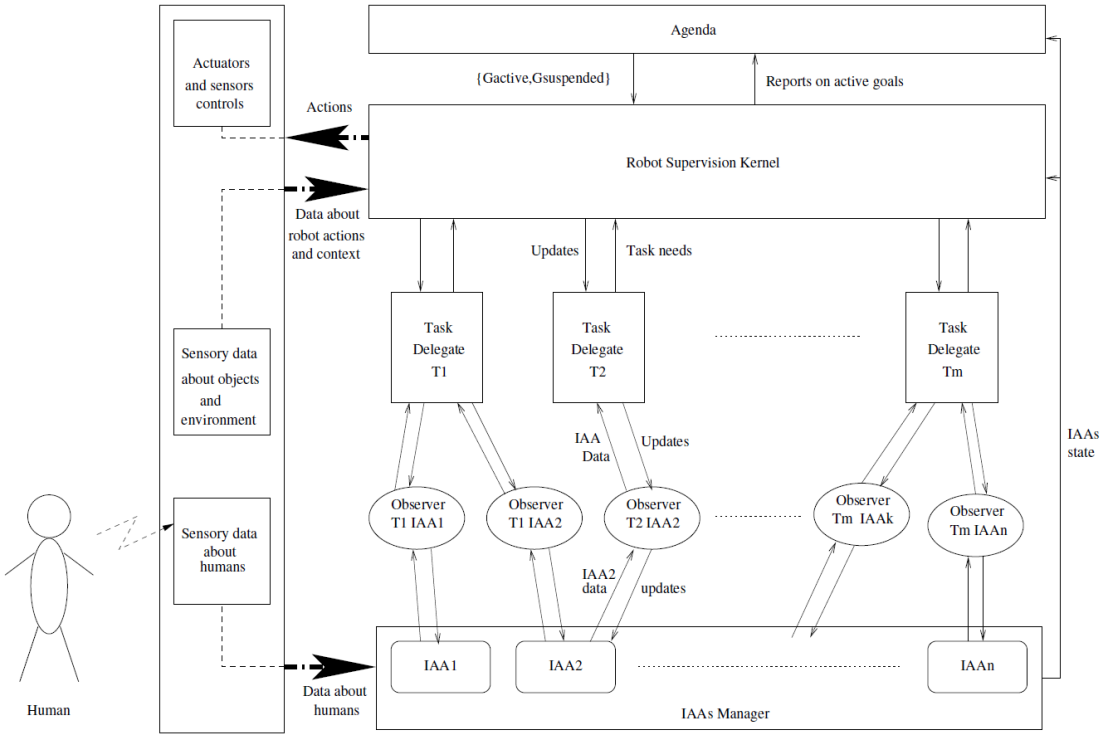


Figure 12.3.: Architecture of the decisional framework for HRI. This illustration is taken from published work (Clodic et al., 2005, p. 545).

and provides the necessary lifecycle elements explicitly highlighting knowledge and learning as important aspects.

Fig. 12.4 presents an overview of the conceptual architecture, which combines and connects different behaviour and intention models with a planning and executing component. The authors argue to use models of human abilities to predict human behaviour and reactions to be able to produce plans that are socially acceptable. Furthermore, the authors state that these models are adjusted through a learning cycle. Nevertheless, the strengths of the system are not the planning techniques or the learning algorithms but the conceptual work of combining two frameworks to facilitate joint human-agent activities. The approach consists of the two components: TRANER (Müller, 2009; Müller et al., 2007) and RoLL (Kirsch, 2008). The first one provides a prefabricated plan library and execution environment for autonomous household agents. The latter provides a mechanism to learn human behaviour during runtime and influences the selection of appropriated plans accordingly. The applicability of a plan to the current situation is evaluated using an integrated simulation module. Therefore, the system enables the agent to continually improve its cooperation behaviour according to the preferences and abilities of the human user. However, the framework also emphasises some shortcomings as it lacks to use individual or dynamic planning techniques. Furthermore, the work lacks details about the used learning techniques, in particular how the simulation com-

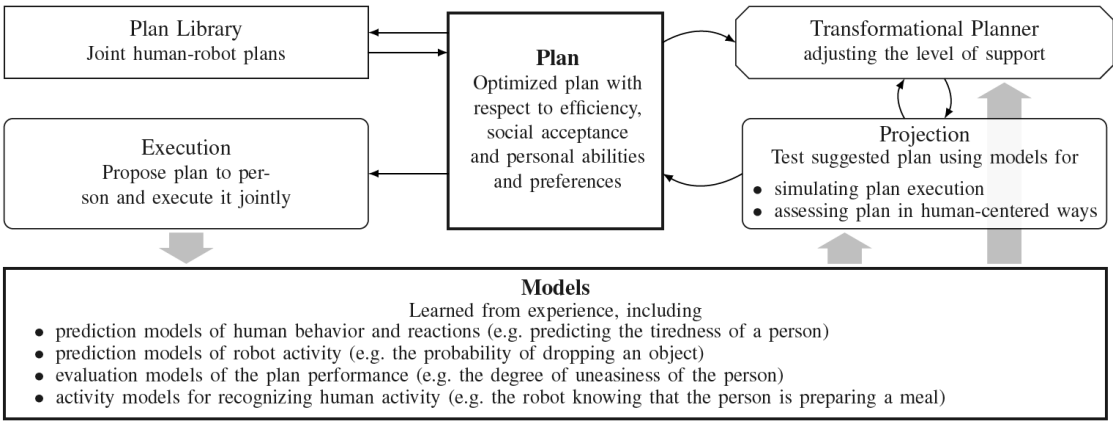


Figure 12.4.: Proposed framework for the plan-based adaptive control of HRI applications. This illustration is taken from published work (Kirsch et al., 2009, Fig. 2).

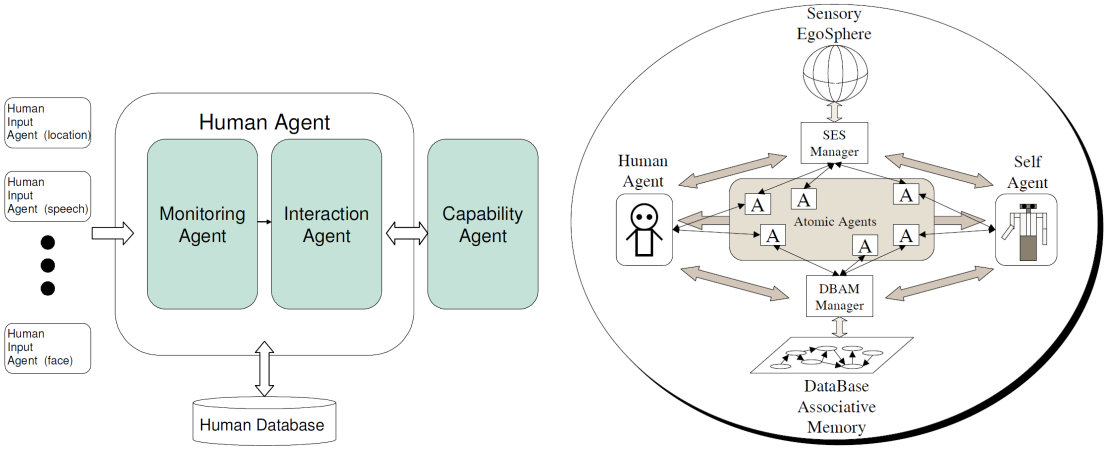
ponent works. Consequently, the authors mark that *"[t]he strength of the system doesn't lie in individual planning techniques or learning algorithms."* (Kirsch et al., 2009, p. 3). In a successive work (Kirsch et al., 2010), the authors describe the requirements for the human-behaviour models in more detail. The work on the proposed framework seems to be inactive; remaining in a conceptual state.

In summary, we can state that the authors recognise the need for managing the state, for being able to (re)plan, for modelling the human behaviour and taking social constraints into account, and for learning about humans to evaluate and optimise the efficiency of the system. However, providing a conceptual work limits the level of detail within the different categories of interest. The plan library itself is prefabricated and thus not a complete planning approach. The authors particularly emphasise the need for models about human behaviour and individual abilities and preferences but do not provide details about the acquisition, representation, or usage of such information. Finally, the framework outlines a learning cycle and highlights the importance of adaptation through learning, though, details on how to observe or assess the efficiency of the system in a human-centred way are missing.

12.2.4. The Human Agent System

Rogers (2003); Rogers et al. (2005); Rogers and Wilkes (2000) present a multi-agent system architecture build for joint human-agent activities between multiple people and service robots; the Human Agent System. This system was build based on the observations that in the HRI domain the interaction between humans and robot should be natural and robust (Rogers and Wilkes, 2000, p. 864). With the work, the authors approach two requirements. The first being the development of a robot with a humanoid shape. The second being the development of naturally communication capabilities. To address these issues, the authors coined the term human-humanoid interaction, which is

rather uncommon in these days.



(a) Internal architecture of a Human Agent. **(b)** Architecture of the entire system.

Figure 12.5.: Architecture of the Human Agent System described by *Rogers et al.* Fig. 12.5a is taken from published work (Rogers et al., 2005, Fig. 2). Fig. 12.5b is taken from published work (Rogers and Wilkes, 2000, Fig. 3).

Fig. 12.5 illustrates the architecture of the presented system with an internal view of the human agent on the left side and an overview of the entire system on the right. In this architecture, the Human Agent takes a role comparable to the concept of IAAs introduced before. It provides the systems internal representation of the humans that are in the environment of the agent, encapsulating different aspects (physical, task-related, and cognitive characteristics) that make out a human. Information about these aspects is provided by a database. The agent itself monitors the environment using other specialised agents observing the environment and pre-processing sensor data. The Interaction Agent is responsible for reasoning about the intentions of the human and for planning in response to the monitoring and intention recognition process. Once the high-level planning is done, the Capability Agent executes the task w.r.t. the platform that is represented. The counterpart of the Human Agent is the Self Agent that represents and monitors the activities of the robot within the environment. From the point of view of a Human Agent, the Self Agent is the Capability Agent, whereas the Atomic Agents in Fig. 12.5b provide basic functionalities like speech or face recognition, i.e. sensor and effector abstractions. The implementation of the system focus on the task of recognising and identifying a human in the environment. In another work, Kawamura et al. (2002) introduce further aspects of the system that are responsible for integrating social rules into the interaction with the human, distinguishing between interaction channel preferences.

The Human Agent System identifies and approaches several of our categories of interest and introduces components responsible for managing the state and planning the

cooperation. Implementation details about the applied techniques are missing. Human behaviour models are integrated in terms of an intention models and intention recognition distinguishing between expressed in inferred intentions. The authors further recognise the need for modelling and recognising cognitive aspects, though, leave these aspects as open research questions. Social constraints are involved in that the system can differentiate between levels of engagement and can alternate its behaviour between searching for interaction, initialising interaction, and interaction with a human. HAI can be adapted w.r.t. the individual human and its preferred interaction channel (GUI vs Voice vs Gesture). The system describes that the identified preferences are stored in a component named Human Database, however, due to the limitations of the described approach we do not acknowledge this as learning.

12.2.5. Others

Talamadupula et al. (2010) describe a planning architecture for human-robot teaming in open worlds, focusing on the technical aspects required for a robot to plan and act in a shared environment. The authors reuse and adapt existing technologies from the planning domain to be used in joint activities where possible. Although it is stated that the work is placed in the human-aware planning domain, it is rather a mixed-initiative approach. One interesting aspect is the introduction of Open World Quantified Goals, which are conditional goals that are tracked and grounded during the execution to react faster to changing environments and unexpected team member behaviour.

In a following work, Chakraborti et al. (2016) present a formal framework for studying human-agent teamwork. Here the authors recognise and emphasise different aspects that make out a humans behaviour and that an optimal plan w.r.t. individual costs is not necessarily the optimal plan when it comes human-aware planning. The work closes without providing a reference model that addresses the metrics of the proposed formal framework.

Rosenthal et al. (2010) emphasise how human-agent cooperation enables a team to accomplish tasks that the team members cannot fulfil on their own in a visitor companion task. The work makes extensive use of plan changes and retasking and also introduces ways to replan in situations where individual agent capabilities are outperformed. Models of human behaviour are not used, though, the human is characterised by its location and its current intention.

Pandey and Alami (2010) describe what they named Mightability Maps, which are models that provide an estimate of what a human might be able to see, reach, or grasp. Thus, the mightability maps provide knowledge about the human and its potential abilities and are updated during runtime to the current positions and physical orientation of the human using visual-spatial reasoning. The maps itself are used for the robots action planning. Although the authors claim to present a framework, the work itself is limited to the actual use case of hand-over tasks.

Trafton et al. (2012) introduce an architecture that simulates the human using the ACT-R (Adaptive Character for Thought-Rational) architecture to reason about the mindset of humans. The reasoning results are used to improve the interaction with teammates. However, the author's definition of teamwork differs from mine, in that the work concentrates on providing human-like robot behaviour in a shared environment without any joint goals. The interesting aspect is the attempt to simulate a humans behaviour online. In doing so, the work differs from the human-behaviour models we discuss in that it simulates the human brain functionalities and does not focus on providing a knowledge base about behaviours.

12.3. Discussion

The related work shows that technical requirements such as the ability to plan, to replan, to act in cooperation with humans, and to monitor the environment and detect failures using different types of sensors can be satisfied using contemporary technologies. At the same time, the related work reveals aspects that are only barely covered by solutions, especially when it comes to models about behavioural aspects of humans and the adaptation of these models during runtime, *i.e.* learning about human characteristics. Furthermore, the majority of work was found in the area of HRI, frequently concentrating on human-aware motion planning.

Table 12.2 summarises the results and classifies the introduced approaches w.r.t. the introduced criteria. At the same time, the table highlights the focus of HPLAN, which will be introduced in the next section. Within the table, I distinguish between different work contributing to the decisional framework for HRI introduced in Section 12.2.2. Furthermore, I provide a more fine-granular classification explicitly differentiating between manage state and failure detection and between behavioural and intention models as the related work revealed the necessity to further differentiate within our categories of interest.

It is emphasised that technical requirements are satisfied by most of the approaches. This is typically done by implementing the 3-layer architecture of dynamic planning components and establishing a lifecycle of planning, executing and monitoring. Additionally, we can conclude that there exists some work related to the integration of specific information about humans. For example, Montreuil et al. (2007) presents a solution that integrates social constraints into the actual planning process. In contrast to social constraints, the integration of information about behaviour (e.g., personality traits) and intentions are not as well advanced. At the same time, different authors involved highlight the need for a novel representation of humans (*cf.* Cirillo et al., 2010; Kirsch et al., 2009; Trafton et al., 2012; Pandey and Alami, 2010). The conceptual work of Kirsch et al. (2009) present a first approach to integrate such representations, which can be adapted at runtime using learning techniques. The concept of InterActionAgents

Table 12.2.: Classification of the presented approaches. The table entries reads as follow: + the approach supports this feature, o the approach does not fully support this feature but it supports it in some (weak) way (e.g. identified as needed, minor implementation of this category), – the approach does not support this feature.

	Cirillo et al. (2010)	Clodic et al. (2005)	Montreuil et al. (2007)	Alili et al. (2009)	Alami et al. (2005)	*Decisional Framework for HRI*	Kirsch et al. (2009)	Rogers and Wilkes (2000)	Kawamura et al. (2002)	Rosenthal et al. (2010)	HPLAN
Manage State	+	+	+	+	o	+	o	+	+	+	+
Replanning	+	+	+	+	o	+	o	+	+	+	+
Failure Detection	o	+	+	o	o	+	o	+	+	+	+
Social Constraints	o	o	+	+	+	+	o	o	o	–	–
Behavioural Models	o	–	o	o	–	o	o	–	o	–	+
Intention Models	+	–	o	o	–	o	o	o	o	o	–
Learning	–	–	–	–	–	–	o	–	–	–	+

introduced by Alami et al. (2005) present a way to encapsulate different information about individuals and is a natural way when developing agent-oriented.

12.4. Conclusion

Within this chapter, I presented a state-of-the-art analysis focusing on frameworks and architectures supporting the development of joint human-agent activities. In doing so, I excluded major agent platforms and general purpose frameworks. Those are surveyed with a focus on top-level requirements by different authors (*cf.* Kravari and Bassiliades 2015; Nikolai and Madey 2009). My particular interest is the usage of human-behaviour models and the use of learning techniques to adapt these models during runtime to identify authors, groups, approaches, and concepts in this field.

Based on the presented analysis and the provided discussion, I can answer the research question that was formulated at the beginning of the chapter: "What is the state-of-the-art for development environments for joint human-agent activities w.r.t. the use of human-behaviour models?"

I achieved to identify four major contributions relevant for this part of the PhD. All of them originate from the HRI community and present thoughts about possible architectures to develop artificial team members. Introducing the contributions and further papers that handle partial aspects of the engineering problem, it appears that the need for further information about humans intention, behaviour, social constructs is present in the community. However, the majority of the work focuses on social constraints like the legibility of plans. Focusing on human-aware motion planing the modelling of physical characteristics of humans is also researched. We can identify mixed results for human-behaviour models that model cognitive characteristics. Although the authors recognise the need for such information, only the work of Kirsch et al. (2009) explicitly considers them. However, the work itself is conceptual and does not provide any insight into how to model or use such human-behaviour models in planning processes for teamwork. At the same time, the contributions of Kirsch et al. (2009, 2010) are the only ones recognising that the artificial team members have to adapt the models about humans they use. This is of surprise, as all of the systems would enable a learning cycle due to the implemented lifecycle where the robots monitor and analyse the environment to detect failures.

In the next chapter, I will introduce HPLAN which makes use of some of the here presented concepts. In particular, I will present an agentfied version of the conceptual model for planning components (*cf.* Ghallab et al., 2004, p. 8); I will adopt the idea of IAAs (*cf.* Alami et al., 2005) and the Human Agent (*cf.* Rogers, 2003) to the agent-framework JIAC V; I will reuse existing technologies where possible as recommended by Talamadupula et al. (2010); and I will integrate a human-behaviour model in the way postulated by Kirsch et al. (2009).

13. The HumanPlan Environment

This chapter introduces the concept, implementation, and evaluation of the HumanPlan (Hplan) environment. Hplan facilitates the implementation of joint human-agent activities by providing a development environment that integrates planning and learning techniques together with our human-behaviour model into a contemporary general purpose agent framework. The goal of this chapter is to answer the question: *How can we integrate our human-personality model into the development cycle of agents?*

By doing so, I will conceptualise a way for developers to provide more information about humans to the planning process of joint human-agent activities in Section 13.1. In Section 13.2, I will describe technical details about the implementation of Hplan within the JIAC V framework, before evaluating the technical feasibility of Hplan using an adapted version of the Blocks World in Section 13.3. Finally, in Section 13.4, I conclude the chapter and discuss the collected experiences.

13.1. Concept

To enable the development of agents that make use of psychological findings on the one hand and account for the individualism of humans on the other comprises at least three components. First, the user model, which defines the human user's personality. Second, the behaviour model, which defines the effects of a personality, *i.e.* the personality preferences for specific behaviours in compliance with psychological findings. Third, a component combining the information provided by the user and behaviour model in a way that is usable for the agents' decision-making.

We refer to the first two components as the human-behaviour model (*cf.* Section 5.3) and provide insights into the modelling of these components in the prior part of the PhD. Thus, the first requirement Hplan has to address, is the possibility to represent a human's personality and to represent psychological finding w.r.t. personality preferences.

The third component uses the human-behaviour model to adapt the agent's behaviour to the human. Here, we can differentiate between the action planning, *i.e.* how can we use the human-behaviour model during the actual planning process, and the observation of the human's behaviour, *i.e.* how is the experience gained during the execution used as a feedback loop to tailor the human-behaviour model. This adds the requirement for Hplan to use an agent-model that is able to plan, to execute the plans, learn from the execution and start over (*cf.* Section 3.3). Therefore it needs the ability to determine

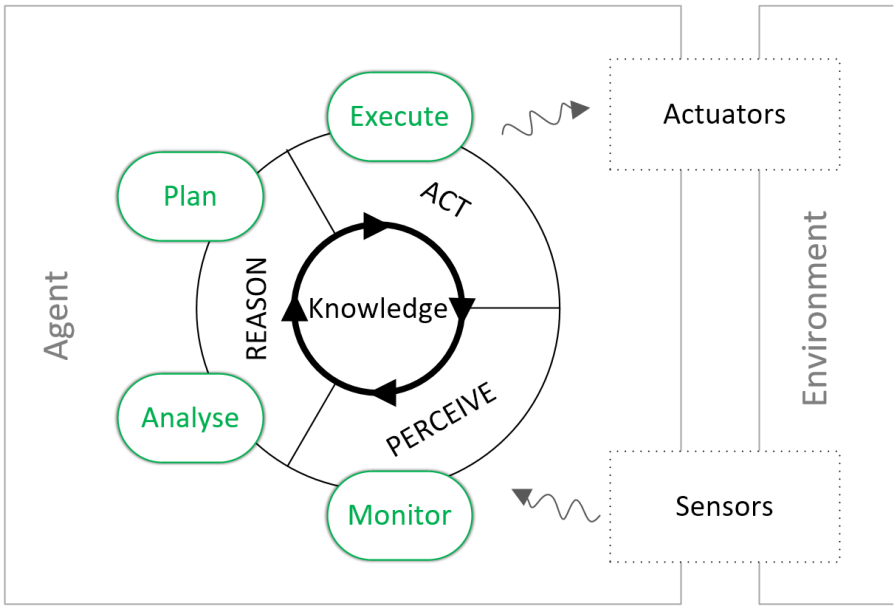


Figure 13.1.: Considerations about the lifecycle of HPLAN agents.

the current state of the environment, to detect failures and to replan if necessary. Furthermore, experiences generated from the execution of actions must be used to improve this process. We already introduced a lifecycle that enables this w.r.t. adaptive systems: The MAPE-K loop (*cf.* Section 5.4). In agent-based systems, there are other terms for the very same type of lifecycle. Each of them identifies the essential parts of monitoring the environment, analysing observations, planning the course of action, and executing actions, e.g. the sense-plan-act cycle or the perceive-reason-act cycle. As these kinds of lifecycles (or control loops) are inherent to most types of agents (Russell and Norvig, 2002, p. 2), we will utilise them in HPLAN and use the terminology of the MAPE-K loop next. The considerations about the lifecycle are visualised in Fig. 13.1. For HPLAN, I assume that each agent, which acts as a team member, follows such a lifecycle to satisfy the requirement that the interaction is ongoing, and that team members act proactively to achieve a joint goal. For instance, to take part in collaborative decision-making as required by Challenge 8 – Collaboration (*cf.* Section 4.2.1).

Based on these initial thoughts, the concept of HPLAN is built around three ideas, which read as follows and are grounded in the related work and experiences collected during this doctoral research study:

- **Cost Estimates** – The idea to more efficiently plan in joint human-agent activities is to forward more information about humans to the planning component. I accomplish this in terms of a more accurate cost estimate for specific capabilities, which is build using the personality of the human, psychological findings about personality preferences, and the actual observations of the human's behaviour. These

costs are used to determine the likelihood that an action is preferred or will be performed next, *i.e.* lower costs indicate a higher likelihood and vice versa. As I do not aim to develop a new planner but reuse existing solutions as far as possible, my approach is to utilise the cost estimates to influence the action selection of a planning process, where the actual planning procedure is a black box. The idea is to integrate the costs into the planning process providing a—roughly speaking—dynamic heuristic about the utility of assigning a task to a given actor/assuming that a task will be performed by a given actor.[1]

- **Actor Agents** – I represent each natural agent as an avatar in the computing system, similar to the concept of InterActionAgents as presented by Alami et al. (2006) and the concept of the HumanAgent presented by Rogers (2003). These avatars are named actor agents. Each actor agent encapsulates the interaction with one human including the monitoring and analysis of the behaviour. Furthermore, it provides the information for the actual planning process to the planning entities by estimating the costs for the available capabilities given knowledge about the personality of the human that is represented, the default behaviour of humans, and the experience that was already gained. The actual execution of actions is left to the developer in that the actor agent opens a possible interaction channel with the human or just assumes that the currently planned action is executed and observes the progress. In doing so, I make use of the modularization of agent-oriented systems as preliminary architecture decision for individualising the interaction.[2]

- **Prior Knowledge** – The tailoring of the human-behaviour model during the interaction with the human is one goal of the thesis and is coined to account for the individualism of humans. In Chapter 8 we learned that agents are, in principle, able to learn about the personality of a human, though, we failed with the task of using this knowledge to more efficiently plan the interaction with the human. The discussion revealed that many more factors are influencing a human's behaviour. Furthermore, work of other authors (*cf.* Du, 2013; Du and Huhns, 2013) substantiated that making predictions of behaviour solely based on personality information may be the wrong way. In Hplan, we address this issue by using reinforcement learning techniques that learn the policy of an actor as state-action values (*cf.* Section 3.2). However, RL from scratch is very costly and is, in consequence, not fast enough in human-agent interaction, as it requires too many learning cycles until it reaches an acceptable behaviour (*cf.* Kaelbling et al., 1996; Knox and Stone, 2010; Kober et al., 2013).[3] The idea to increase the learning speed is to roll out

[1] In the beginning of the thesis, I called this the 'Quality of Behaviour'. Indeed, this term is motivated by the JIAC V framework, which is grounded into agent-oriented software engineering and service-oriented architectures; introducing 'Quality of Service' as one key metric.

[2] The definition for MAS, as presented in Section 3.3.1, is related to the definition of teamwork settings in that problem solving is a cooperative effort by individual entities.

[3] This is characteristic for RL and named *"slowness in convergence"* (Matignon et al., 2006, p. 840). It

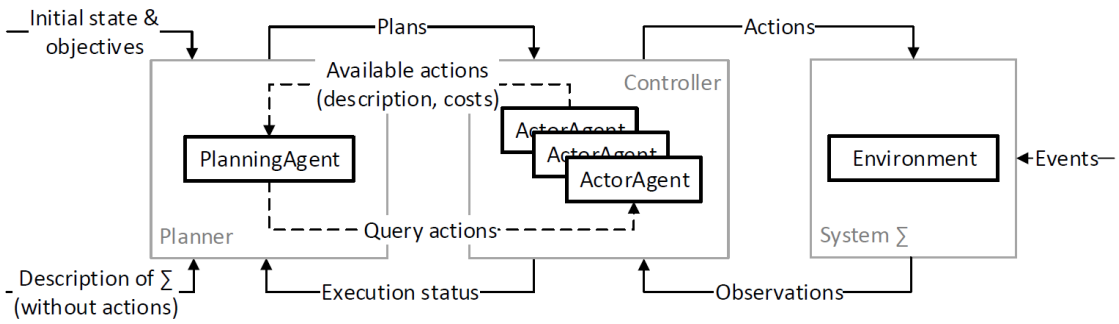


Figure 13.2.: High-level architecture of our approach visualised as part of the conceptual model for planning components (greyed out, Ghallab et al. 2004, p. 8).

domain knowledge using our human-behaviour models, which will be applied as prior knowledge for tabula rasa RL.

In the following, I will describe the architecture of HPLAN (*cf.* Section 13.1.1) and provide details on the learning techniques that can be used to integrate prior knowledge into RL (*cf.* Section 13.1.2).

13.1.1. Architecture

The architecture of HPLAN brings together the idea of actor agents, that serve as individual representatives for a human, an agent, or a robot and the conceptual model for dynamic planning components that enables interleaved planning and acting including the implementation of plan supervision and revision and the possibility to replan (Ghallab et al., 2004, pp. 5–9).

Fig. 13.2 illustrates the architecture of our approach. Here, the controller handles the execution of plans generated by the planner based on an initial state and a set of goals that are provided by an external source. The controller executes actions, processes observations from the environment and informs the planner about the plan execution status. In doing so, the architecture differentiates between offline planning, decoupled from the environment, and online execution, in direct interaction with the environment. This accounts for the fact, that the actual planning process requires a fixed model of the system on the one hand and takes time in which the actual environment evolves on the other.

In our approach, a controller contains a set of actors, each representing a human or an artificial agent that is capable of manipulating the environment. The planning capabilities are provided by a PlanningAgent, which can be an ActorAgent as well. The initial domain description encloses no action descriptions as the available actions are provided by the ActorAgents at runtime. The process starts with a new objective triggering a

is related to the trade-off between exploration and exploitation (*cf.* Section 3.2).

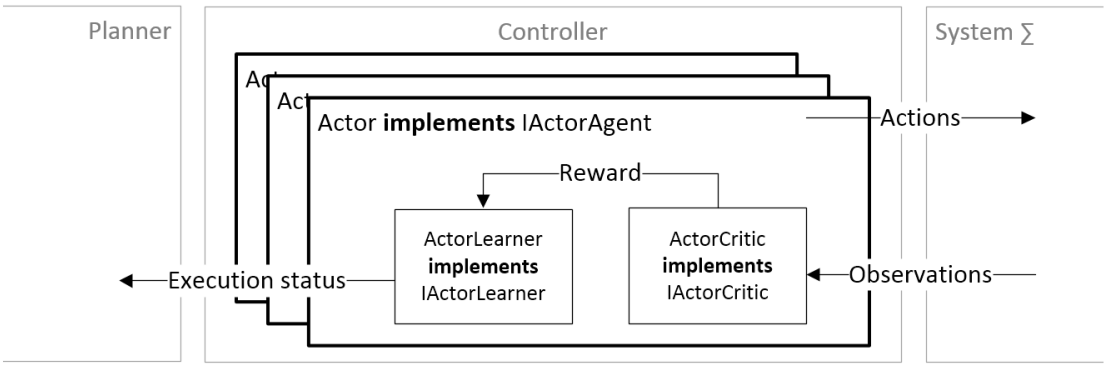


Figure 13.3.: A more detailed view of an actor agent concentrating on the learning cycle.

PlanningAgent to query available actors. Such actors provide action descriptions and the learned costs estimates for each action they want to offer for the planning process. Thus the difference to other planning systems is, that a full domain description is generated at runtime only. This enables the system to react to (dis)appearing actors and new knowledge about the capabilities of actors. Building the complete domain description, the PlanningAgent uses the information to generate a plan in which each task is delegated to the most capable agent. If a task is assigned to an ActorAgent representing a human, the execution can be either:

- passive; i.e. observing whether or not the human actually performs the assigned (predicted) tasks and to which extent;
- interactive, *i.e.* informing the human about the task assignment, e.g. asking for assistance; or
- active, *i.e.* providing assistance in performing a task.

As the implementation of the ActorAgents behaviour depends on the developer, it could also be a mix of these options. During the execution, the observations generated in the environment are evaluated by the associated actor. If a failure occurs, it is reported back to the planner to trigger replanning. Furthermore, the ActorAgents representing a human learn from each execution and adapt its cost estimates accordingly, thus, completing the lifecycle of monitoring, analysing, planning, and executing.

Note that the presented architecture requires to differentiate between planning and acting capabilities. It does neither limit the ActorAgents to act only on behalf of a PlanningAgent nor restricts the MAS to have a single planning entity for the whole system. The persisting requirement is that the planner is decoupled from the controller.

Fig. 13.3 shows a more detailed view of an ActorAgent and introduces additional abstractions, which Hplan needs to provide to developers to fulfil the requirements. The

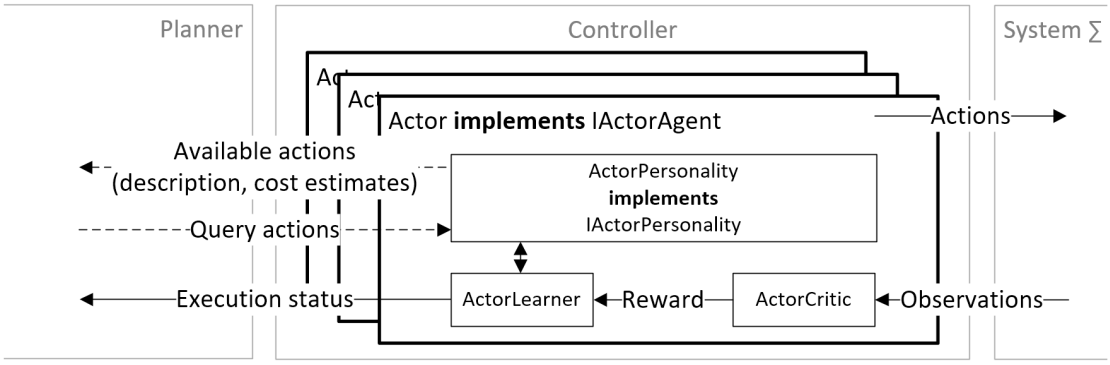


Figure 13.4.: A more detailed view of an actor agent concentrating on the human-behaviour model and its usage within learning and planning.

monitoring of the environment is done via some sensors, that generate observations, which may be raw or preprocessed data. These observations are used to determine the execution status, *i.e.* to inform the PlanningAgent about failures to initiate replannings, and to learn about the human. Therefore, we introduce components called ActorCritic and ActorLearner. The actual learning takes place in the ActorLearner component using an implementation of a RL algorithm. The ActorLearner needs a reward signal as input in order to improve the cost estimates for the actions selected. The ActorCritic is responsible for calculating this rewards giving the observations. In other words, the ActorCritic processes observations, to generate a computable reward signal such as 'failure' or 'success'. This is necessary, as most environments do not provide a numerical reward signal as required by RL algorithms.

The ActorAgent further holds an instance of the human-behaviour model and implements a mapping between the personality of the human and the actual action preferences. For this, I introduce a component named ActorPersonality as illustrated in Fig. 13.4. The ActorPersonality holds the information about the personality of the represented human, *i.e.* the state of the personality. This information can be derived via assessments or learned during the interaction with the human as shown in Chapter 8. The counterpart of the personality information is the correlation between a personality and an action preference, which is given by psychologist. In HPLAN, we expect this correlation to be part of the action/capability description of the agent, besides other information like the precondition and effects of an action. The personality together with these correlations builds our human-behaviour model and is used to calculate the cost estimate for actions. The developer is free to set both vectors in relation depending on domain requirements. The only restriction is that the resulting cost estimate can be used as a metric during the planning process and can be integrated into the action description required by the planning component. In addition to this, the ActorPersonality provide the personality information to the ActorLearner, which uses them to learn about the individual prefer-

ences of a human. In the following, I will introduce learning techniques that justify this connection.

13.1.2. Prior Knowledge to Accelerate Learning

The just presented architecture describes the necessary components to enable interleaved planning and learning and identifies the components necessary for engineers to develop joint human-agent activities. Within this section, I discuss techniques enabling developers to integrate our human-behaviour model as prior knowledge into RL algorithms, providing presumptions on what type of behaviours are effective in our domain.

One can identify different approaches to integrate prior knowledge. We can refer to these approaches as advising the learner, shaping the reward or the policy of the learner, restricting the set of applicable action in a state, and initialising the learner with a 'good' initial set of Q-values. An overview of such techniques is given in the work of Wiewiora et al. (2003), which also provides a good example of the application of prior knowledge in RL: *"The objective of chess is to win a match, and an appropriate reinforcement signal would be based on this. If an agent were to learn chess without prior knowledge, it would have to search for a great deal of time before stumbling onto a winning strategy. We can speed up this process by advising the agent such that it realizes that taking pieces is rewarding and losing pieces is regretful. This advice creates a much richer learning environment but also runs the risk of distracting the agent from the true goal – winning the game."* (Wiewiora et al., 2003, p. 792). This example describes the design and application of different (intermediate) reward functions to advise the learner towards a long-term goal. The nature of real-world domains, which are usually non-deterministic, infinite, influenced by others, and partially-observable, lead to the conclusion that more frequently, immediate reward signals are favourable compared to a long-delayed single reward signal (Mataric, 1994, pp. 183–184). HPLAN makes use of these findings, in that the monitoring of the human's behaviour is assessed for every individual action; allowing developers to implement reward signals for every action of an ActorAgent. Although this is one kind of integrating domain knowledge, it is not a suitable way to integrate correlations between personality and action preferences as those correlations do not deliver knowledge about suitable reward signals. However, the integration of the correlations as prior knowledge can be achieved with other techniques.

Henceforth we refer to our human-behaviour model as prior knowledge using the abbreviation $\hat{H}$. The work of Knox and Stone (2010) introduces and evaluates eight different techniques for integrating $\hat{H}$ to influence the learning process. The evaluation identifies four of them superior. Those read as follows (the used notation is introduced in Section 3.2):

- **Reward Shaping** – $R'(s,a) = R(s,a) + (\beta * \hat{H}(s,a))$ – Reward Shaping increase/decreases the reward signal, thus it influences the actual experience made

and affects the exploration strategy indirectly.

- **Q-Augmentation** – $Q'(s,a) = Q(s,a) + (\beta * \hat{H}(s,a))$ – Q-Augmentation increase/decrease the Q-values every time they are updated. Thus doing both influencing the reward and influencing the action-selection.

- **Control Sharing** – $P(a = argmax_a[\hat{H}(s,a)]) = min(\beta, 1)$ | *otherwise use base RL agent's action selection mechanism* – Control Sharing introduces a second set of Q-values from which the agent should select the maximum, thus directly influencing the action-selection.

- **Action Biasing** – $Q'(s,a) = Q(s,a) + (\beta * \hat{H}(s,a))$ | *only during action selection* – Action Biasing increase/decrease the Q-values only during the action-selection. Thus is directly influences the action-selection, but it neither affects the learned Q-values nor the rewards that are given in an environment.

Although any of these techniques can do the trick, my conclusion is to use Action Biasing. This is because it influences the action selection and not the experience that is made (Knox and Stone, 2012). Furthermore, it is only applied during the actual action selection. Making it for engineers easier to follow the separation of concerns principle. That is, it enables to encapsulate the influence of the prior knowledge on the learning cycle and to control to which extent and how long the action selection should be biased. Finally, within Action Biasing one needs to add correlations for actions that are more beneficial. Those correlations can be directly derived from psychological studies, as we want to use them. Compared with Control Sharing, which also influences the action selection only, Action Biasing is easier to develop, as one only needs to use the correlations provided by psychologists and do not need to construct a second set of Q-values using these correlations.

13.2. Implementation

The concept of HPLAN is implemented as an extension module for the fifth generation of the Java Intelligent Agent Componentware (Hirsch et al., 2009; Lützenberger et al., 2015; Lützenberger et al., 2013) (JIAC V). JIAC V is an open-source general-purpose agent-framework developed at my institute.[4] Within this section, I will provide the necessary information to retrace the realisation and the important design decisions. However, as major parts of the implementation were part of the master thesis presented by Ebert (2013) and Prochnow (2015), supervised by the author, the interested reader is referred to these publications for implementation details such as class diagrams and domain models.

[4] JIAC V is available at `http://www.jiac.de`, last-visited: 2017-09-25

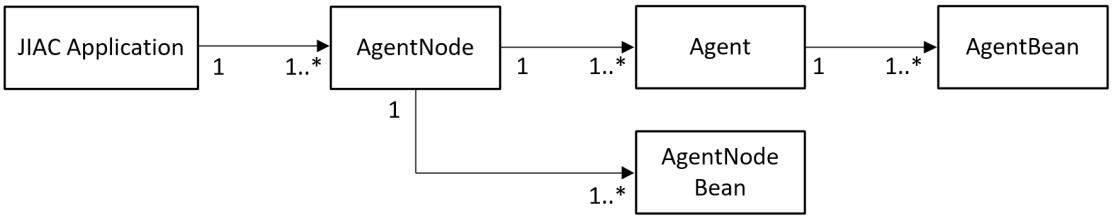


Figure 13.5.: The basic concepts of the JIAC V architecture.

The architecture of a JIAC V application is depicted in Fig 13.5. It introduces different components, starting with the AgentNode, which is the runtime environment for JIAC V agent. An AgentNode is a JVM providing the basic environment for running agents using AgentNodeBeans (e.g., communication infrastructure, service directory, administration interface). Each AgentNode can host an arbitrary number of Agents, which encapsulate their behaviour using AgentBeans. The JIAC V framework provides different default capabilities and abstractions to ease the development of agent applications (*cf.* Lützenberger et al., 2013). For example, it implements a local memory and communication adaptors for each agent. Furthermore, it provides an execution cycle that is responsible for executing the behaviour of the agent. This execution cycle allows developers to set up the lifecycle that is described in Section 13.1 by implementing capabilities for monitoring, analysing, planning, and executing. To enforce this lifecycle, we define an associated interface providing the abstract method declarations, which is named IActorAgent as indicated in Fig. 13.3.

The PhD work of Konnerth (2012) introduces three different types of agents that can be implemented using the default components of JIAC V:[5]

- **$\mathbf{A}_0$ agent** – Describes the basic agent that provides all capabilities to execute actions and to interact with other agents and the environment using the memory, the execution cycle, adaptors (sensor/effector abstraction), communication capabilities, and an optional rule-engine. This agent type is defined as

$$A_0\ agent := \{F, A, T, O, M, R\},$$

  where F are facts stored in the memory of the agent, A refers to all actions an agent can execute, T are triggers that initiate/call an action (e.g. messages, lifecycle, rules), O are the resulting outputs of an executed action, M are all messages, and R are rules, which are executed if a condition is met and can change F, create T, send m.

[5]The interested reader is referred to the work of Konnerth (2012, pp. 79 – 127) for a detailed description of the interaction of the components.

- **A_1 agent** – Describes a BDI agent that extends the capabilities of the A_0 agent introducing intentions and desires. The A_0 agent already introduces beliefs, which are named facts in JIAC V, and are stored in the memory of the agent, building the knowledge base of the agent. This agent type is defined as

$$A_1\ agent := \{F, A, T, O, M, R, D, I\},$$

  where D are the desires of an agent, which are further differentiated into perform desires, state desires, and maintain desires (Konnerth, 2012, p. 95), and I are the intentions of the agent.

- **A_2 agent** – Describes an example extension for the capabilities of the A_2 agent, introducing components for obligation, plan generation, and learning and is defined as

$$A_2\ agent := \{F, A, T, O, M, R, D, I, Ob, P, E\},$$

  where Ob refers to the obligations of an agent, P refers to the plans of an agent, and E marks the experience that is generated while learning about something.

The A_2 agent is introduced to help developers to develop further agent types in JIAC V. For our extension, we have to introduce personality information and define the interaction to clarify how our human-behaviour model is used. To do so, we define an ActorAgent as

$$ActorAgent := \{F, A, T, O, M, R, D, I, Ob, Per, P, E\},$$

where Per represents the personality of the agent. Furthermore, we add an additional transition[6] that applies the experience that is made during the observation of behaviour including the personality correlations as prior knowledge according to Action Biasing, *i.e.* this transition is only applied during the action selection and creates the personality biased cost estimate for an action. Listing 13.1 describes this transition.

To achieve this, the transition requires that the agent knows the action, the costs of the action (experience), the personality of the human, and the personality correlation for the action. The transition further describes that the agent generates a new action description that joins the experience for executing the action and the human-behaviour model. This action description is used during the action selection and removed afterwards. This and the other relevant transitions are defined in the work of Konnerth (2012, pp. 79 – 127)

[6] A transition defines how a system evolves, in that it defines the preconditions in terms of a set of facts that are present (left side) and an add and delete list (right side), *i.e.* facts that are added to respectively deleted from the agents' world model (Konnerth, 2012, pp. 82–83).

A & E(A) & Per & Per(A) → A' /

cond:
Action A is known.
Experience with action A is known (may be empty).
Personality of the human is known.
Personality correlation for action A is known.

add:
A modified action description A' is known, that reflects the experience executing action A taking into consideration the personality preference.

delete:
-

Listing 13.1: Applying the experience that is made creates a new version of the action description that is consumed during the action selection.

and provide the necessary interaction to query for action descriptions, to plan a course of actions, to execute the action, and to learn from the execution of actions.

As all these capabilities are provided by AgentBeans our approach to integrate HPLAN into JIAC V is to provide the identified components as AgentBeans with related interfaces defining the abstract methods that have to be implemented. In doing so, we extend the ability of the A_2 agent by introducing the ActorCriticBean, the ActorLearningBean, and the ActorPersonalityBean, further, we use the PlanningBean as the extension point to implement the planning algorithm.

Fig. 13.6 shows the architecture of an ActorAgent. Typical for JIAC V agents is the decoupling of the components using the execution cycle and the memory, whereby AgentBeans communicate via objects stored in the memory. Within the ActorAgent the ActorCriticBean creates critic objects and writes them to the memory. The ActorLearner subscribes to such objects and is waiting for write-events to process them. The ActorPersonalityBean writes the personality of the represented human to the memory, which can be consumed by any other bean. Furthermore, the ActorPersonalityBean registers a listener listening for a query for action descriptions, which could be sent by the ActorAgents' PlanningBean or another agent in the MAS, which is responsible for the planning. The ActorPersonalityBean is then responsible for creating the action description. The action description is forwarded to the planning entity and includes the actual cost estimates for each action. The cost estimates are created within the ActorAgent implementing one of the abstract methods of the IActorAgent interface.

In the following, I provide details about the used libraries and other extensions implemented for JIAC V to enable the concept of HPLAN.

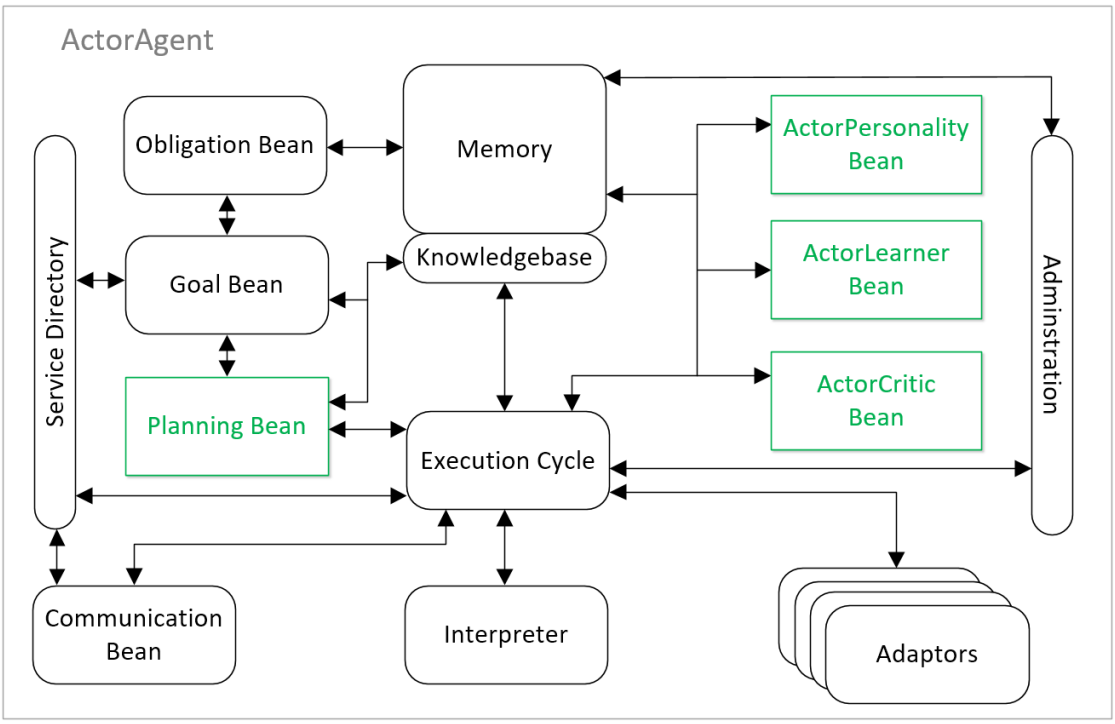


Figure 13.6.: A schematic depiction of an ActorAgents architecture including the learning and planning components (green components). The illustration is an adapted version of the A_2 agent architecture (Konnerth, 2012, pp. 140 – 144).

13.2.1. Planning

For the actual planning, we implemented a planning module using the planning library Planning4J.[7] This approach enables planning with various AI planners, even if they are realised using other programming languages than Java. Within Planning4J the domain description and the actions are described using the planning language PDDL. PDDL is objectified to ease the manipulation of the associated action costs and to support reusability, e.g. to generate the different syntaxes that are required by the supported planners. To induce cost estimates into the action planning, we use the concept of numerical fluents first introduced in PDDL2.1 (Fox and Long, 2003). The thus assigned action costs can be used by the planner as a metric to decide which actor is expected to execute an action. For this, we use the minimisation plan-metric—also first introduced with PDDL2.1—for the quantitative directive of plan creation.

[7] Cerny, M. 2012. Planning4J - Java API for AI planning. Available at `http://code.google.com/p/planning4j`, last-visited: 2017-09-25

```
@Expose(
  name = COPING_SOCIAL_EVENT,
  scope = ActionScope.GLOBAL,
  actor = ActorType.HUMAN,
  descr = COPING_SOCIAL_EVENT_DESCR,
  costEstimate = COPING_SOCIAL_EVENT_COST,
  oceanCorrelations = {0.5,-0.5,0.5,-0.5,0.5}
)
public void setupASocialEvent(){
  // Implementation of user interaction
}
```

Listing 13.2: Example of annotating actions when developing actors that represent humans.

13.2.2. Learning

As a default implementation for the ActorLearner, we apply stateless Q-learning (Claus and Boutilier, 1998). We drop the state dependency to learn the ability of an agent to fulfil a specific task, not to learn the utility of an action in a specific state of the environment. The default implementation is intended to be used as a blueprint for engineers developing joint human-agent activities using HPLAN.

13.2.3. Annotations

We extended the action annotation process used in the agent-framework with the ability to annotate personality information. This satisfies the requirement of HPLAN to add personality correlations derived from psychologists to the action descriptions of agents.

In JIAC V the expose annotation is used to declare an agent's actions. Listing 13.2 shows an annotation for an action named *setupASocialEvent* (*cf.* Chapter 14). At runtime, all relevant information is extracted from the annotation and its attributes. Here, the *name* of an action is used to register it in the dictionary of each AgentNode (the dictionary is a yellow-page service). The *scope* of the action is used to control its visibility. It controls whether the action is visible to all existing agents, to the agents on a single platform, or only to the agent owning such action, e.g. as a service provided by one AgentBean to another. The *actor* defines if an action is provided by an artificial or natural agent. This information is used to determine if there is further information in the action description that can be considered. The *descr* provides the action description in the selected planning language (in our case PDDL). The *costEstimate* attribute holds the current cost estimate for this action by this actor. The *oceanCorrelations* holds a vector that describes the personality action correlation in the order of the OCEAN acronym. To ease the development of this description, the work of Prochnow (2015) introduces a graphical editor, with associated code generation possibilities to Spring configuration files and Java source code.

As indicated in the listing, developers have to provide a description of an action and

the way the actor interacts with the human user when the action is used. Developers are not required to implement the action's logic, as humans will execute the action. However, according to Challenge 5 - Revealing Status and Intention it makes sense for developers to provide an interaction channel that can clarify the intention of the agent system.

13.3. Technical Evaluation

This section presents an in-depth analysis of the concept and the implemented prototype. The evaluation goal is to verify that the approach can fulfil the lifecycle of planning, executing and learning. In particular, we proof if the system can plan, execute, detect failures and replan. Furthermore, we proof if the system's actors can build experience during the execution of their actions and if such additional elaborated information supplied by each actor is used to adapt the task delegation process. In this sense, this section describes a step-by-step evaluation of the technical feasibility of HPLAN. By doing so, I will not focus on human-agent interaction and simulate the human actors. I conclude this section with a discussion of the evaluation results and present some insights into improvement points.

For the evaluation of HPLAN, we will use a classical planning problem—the Blocks World[8] (Gupta and Nau, 1992)—to evaluate the systems ability to plan, execute and learn. A problem in the Blocks World domain includes a finite number of blocks labelled in some way (e.g. letters). Such blocks are exclusively located on each other or on the table, which serves as the surface. A block is clear if there is no other block located on it. To manipulate this environment, an effector exists that can execute one type of action. The action is to move a single clear block to the table or on top of another clear block. Fig. 13.7 illustrates a planning problem in the Blocks World. The goal of an automated planning system is to find a plan that transforms the initial state (to the left) into the goal state (to the right). Taking into account the restrictions of the manipulation action, the first step of each plan would be to move either block F, D or A, which are the ones that are clear.

To fit with the research goal of this thesis, we extend the Blocks World and introduce two types of effectors, namely robots and humans. Each of them is capable of moving single clear blocks. The difference to the classical Blocks World is based on hidden properties introduced by the human actors. For our purpose, those properties can be interpreted as personality characteristics. However, the main point is that the properties are not directly accessible to the planning system, as they are apart of the domain description. They must be taken into account to produce optimal plans regarding costs according to the minimisation plan-metric. In consequence and the planning

[8]The original version of the Blocks World was introduced by Winograd (1972, pp. 117–126) and contains additional symbols and constructs, which are not needed for our purpose.

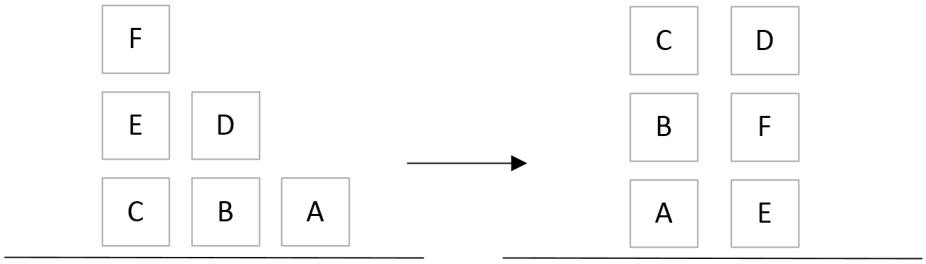


Figure 13.7.: A planning problem in the Blocks World domain is given by an initial state (left-hand side) and a goal state (right-hand side), which have to be solved by an automated planner.

entities point of view, the—so-called Human-Aware Blocks World—domain is not fully observable; making it necessary to estimate the influence of the hidden properties on the action costs of a human actor. One way to achieve this is to build knowledge about the actors using observed behaviours.

We distinguish two types of hidden properties referring to weaknesses differentiating a human actor from the robot actor: Failure at moving a block (denoted as external factor ext_c) and the (not that) timely execution of moving a block (denoted as external factor ext_r). Fig. 13.8 illustrates a problem in the Human-Aware Blocks World and shows that weakness is applied for all boxes that are more than one level of the ground. The goal for an automated planning system is to find a plan that transforms the initial state (to the left) into the goal state (to the right) combining the actions of both actors in the most efficient way. Taking into account the restrictions of the manipulation action and both available actors, the first step of each plan would be to move either block A or E, which are the ones that are clear. Also, the planner has to decide if A or E should be moved by the robot or by the human.[9] This decision is based on the cost estimate of the actors denoted as the efficiency of moving a block (regarding time-steps). The introduced hidden properties ensure that the efficiency of moving a block can vary for each human and is not known in advance. As the hidden properties are not accessible to the planning system and not known to the actor as well, the system is required to learn the efficiency of a specific actor during the execution to minimise the overall plan costs. In consequence and related to our approach, the ActorAgent representing an individual human must observe the environment during the action execution and adapt the cost estimate according to the observation.

In the following, we will give insights into the concrete evaluation goals and how we measure the costs of solving a problem. Afterwards, we will use our approach and build a concrete agent-model for the evaluation domain. The section proceeds with analysing the influence of different configurations of the used learning algorithm in comparison to

[9] Within this section we limit the number of actors to one robot and one human which perform actions consecutively.

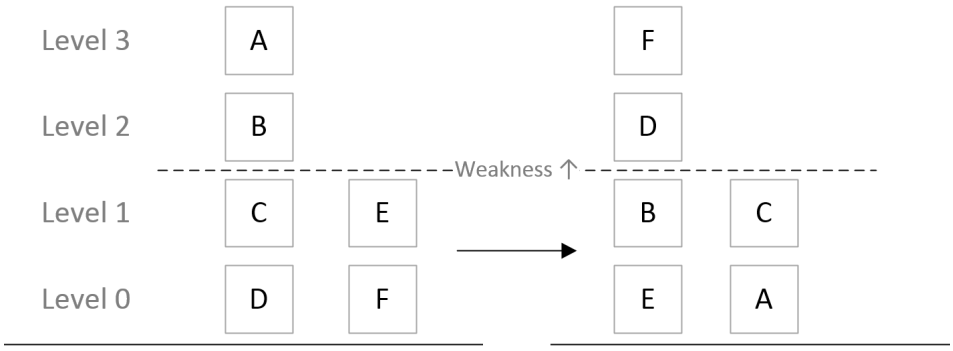


Figure 13.8.: A planning problem in the Human-Aware Blocks World. Each action that is executed by a human actor more than one level of the table triggers weakness. This problem is used in the following as our major scenario.

the sample average method (*cf.* Section 3.2, Eq. 3.2 or Sutton and Barto, 1998, p. 27), which serves as baseline. Subsequently, we analyse the ability to adapt to the individual hidden properties and the combination of both. Furthermore, we analyse if the system can learn problem independent by using the gained experience solving other problems.

13.3.1. Evaluation Goals and Metric

The Human-Aware Blocks World introduces a partial observable version of the Blocks World; adding the requirement to learn. Learning is required when a planner uses the minimisation plan-metric (Fox and Long, 2003, pp. 69–71) for the quantitative directive of generating plans. In particular, it is about the ability to delegate task optimal between the team members in the long run. Using this metric, we are enabled to prove whether our solution can satisfy the following goals or not:

1. To minimise the cost of solving a problem by creating a plan with the least expected cost over time. This addresses the ability to delegate task optimally. Throughout the domain requirement to learn, we expect the system to adapt the planning process during the observation of failures/successes to a point where failures are avoided. One side-effect which we expect is to observe a decreasing amount of replannings. This effect is also important for the second evaluation goal.

2. To minimise the number of weakness affected tasks that are executed by human agents that severely suffer from weakness. This addresses the ability to produce legible behaviour, which is about the safeness and comfort of a human in real world scenarios (Kirsch et al., 2010, p. 224). Using the cost progression is not suitable to show this, as a human has a different point of view on what an optimal plan is. In our domain we expect the system to produce plans that delegates as few weakness

including tasks to a human as possible, once the individual that interacts suffers from weakness. The number of replannings is the corresponding metric.

If the system can satisfy these evaluation goals, we can argue that subordinated requirements such as the ability to plan, to execute, to detect failures, and to replan are fulfilled. That is because these abilities are prerequisites for a planning system able to learn from experience.

To quantify if the cost-effectiveness of solving a problem improves over time a measure is required. For this, we use the average cost C_n^m to solve a problem after n previous experiences averaged over m rounds according to the following formula:

$$C_n^m = \frac{\sum_{i=0}^{m} c_{i,n}}{m} \quad \Big| \quad m, n, i \in \mathbb{N} \tag{13.1}$$

An experience is gathered if one instance of a problem is solved successfully and is given by the costs c required within a run. This includes the execution, the observation of failures and the necessary replannings. This means that the learning takes part in the transition between n and $n + 1$. Each experience is gathered m times to remove statistical variations introduced by the hidden properties ext_c and ext_r. Note that during the technical evaluation we simulate the human's users. To ensure not to learn some random deviations but the real estimate of the hidden properties, we use large numbers of m. If not otherwise stated we will use $n = 1 \ldots 150$ and $m = 100$ for our simulation runs.

13.3.2. Constructing the Agent-Model

The first step to validate if our approach can satisfy the evaluation goals is to use HPLAN to build a solution fitting for the addressed domain. The Human Aware Blocks World introduces the risk of weakness and distinguish the two factors ext_c and ext_r as hidden properties. Whereas the first denotes the likeliness to fail at moving a block, the second one denotes the likeliness that a task will be processed in a timely fashion. In the following, I describe the construction of the agent-model.

To build an estimate of the expected costs of executing an action we apply stateless Q-learning (Claus and Boutilier, 1998, pp. 747–748) as our learning function. We are enabled to drop the state dependency as in the addressed domain an actions utility does not depend on the current state. Thus, we learn the ability of a human to fulfil a specific action (e.g., moving a block), not the utility of an action in a specific state of the environment (e.g. move block D in the initial state of the problem illustrated in Fig. 13.8). We will come back to stateful learning later on.

In Q-learning, the learning agent interacts with its environment by performing actions several times (*cf.* Section 3.2). In sum, a learning agent performs an action a repeatedly

for t times. At each time step t the agent receives a reward $r_t(a)$ and uses this reward to iteratively improve its estimate $Q_t(a)$ of the expected reward for action a. In other words, the agent builds an estimate of the expected costs of executing an action a. This estimate can be iteratively updated using several equations. One way is to build the average value of all rewards received when the action was selected according to the sample-average method (*cf.* Eq. 3.2). However, this approach does not fit well with our problem as it does not react fast enough to changing rewards in the long run. That is through the nature of averaging all values. We require this to test the possibility of the system to adapt its behaviour based on new experiences. Instead, we iteratively update the cost estimate using the following equation, known as the (stateless) Q-learning update rule (Claus and Boutilier, 1998, p. 747):

$$Q_{t+1}(a) \leftarrow Q_t(a) + \alpha \Big(r_t(a) - Q_t(a) \Big). \tag{13.2}$$

Here, the parameter α denotes the learning rate ($0 \leq \alpha \leq 1$), controlling the influence of new observations (rewards) on the currently available estimate. Choosing α always depends on the problem addressed and requires to balance between the influence of a new observation and the importance of the already gathered observations. Therefore the next section includes an analyse of different learning rates.

HPLAN enables us to build knowledge from observations using ActorCritics. Such ActorCritics are responsible for computing the feedback received from the environment when executing actions according to some criteria. One natural approach for such a criteria considers the nature of a reward signal that either can be of qualitative or quantitative form. We will use this criterion to differentiate the received reward signals as follows:

- p_c^k – is learned from a qualitative reward signal in terms of 'failure' and 'success'.
- p_r^k – is learned from a quantitative reward signal in terms of time steps required to execute an action.

Formally, this differentiation is expressed by the twofold Eq. 13.3. Here, the parameter ρ is a constant factor to punish the execution of an action a if the execution has failed. The parameter $c(a) \leftarrow Q_{t_0}(a)$ is the initial cost estimate for action a provided by the developer. The execution time of an action a is denoted as $\Delta_t(a)$.

$$r_t(a) \leftarrow \begin{cases} \rho \times c(a) & \text{if 'failure'} \\ \Delta_t(a) & \text{otherwise} \end{cases}. \tag{13.3}$$

Choosing ρ depends—like the learning rate—on the problem addressed and requires to balance between the influence of an observed 'failure' and prior observations. Therefore the next section also includes an analysis of different punishment factors. One characteristic of Eq. 13.3 is that successfully executed actions are not separately rewarded.

Instead the ActorCritic takes the execution time into account. Thus, if the execution time equals the current estimate the adaptation of the estimate only depends on the learning rate.

The whole learning procedure outlined above is applied by the available actors each time they manipulate the environment (either direct if the actor represents an artificial agent or indirect if the actor accounts for a human). A separate planning agent or one of the actor agents with planning capabilities uses the available actions, and the learned cost estimates to produce plans according to the minimisation plan-metric (targeting minimal overall costs).

Using RL techniques requires further to balance between exploration and exploitation. This comes down to the point that *"[t]he agent has to exploit what it already knows in order to obtain reward, but it also has to explore in order to make better action selections in the future"* (Sutton and Barto, 1998, p. 4). Hence, to ensure that we enable the planner to select the cheapest action available in the long run, we applied the ϵ-greedy policy as the action selection strategy.[10] If not otherwise stated we will use an exploration rate of $\epsilon = 0.1$ for our simulation runs. This action selection strategy guarantees that $Q_t(a)$ converges to $Q^*(a)$ for $t \to \infty$, where $Q^*(a)$ is the mean reward received when a is executed (Sutton and Barto, 1998, pp. 27–30).

We use the ϵ-greedy policy as its effects are well-studied and can be easily analysed during the evaluation. There are several other, and more elaborated action selection strategies and our approach allows us to apply them due to the introduced abstraction levels and since HPLAN does not provide actual implementation but the environment for developers to insert own algorithms. Nevertheless, the evaluation goal is not to analyse the action selection strategy but to analyse the proper function of our approach. Therefore it is beneficial to use an action selection strategy whose influences can be predicted and computed by hand.

13.3.3. Parameter Adjustment

In the following, we will analyse how different values for the learning rate α and the punishment constant ρ affect the learning process. This is done to receive insights into the behaviour of the system using different parametrisations and to elaborate feasible values that are employed in the remainder of this chapter. I further provide a comparison between the Q-learning update rule and the Sample-Average method. This will underline the statement mentioned above that the Sample-Average method does not react fast enough to changes. For the simulation runs, we will use the problem illustrated in Fig. 13.8. We use this problem as it is small enough to calculate minimal plan costs

[10] The interested reader should note that in our approach the agent with planning capabilities is responsible for the action selection and that by all means learner and planner are not the same. That means that the learning agents point of view there is no indication which action to select as there is no knowledge about the other available agents and their capabilities at all.

manually and sufficiently complex to analyse if the learning process works as expected.

To analyse if the system can adapt to dynamic hidden properties we use two scenarios.[11] The first (static) scenario simulates humans with a failure rate of $ext_c = 0.5, ext_r = 1.0$ for the whole simulation run. In the second (dynamic) scenario we change the human behaviour after 50 experiences at $n = 50$ to $ext_c = 0.0, ext_r = 1.0$, simulating a human that does not suffer from weakness. After 50 additional experiences at $n = 101$, we restore the previously used external factor. We can compute that an optimal plan has costs of 12 time-steps if the human suffering from weakness fails to execute 50% of all task triggering weakness. The optimal cost for a human that does not suffer from weakness is 8 time-steps. These costs are the lower limits, and we expect our approach to converge to these limits. The changeover enables us to analyse the ability of the approach to adopt a model that was already learned. Note that the ext_r is static and set to 1.0 during this section. In our simulation ext_r serves as a multiplier and thus has no influence when setting it to 1.0.

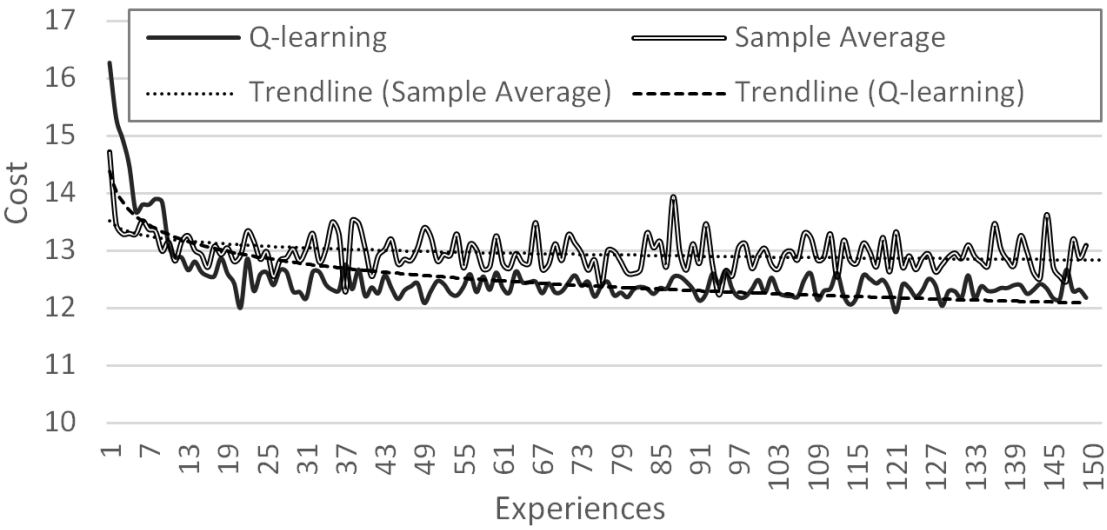
Fig. 13.9 illustrates the difference between the Q-learning update function and the Sample-Average method. In the static scenario shown in Fig. 13.9a it can be seen that both update functions converge to the lower limit, even the Sample-Average converges reasonable slower than the Q-learning. The alternation is explained through the influence of the ϵ-greedy policy as action selection strategy. The effect is more distinct for the Sample-Average method. Both effects are underlined in the dynamic scenario shown in Fig. 13.9b. Here, the third stage between $n = 101 \ldots 150$ depicts a substantial gap between both methods and introduces an even more distinctive alternation. This underlines our decision to use the Q-learning update rule to build an estimate of the costs of an action iteratively.

Additionally, the Q-learning update rule is strongly influenced by the learning rate chosen. Fig. 13.10a illustrates this effect comparing different learning rates. It can be observed that a lower value of α results in a slower adaptation and vice versa. This is an expected behaviour as a higher/lower learning rate increases/decreases the influence of new experiences to the already elaborated cost estimate. It underlines that the prototype works as expected.

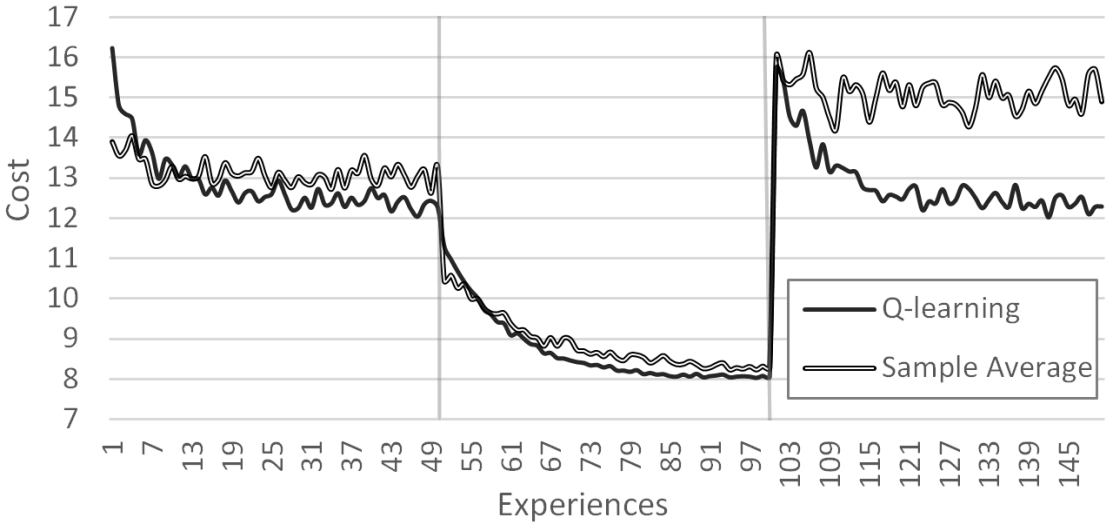
In the following a learning rate of $\alpha = 0.1$ will be used, as the observed behaviour of this α is a good compromise between a fast adaptation to new observations/experiences and an analysable adaptation during the simulation runs. A larger value might prevent us from effectively evaluating the effects of the actual analysed factors as it restricts the learner to the actual observations, whereas a smaller value might require much more experience than provided by a simulation with $n = 150$.

Fig. 13.10b compares the influence of different punishment factors and shows that a lower value of ρ results in a slower adaptation. It can be observed that no relevant

[11] In real-world scenarios the hidden parameters might change during runtime conditioned, e.g. through familiarisation with a problem/action or induced through a changing context like a human who gets tired performing repetitive tasks.

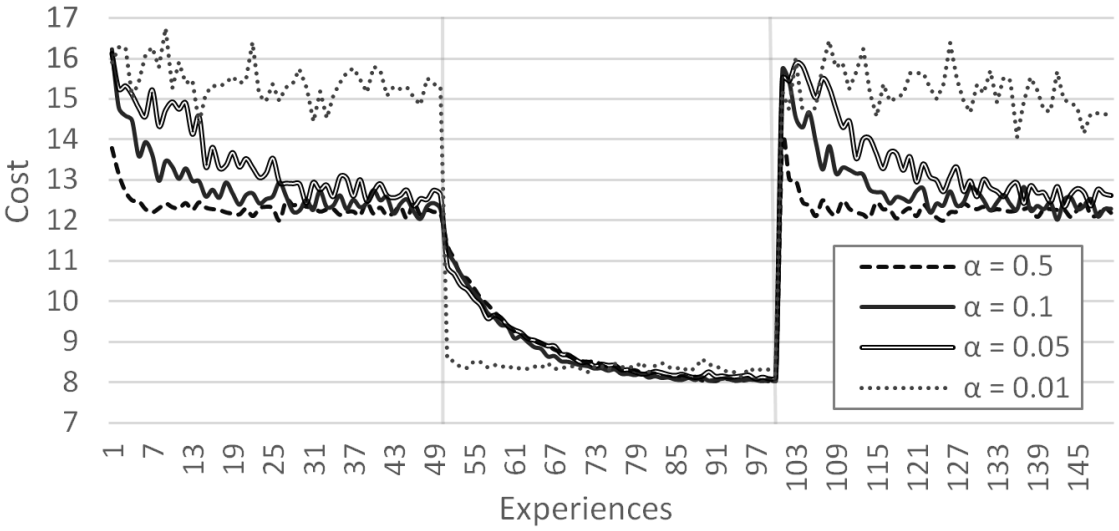


(a) The static scenario with two actors, one robot and one human suffering from weakness ($ext_c = 0.5, ext_r = 1.0$).

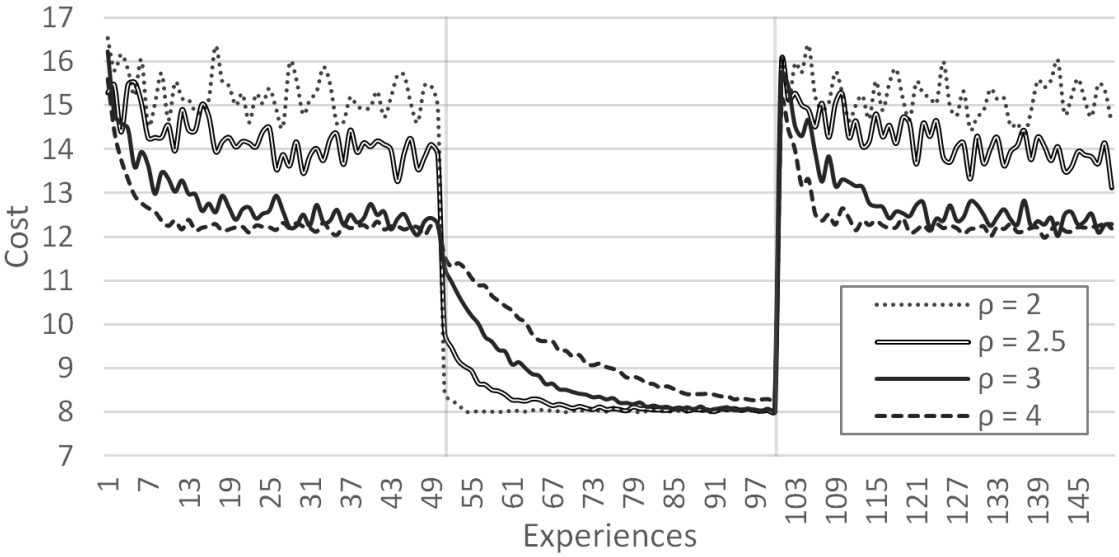


(b) The dynamic scenario with two actors, one robot and one human suffering from weakness ($ext_c = 0.5, ext_r = 1.0$ for $n = 0 \ldots 49$ and $n = 101 \ldots 150$ and $ext_c = 0.0, ext_r = 1.0$ for $n = 50 \ldots 100$).

Figure 13.9.: Comparison of the Q-learning update function ($\alpha = 0.1$) and the Sample-Average method. Both use the same punishment constant $\rho = 3$.



(a) Comparison of different learning rates α ($\rho = 3$, $ext_c = 0.5$, $ext_r = 1.0$, optimal plan costs: $stage_{1,3} = 12, stage_2 = 8$).



(b) Comparison of different punishment constants ρ ($\alpha = 0.1$, $ext_c = 0.5$, $ext_r = 1.0$, optimal plan costs: $stage_{1,3} = 12, stage_2 = 8$).

Figure 13.10.: Comparison of different learning rates and punishment constants using the dynamic scenario.

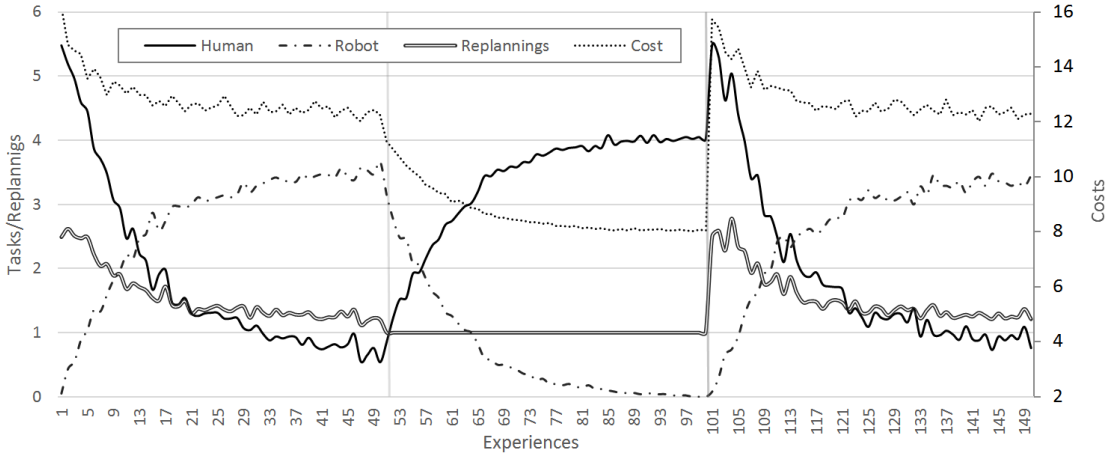
learning effect occurs for $\rho = 2$, as the plan cost instantly drops to the optimum in stage_2. This indicates that the system still predicts the human to be more efficient, even after some failures. The value of ρ needs to be chosen depending on how costly a replanning is and how cumbersome a failure would be for the human. Repeating a task that can not be achieved numerous times could be frustrating or even dangerous for the natural agent. If the goal is to avoid human failure, a high punishment is required. Based on this information, $\rho = 3$ will be used in the remaining evaluation, as it produces an average and consistent progression.

13.3.4. Adapting to the Human

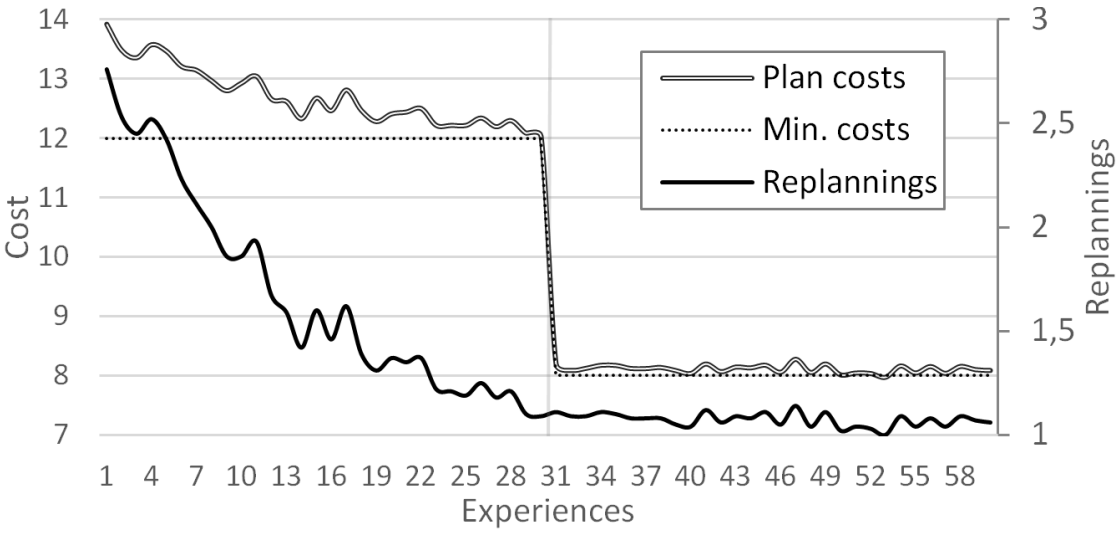
Until now, we have analysed the effects of the different technical aspects of the learning algorithm to select a feasible learning algorithm with comprehensible parameters paying little attention to changes in the human's behaviour. With fixed values for the learning rate and the punishment factor, we now examine how our system performs when varying the hidden parameters of the humans that are represented by our ActorAgent. The dynamic scenario already showed the behaviour when varying the failure rate for moving a block (ext_c), which is the qualitative feedback signal (success/failure). In the following, we will vary both hidden parameters of the human, to also test the ability to adapt to quantitative feedback by altering the time a human actor requires when executing vertigo affected tasks.

Fig. 13.11a evaluates the cost progression for the combination of qualitative and quantitative feedback and the number of tasks that are executed by the human and the robot. At the beginning of stage_1, the human is expected to perform more task as the robot, as we made the human actions cheaper than the robot's action (1:2). However, the system soon adapts to the observed behaviour, *i.e.* that the human fails in one-third of all tasks and requires 1.6-time steps to accomplish a task, and assigns more task to the robot. At the same time, the number of replannings converges to the generation of one plan. This is the behaviour we intended, *i.e.* the amount of weakness affected tasks that the human is expected to perform decreases until the system reaches an efficient balance between both actors. In stage_2, the scenario changes to a perfectly acting human and the system adapts to the changing behaviour of the human by assigning more task to the again cheaper entity. The graph also shows that tasks that are not allocated to the human are executed by the robot, balancing the team behaviour according to the weaknesses that are observed. At the beginning of stage_3, the external factors of the human are again adapted, which becomes visible with the peak in costs and replannings and the changing task assignment and plan cost.

The remaining fixed factor that influences the behaviour of the system is the problem which the actors solve. To show that the system learns problem-independent we tested the use of an already adapted model to solve other problems. To show this, we replace the problem with a different one after $n = 30$. The time-optimal solution plan for this



(a) Number of tasks executed by the robot and the human agent and number of replannings done during the simulation including the overall cost progression ($\alpha = 0.1$, $\rho = 3$, $ext_c = 0.3$, $ext_r = 1.6$, optimal plan costs: $stage_{1,3} = 12, stage_2 = 8$).



(b) Cost progression when changing the problem during runtime ($\epsilon = 0.01$, $\alpha = 0.1$, $\rho = 3$, $ext_c = 0.3$, $ext_r = 1.6$).

Figure 13.11.: Diagrams visualising the results for adapting to a varying human behaviour and a change in the problem.

problem takes 8 time steps. Fig. 13.11b shows, that the system reaches a near time-optimal solution on the second problem without additional adaptation (also underpinned by the number of replannings). We can, therefore, conclude that the system does not simply learn problem specific behaviour but indeed learns the hidden properties of human agents.

As the hidden properties are only accessible via observing the behaviour of the human, we can conclude that the system can fulfil the requirements of being able to plan, to replan, to monitor the execution and to learn from the observations that are made. Also, we can argue that our approach of forwarding information about humans as cost estimates to a planning entity is working as postulated. Therefore, I conclude that we reached the evaluation goals formulated in Section 13.3.1.

13.4. Conclusion

Within this chapter, I presented an agent-based architecture named HPLAN that facilitates the development of joint human-agent activities. It is based on the idea that personality information provides useful correlations to more efficiently plan in joint human-agent activities. HPLAN is strongly related to the conceptual model of dynamic planning components and introduces an agentified version of this model. To account for the individualism of humans, I introduce the concept of ActorAgents that serve as avatars, each one responsible for interacting with one human. The presented technical evaluation shows that our approach fulfils the requirements we formulated and which were used in Chapter 12 to distinguish our work from related work:

- The management of the state is achieved due to the lifecycle we enforce.
- It was shown that plan supervision, replanning, and failure detection is possible due to the separation into controller and planning components.
- The integration of our human-behaviour model is realised as part of the ActorAgent and as part of the action descriptions.
- Learning takes place using this information as prior knowledge for reinforcement learning techniques.

The evaluation results indicate, that the concept of forwarding additional information to the planning component is promising in terms of a more accurate cost estimate at the one hand and that we can reduce the number of failures and replanning using the observations of the human behaviour on the other.

Based on the work presented in this chapter, I can finally answer the research question that was formulated in the beginning, namely: “How can we integrate our human-personality model into the development cycle of agents?”

By separation of concern, we can integrate our model into the development cycle. That is to say:

- The personality of a human can be modelled as part of each agents knowledge base holding a user model. Within the implementation, we introduce a dedicated component called ActorPersonality for this task. The actual data acquisition of the information is left to the developer and could be either implemented as an additional learning task as described in Chapter 8, by using existing data sources, which can provide the information, or by asking the human using assessment instruments, to name but a few.
- The behavioural information, which is the correlation between personality and actions provided by psychologists, can be attached to the action/capability descriptions that are common in agent-based systems. Within the implementation, we introduced an extended version of the action descriptions used in JIAC V that enables developers to annotate this information to each action.
- The usage of the human personality can be realised as prior knowledge in reinforcement learning. Instead of adapting the human-personality model itself, we thus learn the actual state-action values for a human, which at the same time can be used as cost estimates for optimising the planning process. This decision is substantiated by the experience made in Chapter 8. In the implementation, we propose the use of the Action Biasing and left to the developers how the actual personality and the behavioural information are set in relation to each other.

Although the technical evaluation shows in a step-by-step manner that Hplan is a functional development environment providing the necessary components and interfaces to answer our research question, it has to be calibrated and adjusted to the particular application domain. The next chapter presents a real-world case study in the domain of stress assistants were this is done. Here, we will also take a closer look to the correlations derived from psychology and how they are used to bootstrap the learning of individual preferences.

14. Case-Study: The Personality-enabled Stress Assistant

In this chapter, I present the Personality-enabled Stress Assistant (PeSA), which serves as a case study that utilises HPLAN to implement an agent-based application that accounts for the individualism in the stress-coping of humans by utilising the personality of users to personalise its assistants. While, I already established that HPLAN presents a functional development environment, the aim of this chapter is to demonstrate how it can be used to develop a real-world application based on our human-personality model, psychological findings of a specific domain, and experience gained through the observation of the human's behaviour. Complementary to the technical evaluation described in the prior chapter, I will focus on the human-agent interaction and its inherent cooperative and interactive aspects within this chapter. In particular, I will concentrate on receiving and representing the personality of the human (user model), implementing the personality preferences (behaviour model), combining both to bias the learning and planning cycle, and developing this cycle using the corresponding components of HPLAN. This allows us to collect insights from the actual usage of HPLAN and to identify possible extension directions and corrections from software engineering point of view.

To achieve this, I will start the chapter with the motivation and background for developing a stress assistant in Section 14.1. In particular, I provide insights into the importance of stress management, motivate why personalised assistant in this domain is needed, and place the engineering of PeSA into the broader topic of human-agent teamwork and interaction. In Section 14.2, I apply the coactive system design (*cf.* Section 4.2.2) to the task of developing a stress assistant. This is done to structure the development and to identify which actor should perform which task and how the performance in one task depends on the other team members. Afterwards, in Section 14.3, particulars about the implementation of PeSA using HPLAN and other features of JIAC V are provided. I introduce how one can derive the necessary information and how they are applied in the case study. Moreover, I stress the advantage of using an agent-based approach due to the usage of other features that are available in the community. Subsequently, in Section 14.4, I will provide insights into the interaction details of PeSA, to illustrate the implemented demonstrator. Finally, in Section 14.5, I conclude the chapter with a discussion of the experience made with the different features of the development environment.

14.1. Motivation and Background

"Stress is our minds and bodies reaction to a situation that is overwhelming" (Workplaces, 2015).[1] It is a common reaction to various stimuli that can mobilise a temporary extra boost in energy, alertness, and performance, which can be useful in a variety of situations. However, psychological stress is also a well-known trigger for several physical and psychological diseases and has a significant impact on our health and health care costs (*cf.* Hapke et al., 2013). Especially chronic and long-term stress, *i.e.* constant stress, makes it difficult to control own emotions, weakens the immune system, is a significant risk factor for mental disorders, and contributes to processes in premature ageing. Finding these effects on the individual level, the European Agency for Safety and Health at Work estimates the annual costs for stress at work to 25.4 Billion EURO (Hassard et al., 2014, p. 34) and links 50% to 60% of all lost working days (absenteeism) to work-related stress (Milczarek et al., 2009, p. 7) on a population-wide level. This makes the management of stress an important factor for different stakeholders.

The problem that we address with PeSA is that the perception of stress is highly individual and depends on many factors like the age, the sex, and the personality of the individuals (Connor-Smith and Flachsbart, 2007, pp. 1093–1099). As individual as the perception of stress are the actions that may help to relieve stress (Carver and Connor-Smith, 2010, pp. 681–687). Those actions are named coping strategies. Coping strategies are conscious and goal driven acts to adapt the environment or the self as a reaction to stressors such as stressful conditions and situations (Compas et al., 2001, p. 89).[2] With PeSA we want to approach the need for providing personalised guidance in stress management by assisting the human user with coping strategy recommendations and adapting its recommendations to the individuality of the user. In doing so, individual preferences, characteristics, and personalities should be supported. These expectations sketch the requirements for developing such an assistant.

The implementation of this capability is based on a combination of a data-driven and a theory-driven mechanism and directly addresses one of the unsolved human-agent interaction challenges. These challenges are described by (Prada and Paiva, 2014, Section 4), where PeSA particularly contributes to the challenge: *Use data, but also theories.* This challenge states that engineers should create agent models that are based on theories about human behaviour and ground them with data determined by the interaction with the human. The theoretical component is our human-personality model, which acts as an accelerator for learning individual preferences and is based on the user model determining the personality according to the FFM and correlations between personality

[1] The GP Hans Selye introduced the concept of stress in the 1930s and defined it as *"as a non-specific response of the organism to any pressure or demand"* (Selye, 1956, according to Milczarek et al., 2009).

[2] The work of Compas et al. (2001) provides a literature overview about the definitions, dimensions, types, and measures of coping strategies.

and coping strategies as published in different psychological studies. The data-based component learns by receiving rewards from the users and thus grounds the applied human-personality model to the true preferences of the human user.

Beside these capabilities, which are the core concepts of HPLAN, we take advantage of other agent-based techniques. Pesa for itself is a combination of agent-based software engineering and learning algorithms that use the human's feedback as reward signals. From a technical perspective, any PeSA instance is a software agent that autonomously collects data, individually detects situations in which stress-relieving actions are required and adapts its recommendations to the personality of its user. To make the learning process faster, we extended the agents' capability to share knowledge with other PeSA agents. The basic idea behind this is that users with similar personalities prefer similar countermeasures (*cf.* Carver and Connor-Smith, 2010; Connor-Smith and Flachsbart, 2007). Thus, it is possible that relatively new PeSA agents can learn from agents that have already collected experiences in relieving stress levels – given that their users show similar characteristics.

14.2. Cooperation Analysis

To develop a cooperative system assisting the human in coping with stressful situations, we have to identify the tasks that should be performed by the different actors and need to analyse how these tasks depend on each other. The coactive system design defines a cooperation-centred methodology that enables such an analysis focusing on interdependencies (*cf.* Johnson et al., 2014b, pp. 53–60; Johnson, 2014, pp. 72–80; Section 4.2.2). Within this section, I provide insights into this analysis to highlight the cooperative aspects and the approach I took to develop a cooperative stress management.

In the identification process, we analyse the roles, tasks, required capacities, and relationships. For PeSA, three main tasks exist: (1) *identify* – the identification of stressful situations including the measurement of current stress level and the assessment of the long-term stress, (2) *coping* – the search for a feasible coping strategy and the commitment to a coping strategy, and (3) *execution* – the actual execution of the selected coping strategy. Along with the tasks, we identify three different actors: the PeSA agent, the assisted human, and other supporting team members (e.g. fellows, friends, family, chatbots, personal organisers). Required capacities for the identify task are measurements of short and long-term stress (includes sensing and recognising the current and long-term stress level). For the coping task, we require the system to know coping strategies, to be able to identify feasible coping strategies and to assess them to commit to a specific one. For the execution task, the team members are required to perform the coping strategies of different categories.

To analyse the capacities of team members for performing a task a colour coding scheme is used. It provides an instrument to analyse the dependencies of the actors given

Table 14.1.: Team member role alternatives for the tasks of measuring the short-term stress level (1.a. – 1.c.) and long-term stress level (2.a. – 2.c.) of the human including analysis of the possible interdependencies.

	Performer	Supporter		Interpretation
1. IDENTIFY – MEASURE SHORT-TERM STRESS				
a.	Human	PeSA	Others	Humans can measure stress. PeSA agents can visualise stress. Others cannot provide assistance.
b.	PeSA	Human	Others	PeSA agents can measure stress. Humans can contribute to reliability. Others cannot provide assistance.
c.	Others	Human	PeSA	Others cannot measure stress. Human assistance is required. PeSA can improve reliability.
2. IDENTIFY – MEASURE LONG-TERM STRESS				
a.	Human	PeSA	Others	Humans can do it, but reliability is limited for longer periods of time. PeSA can provide long-term assessment tool.
b.	PeSA	Human	Others	PeSA can provide long-term assessment, which needs to be answered by the human.
c.	Others	Human	PeSA	Others cannot assess long-term stress. Human assistance is required. PeSA can improve reliability.

a task and helps to identify the team member role alternative that is most beneficial for implementing a task solution. Table 14.1 lists this analysis for the measuring of short-term stress and the measuring of long-term stress. The performer column emphasises to which extent the team member can fulfil a task (green – alone, yellow – with less than 100% reliability, orange – needs assistance, red – cannot do it). The supporter columns indicate how the supporters can contribute to a task while assisting the performer (green – improves efficiency, yellow – improves reliability, orange – assistance required, red – cannot assist). In the task alternative 1.a. we can see that humans are rated as being able to measure short-term stress on its own. PeSA agents are rated as being able to improve the efficiency of the task. This rating makes an independent operation of this task alternative a viable option. As shown in task alternative 2.a. the human's reliability for making statements over longer periods of time is assumed to be not 100%. This limitation could be solved with the help of the PeSA agent providing relevant assessment methods. Making this a task alternative where the outcome depends on the

cooperation of the team members. This fundamental work now helps to identify the cooperation requirements. For task alternative 1.b. we can, for instance, identify the necessity to observe the human behaviour, e.g. using a smartphone or wearable sensors to determine the current stress level (observability). At the same time, the human has to accept that this task is directed to the PeSA agent (directability). Other tasks, such as the recommendation of coping strategies would require the PeSA agents to predict the strategies that are preferred by the human in a situation (predictability).

Using the coactive design method we analysed all of the above-described tasks, their associated team alternatives, and the cooperation requirements. Given this analysis, we, for instance, selected alternatives 1.a. and 2.a. for the implementation of their respective tasks. The next step is the actual implementation process using HPLAN. The next section describes this process.

14.3. Constructing PeSA

PeSA is implemented as an application for the Android platform[3] with a cloud-based backend that hosts the multi-agent system. Within the Android container, we implement PeSA as a software agent using the Android SDK[4] and the applications inherent lifecycle including an additional service to keep the PeSA agents alive as long as the phone is running. The backend is implemented using the JIAC V framework and the HPLAN extension and hosts mirrors of the mobile agents, enabling the communication within the agent-system. In more detail: The part that interacts with a user and provides recommendations is separated from the one that communicates to other PeSA agents to exchange already learned information about how to treat particular personality profiles. In consequence, a PeSA agent comprises two agents, one mandatory agent on the smartphone, which is named InterfaceAgent and one optional agent in the backend, which is named CommunicationAgent. The reason for this design decision is data security. Exchanging information about a user's psychological profile is critical regarding data security. We offer this feature to accelerate the learning process, yet, only optionally via an opt-in process. Agents on the user's smartphone are capable of learning all required information by themselves. The modular assembly of the PeSA system architecture is illustrated in Fig. 14.1.

We proceed by explaining how PeSA agents can personalise the provided assistants by learning about stress coping preferences (*cf.* Section 14.3.1) and subsequently elaborate on how they exchange their collected knowledge (*cf.* Section 14.3.2).

[3]Android mobile operating system `https://www.android.com/`, last-visited: 2017-04-19
[4]Android SDK: `https://developer.android.com`, last-visited: 2017-09-25

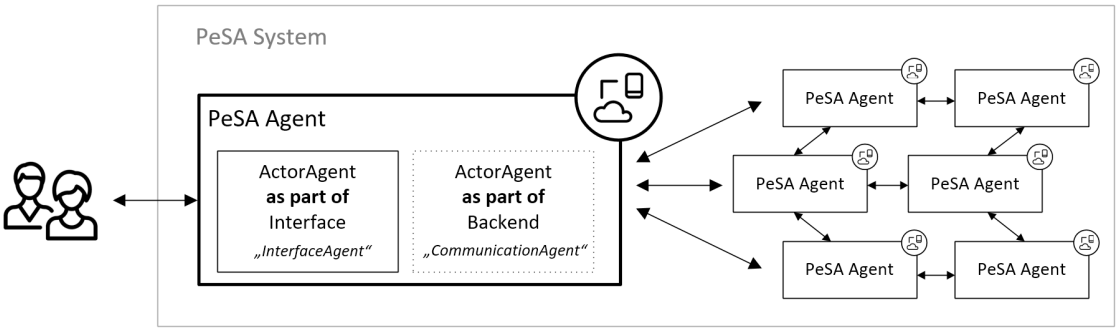


Figure 14.1.: The architecture of PeSA differentiating between the individual user assistants and the multi-agent system.

14.3.1. Personalised Assistants

To enable the PeSA agents to perform stress management and to adapt the provided recommendations to the particular requirements of the assisted users, we implement a reinforcement learning cycle using classical Q-learning (*cf.* Section 3.2) to learn about the action preferences of each user, whereas the actions are the recommended coping strategies. Since tabula-rasa RL is not fast enough in direct interaction with users, we integrate information about the humans' personality and correlations between personality and preferred coping strategies as prior knowledge (*cf.* Section 13.1). As introduced, we refer to this prior knowledge as human-behaviour model $\hat{H}$ and integrate it applying action biasing (*cf.* Section 13.1.2). The resulting learning function reads as follows:

$$Q'(s,a) = Q(s,a) + (\beta * \hat{H}(s,a)) \quad | \quad \text{only during action selection.}$$

To implement this function within our agents, we use the components and structures provided by Hplan. As a result of doing this, we solve the three tasks that we worked on in the prior part of the PhD and which were defined for an adaptive system (*cf.* Section 5.3), namely the data representation, the data acquisition, and the data processing tasks. Fig. 14.2 highlights the relations of the PeSA development to the architecture of adaptive systems and summarises the different steps that contribute to the task of providing personalised coping strategy recommendations. Each task and its realisation within Hplan will be described in detail next.

Data representation

For the data representation, *i.e.* modelling the information structure that holds the user's information, we use the structure provided by the ActorAgent concept of Hplan. As a first step, this solves the data representation task and provides the information structure that holds the users' information:

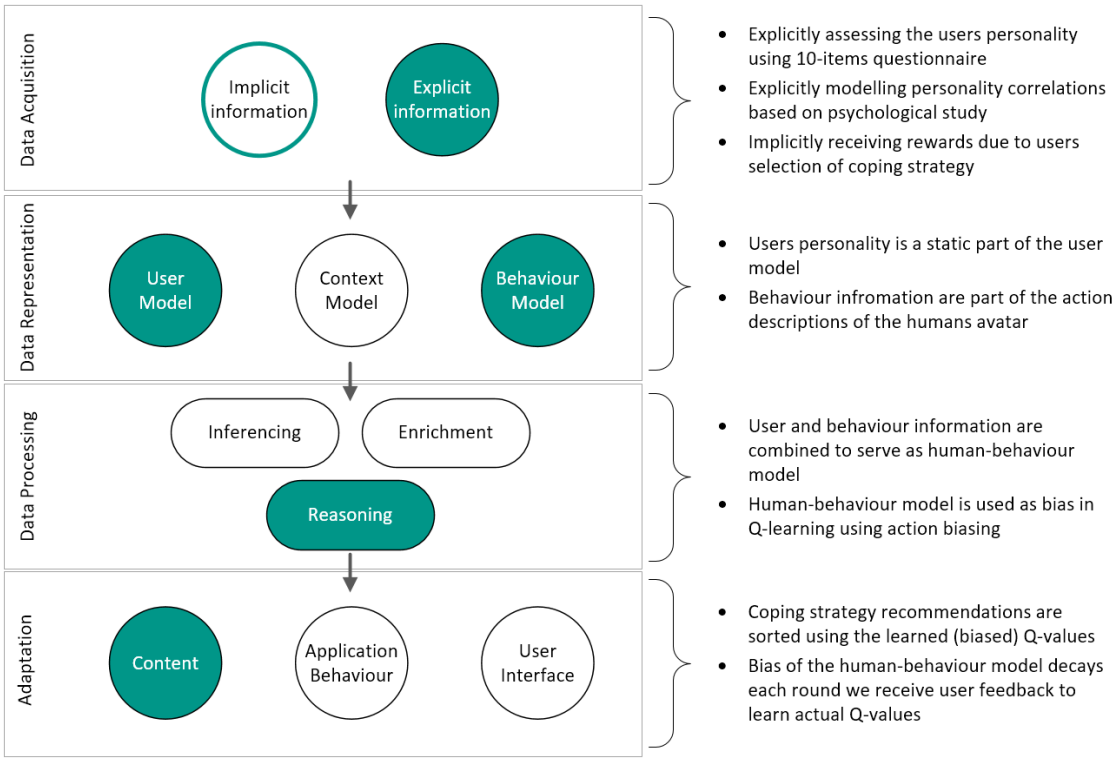


Figure 14.2.: The elements of an adaptive system in relation to PeSA including implementation details.

- Personality of the user – is part of the user model of the ActorPersonality. The personality is stored within the preferences of the agent. We refer to the personality of a human k using the set $P_k = \{p_o^k, p_c^k, p_e^k, p_a^k, p_n^k\}$, where each $p \in P$ represents a personality trait according to the equally named dimension of the FFM.

- Personality coping strategy correlation – is part of the action description of each coping strategy. We refer to the personality correlations for an action a using the set $P_a = \{p_o^a, p_c^a, p_e^a, p_a^a, p_n^a\}$.

Data acquisition

The data acquisition, *i.e.* the task of establishing a modality to collect information about the user, is also twofold. The first step is to determine the personality of the user. In PeSA this is done using a personality assessment during the onboarding process of the application, *i.e.* we acquire the data about the human personality explicitly by asking the user about the required information. The implemented assessment is presented by (Gosling et al., 2003, p. 525) and describes 10-items that can be used to assess the Big-Five personality traits introduced by the FFM: The Ten-Item Personality Inventory (TIPI). The assessment questions are shown in Table 14.2. The user has to indicate

Table 14.2.: The Ten-Item Personality Inventory (TIPI) that is used in PeSA to assess the personality of the user (Gosling et al., 2003, p. 525).

Disagree strongly	Disagree moderately	Disagree a little	Neither agree nor disagree	Agree a little	Agree moderately	Agree strongly
1	2	3	4	5	6	7
			I see myself as...			
			Extraverted, enthusiastic.			
			Critical, quarrelsome.			
			Dependable, self-disciplined.			
			Anxious, easily upset.			
			Open to new experience, complex.			
			Reserved, quiet.			
			Sympathetic, warm.			
			Disorganized, careless.			
			Calm, emotionally stable.			
			Conventional, uncreative.			

to which extent it agrees or disagrees with the statements. Given the answers, the PeSA agent can estimate the users' personality using the TIPI calculation procedure and normalise the results to the range $[0, 1]$ to set it in relation to the personality coping strategy correlations.

To complement the human-behaviour model, we further have to define the coping strategies and its correlations to the personality of a user. Connor-Smith and Flachsbart (2007, Table 7) present these correlations. The correlations are modelled explicitly as part of the action descriptions used in Hplan. Within PeSA we only use a subset of the examined copings, focusing on engagement actions, *i.e.* actions that have a positive attitude to cope with stress (e.g. acceptance, support seeking, planning) in that they actively act towards the stressor.[5] Table 14.3 lists the correlations for the different coping strategy categories available in PeSA and shows the tendencies of a coping action of one category with each trait of the FFM.

Data processing

For the data processing, *i.e.* processing the data derived within the data acquisition, it is necessary to set the personality of the user into relation to the coping strategy correlations. This is necessary as our learning function requires $\hat{H}$ to provide a real number, whereas $\hat{H}$ currently consists out of our two parts. To solve this, Hplan defines a function $\delta : P_k \times P_a \rightarrow \mathbb{R}$ as abstract method within the IActorAgent interface

[5] A complete description of coping categories can be found in the work of (Connor-Smith and Flachsbart, 2007, p. 1082, Table 1).

Table 14.3.: The applied correlations between the personality traits and the preferred coping strategies. Table is adapted from published work (Connor-Smith and Flachsbart, 2007, pp. 1096–1097, Table 7).

Coping strategy	**O**	**C**	**E**	**A**	**N**
Problem solving	0.17	0.34	0.21	0.11	-0.14
Instrumental social support	0.11	0.13	0.29	0.13	0.1
Emotional social support	0.12	0.1	0.29	0.17	0.16
Mixed social support	0.09	0.12	0.29	0.14	0.01
Meditation	-0.03	0.14	0.08	0.2	0.05
Distraction	0.07	0.07	0.09	-0.03	0.21
Cognitive restructuring	0.17	0.22	0.25	0.17	-0.16
Acceptance	0.12	0.12	0.07	0.13	-0.05

of the ActorAgent, which we have to implement. The result of this calculation is used to bias the action selection, thus, our final learning function reads as follows:

$$Q'(s,a,k) = Q(s,a) + (\beta * \delta(P_k, P_a)) \quad | \quad \text{only during action selection,}$$

making the influence of $\hat{H}$ independent of the state s but dependent on an action a of a human k. For the implementation of the relation we build the product of the personality trait with the associated correlation and build the sum of the products. The resulting bias factor is used to determine the initial set of Q-values setting $\beta = 1$. The work of Prochnow (2015) evaluates different definitions of δ and compares the effect of the bias on the coping strategy recommendations. The finally used δ function reads as follows:

$$\delta(P_k, P_a) = \sum_{i=0}^{4} p_i^k * p_i^a$$

To learn the actual preferences of the user, *i.e.* the actual state and action dependent Q-values, β is decayed each time a new experience is made. This is implemented within the interplay of the ActorCritic and ActorLearner. The observable reward signals are whether the user accepts or rejects a recommended coping strategy and the position of an accepted strategy within the list of recommendations. Thus the ActorCritic calculates the reward out of a qualitative and a quantitative signal, increasing the reward conditioned on the distance to the first recommendations that was proposed. The ActorLearner uses the reward implementing the Q-learning update function (*cf.* Section 3.2, Eq. 3.3). PeSA already implements a further possible information source, which is the Perceived Stress Scale (Cohen et al., 1983) (PSS) that provides a long-term reward signal using a 14-item questionnaire. PSS determines the individual perception of stress over

long periods of time and in doing that assess what we introduced as constant stress. In the long run, such a repeated signal could help to identify combinations of coping strategies favourable for a user.

Adaptation

In PeSA, we use the determined Q-values to sort the list of available coping strategies from the most preferred to the least preferred. In consequence, the adaptation that takes place is the personalisation of the content that is provided to the user. The planning problem is restricted to the task of finding the optimal coping strategy for the assisted user in its current context, which aims for an atomar action (one-step plan), in the first place. Using HPLANS' black box approach for planning we can easily implement such a sorting functionality as the PlanningBean abstracts from the underlying planning technique in that it provides methods to customise the planning procedure. The default implementation using the Planning4J library as described in the prior chapter is not used in the current version of PeSA. However, more complex search and planning algorithm could be necessary, when the PeSA system is equipped with a set of agents that can act on behalf of the user to cope with stressors. For instance, agents controlling the environment (the smartphone or the Smart Home), agents that are able to plan and arrange meetings, agents that act as personalised assistants for cognitive restructuring, and respective human counterparts available in the network, e.g. friends and family of a user ready to provide mixed and emotional social support (which, indeed, may be a better tasks for humans instead of chatbots). Making the PeSA agents the first member in a team of artificial and natural agents helping a user to cope with stress. In the prior chapter, I already showed that using HPLAN developers can implement such scenarios providing necessary domain and problem descriptions.

14.3.2. Further Features

Inexperienced PeSA agents are (optionally) allowed to request knowledge that was already learned by existing agents. Knowledge sharing is done to accelerate the learning cycle further while assisting users with similar personality profiles. If knowledge sharing is enabled, each PeSA InterfaceAgent communicates to its respective PeSA CommunicationAgent that is running in the backend. The state of the human-behaviour model, the learned Q-values, and the amount of experience gained is mirrored to the CommunicationAgent. With each new installation, one PeSA agent is deployed in our backend—this agent represents the new user. The agent receives the personality profile from the PeSA InterfaceAgent after the assessment was completed and broadcasts this information to other available agents in the system, asking for similar personalities and already learned preferences. Responses are collected and ranked for compatible profiles. Therefore, the sum of absolute differences over all measured personality traits is calculated and the one,

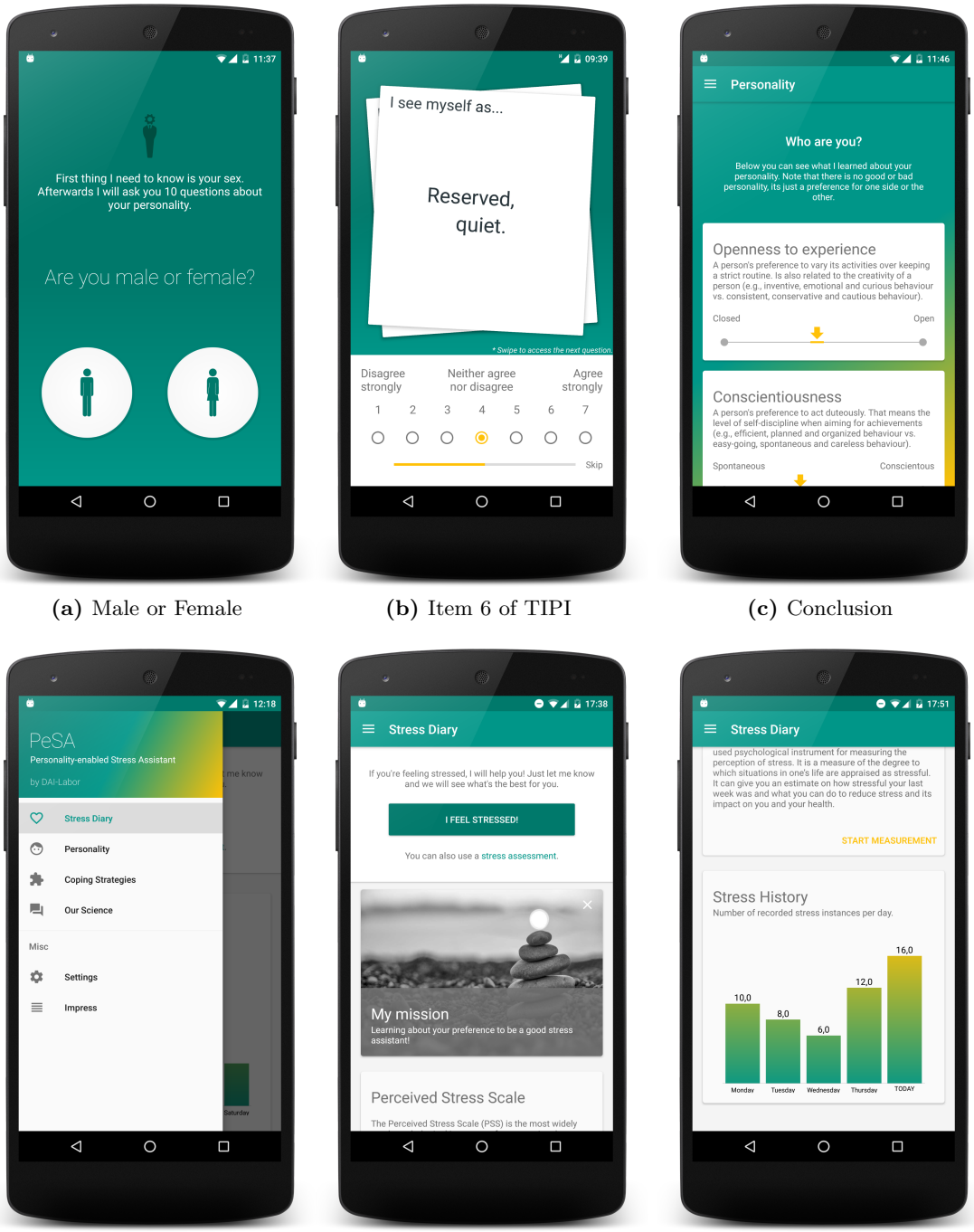
which is closest to zero but below a certain threshold is selected. Based on this ranking, action preferences are updated for the new user, and β is set to the same value that the more experienced agent is using.

The agent paradigms' inherent modular organisation introduces a certain amount of complexity w.r.t. communication aspects. Having these, on the one hand, the agent paradigm helps to ensure privacy due to the distributed representation of information, on the other. That is to say that there exists no single instance that represents the whole user population. This is further enforced by the opt-in process that requires the user to enable data sharing with other agents explicitly. This is of utmost importance in privacy sensible application areas like health applications. Furthermore, the agent-based approach allows us to extend the functionalities of PeSA by including further actors easily. An example for such an extension is described by Breitung (2016) introducing a wearable that senses the stress level of humans and thus provides a jigsaw piece to enhance the currently manual process with an automated solution gradually.

14.4. Interaction Details

Within this section, I will shortly introduce details about the user interface, in order to give the interested reader an impression of PeSA.

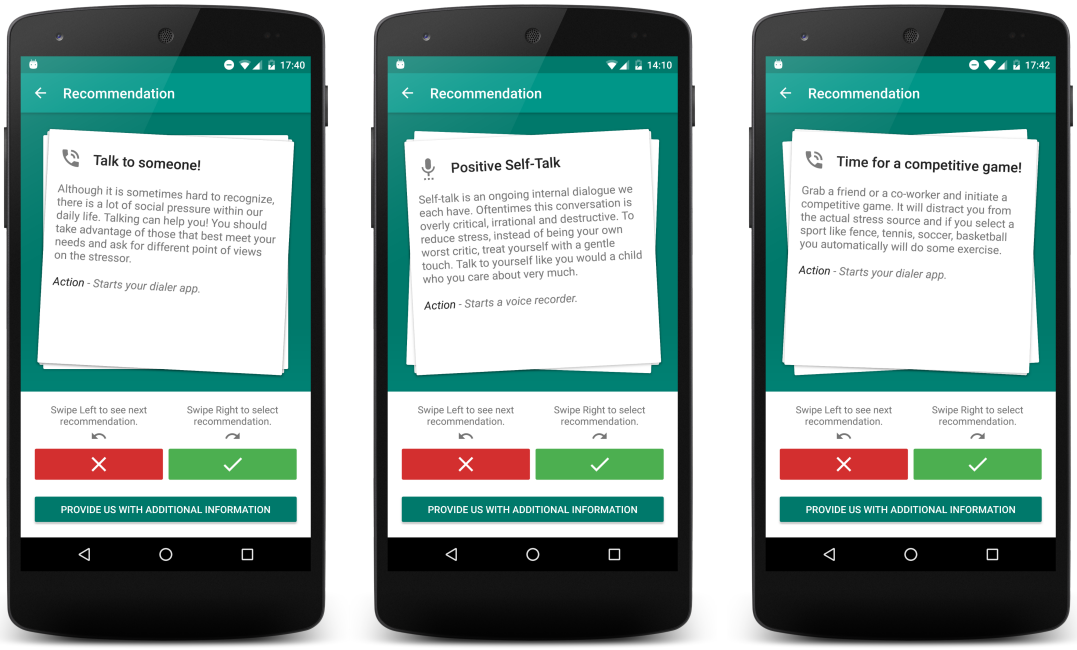
In the first step, PeSA determines a user's personality using the TIPI. Figure 14.3 shows screenshots of this process. The process starts with the question whether the user is male or female (Fig. 14.3a). It proceeds with the TIPI assessing the user's personality (Fig. 14.3b). The process ends with the concluding page showing the user its personality and some informative description (Fig. 14.3c). Based on the user's answers we calculate the personality profile, which is used to determine the initial coping strategy recommendations. The workflow proceeds with the stress diary, which is the landing page of the PeSA app. Here the user receives information about the goal of PeSA and the history of recorded stress events. Fig. 14.3d shows the main navigation of PeSA. The *Stress Diary* part gives users access to the available stress records, the possibility to measure the current stress, and the recommendations that determined for the user (Fig. 14.3e–14.3f). The *Personality* part provides the overview about the personality measurement of the user including explanations and provides the possibility to restart the assessment if skipped in the beginning. The *Coping Strategies* part lists all available coping strategies arranged in categories and provides hints to the user how a category correlates with the personality. The *Our Science* part explains the fundamentals we based our implementation on, including references to the most important papers as further reading points for interested users. The *Settings* part provides access to the applications settings, which includes viewing notifications and shortcuts and enable/disable the cooperative learning. Finally, the *Impress* part hosts the legal information including impress, disclaimer, and privacy statement.



(a) Male or Female (b) Item 6 of TIPI (c) Conclusion

(d) Navigation (e) Landing page (f) Stress history

Figure 14.3.: Screenshots showing parts of the PeSA app.



(a) Emotional social support **(b)** Cognitive restructuring **(c)** Mixed social support

Figure 14.4.: Screenshots showing coping strategy recommendations.

If a stress event is recognised, which is implemented as a call to action, PeSA recommends a list of coping strategies. Figure 14.4 shows screenshots of the recommendation dialogue; listing coping strategies from the emotional social support category (Fig. 14.4a), the cognitive restructuring category (Fig. 14.4b), and the mixed social support category (Fig. 14.4c). Given the correlations listed in Table 14.3, the bias induced by our human-behaviour model would assign a higher value to the emotional and mixed social support if the users personality has a high neuroticism value.[6] The recommendations are organised independently of their categories. For each advice, the user has the option to accept it (swipe right) or reject it (swipe left). These votes are used as reward signals to learn the actual Q-values for each action (coping-strategy). Implementing reward signals in the wild is always challenging, depends on the learning task, and might induce unexpected behaviour as shown in the famous example of learning to drive a bicycle (*cf.* Randløv and Alstrøm, 1998). In its current form, the reward signal is derived from direct interaction with the user. We implemented this swipe-based approach to ensure that we receive a signal.

[6]This is an example, and for simplicity I focus on the neuroticism dimensions as it provides the most distinctive correlation for the three categories.

14.5. Discussion

The management of (dis)stress is an important factor for a long and healthy life. Stress affects people differently, and everyone manages stress in different ways. In this chapter, I introduced PeSA, the Personality-enabled Stress Assistant, an agent-based application that accounts for this individualism. PeSA merges several agent techniques: Reinforcement learning is used to learn about preferences of the users, prior knowledge and knowledge transfer is applied to accelerate the learning process, agent mirroring helps to enable communication and offline functionalities. Based on these mechanisms, PeSA guides through stressful phases by proposing coping strategies that are tailored to the personality of each user. Users can assess these recommendations and thus provide a reward or punishment signal that helps PeSA to improve its suggestions. We utilised the development process of PeSA as a case study for Hplan to collect insights into its usage in a real-world scenario.

The use of Hplan facilitated the development work on PeSA. Hplan provided the necessary components supporting and structuring the implementation of the data representation, acquisition, and processing task. We described how the ActorPersonality, ActorCritic, ActorLearner, and PlanningBean work together and how they can be calibrated and adjusted to the particular application domain. We further showed, what is meant by deriving personality and action preferences from psychological studies. However, the development of PeSA also revealed shortcomings and necessary extensions for Hplan.

For the data representation, we have identified an addition that would further ease the development of applications. That is the representation of the user model, which could be more elaborated, in that real life applications usually require more user information than the personality. One feasible approach would be the integration of user models, like the General User Model Ontology (Heckmann, 2005, pp. 85–99). This addition would further be useful to identify similar user models using techniques from the very same community, e.g. ontology matching techniques (Euzenat and Shvaiko, 2007, Part II: Ontology matching techniques). A possible extension point to integrate this is the user management of the JIAC V framework, which also introduces AAA features (*cf.* Keiser et al., 2011; Lützenberger et al., 2013).

Implementing the data acquisition was straight-forward after identifying the relevant psychological studies that (1) provided an easy to use way to assess the user's personality explicitly and (2) provided the correlations between personalities and coping strategies. Within these studies (*cf.* Carver and Connor-Smith, 2010; Connor-Smith and Flachsbart, 2007; Gosling et al., 2003) we learned that the provided correlations sometimes further differentiate the correlations, for instance, taking into account the sex or age group of humans. Currently, Hplan does not support this fine-granular differentiation. Supporting this in future work requires extending the action descriptions and the associ-

ated components processing the annotated information. In its current form, the software engineer can circumvent this issue implementing specific actions for each category.

For the data processing task, the components for the learning cycles are working as expected and provide the relevant abstractions to preprocess observations and implement the learning function. However, the implementation of the relation between personality and personality coping strategy correlation remains challenging and is intended to be used domain dependent. Part of future work should be an elaboration on how different relations bias the learning cycle to provide respective guidelines or default implementations to developers.

It was shown, that the human-behaviour models can be integrated into the human-agent interaction and provide meaningful information for the decision-making of the agents. At the same time, we have to admit that the direct tailoring of this model is no feasible approach. Instead, the human-behaviour model is used to bias the learning process of the agent to faster converge to the actual preferences of a user. Thus, based on this discussion, I conclude that HPLAN is indeed a helpful extension that facilitates the development of joint human-agent activities.

15. Concluding Remarks

Within this part of the thesis, I worked on the engineering side of joint human-agent activities. In particular, I concentrated on the task of integrating our human-behaviour model into the development cycle of agents. The outcomes of this part are an analysis of the state-of-the-art for development environments for joint human-agent activities, the presentation of a development environment that allows implementing our human-behaviour model and enables interleaved learning and planning, and a case study that uses this development environment to implement a real-world application. In the following, I will summarise this part of the PhD chapter-by-chapter and will lead over to the next part.

In Chapter 12 – Related Work, I analysed the state-of-the-art in agent platforms supporting the development of joint human-agent activities. This was done with the particular focus on the usage of human-behaviour models and the tailoring of such models to the individual during the interaction. The idea was to identify authors and groups working in my field and to determine concepts, in particular architectures, approaches, and technologies that are useful for own work. The analysis justified the development of an own environment in that it identifies gaps regarding the usage and applicability of human-behaviour models, especially related to personality influences, and the usage of learning techniques to adapt the application's behaviour to the individual. The conclusion is, that the existing work can be found in the HRI community and that the main contributions focus on the integration of social constraints. Learning about human characteristics is nearly non-existent. I made use of some of the concepts described in this chapter to ground my work. That is to say; I applied a widely-accepted architecture for dynamic planning solutions, I abstracted from the individual human using the concepts of avatars and made use of available AI planning and learning techniques where possible.

In Chapter 13 – The HumanPlan Environment, I presented Hplan – the HumanPlan environment, developed to facilitate the implementation of joint human-agent activities. Hplan is an extension module to the general purpose agent framework JIAC V, which is developed and maintained by my institute. The chapter describes how our personality model can be integrated into the development cycle of agents that plan and learn in teamwork settings. The concept is based on three core ideas: (1) to more efficiently plan in cooperation with humans I use cost estimates about the human behaviour, enabling the use of available planning components as black boxes; (2) to account for the individuality of humans I introduce the concept of ActorAgents each one representing a specific human user and responsible for providing cost estimates for the capabilities of its hu-

mans; and (3) to estimate the capability costs of an individual I use the human-behaviour model as prior knowledge biasing a classical reinforcement learning process. The chapter describes the concept, implementation, and technical evaluation of the environment, concluding in a step-by-step manner that HPLAN is a functional development environment providing the necessary structures, components, and interfaces to implement joint activities. By doing so, I focused on the technical aspects of interleaved planning and learning simulating the behaviour of a human actor.

In Chapter 14 – The Personality-enabled Stress Assistant, I focused on the usage of HPLAN presenting a case study. Although the former chapter evaluated the general technical feasibility of the environment, it has to be calibrated and adjusted to a particular application domain in order demonstrate how it can be used to develop real-world applications. I do this using the example of a cooperative stress-assistant that is based on our human-personality model, psychological findings of a specific domain, and experience gained through the observation of the human's behaviour. The intention for the case study was to collect insights from the actual usage of HPLAN and to identify possible shortcomings from software engineering point of view. I structured the development of PeSA using the tasks to develop adaptive systems and demonstrated how to use the different components, how they work together, and how developers can satisfy user requirements for joint human-agent activities. The chapter revealed two promising directions for future work: (1) the option to model personality capability correlations more fine granular (e.g. according to the users' sex, age, background), and (2) the development of guidelines for setting the personality in relation to the personality capability correlations.

In the next and final part of the PhD, I will conclude the work described in this document by summarising the contributions discussed in the individual chapters, discussing limitations of the presented approaches, and pointing the interested reader to promising future research directions.

Part V.

Thesis Summary

16. Conclusion

In this thesis, I discussed my contributions to the area of human-agent interaction. Working in this area introduces a multitude of possible directions one can approach during a doctoral research study. Originally coming from the area of agent-based simulation I started my work concentrating on cognitive characteristics therein. In particular, I focused on personality as one of the essential elements influencing the human behaviour. Psychologist describe personality as time and space independent pattern that affects all stage of the decision-making process of a human. Interestingly, there are only a few contributions on integrating personality into agent-based systems (e.g. crowd simulation, virtual humans) when comparing this strand of research to other cognitive characteristics such as emotions in agents. From this point of view, I soon became interested in human-agent collaboration and AI planning. I discussed the requirements and challenges of the former and showed that the affected research areas are not as connected as they should be to contribute to each other. Two questions here, which are surely hard to answer by an individual, are *Why is that?* and *How to solve it?* One answer I can provide to the former question is that I have often seen work talking about the same characteristics using different terms. One consequence hereof, and my contribution to the second question, was the organisation of special sessions, workshops, and demo sessions on the very topic of human-agent-robot teamwork particularly focusing on bringing research communities closer to each other.[1] Part II of the thesis presented a conclusion of the work done here.

At this point, my main focus was still on personality with a little shift towards predictability as an essential requirement for human-agent collaboration. That means the ability of team members to build knowledge about other participants' attitudes, capabilities, the course of action, and so forth. My intention during this time was twofold. First, learning personality information during the interaction with humans. Secondly, using this information to make informed decisions in this interaction. Although I was able to achieve the first objective, I failed for the second. At the same time, Du and Huhns (2013) published a work closely related to the latter objective substantiating that the direct usage of personality information to make informed decisions is no promising approach. Subsequently, I presented a joint formalisation of the concept of personality and the decision-making process of agents. This contribution was motivated by the find-

[1] Special Session on Self-Explaining Agents at PAAMS 2013, PAAMS 2014, PAAMS 2015; Workshop on Human-Aware Planning at INFORMATIK 2014, Demo Session on Human-Agent-Robot Teamwork at MATES 2016

ing mentioned above, that there is a gap between research on personality in agent-based systems and emotional agents. Part III of the thesis described the results of this work.

Lastly, I introduced what is named HPLAN, which is an extension for the agent-framework JIAC V that is developed at my institute. HPLAN brings together the experience made within the prior steps enabling developers to assign human-behaviour models to agents that represent humans. It further abstracts from the concrete planning technique, as a new planning approach was not an intention of this work. HPLAN integrates learning processes into its avatars to adapt the behaviour models towards the user and introduces the multi-layered structure of contemporary planning approaches to the agent-framework. I concluded this part – which is Part IV of the thesis – presenting a case-study, where I made use of HPLAN to build an assistance system with the aim to manage the stress of its user.

Within this chapter, I describe the individual achievements of this PhD (*cf.* Section 16.1). Therefore, I will readopt the research questions and contribution outlined in Chapter 1. Afterwards, I discuss the limitations of this work with respect to the contributions (*cf.* Section 16.2). Finally, I will outline possible future work directions before closing the work with some concluding remarks in Section 16.3.

16.1. Achievements

The goal of this thesis is to prove that we can derive information from social and psychological studies, integrate them as human-behaviour model into agent-based systems, make use of the information to improve the efficiency of human-aware planning components and tailor these models to the individual preferences and habits of a particular human during the interaction. In particular, I focused myself on using personality as the affective phenomenon. In the individual chapters, I approached different aspects of the goal and formulated a set of questions that were put forward and will each be answered next based on the results described in the corresponding chapters.

1. What is the state-of-the-art for integrating personality, in particular, the FFM, as an affective phenomenon in agent-based research?

   In short: Limited to specific use cases and environments applying either the MBTI or a subset of the FFM.

   I analyse the state-of-the-art in Chapter 7. The majority of contributions is found in the area of agent-based simulation, in particular, agent-based traffic and crowd simulations. Only a couple of the considered contributions provide justifications to apply either the FFM, the MBTI, or a simplified representation of personality. This shortcoming is problematic given the task of reflecting psychological findings within computer-processable models. The examined contributions introduce agent-models streamlined for the specific use-case and environment. Work that provides more

generalised investigations on personality in agents is rare, and the most advanced applied the MBTI or a subset of the FFM dimensions, justifying the reason to present an own complete model.

2. How can we derive an agent-model representing the effects of personality with respect to the FFM?

   In short: By using the results of studies doing factor analyses and interpreting the personality trait influences in a given stage of an agents lifecycle.

   I present an approach that integrates the dimensions of the FFM into each stage of an agents' lifecycle according to the BDI paradigm in Chapter 7. The approach provides the parts that we defined for a human-behaviour model, in that the FFM represents the user model and BDI represents the behaviour model. I provided evidence substantiating that personality influences all stages of the BDI lifecycle, though, some traits are more influential in a specific stage. Also, I substantiated how the personality traits are interpreted in a specific stage and provide examples for the implementation of the influence. The implementation grounded the agent-model in a context, complementing the last part of the human-behaviour model. The evaluation showed that the agent-model can simulate personality-related behaviour according to the FFM, lifting the existing results to a state-of-the-art personality theory and all its dimensions. Furthermore, the evaluation provides evidence that a personality-specific task-assignment is beneficial when facing different kinds of tasks as it improves the overall performance reached in the simulated environment – the same observation that is made by psychologist analysing human behaviour.

3. Can agents use our model to learn the personality traits of a human during the interaction with this human?

   In short: Yes, but we need further experiments (more participants, other environments) to generalise the results.

   In Chapter 8 I approach the task of Automatic Personality Recognition, namely recognising the true personality of an individual. I show that the related work concentrates on learning personality traits using supervised approaches, though, in several domains the requirement of having labelled training data sets is not satisfied. Agent-based research on this topic further focuses on learning aspects of personalities simulated by other agents, e.g. to improve negotiation outcomes or to study effects on the social welfare of groups of agents. I extend the state-of-the-art in that I do not follow a supervised learning approach requiring existing data sets, but describe two agent models able to learn about the personality of humans in direct interaction with these humans during repeatedly played rounds in the Colored Trails Game. The results show that some characteristics of a personality

can be learned more accurately/easily than others. Other works substantiate that these results are not specific to the environment, but that we need individual solutions for the respective observable behaviour and personality facet. I finally provide a discussion on factors that affect the presented results.

4. Can we use personality information directly to make informed decisions about the potential behaviour of human?

   In short: No, at least not using the described utility functions in the Colored Trails Game.

   Using the above-described agent-models, I provided results indicating that we can learn about a humans personality using relatively sparse data. Within Chapter 8 I further used the agent-models and the human-behaviour model to determine the possible course of action of the human. Therefore, I defined utility functions controlling the agent behaviour. Although the agents outperformed the humans in average, the analysis of the results also revealed that these observations are not related to the personality. The provided discussion reveals several rationales, reaching from the applied utility function to the CT environment that might not be suitable for the task. In conclusion, I can only negate this question, at least not with the utility functions applied. I substantiated this conclusion referring to another author that approached the same question, which finally led to the decision to use the personality information not exclusively as guidance for the decision-making process but as prior knowledge in the learning cycle.

5. How to represent the state and effect of personality in BDI logics?

   In short: By introducing personality as new modality beside the belief, desire, intention modalities.

   I formalised personality in compliance with the FFM as a new modality besides the belief, desire, intention modalities in Chapter 9. Therefore, I extend the syntax and semantics of the BDI logic $\mathcal{LORA}$ to represent the personality modality. I substantiate the decision for a new modality by the time and space independent nature of the personality phenomenon, the finding that it influences each stage of the decision-making process of humans, and the finding that personality is a determinant for emotional reactions, which are defined via combinations of the existing modalities. I introduce the personality modality with the corresponding state formulas and operators to compare different personalities and described the representation of the state and effects using a set of prior defined statements.

6. Which relations between the personality modality and the belief, desire, and intention modality are meaningful for reasoning about rational behaviour?

   In short: The optional ones, i.e. personality contributes to consistent behaviour in the long-run, but can be overwritten by other concerns in every situation.

To answer this question, I discuss the intuitive understanding of the interplay of the newly introduced personality modality with beliefs, desires, and intentions in Chapter 9. I do this to clarify how the properties influence the behaviour of an agent and which characteristics are meaningful/reasonable for analysing (ir)rational behaviour. I can conclude that strong relations, *i.e.* relations requiring personality to be the main concern in every situation, are an unacceptably strong relation for most agents. In contrast, optional relations that enable the agent to act in compliance with its personality and drop this compliance if needed are more acceptable. With these optional relations, we can capture personality as an influential factor that contributes to a consistent long-term behaviour, which corresponds to the personality definition provided in Section 3.4.

7. What is the state-of-the-art for development environments for joint human-agent activities w.r.t. the use of human-behaviour models?

   In short: Coming from the HRI community and concentrating on social constraints without learning about human characteristics.

   I analyse the state-of-the-art in Chapter 12. The specific purpose of this analysis is the usage of human-behaviour models and the tailoring of such models to the individual during the interaction. I identified several contributions provided by the HRI community introducing concepts useful for own work, namely the conceptual model for planning components, the concept of avatars, and available AI planning and learning solutions. The examined contributions focus on the integration of social constraints and intention models into teamwork development. At the same time, some authors recognised the need for more elaborate representation of human characteristics and models. The usage of learning techniques is nearly non-existent.

8. How can we integrate our human-personality model into the development cycle of agents?

   In short: By separation of concern, i.e. adding a user model, extending capability descriptions to model behaviour, and representing humans using avatars.

   To answer this question, I presented the concept, implementation, and evaluation of the HPLAN environment in Chapter 13. The basic ideas are to forward information about humans to planning components using cost estimates, to refer to each human user using dedicated avatars responsible for building the cost estimates, and to use the human-behaviour model as prior knowledge to accelerate the learning of the cost estimates. HPLAN provides the necessary structures and interfaces to integrate our human-behaviour model. The developer implementing the avatars can refer to the personality of the user using a common user model. They can integrate the behaviour information as part of the capability descriptions of the agents and can provide a relation between these parts to calculate a bias that is used as prior

knowledge during the integrated planning and learning cycle. In Chapter 14, I used HPLAN to implement a real-world application to gain insights into the technical maturity of the environment. The case study showed the practicality and usefulness of HPLAN, though, it also identified future work.

16.2. Limitations

In the following, I will provide some notes on the limitations of the approaches and results presented in the individual chapters of the PhD. Notable and maybe the main limitation is that my background is computer science (communication and operation systems, data security and privacy) and not psychology. I approached this weakness by reading, and in consequence referencing, much work from psychology. Also, I discussed the approaches and results with psychologists to verify the conclusion I have drawn from literature work. By doing so, I learned a lot about the cognitive processes controlling the human behaviour. However, my background is computer science and software engineering, making my knowledge in this area limited.

Chapter 7 – Effects of Personality The described agent-model is implemented and evaluated in an agent-based simulation environment providing a crowd simulation. In the evaluation, I show that personality indeed influences all stages of the decision-making process, though, the evaluation is limited to a comparison between formulated expectations and the observable behaviour of the crowd. Although the expectations are grounded in psychological literature, the agent-model should be evaluated empirically. For this, one could apply factor analysis techniques and let humans describe the visible behaviour using personality-descriptive verbs.

Chapter 8 – Learning Personalities Recognising the true personality of an individual during interaction is an interesting problem. In contrast to existing work, I presented an approach that does not use trained classifiers but builds its personality estimates from scratch during the interaction. The achieved results are mixed. They show that we can learn about the personality aspects and that the linkage between the observable actions and the expressible and interpretable distal cues is crucial to this task. As very little is know about the approached task, the results offer valuable insights that should be used to extend the experiments into two directions: evaluating the agent-models with bigger user groups to improve the significance and expand the approach to other environments.

Chapter 9 – Reasoning about Personality Representing the state and effect of personality and discussing the relations that are meaningful for reasoning about rational behaviour are the first steps to formalise this affective phenomenon. The presented results are foundational work, that should be extended with statements derived from social and

psychological studies that enable reasoning between personality and behaviour. Also, the presented results should be used within a bigger case study to verify the conclusions derived from the discussion of personality characteristics.

Chapter 13 – The HumanPlan Environment The presented development environment introduces the components to implement our human-behaviour model and use in agent-system that plan and act in cooperation with humans. The technical evaluation showed that HPLAN satisfies the asked requirements. The described case study revealed that the current human-behaviour models should be more fine-granular in terms of using more information from the user model (e.g. age, sex, cultural background). Furthermore, implementing the relations between the users' personality to the personality capability correlations is not straightforward. An elaboration on how different relations bias the results in real-world settings should be done to distil respective guidelines or default implementations for developers.

Chapter 14 – The Personality-enabled Stress Assistant The case study presented to receive insights into the usage of HPLAN revealed the potential of using human-behaviour models in teamwork settings. It showed how to derive the required information and how to implement them using HPLAN and achieved its intended purpose of identifying the above-described limitations of HPLAN. However, as PeSA was not released, I could not provide empirical results for the effects on the human-agent interaction. This should be done to build evidence that the human-behaviour models indeed improve the efficiency of the teamwork, though, evaluating this in the wild introduces numerous biases, e.g. user expectations related to the content, behaviour, interaction and visual embodiment.

16.3. Future Work

This document intends to give the reader a comprehensive overview of the research I have performed during my doctoral studies. Based on this overview and the just summarised achievements and limitations, we can identify further promising research directions and potential future work. First of all these directions are given by the limitations, at which the extension of the work on Automatic Personality Recognition is the most interesting for me. This is because the related work shows no 'bottom-up' approaches, though, in several domains the requirement of having labelled training data sets is not satisfied. Furthermore, big data based approaches like presented by Cambridge Analytica[2] illustrate the interest of different domains (market analysis, marketing, search engines) in determining characteristics of humans, in particular, personality information. The claims made by the big data analysis companies are frequently unproven, and applications like targeted marketing can benefit from 'bottom-up' approaches complementing

[2]Cambridge Analytics: `https://cambridgeanalytica.org/`, last-visited: 2017-11-13

the big data analyses.

Another and even broader direction for future work is the combination of different affective phenomena in agent-based systems. The concentration of the community on emotional agents is comprehensible as emotions are an intense and powerful influence. Furthermore, they can be observed relatively easily in everyday life in the behaviour, mimic, gestic, and interaction of humans. With the OCC model of emotions (*cf.* Ortony et al., 1988) psychologists provide a structure that was frequently used by agent-researchers to model emotional agents. A first next step should be to provide a model that combines personality with existing work on emotional agents. Modelling both together allows determining the occurrence, intensity, and duration of emotional reaction more realistic, e.g. to reason about facial reactions (*cf.* Resseguier et al., 2016) or to reason about the influence on the decision-making (*cf.* Zelenski, 2007). A further step is to join the formalisation of emotional agents with the formalisation of personality. This could be done by extending the work of Adam (2007) integrating the personality modality into the defined emotional reactions.

From a practical perspective, I intend to include HPLAN into the public available version of the JIAC V agent-framework, making it available as an open-source tool that can be used by the community. Also, I will use HPLAN and parts of PeSA in an upcoming project to implement an assistant in the health domain. Here a field study is scheduled, which I plan to use to gather empirical data about the usage and effects of the human-behaviour model.

16.4. Concluding Remarks

The results of this thesis contribute to the area of human-agent interaction. Besides the above-described achievements, potentials, and limitations, the work provides additional contributions and further reading points in different aspects. Those range from human-personality theories and their usage in agent-based systems to elements and challenges in cooperative systems to the classification and definition of joint human-agent activities and human-aware planning to a more-detailed analysis of predictability as an essential characteristic of 'good' teams. By doing this, I tried to highlight the connections between the different communities. In fact, connecting the research areas and communities promise the necessary momentum to achieve the envisioned long-term goal of making agents team players in joint human-agent activities.

Part VI.

Bibliography

Bibliography

Adam, C. (2007). *Emotions: from psychological theories to logic formalization and Implementation in a BDI agent*. PhD thesis, Institut de Recherche en Informatique de Toulouse, L' Institute National Polytechnique de Toulouse, Toulouse, France.

Ahrndt, S. (2012). Exploring self-explanation: The human side. DAI-Labor, Technische Universität Berlin. Contribution to the Doctoral Mentoring programm of the 10th German Conference on Multiagent System Technologies (MATES 2012), `http://dainas.dai-labor.de/~ahrndt@dai/mates2012dc_abstract.pdf`.

Ahrndt, S. (2013). Improving human-aware planning. In Klusch, M., Thimm, M., and Paprzycki, M., editors, *Multiagent System Technologies*, number 8076 in Lecture Notes in Artificial Intelligence, pages 400–403. Springer Berlin Heidelberg, Berlin, Germany.

Ahrndt, S. and Albayrak, S. (2016). Joint human-agent activities: Challenges and definition. In Klusch, M., Unland, R., Shehory, O., Pokahr, A., and Ahrndt, S., editors, *Multiagent System Technologies*, number 9872 in Lecture Notes in Artificial Intelligence, pages 105–112. Springer International Publishing, Cham, Switzerland.

Ahrndt, S. and Albayrak, S. (2017). Learning about human personality. In Berndt, J. O., Petta, P., and Unland, R., editors, *Multiagent System Technologies*, Lecture Notes in Artificial Intelligence, pages 1–18. Springer International Publishing, Cham, Switzerland.

Ahrndt, S., Aria, A., Fähndrich, J., and Albayrak, S. (2015a). Ants in the OCEAN: Modulating agents with personality for planning with humans. In Bulling, N., editor, *Multi-Agent Systems*, number 8953 in Lecture Notes in Artificial Intelligence, pages 3–18. Springer International Publishing, Cham, Switzerland.

Ahrndt, S., Breitung, B., Fähndrich, J., and Albayrak, S. (2015b). Predictability in human-agent cooperation: Adapting to humans' personalities. In *SAC '15, Proceedings of the 30th Annual ACM Symposium on Applied Computing 2015*, volume 1: Artificial Intelligence & Agents, Distributed Systems, and Information Systems, pages 474–479. ACM Press, New York, NY, USA.

Ahrndt, S., Ebert, P., Fähndrich, J., and Albayrak, S. (2014a). Hplan: Facilitating the implementation of joint human-agent activities. In Demazeau, Y., Zambonelli, F., Corchado, J. M., and Bajo, J., editors, *Advances in Practical Applications of Heterogeneous Multi-Agent Systems. The PAAMS Collection.*, volume 8473 of *Lecture Notes in Computer Science*, pages 1–12. Springer International Publishing, Cham, Switzerland.

Ahrndt, S., Fähndrich, J., and Albayrak, S. (2013a). Agents vote against falls: The agent perspective in EPRs. In Demazeau, Y., Ishida, T., Corchado, J., and Bajo,

J., editors, *Advances on Practical Applications of Agents and Multi-Agent Systems*, volume 7879 of *Lecture Notes in Computer Science*, pages 263–266. Springer Berlin Heidelberg, Berlin, Germany.

Ahrndt, S., Fähndrich, J., and Albayrak, S. (2013b). Preventing elderly from falls: The agent perspective in EPRs. In Demazeau, Y., Ishida, T., Corchado, J. M., and Bajo, J., editors, *Advances on Practical Applications of Agents and Multi-Agent Systems*, volume 7879 of *Lecture Notes in Computer Science*, pages 1–12. Springer Berlin Heidelberg, Berlin, Germany.

Ahrndt, S., Fähndrich, J., and Albayrak, S. (2014b). Human-aware planning: A survey related to joint human-agent activities. In Bajo, J., Corchado, J. M., Mathieu, P., Campbell, A., Ortega, A., Adam, E., Navarro, E. M., Ahrndt, S., Moreno, M., and Julián, V., editors, *Trends in Practical Applications of Heterogeneous Multi-Agent Systems. The PAAMS Collection.*, volume 293 of *Advances in Intelligent Systems and Computing*, pages 95–102. Springer International Publishing, Cham, Switzerland.

Ahrndt, S., Fähndrich, J., and Albayrak, S. (2015c). Modelling of personality in agents: From psychology to implementation. In Bordini, R. H., Elkind, E., Weiss, G., and Yolum, P., editors, *Proceedings of the Fourth International Workshop on Human-Agent Interaction Design and Models (HAIDM 2015), co-located with AAMAS 2015*, pages 1–16. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, USA.

Ahrndt, S., Fähndrich, J., and Albayrak, S. (2016a). Human-agent teamwork: What is predictability, why is it important? In *SAC '16, Proceedings of the 31st Annual ACM Symposium on Applied Computing*, volume 1, pages 284–286, New York, NY, USA. ACM SIGAPP, ACM.

Ahrndt, S., Fähndrich, J., Lützenberger, M., and Albayrak, S. (2015d). Modelling of personality in agents: From psychology to logical formalisation and implementation. In Bordini, R. H., Elkind, E., Weiss, G., and Yolum, P., editors, *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '15, pages 1691–1692. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, USA.

Ahrndt, S., Fähndrich, J., Lützenberger, M., Rieger, A., and Albayrak, S. (2012a). An agent-based augmented reality demonstrator in the domestic energy domain. In Demazeau, Y., Müller, J. P., Rodríguez, J. M. C., and Pérez, J. B., editors, *Advances on Practical Applications of Agents and Multi-Agent Systems — 10th International Conference on Practical Applications of Agents and Multi-Agent Systems*, volume 155 of *Advances in Intelligent and Soft Computing*, pages 225–228. Springer Berlin Heidelberg, Berlin, Germany.

Ahrndt, S., Lützenberger, M., Heßler, A., and Albayrak, S. (2011). HAI – a human agent interface for JIAC. In Klügl, F. and Ossowski, S., editors, *Multiagent System Technologies*, volume 6973 of *Lecture Notes in Computer Science*, pages 149–156. Springer Berlin Heidelberg, Berlin, Germany.

Ahrndt, S., Lützenberger, M., and Prochnow, S. M. (2016b). Using personality models as prior knowledge to accelerate learning about stress-coping preferences: (demonstration). In Thangarajah, J., Tuyls, K., Jonker, C. M., and Marsella, S. C., editors, *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016)*, pages 1485–1487. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, USA.

Ahrndt, S., Rieger, A., and Albayrak, S. (2012b). Entwicklung einer mobilen elektronischen Patientenakte für die ambulante Versorgung in ländlichen Regionen (Development of a mobile electronic patient record for the ambulantory treatment in rural regions of Germany). In Goltz, U., Magnor, M., Appelrath, H.-J., Matthies, H., Balke, W.-T., and Wolf, L., editors, *INFORMATIK 2012: Was bewegt uns in der/die Zukunft?*, number 208 in Lecture Notes in Informatics, pages 1167–1181, Berlin, Germany. Bonner Köllen Verlag.

Ahrndt, S., Rieger, A., and Albayrak, S. (2014c). A mobile electronic patient record system for a new healthcare professional in Germany: The agnes ^{zwei} app. Technical report, DAI-Labor, Technische Universität Berlin.

Ahrndt, S., Rieger, A., Eryilmaz, E., and Albayrak, S. (2014d). Evaluation of a mobile EPR system for a new healthcare professional in Germany. In Lovis, C., Seroussi, B., Hasman, A., Pape-Haugaard, L., Saka, O., and Andersen, S. K., editors, *e-Health – For Continuity of Care*, volume 205 of *Studies in Health Technology and Informatics*, page 1270, Amsterdam, The Netherlands. European Federation for Medical Informatics, IOS Press.

Ahrndt, S., Roscher, D., Lützenberger, M., Rieger, A., and Albayrak, S. (2012c). Applying model-based techniques to the development of uis for agent systems. In Rodríguez, J. M. C., Pérez, J. B., Golinska, P., Giroux, S., and Corchuelo, R., editors, *Trends in Practical Applications of Agents and Multiagent Systems*, volume 157 of *Advances in Intelligent and Soft Computing*, pages 1–8. Springer International Publishing, Cham, Switzerland.

Ahrndt, S., Trollmann, F., Fähndrich, J., and Albayrak, S. (2016c). Personality and agents: Formalising state and effects. In Klusch, M., Unland, R., Shehory, O., Pokahr, A., and Ahrndt, S., editors, *Multiagent System Technologies*, number 9872 in Lecture Notes in Artificial Intelligence, pages 18–26. Springer International Publishing, Cham, Switzerland.

Alami, R., Clodic, A., Montreuil, V., Sisbot, E. A., and Chatila, R. (2005). Task planning for human-robot interaction. In Bailly, G., Crowley, J. L., and Privat, G., editors, *Proceedings of the 2005 Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-aware Services: Usages and Technologies*, sOc-EUSAI '05, pages 81–85, New York, NY, USA. ACM Press.

Alami, R., Clodic, A., Montreuil, V., Sisbot, E. A., and Chatila, R. (2006). Toward human-aware robot task planning. In Fong, T., editor, *AAAI Spring Symposium To boldly go where no human-robot team has gone before*, pages 1–8. The AAAI Press, Menlo Park, California, USA. Technical Report, SS-06-07.

Alili, S., Warnier, M., Ali, M., and Alami, R. (2009). Planning and plan-execution for human-robot cooperative task achievement. In *Proceedings of the 19th International Conference on Automated Planning & Scheduling*, pages 1–6.

Allbeck, J. and Badler, N. (2002). Toward representing agent behavior modified by personality and emotion. In *Proceedings of the Workshop on Embodied Conversational Agents at the 1st International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. ACM Press, New York, NY, USA.

Allen, J. and Ferguson, G. (2002). Human-machine collaborative planning. In *Proceedings of the Third International NASA Workshop on Planning and Scheduling in Space*, pages 1–10.

André, E., Klesen, M., Gebhard, P., Allen, S., and Rist, T. (2000). Integrating models of personality and emotions into lifelike characters. In Paiva, A., editor, *Affective Interactions*, volume 1814 of *Lecture Notes in Computer Science*, pages 150–165. Springer-Verlag New York, Inc., New York, NY, USA.

Aria, A. (2014). Integration von psychologischen Verhaltensmodellen auf Agenten (engl.: Integration of a psychological behaviour model into agents). Bachelor's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin.

Baarslag, T., Hendrikx, M. J., Hindriks, K. V., and Jonker, C. M. (2015). Learning about the opponent in automated bilateral negotiation: A comprehensive survey of opponent modeling techniques. *Autonomous Agents and Multi-Agent Systems*, pages 1–50.

Bainbridge, L. (1983). Ironies of automation. *Automatica*, 19(6):775–779.

Balke, T. and Gilbert, N. (2014). How do agents make decision? a survey. *Journal of Artificial Societies and Social Simulation*, 4(17):13.

Baumert, A. and Schmitt, M. (2012). Personality and information processing. *European Journal of Personality*, 26(2):87–89.

Bevacqua, E., de Sevin, E., Pelachaud, C., McRorie, M., and Sneddon, I. (2010). Building credible agents: Behaviour influenced by personality and emotional traits. In Lèvy, P., Bouchard, C., Yamanaka, Y., and Aoussat, A., editors, *The Proceedings of the Kansei Engineering and Emotion Research International Conference 2010 - KEER 2010*, pages 1–10, Paris, France.

Blackburn, P., van Benthem, J., and Wolter, F. (2006). *Handbook of Modal Logic*, volume 3 of *Studies in Logic and Practical Reasoning*. Elsevier, Amsterdam, The Netherlands.

Bradshaw, J. M., Dignum, V., Jonker, C. M., and Sierhuis, M. (2012). Human-agent-robot teamwork. *IEEE Intelligent Systems*, 27:8–13.

Bradshaw, J. M., Feltovich, P., Johnson, M., Breedy, M., Bunch, L., Eskridge, T., Jung, H., Lott, J., Uszok, A., and Diggelen, J. (2009). From tools to teammates: Joint activity in human-agent-robot teams. In Kurosu, M., editor, *Human Centered Design*,

volume 5619 of *Lecture Notes in Computer Science*, pages 935–944. Springer Berlin Heidelberg, Berlin, Germany.

Bradshaw, J. M., Feltovich, P. J., Johnson, M. J., Bunch, L., Breedy, M. R., Eskridge, T., Jung, H., Lott, J., and Uszok, A. (2008). Coordination in human-agent-robot teamwork. In *2008 International Symposium on Collaborative Technologies and Systems*, pages 467–476.

Bradshaw, J. M., Hoffman, R. R., Johnson, M. J., and Woods, D. D. (2013). The seven deadly myths of "autonomous systems". *IEEE Intelligent Systems*, 28(3):54–61.

Bratman, M. (1987). *Intention, plans, and practical reason.* Harvard University Press.

Bratman, M. E. (1992). Shared cooperative activity. *The Philosophical Review*, 101(2):327–341.

Braun, S. (2011). Konzeption, Implementierung und Dokumentation einer Kontextmenü API für die Google Maps API (engl.: Concept, implementation and documentation of a conext-menu api for the google maps api). Bachelor's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin.

Breitung, B. (2014). Modulating human-aware agent behavior in the colored trails game. Bachelor's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin.

Breitung, B. (2016). Personality-enabled stress assistant: A wearable to manage stress. Master's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin.

Brennan, S. E. (1998). The grounding problem in conversations with and through computers. In Fussell, S. R. and Kreuz, R. J., editors, *Social and Cognitive Psychological Approaches to Interpersonal Communicaton*, pages 201–225. Lawrence Erlbaum Associates.

Byrne, K. A., Silasi-Mansat, C. D., and Worthy, D. A. (2015). Who chokes under pressure? the big five personality traits and decision-making under pressure. *Personality and Individual Differences*, 74:22–28.

Cacheda, F., Carneiro, V., Fernandez, D., and Formoso, V. (2011). Comparison of collaborative filtering algorithms: Limitations of current techniques and proposals for scalable, high-performance recommender systems. *ACM Transactions on the Web*, 5(1):2:1–2:33.

Campos, A., Dignum, F., Dignum, V., Signoretti, A., Magály, A., and Fialho, S. (2009). A process-oriented approach to model agent personality. In Sierra, C., Castelfranchi, C., Decker, K. S., and Sichman, J. S., editors, *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems - Volume 2*, pages 1141–1142, Richland, SC, USA. International Foundation for Autonomous Agents and Multiagent Systems.

Canuto, A. M., Campos, A. M., M.Santos, A., Moura, E. C., Santos, E. B., Soares, R. G., and Dantas, K. A. (2006). Simulating working environments through the use of personality-based agents. In Sichman, J. S., Coelho, H., and Rezende, S. O., editors, *Advances in Artificial Intelligence - IBERAMIA-SBIA 2006*, number 4140 in LNAI, pages 108–117. Springer.

Canuto, A. M. P., Campos, A. M. C., Alchiere, J. C., Moura, E. C. M., M.Santos, A., Santos, E. B., and Soares, R. G. (2005). A personality-based model of agents for representing individuals in working organizations. In *Proceedings of the 2005 IEEE/WIC/ACM International Conference on Intelligent Agent Technology (IAT'05)*, pages 65–71. IEEE.

Carver, C. S. and Connor-Smith, J. (2010). Personality and coping. *Annual Review of Psychology*, 61:679–704.

Caspi, A., Roberts, B. W., and Shiner, R. L. (2005). Personality development: Stability and change. *Annual Review of Psychology*, (56):453–484.

Castelfranchi, C., de Rosis, F., Falcone, R., and Pizzutilo, S. (1998). Personality traits and social attitudes in multi-agent cooperation. *Applied Artificial Intelligence*, 12(7-8):649–675. Special Issue on 'Socially Intelligent Agents'.

Chakraborti, T., Talamadupula, K., Zhang, Y., and Kambhampati, S. (2016). A formal framework for studying interaction in human-robot societies. In *AAAI 2016 Workshop on Symbiotic Cognitive Systems*.

Chen, L. and Pu, P. (2004). Survey of preference elicitation methods. techreport EPFL Technical Report IC/2004/67, Ecole Politechnique Federale de Lausanne, Lausanne, Switzerland.

Cheng, B. H., Lemos, R., Giese, H., Inverardi, P., Magee, J., Andersson, J., Becker, B., Bencomo, N., Brun, Y., Cukic, B., Marzo Serugendo, G., Dustdar, S., Finkelstein, A., Gacek, C., Geihs, K., Grassi, V., Karsai, G., Kienle, H. M., Kramer, J., Litoiu, M., Malek, S., Mirandola, R., Müller, H. A., Park, S., Shaw, M., Tichy, M., Tivoli, M., Weyns, D., and Whittle, J. (2009). Software engineering for self-adaptive systems: A research roadmap. In Cheng, B. H., Lemos, R., Giese, H., Inverardi, P., and Magee, J., editors, *Software Engineering for Self-Adaptive Systems*, pages 1–26. Springer-Verlag, Berlin, Heidelberg.

Chi, D. M., Costa, M., Zhao, L., and Badler, N. (2000). The EMOTE model for effort and shape. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques SIGGRAPH '00*, pages 173–182, New York, NY, USA. ACM Press.

Christoffersen, K. and Woods, D. D. (2002). How to make automated systems team players. *Advances in Human Performance and Cognitive Engineering Research*, 2:1–12.

Cirillo, M. (2010). *Planning in Inhabited Environments: Human-Aware Task Planning and Activity Recognition*. PhD thesis, Mobile Robotics Lab, School of Science and Technology, Örebro University, Örebro, Sweden.

Cirillo, M., Karlsson, L., and Saffiotti, A. (2008). A framework for human-aware robot planning. In *Tenth Scandinavian Conference on Artificial Intelligence*, pages 52–59, Amsterdam, The Netherlands. IOS Press.

Cirillo, M., Karlsson, L., and Saffiotti, A. (2010). Human-aware task planning: An application to mobile robots. *ACM Transactions on Intelligent Systems and Technology*, 1(2):15:1–15:26.

Cirillo, M., Karlsson, L., and Saffiotti, A. (2012). Human-aware planning for robots embedded in ambient ecologies. *Pervasive and Mobile Computing*, 8:542–561.

Clark, H. H. (1996). *Using Language*. Cambridge University Press, Cambridge, UK.

Claus, C. and Boutilier, C. (1998). Thy dynamics of reinforcement learning in cooperative multiagent systems. In Mostow, J. and Rich, C., editors, *Proceedings of the 15th National Conference on Artificial Intelligence (AAAI-98)*, pages 746–752, Menlo Park, CA, USA. American Association for Artificial Intelligence.

Clodic, A., Alami, R., and Chatila, R. (2014). Key elements for human-robot joint action. In Hakli, R. and Seibt, J., editors, *Sociable Robots and the Future of Social Relations: Proceedings of Robo-Philosophy 2014*, Studies in the Philosophy of Sociality, pages 23–33. Springer International Publishing, Cham, Switzerland.

Clodic, A., Cao, H., Alili, S., Montreuil, V., Alami, R., and Chatila, R. (2009). Shary: A supervision system adapted to human-robot interaction. In Khatib, O., Kumar, V., and Pappas, G. J., editors, *Experimental Robotics*, volume 54 of *Springer Tracts in Advanced Robotics*, pages 229–238. Springer Berlin Heidelberg, Berlin, Germany.

Clodic, A., Montreuil, V., Alami, R., and Chatila, R. (2005). A decisional framework for autonomous robots interacting with humans. In *2005 IEEE International Workshop on Robots and Human Interactive Communication*, pages 543–548. IEEE.

Cohen, P. R. and Levesque, H. J. (1990). Intention is choice with commitment. *Artifical Intelligence*, 42:213–261.

Cohen, P. R., Levesque, H. J., and Smith, I. (1997). On team formation. In Holmstrom-Hintikka, G. and Tuomela, R., editors, *Contemporary Action Theory Volume 2: Social Action*, pages 87–114. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Cohen, S., Kamarck, T., and Mermelstein, R. (1983). A global measure of perceived stress. *Journal of Health and Social Behavior*, pages 385–396.

Compas, B. E., Connor-Smith, J., Saltzman, H., Thomson, A. H., and Wadsworth, M. E. (2001). Coping with stress during childhood and adolescence: Problems, progress, and potential in theory and research. *Psychological Bulletin*, 127(1):87–127.

Connor-Smith, J. and Flachsbart, C. (2007). Relations between personality and coping: A meta-analysis. *Journal of Personality and Social Psychology*, 93(6):1080–1107.

Corr, P. J. and Matthews, G., editors (2009). *The Cambridge Handbook of Personality Psychology*. Cambridge University Press, Cambridge and New York and Melbourne.

Dastani, M. and Lorini, E. (2012). A logic of emotions: from appraisal to coping. In Conitzer, V., Winikoff, M., Padgham, L., and van der Hock, W., editors, *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 2*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, USA.

de Raad, B., Mulder, E., Kloosterman, K., and Hofstee, W. K. (1988). Personality-descriptive verbs. *European Journal of Personality*, 2:81–96.

de Silva, L., Sardiña, S., and Padgham, L. (2009). First principles planning in BDI systems. In Sierra, C., Castelfranchi, C., Decker, K. S., and Sichman, J. S., editors, *Proceedings of The 8th International Conference on Autonomous Agents and Multi-agent Systems - Volume 2*, volume 2, pages 1105–1112. International Foundation for Autonomous Agents and Multiagent Systems.

de Winter, J. C. and Dodou, D. (2014). Why the Fitts list has persisted throughout the history of function allocation. *Cognition, Technology & Work*, 16:1–11.

de Winter, J. C. and Hancock, P. A. (2015). Reflections on the 1951 fitts list: Do humans believe now that machines surpass them? *Procedia Manufacturing*, 3:5334–5341.

Dean, T. L. and Wellmann, M. P. (1991). *Planning and Control*. Morgan Kaufmann.

Devin, S., Clodic, A., and Alami, R. (2017). About decisions during human-robot shared plan achievements: Who should act and how? In *Social Robotics*, number 10652 in Lecture Notes in Computer Science, pages 453–463, Cham, Switzerland. Springer.

Dickinson, I. and Wooldridge, M. (2005). Agents are not (just) web services: considering BDI agents and web services. In *Proceedings of the 2005 Workshop on Service-Oriented Computing and Agent-Based Engineering (SOCABE'2005), Utrecht, The Netherlands*.

Doce, T., Dias, J., Prada, R., and Paiva, A. (2010). Creating individual agents through personality traits. In Allbeck, J., Badler, N., Bickmore, T., Pelachaud, C., and Safonova, A., editors, *Intelligent Virtual Agents — 10th International Conference, IVA 2010, Philadelphia, PA, USA, September 20-22, 2010. Proceedings*, volume 6356 of *Lecture Notes in Computer Science Volume*, pages 257–264, Berlin, Germany. Springer Berlin Heidelberg.

Doran, J. E., Franklin, S., Jennings, N. R., and Norman, T. J. (1997). On cooperation in multi-agent systems. *The Knowledge Engineering Review*, 12(3):309–314.

Dryer, C. (1999). Getting personal with computers: How to design personalities for agents. *Applied Artificial Intelligence*, 13(3):273–295.

Du, H. (2013). The effects of human personality on human-agent interactions. In Ito, T., Jonker, C. M., Gini, M., and Shehory, O., editors, *Proceedings of the 2013 International Conference on Autonomous Agents and Multi-agent Systems (AAMAS 2013)*, pages 1427–1428. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, USA.

Du, H. and Huhns, M. N. (2013). Determining the effect of personality types on human-agent interactions. In *Proceedings of the 2013 IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT) - Volume 02*, volume 2 of *WI-IAT '13*, pages 239–244, Washington, DC, USA. IEEE Computer Society.

Durfee, E. H. and Lesser, V. R. (1989). Negotiating task decomposition and allocation using partial global planning. In Gasser, L. and Huhns, M. N., editors, *Distributed Artificial Intelligence*, volume 2, chapter 10, pages 229–243. Pitman Publishing Ltd.

Durupinar, F., Allbeck, J., Pelechano, N., and Badler, N. (2008). Creating crowd variation with the OCEAN personality model. In Padgham, L., Parkes, D., Müller, J., and Parsons, editors, *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 3*, pages 1217–1220. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, USA.

Durupinar, F., Pelechano, N., Allbeck, J. M., Güdükbay, U., and Badler, N. I. (2011). How the ocean personality model affects the perception of crowds. *IEEE Comput Graph Appl*, 31(2).

Dyan, P. and Niv, Y. (2008). Reinforcement learning: The good, the bad and the ugly. *Current Opinion in Neurobiology*, 18(2):185–196.

Ebert, P. (2013). Improving human-aware planning through reinforcement learning – a multi-agent based approach. Master's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin.

Egges, A., Kshirsagar, S., and Magnenat-Thalmann, N. (2004). Generic personality and emotion simulation for conversational agents. *Computer Animation and Virual Worlds*, 15:1–13.

Euzenat, J. and Shvaiko, P. (2007). *Ontology Matching*. Springer-Verlag Berlin Heidelberg.

Fasli, M. (2001a). Heterogeneous BDI agents I: Bold agents. In *FLAIRS Conference*, pages 195–199.

Fasli, M. (2001b). Heterogeneous BDI agents II: Circumspect agents. In *Intelligent Agent Technology: Research and Development*, pages 74–79. World Scientific.

Fasli, M. (2003). Interrelations between the BDI primitives: Towards heterogeneous agents. *Cognitive Systems Research*, 4:1–22.

Feist, J., Feist, G. J., and Roberts, T.-A. (2012). *Theories of Personality*. McGraw-Hill Education, New York, NY, USA.

Ficici, S. G. and Pfeiffer, A. (2008). Modeling how humans reason about others with partial information. In Padgham, L., Parkes, A. J., Müller, J. P., and Parsons, S., editors, *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 1*, volume 1, pages 315–322. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, USA.

Fitts, P. M. (1951). Human engineering for an effective air-navigation and traffic-control system.

Fox, M. and Long, D. (2003). PDDL2.1: An extension to PDDL for expressing temporal planning domains. *Journal of Artifical Intelligence Research*, 20:61–124.

Franklin, S. and Graesser, A. (1996). Is it an agent, or just a program?: A taxonomy for autonomous agents. In *Proceedings of the Workshop on Intelligent Agents III, Agent Theories, Architectures, and Languages*, pages 21–35, London, UK. Springer-Verlag.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review*, 102(4):652–670.

Furnham, A. (1996). The big five versus the big four: The relationship between the myers-briggs type indicator (mbti) and neo-pi five factor model of personality. *Personality and Individual Differences*, 21(2):303–307.

Gal, Y., Pfeffer, A., Marzo, F., and Grosz, B. J. (2004). Learning social preferences in games. In *Proceedings of the 19th National Conference on Artifical Intelligence*, pages 226–231, Palo Alto, California, USA. AAAI Press.

Gal, Y. K., Grosz, B. J., Kraus, S., Pfeffer, A., and Shieber, S. M. (2010). Agent decision-making in open-mixed networks. *Artificial Intelligence*, 174(18):1460–1480.

Ghallab, M., Nau, D., and Traverso, P. (2004). *Automated Planning: Theory & Practice*. Morgan Kaufmann Series in Artificial Intelligence. Morgan Kaufmann, San Francisco, CA, USA.

Gmytrasiewicz, P. J. and Lisetti, C. L. (2002). Emotions and personality in agent design and modeling. In Meyer, J.-J. C. and Tambe, M., editors, *International Workshop on Agent Theories, Architectures, and Languages*, number 2333 in Lecture Notes in Computer Science, pages 21–31, Berlin, Germany. Springer Berlin Heidelberg.

Gosling, S. D., Rentfrow, P. J., and Swann Jr., W. B. (2003). A very brief measure of the big-five personality domains. *Journal of Research in Personality*, 37:504–528.

Grosz, B. J. and Kraus, S. (1996). Collaborative plans for complex group actions. *Artifical Intelligence*, 86(2):269–357.

Grosz, B. J., Kraus, S., Talman, S., Stossel, B., and Havlin, M. (2004). The influence of social dependencies on decision-making: Initial investigations with a new game. In *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems - Volume 2*, volume 2, pages 782–789. IEEE Computer Society, Washington, DC, USA.

Grosz, B. J. and Sidner, C. (1990). *Plans for discourse*, chapter 20, pages 417–444. MIT Press.

Gupta, N. and Nau, D. S. (1992). On the complexity of blocks-world planning. *Artifical Intelligence*, 56(2–3):223–254.

Guy, S. J., Kim, S., Lin, M. C., and Manocha, D. (2011). Simulating heterogeneous crowd behaviors using personality trait theory. In *Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, SCA '11, pages 43–52, New York, NY, USA. ACM.

Haim, G., Gal, Y. K., Gelfand, M., and Kraus, S. (2012). A cultural sensitive agent for human-computer negotiation. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 1*, AAMAS '12, pages 451–458, Richland, SC, USA. International Foundation for Autonomous Agents and Multiagent Systems.

Hampson, S. E. and Goldberg, L. R. (2006). A first large-cohort study of personality-trait stability over the 40 years between elementary school and midlife. *Journal of Personality and Social Psychology*, (91):763–779.

Hanna, N. H. (2016). *Human-Agent Teamwork in Collaborative Virtual Environments*. phdthesis, Department of Computing, Faculty of Science and Engineering, Macquarie University, Sydney, Australia.

Hapke, U., Maske, U., Scheidt-Nave, C., Bode, L., Schlack, R., and Milde-Busch, A. (2013). Chronic stress among adulty in Germany. *Bundesgesundheitsblatt*, (5/6):1–5.

Hassard, J., Teoh, K., Cox, T., Dewe, P., Cosmar, M., Gründler, R., Flemming, D., Cosemans, B., and den Broek, K. V. (2014). *Calculating the costs of work-related stress and psychosocial risks*. European Agency for Safety and Health at Work, Bilbao, BI, Spain.

Hatemi, P. K. and Verhulst, B. (2015). Political attitudes develop independently of personality traits. *PLOS ONE*, 10:1–24.

Heckmann, D. (2005). *Ubiquitous User Modeling*. PhD thesis, Naturwissenschaftlich-Technischen Fakultäten, Universität des Saarlandes, Saarbrücken, Germany.

Herpers, R., Becker, P., Seele, S., Scherfgen, D., and Saitov, T. (2015). Agentenbasierte Verkehrssimulation mit psychologischen Persönlichkeitsprofilen (AVeSi) (engl.: Agent-based traffic simulation with psychological personality profiles). Technical report, University of Applied Sciences Bonn-Rhein-Sieg, Department of Computer Science, Sankt Augustin, Germany.

Herzig, A. and Longin, D. (2004). C&l intention revisited. In Dubois, D., Welty, C., and Williams, M.-A., editors, *Principles of Knowledge Representation and Reasoing: Proc. of the Ninth Int. Conf. (KR2004)*, pages 527–535. AAAI Press, Palo Alto, California, USA.

Hirsch, B., Konnerth, T., and Heßler, A. (2009). Merging agents and services – the JIAC agent platform. In Bordini, R. H., Dastani, M., Dix, J., and Amal, E. F. S., editors, *Multi-Agent Programming: Languages, Tools and Applications*, pages 159–185. Springer, New York, USA.

Hoffman, G. and Breazeal, C. (2007a). Cost-based anticipatory action selection for human-robot fluency. *IEEE Transactions on Robotics*, 5(23):952–961.

Hoffman, G. and Breazeal, C. (2007b). Effects of anticipatory action on human-robot teamwork – efficiency, fluency, and perception of team. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*, HRI '07, pages 1–8, New York, NY, USA. ACM.

Hoog, L. M. and Jennings, N. R. (2001). Socially intelligent reasoning for autonomous agents. *IEEE Transactions on Systems Man and Cybernetics*, 31(5):381–393.

Hunter, J. E. and Schmidt, F. L. (2004). *Methods of Meta-Analysis*. Sage Publications Inc., 2nd edition.

IBM Corporation (2005). An architectural blueprint for autonomic computing. IBM Corporation.

Jameson, A. (2001). Modeling both the context and the user. *Personal Ubiquitous Comput.*, 1(5).

Jennings, N. R. (1993). Commitments and conventions: The foundation of coordination in multi-agent systems. *The Knowledge Engineering Review*, 8(3):223–250.

Jennings, N. R. (2001). An agent-based approach for building complex software systems. *Communications of the ACM, Forthcoming*, 44(4):35–41.

Jiang, H., Vidal, J. M., and Huhns, M. N. (2007). Ebdi: An architecture for emotional agents. In *Proceedings of the 6th International Joint Conference on Autonomous Agents and Multiagent Systems*, AAMAS '07, pages 11:1–11:3, New York, NY, USA. ACM.

Joe, J. C., O'Hara, J., Medema, H. D., and Oxstrand, J. H. (2014). Identifying requirements for effective human-automation teamwork. In *Proceedings of the 12th International Conference on Probabilistic Safety Assessment and Management (PSAM 12)*, pages 1–12.

John, O. P. and Srivastava, S. (1999). The big-five trait taxonomy: History, measurement, and theoretical perspectives. In Pervin, L. A. and John, O. P., editors, *Handbook of Personality: Theory and Research*, volume 2, pages 102–138. The Guilford Press.

Johnson, M., Bradshaw, J. M., Hoffmann, R. R., Feltovich, P. J., and Woods, D. D. (2014a). Seven cardinal virtues of human-machine teamwork: Examples from the DARPA robotic challenge. *IEEE Intelligent Systems*, 29(6):74–80.

Johnson, M. J. (2014). *Coactive Design Desining Support for Interdependence in Human-Agent Teamwork*. phdthesis, Intelligent Systems, Electrical Engineering, Mathematics and Computer Science, Technische Universität Delft, Delft, The Netherlands.

Johnson, M. J., Bradshaw, J. M., Feltovich, P. J., Jonker, C. M., van Riemsdijk, B., and Sierhuis, M. (2011). The fundamental principle of coactive design: Interdependence must shape autonomy. In Vos, M. D., Fornara, N., Pitt, J. V., and Vouros, G., editors, *Coordination, Organizations, Institutions, and Norms in Agent Systems VI*, number 6541 in Lecture Notes in Computer Science, pages 172–191, Berlin, Heidelberg. Springer-Verlag.

Johnson, M. J., Bradshaw, J. M., Feltovich, P. J., Jonker, C. M., van Riemsdijk, B., and Sierhuis, M. (2012). Autonomy and interdependence in human-agent-robot teams. *IEEE Intelligent Systems*, 27(2):43–51.

Johnson, M. J., Bradshaw, J. M., Feltovich, P. J., Jonker, C. M., van Riemsdijk, M. B., and Sierhuis, M. (2014b). Coactive design: Designing support for interdependence in joint activity. *Journal of Human Robot Interaction*, 3(1):43–69.

Jones, H., Saunier, J., and Lourdeaux, D. (2009). Personality, emotions and physiology in a bdi agent architecture: The pep ->> bdi model. In *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, pages 263–266, Washington, DC, USA. IEEE Computer Society.

Jung, C. G. (1971). *Psychological Types*, volume 6 of *The collected works of C.G. Jung*. Princeton University Press. originally published in 1921.

Kaelbling, L. P., Littman, M. L., and Andreaw (1996). Reinforcement learning: A survey. *Journal of Artificial Intelligence Research*, 4:237–285.

Kaisers, M. (2012). *Learning against Learning: Evolutionary Dynamics of Reinforcement Learning Algorithms in Strategic Interaction*. doctoral thesis, Maastricht University, Maastricht, The Netherlands.

Kawamura, K., Rogers, T. E., and Ao, X. (2002). Development of a cognitive model of humans in a multi-agent framework for human-robot interaction. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems: Part 3*, AAMAS '02, pages 1379–1386. ACM, New York, NY, USA.

Keiser, J., Glass, J., Masuch, N., Lützenberger, M., and Albayrak, S. (2011). A distributed multi-operator W2V2G management approach. In *2011 IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 273–278. IEEE, Piscataway, NJ, USA.

Kephart, J. O. (2011). Autonomic computing: The first decade. In *Proceedings of the 8th ACM International Conference on Autonomic Computing*, ICAC '11, pages 1–2, New York, NY, USA. ACM.

Kephart, J. O. and Chess, D. M. (2003). The vision of autonomic computing. *Computer*, 36(1):41–50.

Kirsch, A. (2008). *Integration of Programming and Learning in a Control Language for Autonomous Robots Performing Everyday Activities*. doctoral thesis, Lehrstuhl für Bildverstehen und wissensbasierte Systeme, Institut für Informatik, Technische Universtität München, München, Germany.

Kirsch, A., Kruse, T., and Mösenlechner, L. (2009). An integrated planning and learning framework for human-robot interaction. In *4th Workshop on Planning and Plan Execution for Real-World Systems (held in conjuction with ICAPS 09)*, pages 1–6, Thessaloniki, Greece.

Kirsch, A., Kruse, T., Sisbot, E. A., Alami, R., Lawitzky, M., Brscic, D., Hirche, S., Basili, P., and Glasauer, S. (2010). Plan-based control of joint human-robot activities. *KI – Künstliche Intelligenz*, 24(3):223–231.

Klein, G., Feltovich, P. J., Bradshaw, J. M., and Woods, D. D. (2005). Common ground and coordination in joint activitiy. In Rouse, W. B. and Boff, K. R., editors, *Organizational Simmulation*, chapter 6, pages 139–184. John Wiley & Sons.

Klein, G., Woods, D. D., Bradshaw, J. M., Hoffmann, R. R., and Feltovich, P. J. (2004). Ten challenges for making automation a "team player" in joint human-agent activity. *Human-Centered Computing*, 19(6):91–95.

Klein, G. and Wright, C. (2016). Macrocognition: From theory to toolbox. *Frontiers in Psychology*, 7(54):54:1–54:5.

Knox, W. B. and Stone, P. (2010). Combining manual feedback with subsequent mdp reward signals for reinforcement learning. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1 - Volume 1*, AAMAS '10, pages 5–12, Richland, SC, USA. International Foundation for Autonomous Agents and Multiagent Systems.

Knox, W. B. and Stone, P. (2012). Reinforcement learning from simultaneous human and MDP reward. In Conitzer, V., Winikoff, M., Padgham, L., and van der Hoek, W., editors, *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2012)*, pages 475–482. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, USA.

Kober, J., Bagnell, J. A., and Peters, J. (2013). Reinforcement learning in robotics: A survey. *The International Journal of Robotics Research*, (32):1238 – 1274.

Kochanowicz, J., Tan, A.-H., and Thalmann, D. (2015). Beyond traits: Social context based personality model. In Bordini, R. H., Elkind, E., Weiss, G., and Yolum, P., editors, *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015)*, pages 1529–1538. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, USA.

Kokinov, B. (1995). A dynamic approach to context modeling. In *Proceedings of the IJCAI-95 Workshop on Modeling Context in Knowledge Representation and Reasoning*, volume 95, pages 199–209.

Konnerth, T. (2012). *An Agent-Based Approach to Service-Oriented Architectures*. doctoral thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin, Berlin, Germany.

Kravari, K. and Bassiliades, N. (2015). A survey of agent platforms. *Journal of Artificial Societies and Social Simulation*, 18(1):11.

Kruse, T., Pandey, A. K., Alami, R., and Kirsch, A. (2013). Human-aware robot navigation: A survey. *Robotics and Autonomous Systems*, 61(12):1726–1743.

Kücübayraktar, E. (2012). SOA, AOSE, organic computing and friends: Which fits best for ubicomp? Bachelor's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin.

Lallement, R., de Silva, L., and Alami, R. (2014). HATP: An HTN planner for robotics. In *Proceedings of the 24th International Conference on Automated Planning and Scheduling (ICAPS 2014)*.

Lemaignan, S., Warnier, M., Sisbot, E. A., and Alami, R. (2014). Human-robot interaction: Tackling the AI challenges. *Artificial Intelligence*.

Lin, R. and Kraus, S. (2010). Can automated agents proficiently negotiate with humans? *Communications of the ACM*, 53(1):78–88.

Lützenberger, M. (2014). A driverŠs mind: Psychology runs simulation. In Janssens, D., Yasar, A.-U.-H., and Knapen, L., editors, *Data Science and Simulation in Transportation Research*, pages 182–205. IGI Global.

Lützenberger, M. and Albayrak, S. (2014). Current frontiers in reproducing human driver behavior. In *Proceedings of the 2014 Summer Simulation Multiconference*, pages 71:1–71:8. Society for Computer Simulation International, San Diego, CA, USA.

Lützenberger, M., Konnerth, T., and Küster, T. (2015). Programming of multiagent applications with jiac. In Leitao, P. and Karnouskos, S., editors, *Industrial Agents – Emerging Applications of Software Agents in Industry*, pages 381–400. Elsevier, Amsterdam, The Netherlands.

Lützenberger, M., Küster, T., Konnerth, T., Thiele, A., Masuch, N., Heßler, A., Burkhardt, M., Tonn, J., Kaiser, S., and Keiser., J. (2013). Engineering industrial multi-agent systems — the JIAC V approach. In Cossentino, M., Seghrouchni, A. E. F., and Winikoff, M., editors, *Proceedings of the 1st International Workshop on Engineering Multi-Agent Systems (EMAS 2013)*, pages 160–175.

Lützenberger, M., Küster, T., Konnerth, T., Thiele, A., Masuch, N., Heßler, A., Keiser, J., Burkhardt, M., Kaiser, S., and Albayrak, S. (2013). JIAC V — A MAS framework for industrial applications (extended abstract). In Ito, T., Jonker, C., Gini, M., and Shehory, O., editors, *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems, Sait Paul, Minnesota, The USA.*, pages 1189–1190.

Ly, T.-A. (2015). Konzeption in Mensch-Agent Beziehungen unter Berücksichtigung von Persönlichkeit und 'Predictability' (engl.: Taking personality and predictability into consideration for human-agent relations). Bachelor's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin.

MacMillan, J., Entin, E. E., and Serfaty, D. (2004). Communication overhead: The hidden cost of team cognition. In Salas, E. and Fiore, S. M., editors, *Team Cognition: Understanding the Factor that Drive Process and Performance*, pages 61–82. US: American Psychological Association.

Manktelow, K. I. (2004). Reasoning and rationality: the pure and the practical. In Manktelow, K. I. and Chun, M. C., editors, *Psychology of Reasoning: Theoretical and Historical Perspectives*. Hove: Psychology Press, Hove, UK.

Maslow, A. H. (1943). A theory of human motivation. *Psychological Review*, (50):370–396.

Mataric, M. J. (1994). Reward functions for accelerated learning. In *Proceedings of the Eleventh International Conference on Machine Learning*, pages 181–189.

Matignon, L., Laurent, G. J., and Fort-Piat, N. L. (2006). Reward function and initial values: Better choices for accelerated goal-directed reinforcement learning. In *Artificial Neural Networks Ű ICANN 2006. ICANN 2006*, Lecture Notes in Computer Science, pages 840–849. Springer, Berlin, Heidelberg.

McCrae, R. R. and Costa, P. (1989). Reinterpreting the myers-briggs type indicators from the perspective of the five-factor model of personality. *Journal of Personality*, 57(1):17–40.

McCrae, R. R. and Costa Jr., P. T. (2006). *Personality in Adulthood A Five-Factor Theory Perspective*. The Guilford Press, New York, NY, USA, 2 edition.

McCrae, R. R. and Costa Jr., P. T. (2010). The five-factor theory of personality. In John, O. P., Robins, R. W., and Pervin, L. A., editors, *Handbook of Personality: Theory and Research*, chapter 5, pages 159–181. The Guilford Press, New York, NY, USA.

McCrae, R. R. and John, O. P. (1992). An introduction to the five-factor model and its applications. *Journal of Personality*, 60(2):175–215.

McCrae, R. R., Kurtz, J. E., Yamagata, S., and Terracciano, A. (2011). Internal consistency, retest reliability, and their implications for personality scale validity. *Personality and Social Psychology Review*, 15(1):28–50.

McDonald, J. H. (2014). *Handbook of Biological Statistics*. Sparky House Publishing, 3rd edition.

Milczarek, M., Schneider, E., and Gonzalez, E. R. (2009). *OSH in figures: stress at work – facts and figures*. European Agency for Safety and Health at Work, Bilbao, BI, Spain.

Millot, P. and Boy, G. A. (2012). Human-machine cooperation: A solution for life-critical systems? *Work*, 41:4552–4559.

Mitchell, T. M. (1997). *Machine Learning*. McGraw-Hill Science.

Mohammed, S. and Angell, L. C. (2003). Personality heterogeneity in teams: Which differences make a difference for team performance? *Small Group Research*, 34(6):651–677.

Mohammed, S., Ferzandi, L., and Hamilton, K. (2010). Metaphor no more: A 15-year review of the team mental model construct. *Journal of Management*, 36(4):876–910.

Mohammed, S., Hamilton, K., Tesler, R., Mancuso, V., and McNeese, M. (2015). Time for temporal team mental models: Expanding beyond what and how to incorporate when. *European Journal of Work and Organizational Psychology*, 24:693–709.

Montreuil, V., Clodic, A., and Alami, R. (2007). Planning human centered robot activities. In *IEEE International Conference on Systems, Man and Cybernetics*, ISIC '2007, pages 2618 –2623. IEEE Computer Society.

Müller, A. (2009). *TRANER – Transformational Planner: Transformational Planning for Autonomous Household Robots Using Libraries of Robust and Flexible Plans*. Suedwestdeutscher Verlag.

Müller, A., Kirsch, A., and Beetz, M. (2007). Transformational planning for everyday activity. In Boddy, M. S., Fox, M., and Thiébaux, S., editors, *Proceedings of the 17th International Conference on Automated Planning and Scheduling (ICAPS'07)*, pages 248–255, Providence, Rhode Island, USA. AAAI Press.

Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*, volume 1 of *Adaptive Computation and Machine Learning Series*. The MIT Press.

Myers, I. B. and Byers, P. B. (1995). *Gifts Differing: Understanding Personality Type*. Nicholas Brealey Publishing, 2 edition.

Myers, K. L. and Morley, D. N. (2001). Human directability of agents. In *Proceedings of the 1st International Conference on Knowledge Capture, K-CAP '01*, pages 108–115. ACM Press, New York, NY, USA.

Nebel, B. (2000). On the compilability and expressive power of propositional planning formalisms. *Journal of AI Research*, 12:271–315.

Neto, A. F. B. and da Silva, F. S. C. (2010). On the construction of synthetic characters with personality and emotion. In da Rocha Costa, A. C., Vicari, R., and Tonidandel, F., editors, *Advances in Artificial Intelligence Ű SBIA 2010*, volume 6404 of *Lecture Notes in Computer Science*, pages 102–111. Springer Berlin Heidelberg, Berlin, Germany.

Nikolai, C. and Madey, G. (2009). Tools of the trade: A survey of various agent based modeling platforms. *Journal of Artificial Societies and Social Simulation*, 2(2):1460–7425.

Oatley, K. and Jenkins, J. M. (1996). *Understanding Emotions*. Blackwell Publishing, Oxford, UK.

O'Connor, B. P. (2002). A quantitative review of the comprehensiveness of the five-factor model in relation to popular personality inventories. *Assessment*, 9(2):188–203.

Ortony, A., Clore, G. L., and Collins, A. (1988). *The Cognitive Structure of Emotions*. Cambridge University Press, Cambridge, UK.

Ortony, A., Norman, D., and Revelle, W. (2005). Affect and proto-affect in effective functioning. In Fellous, J.-M. and Arbib, M. A., editors, *Who needs Emotions?*, Series in Affective Science, pages 173–202. Oxford University Press.

Ozer, D. J. and Benet-Martínez, V. (2006). Personality and the prediction of consequential outcomes. *Annu. Rev. Psychol.*, 57:401–421.

Pandey, A. K. and Alami, R. (2010). Mightability maps: A perceptual level decisional framework for co-operative and competitive human-robot interaction. In *The 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5842–5848. IEEE.

Paunonen, S. V. and Jackson, D. N. (2000). What is beyond the big five? plenty! *Journal of Personality*, 68(5):821–835.

Pereira, D., Oliveria, E., Nelma, and Sarmento, L. (2005). Towards an architecture for emotional BDI agents. In *EPIA'05: Proceedings of 12th Portuguese Conference on Artificial Intelligence*, page 7. Springer.

Pianesi, F. (2013). Seasearch for personality. *IEEE Signal Processing Magazine*, 1(30):146–158.

Pittenger, D. J. (2005). Cautionary comments regarding the myers-briggs type indicator. *Consulting Psychology Journal: Practice and Research*, 57(3):210–221.

Plumbaum, T. (2015). *User Modeling in the Social Semantic Web*. doctoral thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin, Berlin, Germany.

Pockrandt, T. (2014). Human-Aware Planning und Joint Human-Agent Activities: Eine Übersicht des aktuellen Stands der Technik (engl.: Human-aware planning and joint human-agent activities: An overview of the current state-of-the-art). Bachelor's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin.

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., and Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and reommended remmedies. *Journal of Applied Psychology*, 88(5):879–903.

Pommeranz, A., Broekens, J., Wiggers, P., Brinkmann, W.-P., and Jonker, C. M. (2012). Designing interfaces for explicit preference elicitation: a user-centered investigation of preference srepresentation and elicitation process. *User Modeling and User-Adapted Interaction*, 22(4–5):357–397.

Prada, R. and Paiva, A. (2014). Human-agent interaction: Challenges for bringing humans and agents together. In *Third Internatioanl Workshop on Human-Agent Interaction Design and Models (HAIDM 2014) at the Thirteens International Conference on Agent and Multi-Agent Systems (AAMAS 2014)*, pages 1–10.

Prochnow, S. (2015). Enabling personality modeling in multi-agent systems through reinforcement learning. Master's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin.

Radke, L. (2018). Digitalisierung eines Hypercholesterinämie Behandlungspfades (engl.: Digitalisation of a disease management program for hypercholesteremia patients). Bachelor's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin.

Randløv, J. and Alstrøm, P. (1998). Learning to drive a bicycle using reinforcement learning and shaping. In *Proceedings of the Fifteenth International Conference on Machine Learning*, ICML '98, pages 463–471, San Francisco, CA, USA. Morgan Kaufmann Publishers Inc.

Rao, A. S. and Georgeff, M. P. (1991). Modeling rational agents within a BDI-architecture. In Allen, J., Fikes, R., and Sandewall, E., editors, *Proc. of the 2nd International Conference on Principles of Knowledge Representation and Reasoning*, pages 473–484. Morgan Kaufmann Publishers Inc., San Mateo, CA, USA.

Rao, A. S. and Georgeff, M. P. (1995). BDI agents: From theory to practice. In Lesser, V. and Gasser, L., editors, *Proceedings of the First International Conference on Mutiagent Systems (ICMAS 1995)*, pages 312–319, Palo Alto, California, USA. AAAI Press.

Rao, A. S. and Georgeff, M. P. (1998). Decision procedures for BDI logics. *Journal of Logic and Computation*, 8(3):293–343.

Resseguier, B., Léger, P.-M., Sénécal, S., Bastarache-Roberge, M.-C., and Courtemanche, F. (2016). *The Influence of Personality on Users' Emotional Reactions*, pages 91–98. Springer International Publishing, Cham, Switzerland.

Revelle, W. and Scherer, K. R. (2010). Personality and emotion. In Sander, D. and Scherer, K., editors, *The Oxford Companion to Emotion and the Affective Sciences*, Series in Affective Science, pages 447–448. Oxford University Press. Oxford Companion to the Affective Sciences.

Rizzo, P., Veloso, M., Miceli, M., and Cesta, A. (1999). Goal-based personalities and socialbehaviors in believable agents. *Applied Artificial Intelligence*, 13:239–272.

Rogers, T. E. (2003). *The Human Agent: A model for human-robot interaction.* PhD thesis, Vanderbilt University, Nashville, TN, USA.

Rogers, T. E., Peng, J., and Zein-Sabatto, S. (2005). Modeling human-robot interaction for intelligent mobile robotics. In *Robot and Human Interactive Communication, 2005. ROMAN 2005. IEEE International Workshop on.* IEEE, Piscataway, New Jersey, USA.

Rogers, T. E. and Wilkes, M. (2000). The human agent: a work in progress toward human-humanoid interaction. In *Systems, Man, and Cybernetics, 2000 IEEE International Conference on*, pages 864–869, Piscataway, New Jersey, USA. IEEE.

Rosenthal, S., Biswas, J., and Veloso, M. (2010). An effective personal mobile robot agent through symbiotic human-robot interaction. In Luck, M. and Sen, S., editors, *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: Volume 1 - Volume 1*, volume 1, pages 915–922, Richland, SC, USA. International Foundation for Autonomous Agents and Multiagent Systems.

Russell, S. and Norvig, P. (2002). *Artifical Intelligence: A Modern Approach*, volume 2. Prentice Hall, Upper Saddle River, New Jersey, USA.

Saberi, M. (2016). Personality-based cognitive design of characters in virtual environments. In Turner, J. O., Nixon, M., Bernardet, U., and DiPaola, S., editors, *Integrating Cognitive Architectures into Virtual Character Design*, pages 124–150. IGI Global, Hershey, PA, USA.

Salvit, J. and Sklar, E. (2011). Toward a myers-briggs type indicator model of agent behavior in multiagent teams. In Bosse, T., Geller, A., and Jonker, C. M., editors, *Multi-Agent-Based-Simulation XI, International Workshop, MABS 2010, Toronto, Canada, May 2010, Revised Selected Papers*, number 6532 in Lecture Notes in Artificial Intelligence, pages 28–43, Berlin, Germany. Springer Berlin Heidelberg.

Salvit, J. and Sklar, E. (2012). Modulating agent behavior using human personality type. In *Workshop on Human-Agent Interaction Design and Models (HAIDM) at Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 145–160.

Santos, R., Marreiros, G., Ramos, C., Neves, J., and Bulas-Cruz, J. (2010). Using personality types to support argumentation. In McBurney, P., Rahwan, I., Parsons, S., and Maudet, N., editors, *Argumentation in Multi-Agent Systems*, pages 292–304, Berlin, Germany. Springer Berlin Heidelberg.

Santos, R., Marreiros, G., Ramos, C., Neves, J., and Bulas-Cruz, J. (2011). Personality, emotion, and ood in agent-based group decision decision-making. *IEEE Intelligent Systems*, 26(6):58–66.

Satow, L. (2012). Big-Five-Persönlichkeitstest (B5T): Testmanual und Normen. `www.drsatow.de`, last-visited: 2017-09-01.

Schenk, A. K. (2012). Entwicklung einer Sturzrisikoerhebung für die Verwendung in elektronischen Patientenakten (engl.: Deveploment of a fall-risk assessment tool for electronic patient records). Bachelor's thesis, Evangelische Hochschule Berlin.

Schmuckler, M. A. (2001). What is ecological validity? a dimensional analysis. *Infancy*, 2(4):419–436.

Selye, H. (1956). *The Stress of Life*. McGraw-Hill Book Company, New York, NY, USA.

Sen, S. and Dutta, P. S. (2002). The evolution and stability of cooperative traits. In *Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems: Part 3*, New York, NY, USA. ACM.

Sheridan, T. B. (1992). *Telerobotics, Automation, and Human Supervisory Control*. The MIT Press, Cambridge, Massachusetts, USA.

Shoham, Y. (1993). Agent-oriented programming. *Artifical Intelligence*, 60(1):51–92.

Smith-Jentsch, K. A., Cannon-Bowers, J. A., Tannenbaum, S. I., and Salas, E. (2008). Guided team self-correction impacts on team mental models, processes, and effectiveness. *Small Group Research*, 39(3):303–327.

Sreedharan, S., Chakraborti, T., and Kambhampati, S. (2017). Balancing explicability and explanation in human-aware planning. In *AAAI 2017 Fall Symposium on Artificial Intelligence for Human-Robot Interaction (AI-for-HRI)*, pages 1–7. Association for the Advancement of Artificial Intelligence.

Steunebrink, B. R. (2010). *The Logical Structure of Emotions*. PhD thesis, Utrecht University, Utrecht, The Netherlands.

Sukthankar, G., Geib, C., Bui, H., Pynadath, D., and Goldman, R. P. (2014). *Plan, Activity, and Intent Recognition: Theory and Practice*. Morgen Kaufmann, San Francisco, CA, USA, 1 edition.

Sutton, R. S. and Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Adaptive Computation and Machine Learning. The MIT Press, Cambridge, Massachusetts, USA.

Sycara, K. P. (1998). Multiagent systems. *AI Magazine*, 19(2):79–92.

Sycara, K. P. and Lewis, M. (2004). Integrating intelligent agents into human teams. In Salas, E. and Fiore, S. M., editors, *Team Cognition: Understanding the Factors that Drive Process and Performance*, pages 1–36, Washington, DC, USA. American Psychological Association.

Sycara, K. P. and Sukthankar, G. (2006). Literature review of teamwork models. Technical Report CMU-RI-TR-06-50, Robotics Institute, Pittsburgh, PA.

Talamadupula, K., Kambhampti, S., Schermerhorn, P., Benton, J., and Scheutz, M. (2010). Planning for human-robot teaming in open worlds. *ACM Transactions on Intelligent Systems and Technology*, 1(2):14:1–14:24.

Talman, S., Hadad, M., Gal, Y., and Kraus, S. (2005). Adapting to agents' personalities in negotiation. In Pechoucek, M., Steiner, D., and Thompson, S., editors, *Proceedings of the Fourth International Joint Conference on Autonomous Agents and Multiagent Systems*, pages 383–389, New York, NY, USA. ACM.

Tapus, A., Matarić, M. J., and Scassellati, B. (2007). The grand challenges in socially assistive robotics. *IEEE Robotics and Automation Magazin*, 14(1):35–42.

Terracciano, A., Costa Jr., P. T., and McCrae, R. R. (2006). Personality plasticity after age 30. *Personality and Social Psychology Bulletin*, 32(8):999–1009.

Tomic, S., Pecora, F., and Saffiotti, A. (2014). Too cool for school — adding social constraints in human aware planning. In *CogRob 2014, The 9th International Workshop on Cognitive Robotics*, pages 1–6, Prague, Czech Republic.

Torre, I. (2009). Adaptive systems in the era of the semantic and social web, a survey. *User Modeling and User-Adapted Interaction*, 19(5):433–48.

Trafton, J. G., Hiatt, L. M., Harrison, A. M., Tramborello, F. P., Khemlani, S. S., and Schultz, A. C. (2012). ACT-R/E: An embodied cognitive architecture for human-robot interaction. *Jorurnal of Human-Robot Interaction*, 1(1):78–95.

van der Linden, D., te Nijenhuis, J., and Bakker, A. B. (2010). The general factor of personality: A meta-analysis of big five intercorrelations and a criterion-related validity study. *Journal of Research in Personality*, (44):315–327.

van Wissen, A., van Diggelen, J., and Dignum, V. (2009). The effects of cooperative agent behaviour on human cooperativeness (extended abstract). In *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems - Volume 2*, volume 2, pages 1179–1180, Richland, SC, USA. International Foundation for Autonomous Agents and Multiagent Systems. A full version of the paper was presented at the AAMAS 2009 Workshop on Mixed-Initiative Multiagent Systems (MIMS) using the same title.

Vered, M., Kaminka, G. A., and Biham, S. (2016). Online goal recognition through mirroring: Humans and agents. In *Proceedings of the Fourth Annual Conference on Advances in Cognitive Systems*, Evanston, Illinois, USA. Cognitive Systems Foundation.

Vinciarelli, A. (2014). More personality in personality computing. *IEEE Transactions on Affective Computing*, 5:297–300.

Vinciarelli, A. and Mohammadi, G. (2014). A survey of personality computing. *IEEE Transactions on Affective Computing*, 5:273–291.

von der Pütten, A. M., Krämer, N. C., and Gratch, J. (2010). How our personality shapes our interactions with virtual characters — implications for research and development. In Allbeck, J., Badler, N., Bickmore, T., Pelachaud, C., and Safonova, A., editors, *Intelligent Virtual Agents*, volume 6356 of *Lecture Notes in Computer Science*, pages 208–221. Springer Berlin Heidelberg, Berlin, Germany.

Walczak, A., Braubach, L., Pokahr, A., and Lamersdorf, W. (2007). Augmenting bdi agents with deliberative planning techniques. In Bordini, R. H., Dastani, M., Dix, J., and Seghrouchni, A., editors, *Programming Multi-Agent Systems*, volume 4411 of *Lecture Notes in Computer Science*, pages 113–127. Springer Berlin Heidelberg, Berlin, Germany.

Watkins, C. and Dayan, P. (1992). Q-learning. *Machine Learning*, 8:279–292.

Watkins, C. J. C. H. (1989). *Learning from Delayed Rewards*. PhD thesis, King's College, London, UK.

Weiß, G. (2001). Agentenorientiertes Software Engineering (engl.: Agent-oriented software engineering). *Informatik Spektrum*, 24(2):98–101.

Wiewiora, E., Cottrell, G., and Elkan, C. (2003). Principled methods for advising reinforcement learning agents. In *Proceedings of the Twentieth International Conference on International Conference on Machine Learning*, ICML'03, pages 792–799, Palo Alto, California, USA. AAAI Press.

Wilks, L. (2009). The stability of personality over time as a function of personality trait dominance. *Griffith University Undergraduated Student Psychology Journal*, 1:1–9.

Wilson, B., Zuckerman, I., and Nau, D. (2011). Modeling social preferences in multiplayer games. In Sonenberg, L., Stone, P., Tumer, K., and Yolum, P., editors, *The 10th International Conference on Autonomous Agents and Multiagent Systems - Volume 1*, volume 1–3, pages 337–344. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, USA.

Winograd, T. (1972). *Understanding Natural Language*, volume 3 of *Cognitive Psychology*. Academic Press, Inc, New York, NY, USA.

Wooldridge, M. (2000). *Reasoning about Rational Agents*. Intelligent Robotics and Autonomous Agents. The MIT Press, Cambridge, Massachusetts, USA.

Wooldridge, M. (2009). *An Introduction to MultiAgent Systems*. John Wiley & Sons, 2nd edition.

Wooldridge, M. and Jennings, N. R. (1995). Intelligent agents: theory and practice. *The Knowledge Engineering Review*, 10(2):115–152.

Wooldridge, M. J. (2002). Intelligent agents: The key concepts. In Marik, V., Stepankova, O., Krautwurmova, H., and Luck, M., editors, *Proceedings of the 9th ECCAI-ACAI/EASSS 2001, AEMAS 2001, HoloMAS 2001 on Multi-Agent-Systems and Applications II - Selected Revised Papers*, pages 3–43, London, UK. Springer Verlag.

Workplaces, H. (2015). Healthy workplaces manage stress 2014–2015. European Agency for Safety and Health at Work, `http://www.epsu.org/sites/default/files/article/gallery/EU-OSHA-HWC-2014_2015-Infographic-1-EN.jpg`, last-visited: 2017-09-01.

Wright, A. G. (2014). Current directions in personality science and the potential for advances through computing. *IEEE Transactions on Affective Computing*, 5(3):292–296.

Xu, M.-L. and Deng, X.-G. J. N. Z. (2014). Crowd simulation and its application: Recent advances. *Journal of Computer Science and Technology*, 5(92).

Zelenski, J. M. (2007). The role of personality in emotion, judgment, and decision making. In Vohs, K. D., Baumeister, R. F., and Loewenstein, G., editors, *Do Emotions Help or Hurt Decisionmaking?: A Hedgefoxian Perspective*, chapter 5, pages 117–132. Russell Sage Foundation, New York, NY, USA.

List of Publications

Parts of this work have been published within a subset of the following peer-reviewed publications, which represent the first-author papers published during the authors PhD studies. For a mapping between published work and the work presented in this document, the interested reader is referred to Section 2.2.

- Ahrndt, S., Lützenberger, M., Heßler, A., and Albayrak, S. (2011). HAI – a human agent interface for JIAC. In Klügl, F. and Ossowski, S., editors, *Multiagent System Technologies*, volume 6973 of *Lecture Notes in Computer Science*, pages 149–156. Springer Berlin Heidelberg, Berlin, Germany

- Ahrndt, S., Roscher, D., Lützenberger, M., Rieger, A., and Albayrak, S. (2012c). Applying model-based techniques to the development of uis for agent systems. In Rodríguez, J. M. C., Pérez, J. B., Golinska, P., Giroux, S., and Corchuelo, R., editors, *Trends in Practical Applications of Agents and Multiagent Systems*, volume 157 of *Advances in Intelligent and Soft Computing*, pages 1–8. Springer International Publishing, Cham, Switzerland

- Ahrndt, S., Fähndrich, J., Lützenberger, M., Rieger, A., and Albayrak, S. (2012a). An agent-based augmented reality demonstrator in the domestic energy domain. In Demazeau, Y., Müller, J. P., Rodríguez, J. M. C., and Pérez, J. B., editors, *Advances on Practical Applications of Agents and Multi-Agent Systems — 10th International Conference on Practical Applications of Agents and Multi-Agent Systems*, volume 155 of *Advances in Intelligent and Soft Computing*, pages 225–228. Springer Berlin Heidelberg, Berlin, Germany

- Ahrndt, S. (2012). Exploring self-explanation: The human side. DAI-Labor, Technische Universität Berlin. Contribution to the Doctoral Mentoring programm of the 10th German Conference on Multiagent System Technologies (MATES 2012), `http://dainas.dai-labor.de/~ahrndt@dai/mates2012dc_abstract.pdf`

- Ahrndt, S., Rieger, A., and Albayrak, S. (2012b). Entwicklung einer mobilen elektronischen Patientenakte für die ambulante Versorgung in ländlichen Regionen (Development of a mobile electronic patient record for the ambulantory treatment in rural regions of Germany). In Goltz, U., Magnor, M., Appelrath, H.-J., Matthies, H., Balke, W.-T., and Wolf, L., editors, *INFORMATIK 2012: Was bewegt uns in der/die Zukunft?*, number 208 in Lecture Notes in Informatics, pages 1167–1181, Berlin, Germany. Bonner Köllen Verlag

- Ahrndt, S., Fähndrich, J., and Albayrak, S. (2013b). Preventing elderly from falls: The agent perspective in EPRs. In Demazeau, Y., Ishida, T., Corchado, J. M.,

and Bajo, J., editors, *Advances on Practical Applications of Agents and Multi-Agent Systems*, volume 7879 of *Lecture Notes in Computer Science*, pages 1–12. Springer Berlin Heidelberg, Berlin, Germany

- Ahrndt, S., Fähndrich, J., and Albayrak, S. (2013a). Agents vote against falls: The agent perspective in EPRs. In Demazeau, Y., Ishida, T., Corchado, J., and Bajo, J., editors, *Advances on Practical Applications of Agents and Multi-Agent Systems*, volume 7879 of *Lecture Notes in Computer Science*, pages 263–266. Springer Berlin Heidelberg, Berlin, Germany

- Ahrndt, S. (2013). Improving human-aware planning. In Klusch, M., Thimm, M., and Paprzycki, M., editors, *Multiagent System Technologies*, number 8076 in Lecture Notes in Artificial Intelligence, pages 400–403. Springer Berlin Heidelberg, Berlin, Germany

- Ahrndt, S., Ebert, P., Fähndrich, J., and Albayrak, S. (2014a). Hplan: Facilitating the implementation of joint human-agent activities. In Demazeau, Y., Zambonelli, F., Corchado, J. M., and Bajo, J., editors, *Advances in Practical Applications of Heterogeneous Multi-Agent Systems. The PAAMS Collection.*, volume 8473 of *Lecture Notes in Computer Science*, pages 1–12. Springer International Publishing, Cham, Switzerland

- Ahrndt, S., Fähndrich, J., and Albayrak, S. (2014b). Human-aware planning: A survey related to joint human-agent activities. In Bajo, J., Corchado, J. M., Mathieu, P., Campbell, A., Ortega, A., Adam, E., Navarro, E. M., Ahrndt, S., Moreno, M., and Julián, V., editors, *Trends in Practical Applications of Heterogeneous Multi-Agent Systems. The PAAMS Collection.*, volume 293 of *Advances in Intelligent Systems and Computing*, pages 95–102. Springer International Publishing, Cham, Switzerland

- Ahrndt, S., Rieger, A., Eryilmaz, E., and Albayrak, S. (2014d). Evaluation of a mobile EPR system for a new healthcare professional in Germany. In Lovis, C., Seroussi, B., Hasman, A., Pape-Haugaard, L., Saka, O., and Andersen, S. K., editors, *e-Health – For Continuity of Care*, volume 205 of *Studies in Health Technology and Informatics*, page 1270, Amsterdam, The Netherlands. European Federation for Medical Informatics, IOS Press

- Ahrndt, S., Rieger, A., and Albayrak, S. (2014c). A mobile electronic patient record system for a new healthcare professional in Germany: The agnes ^{zwei} app. Technical report, DAI-Labor, Technische Universität Berlin

- Ahrndt, S., Aria, A., Fähndrich, J., and Albayrak, S. (2015a). Ants in the OCEAN: Modulating agents with personality for planning with humans. In Bulling, N., editor, *Multi-Agent Systems*, number 8953 in Lecture Notes in Artificial Intelligence, pages 3–18. Springer International Publishing, Cham, Switzerland

- Ahrndt, S., Breitung, B., Fähndrich, J., and Albayrak, S. (2015b). Predictability in human-agent cooperation: Adapting to humans' personalities. In *SAC '15, Proceedings of the 30th Annual ACM Symposium on Applied Computing 2015*, volume

1: Artificial Intelligence & Agents, Distributed Systems, and Information Systems, pages 474–479. ACM Press, New York, NY, USA

- Ahrndt, S., Fähndrich, J., Lützenberger, M., and Albayrak, S. (2015d). Modelling of personality in agents: From psychology to logical formalisation and implementation. In Bordini, R. H., Elkind, E., Weiss, G., and Yolum, P., editors, *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '15, pages 1691–1692. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, USA

- Ahrndt, S., Fähndrich, J., and Albayrak, S. (2015c). Modelling of personality in agents: From psychology to implementation. In Bordini, R. H., Elkind, E., Weiss, G., and Yolum, P., editors, *Proceedings of the Fourth International Workshop on Human-Agent Interaction Design and Models (HAIDM 2015), co-located with AAMAS 2015*, pages 1–16. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, USA

- Ahrndt, S., Fähndrich, J., and Albayrak, S. (2016a). Human-agent teamwork: What is predictability, why is it important? In *SAC '16, Proceedings of the 31st Annual ACM Symposium on Applied Computing*, volume 1, pages 284–286, New York, NY, USA. ACM SIGAPP, ACM

- Ahrndt, S., Lützenberger, M., and Prochnow, S. M. (2016b). Using personality models as prior knowledge to accelerate learning about stress-coping preferences: (demonstration). In Thangarajah, J., Tuyls, K., Jonker, C. M., and Marsella, S. C., editors, *Proceedings of the 15th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2016)*, pages 1485–1487. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, USA

- Ahrndt, S., Trollmann, F., Fähndrich, J., and Albayrak, S. (2016c). Personality and agents: Formalising state and effects. In Klusch, M., Unland, R., Shehory, O., Pokahr, A., and Ahrndt, S., editors, *Multiagent System Technologies*, number 9872 in Lecture Notes in Artificial Intelligence, pages 18–26. Springer International Publishing, Cham, Switzerland

- Ahrndt, S. and Albayrak, S. (2016). Joint human-agent activities: Challenges and definition. In Klusch, M., Unland, R., Shehory, O., Pokahr, A., and Ahrndt, S., editors, *Multiagent System Technologies*, number 9872 in Lecture Notes in Artificial Intelligence, pages 105–112. Springer International Publishing, Cham, Switzerland

- Ahrndt, S. and Albayrak, S. (2017). Learning about human personality. In Berndt, J. O., Petta, P., and Unland, R., editors, *Multiagent System Technologies*, Lecture Notes in Artificial Intelligence, pages 1–18. Springer International Publishing, Cham, Switzerland

List of Supervised Theses

The following bachelor and master theses have been supervised by the author during the PhD studies.

- Braun, S. (2011). Konzeption, Implementierung und Dokumentation einer Kontextmenü API für die Google Maps API (engl.: Concept, implementation and documentation of a conext-menu api for the google maps api). Bachelor's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin

- Schenk, A. K. (2012). Entwicklung einer Sturzrisikoerhebung für die Verwendung in elektronischen Patientenakten (engl.: Deveploment of a fall-risk assessment tool for electronic patient records). Bachelor's thesis, Evangelische Hochschule Berlin

- Kücübayraktar, E. (2012). SOA, AOSE, organic computing and friends: Which fits best for ubicomp? Bachelor's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin

- Ebert, P. (2013). Improving human-aware planning through reinforcement learning – a multi-agent based approach. Master's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin

- Breitung, B. (2014). Modulating human-aware agent behavior in the colored trails game. Bachelor's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin

- Aria, A. (2014). Integration von psychologischen Verhaltensmodellen auf Agenten (engl.: Integration of a psychological behaviour model into agents). Bachelor's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin

- Pockrandt, T. (2014). Human-Aware Planning und Joint Human-Agent Activities: Eine Übersicht des aktuellen Stands der Technik (engl.: Human-aware planning and joint human-agent activities: An overview of the current state-of-the-art). Bachelor's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin

- Ly, T.-A. (2015). Konzeption in Mensch-Agent Beziehungen unter Berücksichtigung von Persönlichkeit und 'Predictability' (engl.: Taking personality and predictability into consideration for human-agent relations). Bachelor's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin

- Prochnow, S. (2015). Enabling personality modeling in multi-agent systems through reinforcement learning. Master's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin

- Breitung, B. (2016). Personality-enabled stress assistant: A wearable to manage stress. Master's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin
- Radke, L. (2018). Digitalisierung eines Hypercholesterinämie Behandlungspfades (engl.: Digitalisation of a disease management program for hypercholesteremia patients). Bachelor's thesis, Fakultät IV – Elektrotechnik und Informatik, Technische Universität Berlin

Part VII.

Appendix

Appendix A – List of Abbreviations

ACT-R	**A**daptive **C**haracter for **T**hought-**R**ational
ADL	**A**ctivities of **D**aily **L**iving
AI	**A**rtifical **I**ntelligence
AOSE	**A**gent-**O**riented **S**oftware **E**ngineering
API	**A**pplication **P**rogramming **I**nterface
BDI	**B**elief **D**esire **I**ntention
CT	**C**olored **T**rails
DHHB	**D**ynamic **H**euristic of **H**uman **B**ehaviour
IAA	**I**nter**A**ction**A**gent
HCI	**H**uman **C**omputer **I**nteraction
HRI	**H**uman **A**gent **I**nteraction
HAP	**H**uman-**A**ware **P**lanning
HATP	**H**uman **A**ware **T**ask **P**lanner
HPLAN	**H**uman-**PLAN**
HRI	**H**uman **R**obot **I**nteraction
HTN	**H**irachical **T**ask **N**etwork
IPIP	**I**nternational **P**ersonality **I**tem **P**ool
JIAC	**J**ava-based **I**ntelligent **A**gent **C**omponentware
LORA	**L**ogic **O**f **R**ational **A**gents
FFM	**F**ive-**F**actor **M**odel
MAPE-K	**M**onitor **A**analyse **P**lan **E**xecute - **K**nowledge
MAS	**M**ulti-**A**gent **S**ystem
MBTI	**M**yers-**B**riggs **T**ype **I**ndicator
MDP	**M**arkov **D**ecision **P**rocess
OCEAN	**O**penness **C**onscientiousness **E**xtraversion **A**greeableness **N**euroticism
PDDL	**P**lanning **D**omain **D**escription **L**anguage
PeSA	**P**ersonality-**e**nabled **S**tress **A**ssistant

PSS	**P**erceived **S**tress **S**cale
RL	**R**einforcement **L**earning
SDK	**S**oftware **D**evelopment **K**it
TIPI	**T**en-**I**tem **P**ersonality **I**nventory
XML	E**X**tensible **M**arkup **L**anguage

Appendix B – A 100-Item Set to Determine the Big-Five Factors

To determine the personality of an individual the questionnaire listed in the following table was used. It is derived from the IPIP (International Personality Item Pool, `http://ipip.ori.org/`) and shows a Cronbachs Alpha between .88 and .91.[3]

#	Item	Trait	Direction
1	AmÂăthe life of the party.	E	+
2	Insult people.	A	-
3	Am always prepared.	C	+
4	Get stressed out easily.	N	-
5	Have a rich vocabulary.	O	+
6	Often feel uncomfortable around others.	E	-
7	Am interested in people.	A	+
8	Leave my belongings around.	C	-
9	Am relaxed most of the time.	N	+
10	Have difficulty understanding abstract ideas.	O	-
11	Feel comfortable around people.	E	+
12	Am not interested in other people's problems.	A	-
13	Pay attention to details.	C	+
14	Worry about things.	N	-
15	Have a vivid imagination.	O	+
16	Keep in the background.	E	-
17	Sympathize with others' feelings.	A	+
18	Make a mess of things.	C	-
19	Seldom feel blue.	N	+
20	Am not interested in abstract ideas.	O	-
21	Start conversations.	E	+
22	Feel little concern for others.	A	-
23	Get chores done right away.	C	+
24	Am easily disturbed.	N	-
25	Have excellent ideas.	O	+
26	Have little to say.	E	-
27	Have a soft heart.	A	+
28	Often forget to put things back in their proper place.	C	-
29	Am not easily bothered by things.	N	+

continued on next page...

[3]Last-visited: 2017-02-16, `http://ipip.ori.org/newBigFive5broadTable.htm`

...continued from previous page

#	**Item**	**Trait**	**Direction**
30	Do not have a good imagination.	O	-
31	Talk to a lot of different people at parties.	E	+
32	Am not really interested in others.	A	-
33	Like order.	C	+
34	Get upset easily.	N	-
35	Am quick to understand things.	O	+
36	Don't like to draw attention to myself.	E	-
37	Take time out for others.	A	+
38	Shirk my duties.	C	-
39	Rarely get irritated.	N	+
40	Try to avoid complex people.	O	-
41	Don't mind being the center of attention.	E	+
42	Am hard to get to know.	A	-
43	Follow a schedule.	C	+
44	Change my mood a lot.	N	-
45	Use difficult words.	O	+
46	Am quiet around strangers.	E	-
47	Feel others' emotions.	A	+
48	Neglect my duties.	C	-
49	Seldom get mad.	N	+
50	Have difficulty imagining things.	O	-
51	Make friends easily.	E	+
52	AmÂăindifferent to the feelings of others.	A	-
53	AmÂăexacting in my work.	C	+
54	Have frequent mood swings.	N	-
55	Spend time reflecting on things.	O	+
56	Find it difficult to approach others.	E	-
57	Make people feel at ease.	A	+
58	Waste my time.	C	-
59	Get irritated easily.	N	-
60	Avoid difficult reading material.	O	-
61	Take charge.	E	+
62	Inquire about others' well-being.	A	+
63	Do things according to a plan.	C	+
64	Often feel blue.	N	-
65	Am full of ideas.	O	+
66	Don't talk a lot.	E	-
67	Know how to comfort others.	A	+
68	Do things in a half-way manner.	C	-
69	Get angry easily.	N	-
70	Will not probe deeply into a subject.	O	-
71	Know how to captivate people.	E	+
72	Love children.	A	+

continued on next page...

...continued from previous page

#	Item	Trait	Direction
73	Continue until everything is perfect.	C	+
74	Panic easily.	N	-
75	Carry the conversation to a higher level.	O	+
76	Bottle up my feelings.	E	-
77	Am on good terms with nearly everyone.	A	+
78	Find it difficult to get down to work.	C	-
79	Feel threatened easily.	N	-
80	Catch on to things quickly.	O	+
81	Feel at ease with people.	E	+
82	Have a good word for everyone.	A	+
83	Make plans and stick to them.	C	+
84	Get overwhelmed by emotions.	N	-
85	Can handle a lot of information.	O	+
86	Am a very private person.	E	-
87	Show my gratitude.	A	+
88	Leave a mess in my room.	C	-
89	Take offense easily.	N	-
90	Am good at many things.	O	+
91	Wait for others to lead the way.	E	-
92	Think of others first.	A	+
93	Love order and regularity.	C	+
94	Get caught up in my problems.	N	-
95	Love to read challenging material.	O	+
96	Am skilled in handling social situations.	E	+
97	Love to help others.	A	+
98	Like to tidy up.	C	+
99	Grumble about things.	N	-
100	Love to think up new ways of doing things.	O	+

The attendees were asked to provide answers to the question *'How Accurately Can You Describe Yourself?'* for the respective item. The introduction contained the following statement:

> *Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age. So that you can describe yourself in an honest manner, your responses will be kept in absolute confidence. Indicate for each statement whether it is 1. Very Inaccurate, 2. Moderately Inaccurate, 3. Neither Accurate Nor Inaccurate, 4. Moderately Accurate, or 5. Very Accurate as a description of you.*

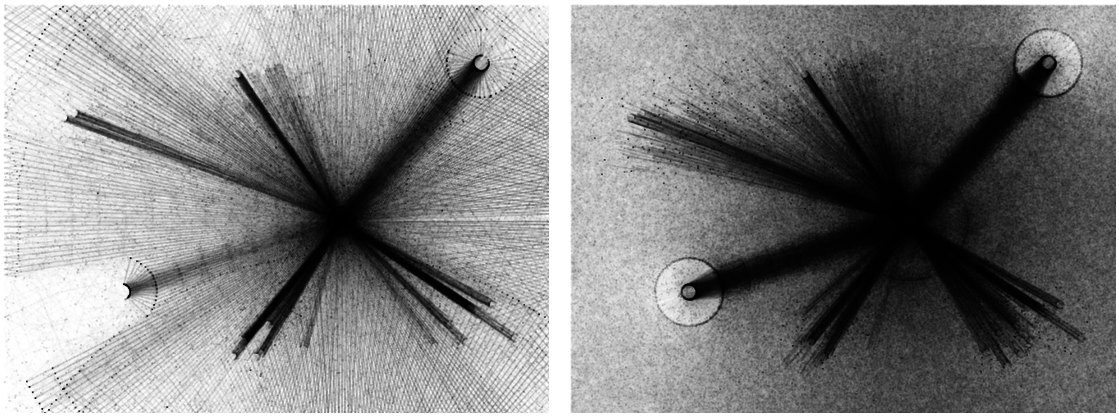
To build the scores for each personality trait the items are marked with the trait they belong to and the direction this item influences the final score. The direction reads as shown in the following Table.

Answer	**Direction**	**Score**
1. Very Inaccurate	+	1
	-	5
2. Moderately Inaccurate	+	2
	-	4
3. Neither Inaccurate nor Accurate	+	3
	-	3
4. Moderately Accurate	+	4
	-	2
5. Very Accurate	+	5
	-	1

Such information is then used to build the total scores for the personality traits by summing all the values for the associated item. The results are a score for each personality trait in the range $[20 \ldots 100]$, which was normalised in the evaluation.

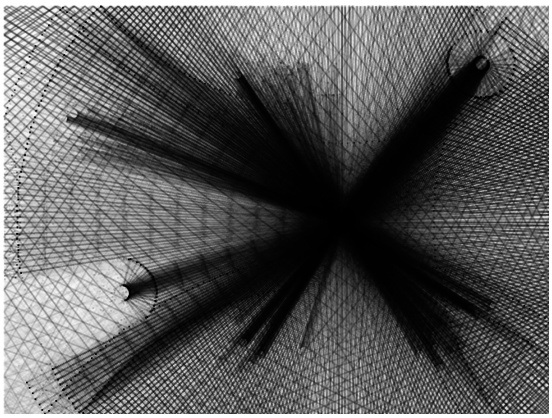
Appendix C – Ants in the OCEAN: Paths

One of the questions answered in this document is if different personalities lead to variations in the interpretation of inputs, the decision-making process, and the generation of outputs of agents. This was done using the AntMe! simulation environment. Below the interested reader can find all cumulated paths of the simulated ant populations. The declared characteristic of the FFM traits is in the order OCEAN.

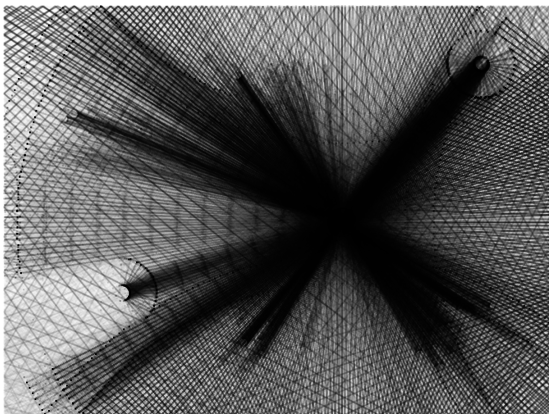


(a) (0,0,0,0,0) **(b)** (1,1,1,1,1)

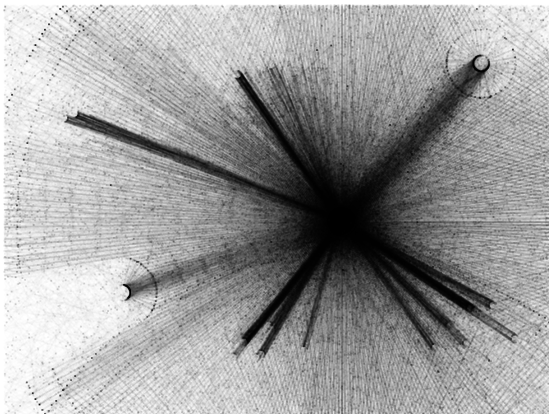
Figure 1.: Path heat-maps of the ant populations with all traits at minimal values (left side) and all traits at maximal values (right side).



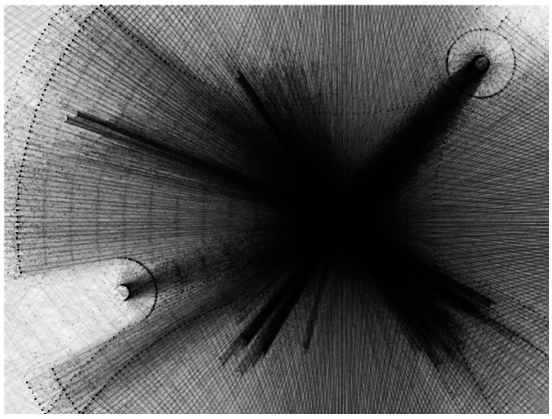
(a) (0,0,0,0,1)



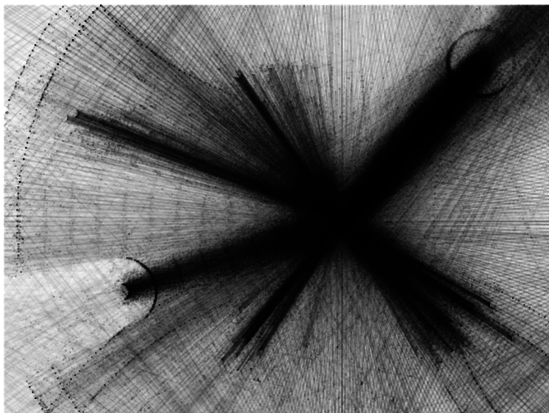
(b) (0,0,0,1,1)



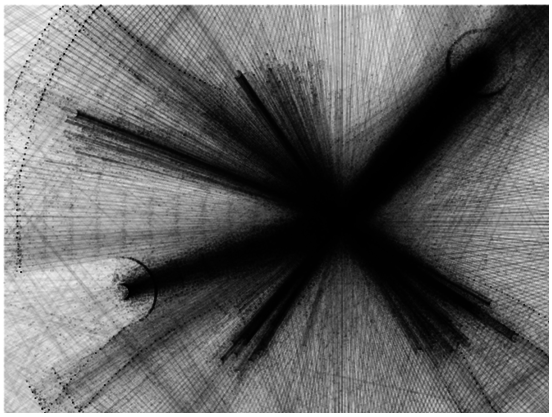
(c) (0,0,1,0,0)



(d) (0,0,1,0,1)

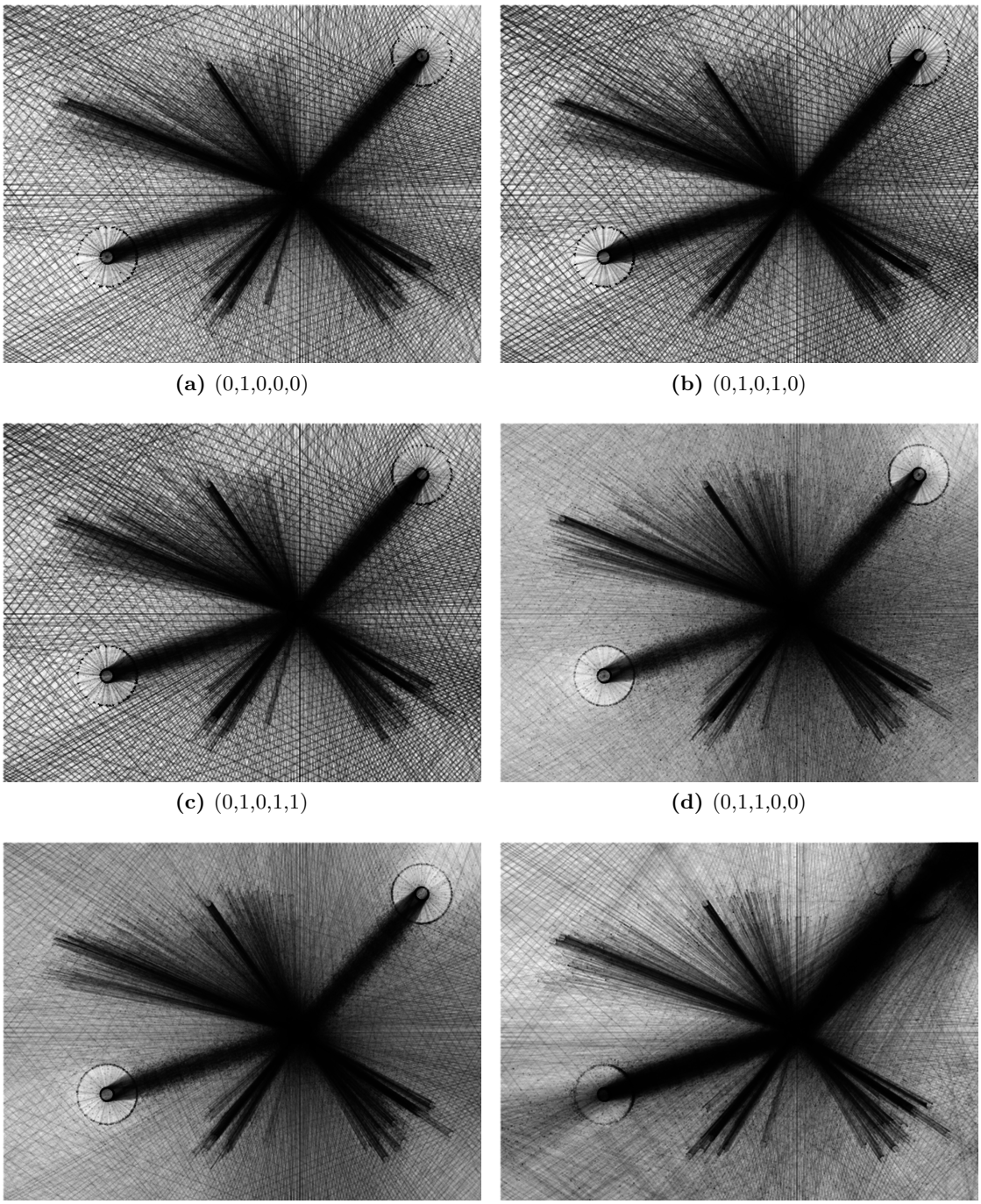


(e) (0,0,1,1,0)



(f) (0,0,1,1,1)

Figure 2.: Path heat-maps of different ant-populations.

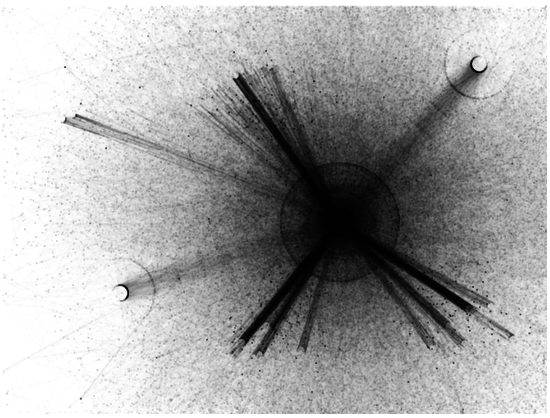


(a) (0,1,0,0,0) **(b)** (0,1,0,1,0)

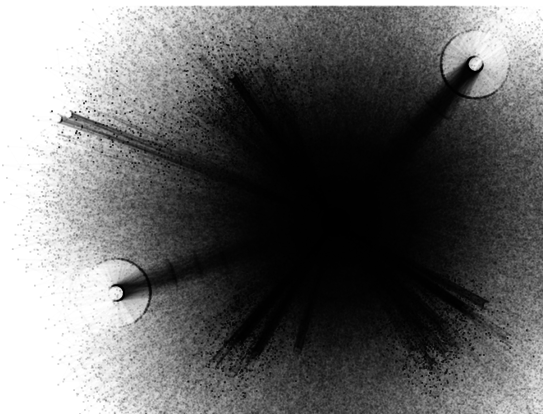
(c) (0,1,0,1,1) **(d)** (0,1,1,0,0)

(e) (0,1,1,0,1) **(f)** (0,1,1,1,1)

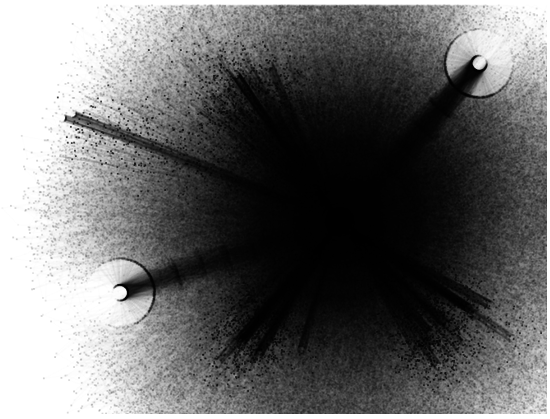
Figure 3.: Path heat-maps of different ant-populations.



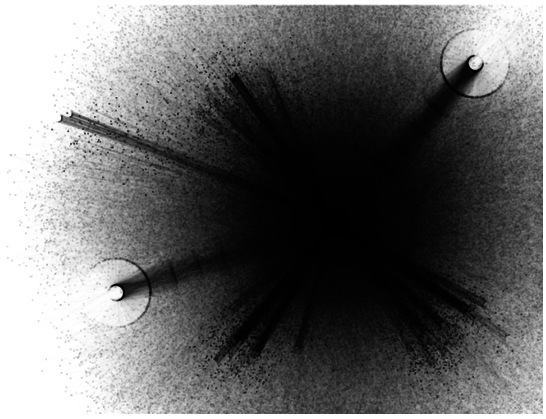
(a) (1,0,0,0,0)



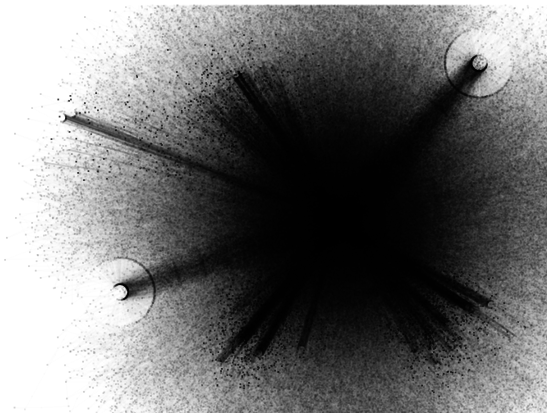
(b) (1,0,0,0,1)



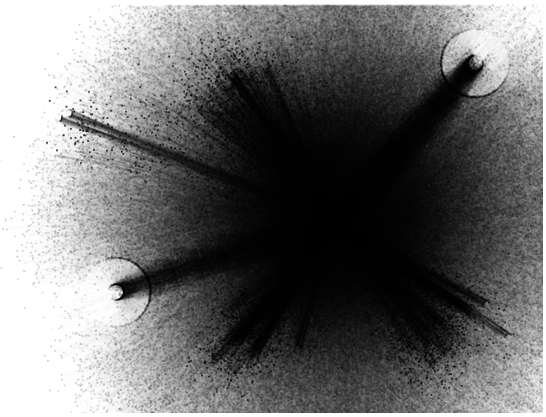
(c) (1,0,0,1,0)



(d) (1,0,0,1,1)

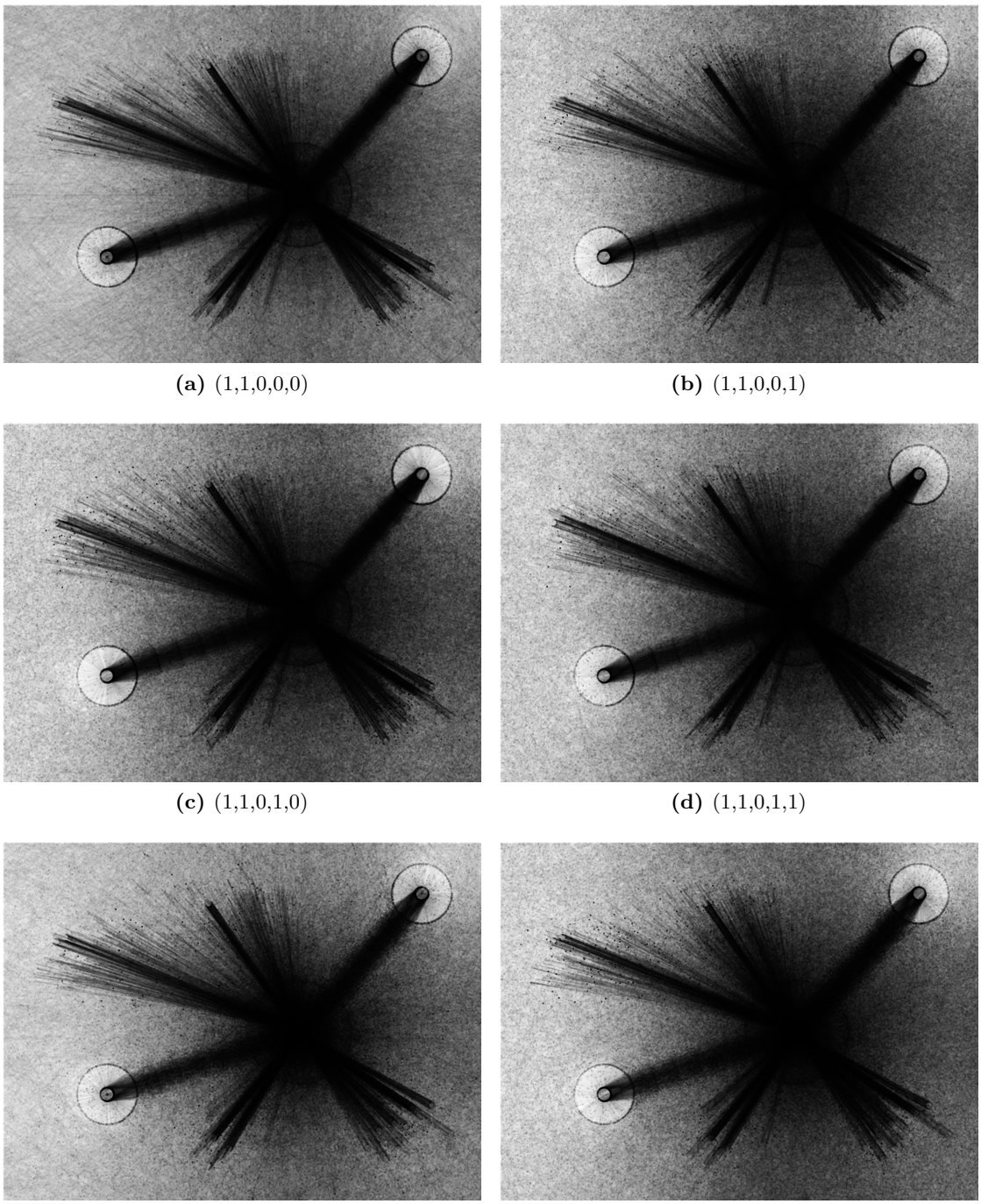


(e) (1,0,1,0,1)



(f) (1,0,1,1,0)

Figure 4.: Path heat-maps of different ant-populations.

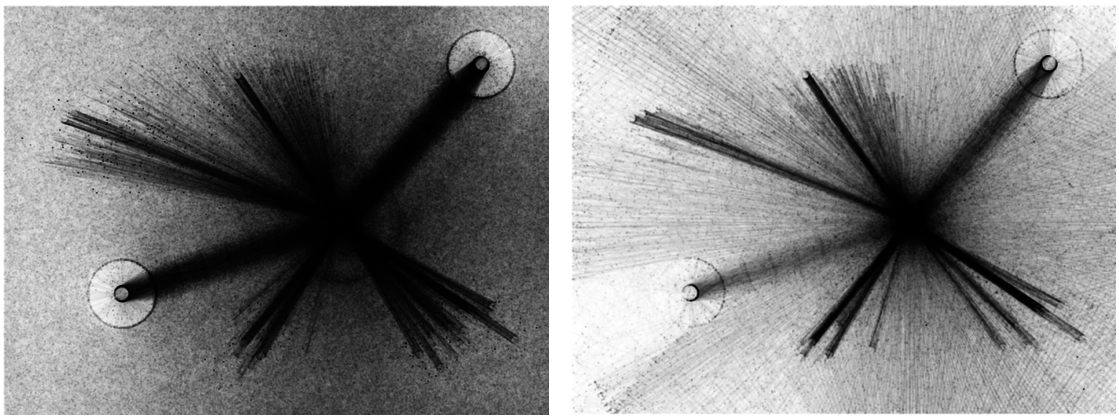


(a) (1,1,0,0,0) **(b)** (1,1,0,0,1)

(c) (1,1,0,1,0) **(d)** (1,1,0,1,1)

(e) (1,1,1,0,0) **(f)** (1,1,1,0,1)

Figure 5.: Path heat-maps of different ant-populations.



(a) (1,1,1,1,0) **(b)** $(\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2},\frac{1}{2})$

Figure 6.: Path heat-maps of different ant-populations.

Appendix D – KD-Proofs for Personality Modality

The personality modality introduced in Chapter 9 has a logic that corresponds to the normal modal system KD (*cf.* Blackburn et al., 2006, pp. 86–138). That means the following terms hold (proves are integrated for the sake of completeness and can be looked up in the literature as well):

1. $\models_{\mathcal{S}}$ (Per i $(\varphi \Rightarrow \psi)$) $\Rightarrow$ ((Per i φ) $\Rightarrow$ (Per i ψ)).

   To prove this we suppose that $\langle M, V, w, t\rangle \models_{\mathcal{S}}$ (Per i $(\varphi \Rightarrow \psi)$) and $\langle M, V, w, t\rangle \models_{\mathcal{S}}$ (Per i φ) holds for any $\langle M, V, w, t\rangle$. From the semantics of Per, we know that $\langle M, V, w, t\rangle \models_{\mathcal{S}} \varphi \Rightarrow \psi$ and $\langle M, V, w, t\rangle \models_{\mathcal{S}} \varphi$ for all $w' \in \mathcal{P}_t^w([\![i]\!])$. From this it follows that, $\langle M, V, w', t\rangle \models_{\mathcal{S}} \psi$ and thus $\langle M, V, w', t\rangle \models_{\mathcal{S}}$ (Per i (φ).

2. $\models_{\mathcal{S}}$ (Per i φ) $\Rightarrow \neg$(Per i $\neg\varphi$).

   Suppose $\not\models_{\mathcal{S}}$ (Per i φ) $\Rightarrow \neg$(Per i $\neg\varphi$). Then for some $\langle M, V, w, t\rangle$ it applies both: $\langle M, V, w, t\rangle \models_{\mathcal{S}}$ (Per i φ) and $\langle M, V, w, t\rangle \models_{\mathcal{S}}$ (Per i $\neg\varphi$). As we require the personality accessibility relation to be serial, we can follow that there exist at least one $w' \in \mathcal{P}_t^w([\![i]\!])$. Mapping this to the semantics of Per (*cf.* Eq. 9.5), we can follow that $\langle M, V, w', t\rangle \models_{\mathcal{S}} \varphi$ and $\langle M, V, w', t\rangle \models_{\mathcal{S}} \neg\varphi$. Given this a contradiction follows as $\langle M, V, w', t\rangle \models_{\mathcal{S}} \varphi$ and $\langle M, V, w', t\rangle \models_{\mathcal{S}} \varphi$ applies at the same time; implying that $\models_{\mathcal{S}}$ (Per i φ) $\Rightarrow \neg$(Per i $\neg\varphi$) holds.

3. If $\models_{\mathcal{S}} \varphi$ then $\models_{\mathcal{S}}$ (Per i φ).

   To prove this we assume that $\models_{\mathcal{S}} \varphi$. Then we have to show that $\langle M, V, w, t\rangle \models_{\mathcal{S}}$ (Per i φ) holds for all $\langle M, V, w, t\rangle$. Deduced from $\models_{\mathcal{S}} \varphi$ it applies that φ is satisfied by all interpretation structures. In particular this means that $\langle M, V, w', t\rangle \models_{\mathcal{S}} \varphi$ holds for all $w' \in \mathcal{P}_t^w([\![i]\!])$. It follows that $\langle M, V, w, t\rangle \models_{\mathcal{S}}$ (Per i φ) (*cf.* Eq. 9.5).