Halbrügge, M. (2016). *Towards the Evaluation of Cognitive Models using Anytime Intelligence Tests.* In D. Reitter & F. E. Ritter (eds.), Proceedings of the 14th international conference on cognitive modeling (pp. 261–263). University Park, PA: Penn State.

Towards the Evaluation of Cognitive Models using Anytime Intelligence Tests

Marc Halbrügge (marc.halbruegge@tu-berlin.de)
Quality & Usability Lab, Telekom Innovation Laboratories
Technische Universität Berlin
Ernst-Reuter-Platz 7, 10587 Berlin

Abstract

Cognitive models are usually evaluated based on their fit to empirical data. Artificial intelligence (AI) systems on the other hand are mainly evaluated based on their performance. Within the field of artificial general intelligence (AGI) research, a new type of performance measure for AGI systems has recently been proposed that tries to cover both humans and artificial systems: Anytime Intelligence Tests (AIT; Hernández-Orallo & Dowe, 2010). This paper explores the viability of the AIT formalism for the evaluation of cognitive models based on data from the ICCM 2009 "Dynamic Stocks and Flows" modeling challenge.

Keywords: Anytime Intelligence Test; Model Evaluation; Decision Making; Stock-Flow Problems;

Introduction

Cognitive modeling as a field, although being rooted in AI, has diverged from AI research in recent years because both fields pursue different goals. While modelers try to understand human behavior by creating systems that act as human-like as possible, AI researchers strive for systems that act as perfect as possible, or in Legg and Hutter (2007)'s words: *universal intelligence*. At the same time, parts of the cognitive modeling community are suggesting to direct the field towards more generic models of human behavior ("cognitive supermodels"; Salvucci, 2010) as opposed to task-specific models. Such supermodels did not come into existence yet, but given the methodological advances in the AI field, it may be worthwhile to think about how the abilities (i.e., intelligence) of generic cognitive models should be evaluated. A promising approach to this question are *anytime intelligence tests* (AIT; Hernández-Orallo & Dowe, 2010).

Anytime Intelligence Tests

These intelligence tests are crossing the boundaries between the modeling and the AI field because they are targeting both biological and artificial systems. Based on the work of Legg and Hutter (2007), they intend to measure intelligence of an agent (i.e., 'model' in the terms of the 'other' field) as the accumulated amount of reward[1] r it receives through interaction with a set of (deterministic) environments of varying (computational) complexity. The validity of this accumulated reward is achieved through several means: a) The reward is bound to the range [-1;1]; b) All environments must be balanced, i.e., a random agent will on average receive a reward of zero; and c) The aggregated reward is scaled by the computational complexity of the transition function of the environment. An example of such an intelligence test that was applied to both humans and AI agents can be found in Insa-Cabrera, Dowe, España-Cubillo, Hernández-Lloreda, and Hernández-Orallo (2011).

Besides the construction of new environments that follow the AIT formalism, one can try to analyze published data and models from the literature. This way, potential insights from the AIT procedure can be compared to fit-based evaluations that have been performed before. It is often possible to transform existing tasks into AIT environments by constructing new reward functions for the tasks.

Dynamic Stock and Flow Task

A promising candidate for such a reward reconstruction is the Dynamic Stock and Flow task (DSF; Dutt & Gonzalez, 2007) that has been used for the modeling challenge of the same name (Lebiere, Gonzalez, & Warwick, 2009). The task for this challenge was to maintain the level (i.e., stock) in a water tank at a given target value in the presence of dynamically changing water in- or outflow from an external source. There were four training conditions with monotonously changing inflow (Lin-, Lin+, NonL-, NonL+) and five transfer conditions. Two of these featured linearly increasing inflow (like Lin+), but the agents' actions were delayed by one (Del2) or two (Del3) additional time steps. The remaining conditions featured two (Seq2, Seq2Nos) or four (Seq4) time steps long repeating sequences of inflow; in case of Seq2Nos the pattern was masked by additional noise. The evaluation of the models in the competition was based on the goodness-of-fit to human data in all nine conditions. The crucial variable for this fit was the time-dependent water level.

Besides convenience (i.e., availability of data and models), the DSF task is especially suited as an AIT because it is deterministic and open-ended (in contrast to, e.g., robotic soccer), and the computational complexity of the environment should be both easily scalable and easily quantifiable. Whether and how this task can be transformed to an AIT will now be reviewed regarding possible reward functions for the task. The question of the complexity of the different task conditions will be discussed in a later contribution.

Possible Reward Functions

In the following, r_t denotes the reward for time step t, $amount_t$ denotes the water level for t, env_t the external inflow for t, and $goal$ denotes the target water level.

[1]Note: In reinforcement learning, reward functions are a crucial part of the learning agents themselves (Singh, Lewis, & Barto, 2009). In the context of this paper, rewards come from an external 'critic' and were not available to the agents (i.e., human participants and cognitive models) during exploration and learning.

Scaled Absolute Difference. For every time step, the absolute difference to the target water level is multiplied with some constant c and mapped to the range from -1 to 1.

$$r_t = \max(-1, 1 - c|amount_t - goal|)$$

This is the most straightforward solution with the highest face-validity. The agent's proximity to the target level is represented very well. On the other hand, c is arbitrary. Most problematic is that the function is not balanced. Because all task conditions feature external water inflow, random agents would receive an accumulated reward close to the lower boundary of -1.

Relative Progress. A balanced environment could be created by concentrating on the relative progress to the target level. The most simple option is a binary decision whether the water level has improved.

$$r_t = \begin{cases} 1 & \text{if } |amount_t - goal| \leq |amount_{t-1} - goal| \\ -1 & \text{otherwise} \end{cases}$$

This solution has the downside of the current water level being underrepresented. Getting from anywhere to the exact target level would be as good as getting an arbitrarily small amount closer to it. At the same time, environments with external in- or outflow would still not result in random agents receiving a reward of zero.

Relative Progress with Weighting. This can be solved using the following rationale: A perfect agent would always bring the water level to the goal amount in the next step. If this maps to a reward of 1 and no action maps to 0, then the agent should be awarded a reward that is proportional to the stock change made by the agent compared to the two extremes. The result is clipped at -1 in order to stick to the properties of AIT.

$$r_t = \max(-1, \frac{|amount_t - goal|}{|amount_{t-1} + env_t - goal|})$$

Boxplots of the human data collected for the DSF challenge recoded using the proposed reward functions together with the results achieved by a random agent ($flow \sim \mathrm{N}(0,5)$), a null agent ($flow = 0$), and an ACT-R model[2] (Halbrügge, 2010) are given in Figure 1. Of the three proposed reward functions, 'weighted relative progress' provides the best fit to the AIT requirement of balanced rewards.

Discussion and Conclusions

The accumulated reward for the human sample provides interesting evidence about the different difficulty of the nine task conditions. While the four monotonous conditions on the left

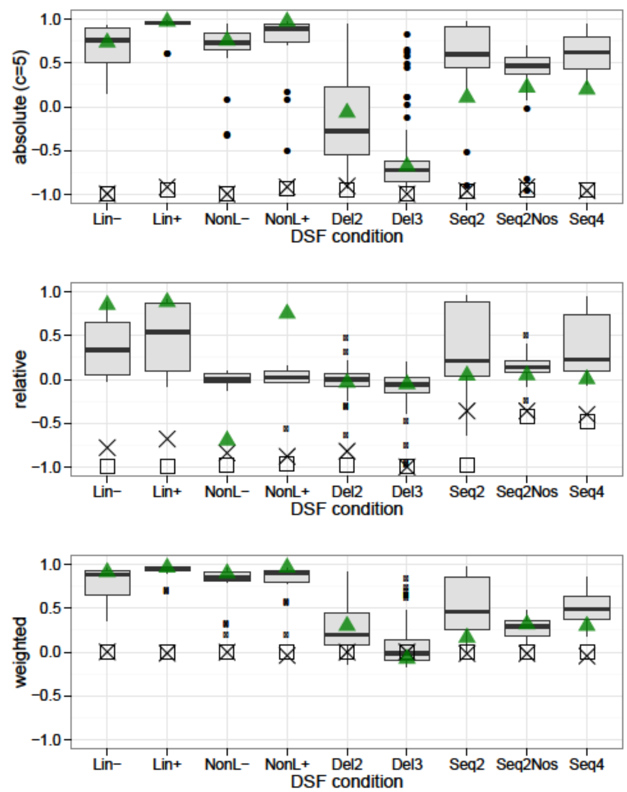

Figure 1: Average reward in the DSF task after recoding using the three proposed reward functions. Boxplots: Human Data. Squares: Null Agent (no action). Crosses: Random Agent. Triangles: ACT-R model (Halbrügge, 2010).

(Lin- to NonL+) are all comparatively easy, the delay conditions are very hard. In the Del3 condition, the median of the human sample is close to random performance. The difficulty of the sequence conditions lies between the monotonous and the delay conditions.

The performance of both humans and the model will be compared to the computational complexity of the respective environments. Then, the models that had entered the competition can be evaluated with respect to their intelligence as opposed to their fit to the human data.[3] Such an evaluation should consider the computational complexity of the models as well (Halbrügge, 2007). Complexity metrics based on the source code could be accompanied by Model Flexibility Analysis (Veksler, Myers, & Gluck, 2015), which tries to estimate the range of possible model behavior through simulation (see also Gluck, Stanley, Moore, Reitter, & Halbrügge, 2010). Together with the AIT formalism, this could lead to a new evaluation criterion that would be complementary to fit measures like R^2 and RMSE and could also provide a step towards reuniting the fields of cognitive modeling and artificial (general) intelligence.

[2] The source code of the cognitive model is available for download at http://dx.doi.org/10.14279/depositonce-5163

[3] Especially the delay conditions often lead to oscillating stock levels, which renders averaging across participants questionable.

References

Dutt, V., & Gonzalez, C. (2007). Slope of inflow impacts dynamic decision making. In *Proceedings of the 25th international conference of the system dynamics society* (p. 79). Boston, MA.

Gluck, K. A., Stanley, C. T., Moore, L. R., Reitter, D., & Halbrügge, M. (2010). Exploration for understanding in cognitive modeling. *Journal of Artificial General Intelligence*, *2*(2), 88–107. doi: 10.2478/v10229-011-0011-7

Halbrügge, M. (2007). Evaluating cognitive models and architectures. In G. A. Kaminka & C. R. Burghart (Eds.), *Evaluating architectures for intelligence. papers from the 2007 aaai workshop* (p. 27-31). Menlo Park, California: AAAI Press.

Halbrügge, M. (2010). Keep it simple – a case study of model development in the context of the dynamic stocks and flows (dsf) task. *Journal of Artificial General Intelligence*, *2*(2), 38–51. doi: 10.2478/v10229-011-0008-2

Hernández-Orallo, J., & Dowe, D. L. (2010). Measuring universal intelligence: Towards an anytime intelligence test. *Artificial Intelligence*, *174*(18), 1508–1539. doi: 10.1016/j.artint.2010.09.006

Insa-Cabrera, J., Dowe, D. L., España-Cubillo, S., Hernández-Lloreda, M. V., & Hernández-Orallo, J. (2011). Comparing humans and ai agents. In *Artificial general intelligence* (pp. 122–132). Springer. doi: 10.1007/978-3-642-22887-2_13

Lebiere, C., Gonzalez, C., & Warwick, W. (2009). A comparative approach to understanding general intelligence: Predicting cognitive performance in an open-ended dynamic task. In *Proceedings of the second conference on artificial general intelligence* (pp. 103–107). Amsterdam: Atlantis Press.

Legg, S., & Hutter, M. (2007). Universal intelligence: A definition of machine intelligence. *Minds and Machines*, *17*(4), 391–444. doi: 10.1007/s11023-007-9079-x

Salvucci, D. D. (2010). *Cognitive supermodels.* (Invited talk, European ACT-R Workshop, Groningen, The Netherlands)

Singh, S., Lewis, R. L., & Barto, A. G. (2009). Where do rewards come from? In *Proceedings of the annual conference of the cognitive science society* (pp. 2601–2606).

Veksler, V. D., Myers, C. W., & Gluck, K. A. (2015). Model flexibility analysis. *Psychological Review*, *122*, 755-769. doi: 10.1037/a0039657