Modelling of Personality in Agents: From Psychology to Logical Formalisation and Implementation

(Extended Abstract)

Sebastian Ahrndt, Johannes Fähndrich, Marco Lützenberger, and Sahin Albayrak
DAI-Labor, Technische Universität Berlin
Faculty of Electrical Engineering and Computer Science
Berlin, Germany
sebastian.ahrndt@dai-labor.de

ABSTRACT

This paper aims to motivate the impact of personality on essential elements of the behaviour of agents (e.g. decision-making processes, emotions, moods, or coping strategies). We show that available works on agent behaviour and works 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. We argue that such a connection requires a formalisation that specifies the concept of personality and the concept of decision-making processes jointly.

Categories and Subject Descriptors

H.1.2 [**Information Systems**]: User/Machine Systems—*Human factors,Software psychology*

General Terms

Human Factors, Algorithms, Theory

Keywords

Logics for agents and multi-agent systems, Belief-Desire-Intention theories and models

1. MOTIVATION AND PROBLEM

Over the last years, the agent community presented several approaches to bring emotions to artificial agents. Available solutions reach from modelling and applying emotions to (completely axiomatised) logics of emotions. Yet, when looking into a similar branch of behaviour-engineering, that is to say those works that aim to conceptualise an agent's personality, it becomes obvious that there is actually little connection between personality models and models for emotions (except works that build these connection discussing architectural considerations from the software engineering perspective). The lack of personality concepts in works on emotions is striking and surprising at the same time—Is it not that our personality affects our emotions and determines our entire behaviour? Ozer and Benet-Martínez [9] argue that personality is indeed a significant factor for human behaviour and determines the individual outcome of essential behavioural processes, e.g. cognition and emotional reactions.

Furthering this opinion, we present our approach to integrate the impact of personality into an agent's decision-making processes and outline how such integration can be formalised and implemented. Thus, introducing the first step towards the integration of personality and emotions in agents. Using this implementation we were able to show that different personalities indeed cause variations in the interpretation of inputs, the decision-making process, and the generation of outputs.

Appears in: *Proceedings of the 14th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2015), Bordini, Elkind, Weiss, Yolum (eds.), May 4–8, 2015, Istanbul, Turkey.*

2. PERSONALITY

Human factor psychology describes a human's personality by means of traits or types. What these approaches have in common is that traits or types are characteristic features of human beings, and that the human's behaviour and motives can be explained along these behavioural patterns. Today, there are two well-established theories about human personality, namely: the *Five-Factor Model* of personality [7] and the *Myers-Briggs Type Indicator* [8]. We already discussed [1] the differences between both approaches and showed [1] that psychologists commonly favour the former as a conceptual framework for describing personality.

3. PERSONALITY AND LOGIC

Although there are several works on personality in agent-based system (*cf.* [10]), we are not aware of any approaches that formalise this concept. The development of such formalism, however, is motivated by an on-going discussion between psychologists, about the existence and definition of personality traits. Yet, this discussion all too often leads to subjective explanations of fundamental terms. The agent-community, on the other hand, requires clear definitions and semantics in order to develop executable models. For this reason we decided to include the concept of personality into an established and formalised agent-behaviour concept, namely *Belief Desire Intention logics*, (or short, *BDI logics* see e.g. [12]).

Our 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 [12, pp. 69]. We minimally extend the syntax and semantics of this logic by a new modal connectivity, representing the personality of

an agent and thus establish the foundation for a complete formalisation of concrete personality traits, which we aim to develop in the future.

In order to define the semantic of the personality, we apply the same argumentation as Wooldridge [12, pp. 74] provided when introducing the Bel, Des, Int modalities. Consequently, the personality of each agent is given by the personality modality Per, which, for itself, is characterised using the function $\mathcal{P}$, which is defined as follows:

$$\mathcal{P} : D_{Ag} \rightarrow \wp(W \times T \times W). \quad (1)$$

The function $\mathcal{P}$ is also referred to as 'personality accessibility relation'. In more detail, the function can be used to determine the set of worlds, which is accessible for an agent i in a specific situation $\langle w, t\rangle$, where $w \in W$ and $t \in T$. We define this as follows:

$$\mathcal{P}_t^w(i) = \{w' | \langle w, t, w'\rangle \in \mathcal{P}(i)\} \quad (2)$$

For the semantics of state formulae in $\mathcal{LORA}$ this implies a new rule, which is defined by $\langle M, V, w, t\rangle \models_{\mathcal{S}} (Per\ i\ \varphi)$ iff $\forall w' \in W$, if $w' \in \mathcal{P}_t^w([\![i]\!])$, then $\langle M, V, w', t\rangle \models_{\mathcal{S}} \varphi$. As well as 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. In more detail, 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'' [12, p. 74]. Formally specified, this means that $w' \in P_t^w(i)$ implies $t \in w$ and $t \in w'$. Moreover, it is ensured that the personality modality has a logic that corresponds to the normal modal system KD [3]. 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.

4. PERSONALITY AND ALGORITHMS

We implemented such a model by means of *AntMe!*[1], an agent-based simulation framework, which provides a completely adaptable test-bed for behavioural studies. Despite the fact that we simulated ants, we were able to show that personality affects all relevant phases of the decision-making processes. To accomplish this, we extended the different phases of the life-cycle of BDI agents by integrating personality as influential characteristic. As an example, 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. To substantiate this interpretation, we used works of different authors that investigate the relation between personalities and behaviour types (cf. [5, 11]). We also used results from experiments that examined the impact of personalities on specific stages of the decision cycle (e.g. effects on coping strategies [4] or effects on information processing [2]). A detailed discussion of the influences using the example of a naive BDI algorithm can be found in prior work [1].

Taking the experimental results into account we can state that the parameters we added to the BDI life-cycle can be interpreted as personality traits and the resulting behavioural change of the agents can be interpreted as personality. In addition, we were able to show that different personality traits affect the result of the simulation and that certain personalities are better suited for particular tasks than others. This extends available work [6] on this topic to the complete set of personality traits available through the Five-Factor Model. We also demonstrated that such parameters can influence the behaviour of agents in a domain independent way and that one subject to research is the task-dependent interpretation of the effect of a personality. Finally, our experiment confirmed the findings of Salvit and Sklar [10] with respect to the Five-Factor Model. Namely, that the interpretation of the parameters as personality traits results in (personality-)consistent behaviour of agents.

[1] Further information: `http://www.antme.net/`.

REFERENCES

[1] S. Ahrndt, A. Aria, J. Fähndrich, and S. Albayrak. Ants in the OCEAN: Modulating agents with personality for planning with humans. In N. Bulling, editor, *Multi-Agent Systems (EUMAS 2014)*, Lecture Notes in Artificial Intelligence, pages 1–16. Springer, 2014.

[2] A. Baumert and M. Schmitt. Personality and information processing. *Eur J Per*, 26(2):87–89, 2012.

[3] P. Blackburn, J. van Benthem, and F. Wolter, editors. *Handbook of Modal Logic*, volume 3 of *Studies in Logic and Practical Reasoning*. Elsevier Science, 2006.

[4] J. K. Connor-Smith and C. Flachsbart. Relations between personality and coping: A meta-analysis. *J Pers Soc Psychol*, 93(6):1080–1107, 2007.

[5] H. Du and M. N. Huhns. Determining the effect of personality types on human-agent interactions. In *Web Intelligence (WI) and Intelligent Agent Technologies (IAT), 2013 IEEE/WIC/ACM Int. Joint Conf. on*, volume 2, pages 239–244. IEEE, 2013.

[6] F. Durupinar, J. Allbeck, N. Pelechano, and N. Badler. Creating crowd variation with the OCEAN personality model. In Padgham, Parkes, Müller, and Parsons, editors, *Proc. of the 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, pages 1217–1220. IFAAMS, 2008.

[7] R. R. McCrea and O. P. John. An introduction to the five-factor model and its applications. *J Pers*, 60(2):175–215, 1992.

[8] I. B. Myers and P. B. Byers. *Gifts Differing: Understanding Personality Type*. Nicholas Brealey Publishing, 2 edition, 1995.

[9] D. J. Ozer and V. Benet-Martínez. Personality and the prediction of consequential outcomes. *Annu Rev Psychol*, 57:401–421, 2006.

[10] J. Salvit and E. Sklar. Modulating agent behavior using human personality type. In *Proc. of the Workshop on Human-Agent Interaction Design and Models (HAIDM) at Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 145–160, 2012.

[11] A. M. von der Pütten, N. C. Krämer, and J. Gratch. How our personality shapes our interactions with virtual characters — implications for research and development. In J. Allbeck, N. Badler, T. Bickmore, C. Pelachaud, and A. Safonova, editors, *Intelligent Virtual Agents*, volume 6356 of *Lecture Notes in Computer Science*, pages 208–221. Springer, 2010.

[12] M. Wooldridge. *Reasoning about Rational Agents*. Intelligent Robotics and Autonomous Agents. The MIT Press, July 2000.