# Incorporating BDI agents into human-agent decision making research*

Bart Kamphorst Arlette van Wissen Virginia Dignum

*Institute of Information and Computing Sciences*
*Utrecht University, the Netherlands*

## 1 Introduction

Decision making has since long been an area of interest to scholars from all kinds of disciplines: psychology, sociology, economics and more recently, computer science. A lot of research focuses on finding, isolating and formalizing the factors that are involved in decision making processes of both humans and software agents. The Colored Trails (CT) framework [2] is designed to aid researchers in this purpose. CT is a testbed developed (i) to study interaction between multiple actors (humans or software agents) in a dynamic environment and (ii) to study and model both human and agent decision making. In the current implementation of CT, agents lack the explanatory power to help understand the reasoning processes involved in decision making. In order to gain more insights into the actual reasoning processes that lie behind a decision of a software agent, the agents must be endowed with a richer model of reasoning. Agents can be constructed to reason with abstract concepts such as beliefs, goals, plans and events. These types of agents are often referred to as Belief, Desire and Intention (BDI) agents. 2APL (pronounced double-a-p-l) is a practical agent programming language designed to implement BDI agents [1]. The 2APL platform, used to evaluate 2APL agents in agent-agent scenarios, is however not very suitable for human-agent interaction. This paper presents middleware called CTAPL that combines the strengths of both the 2APL platform and the CT framework by letting BDI agents that are written in 2APL interact with humans and other software agents in CT. CTAPL enables researchers to study interaction between software agents and humans in a broad range of domains and to take advantage of the explanatory power the BDI approach offers.

## 2 CTAPL

CTAPL is a platform designed for the implementation of various interaction scenarios between BDI agents, algorithmic agents, humans and heterogeneous groups. CTAPL extends the existing CT framework by making it possible to incorporate BDI based software agents written in the practical programming language 2APL into any CT scenario. This is beneficial for researchers in the field of decision making because BDI models allow clear functional decompositions with clear and retractable reasoning patterns. Additionally, BDI agents use 'mental attitudes' such as beliefs and intentions, resembling the kind of reasoning that humans use in our everyday lives. Taken together, BDI agents provide more helpful feedback and more explanatory power than agents lacking a BDI structure. Furthermore, the BDI approach has proved valuable for the design of agents that operate in dynamic environments. It offers a higher level of abstraction by explicitly allowing beliefs to have a direct impact upon the agents behavior. This ensures that the agents can respond flexibly to changing circumstances despite incomplete information about the state of the world and the agents in it.

Figure 1 shows the conceptual design of CTAPL. The top layer represents the 2APL platform. It consists of a server and one or more BDI agents ($A_1 \ldots A_i$). The bottom layer represents the CT framework, with software agents and human actors. $GA_1 \ldots GA_i$ are hooks in CT that allow the BDI agents to communicate with the CT environment. Each agent $A_n$ thus corresponds with hook $GA_n$. The 2APL platform is extended

*For the full paper, please refer to [3].

by an external environment that instantiates (i) the hooks for each 2APL agent and (ii) a Java Thread that continually listens whether agents have received any new messages from the server.

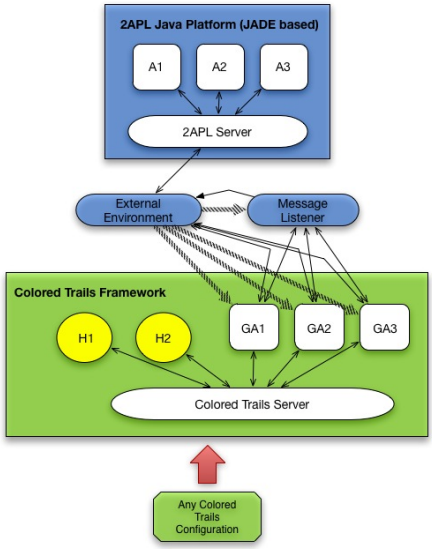


Figure 1: The conceptual design of CTAPL

# 3 Conclusions and Future Work

Typically in BDI models, concepts such as 'utility' and 'uncertainty' are not included. Future work will focus on deploying BDI agents in CT scenarios in which agents have to deal with such concepts. Future research with CTAPL will include (i) building BDI agents that model human decision making processes in team formation with self-interested agents and (ii) improving the planning mechanism of agents in a collaborative setting with uncertainty.

The authors have proposed a technical solution for dealing with the explanatory gap that exists when algorithmic CT agents are used to investigate decision making. We have argued that BDI based agents can assist in filling the gap because they use clear and retractable reasoning patterns. This paper has described new middleware called CTAPL that is designed to combine the strengths of a BDI based agent approach with the CT testbed for decision making. CTAPL makes three major contributions. First, CTAPL lets BDI researchers explore existing research domains developed in CT for agent-agent interaction. Secondly, it gives BDI researchers the opportunity to have BDI agents interact with human players and to use observations from these interactions to improve the agent models. Lastly, it makes it possible for CT researchers to write agents that can qualitatively reason in terms of beliefs, goals and plans by using the 2APL agent programming language.

***Acknowledgements.*** We thank Ya'akov (Kobi) Gal and Maarten Engelen for helpful comments and assistance with the initial setup of CTAPL. This research is funded by the Netherlands Organization for Scientic Research (NWO), through Veni-grant 639.021.509.

# References

[1] M. Dastani. 2apl: a practical agent programming language. *Autonomous agents and multi-agent systems*, 16(3):214–248, 2008.

[2] B. Grosz, S. Kraus, S. Talman, B. Stossel, and M. Havlin. The influence of social dependencies on decision-making: Initial investigations with a new game. *AAMAS*, 2004.

[3] B. Kamphorst, A. van Wissen, and V. Dignum. Incorporating bdi agents into human-agent decision making research (forthcoming). *Tenth International Workshop "Engineering Societies in the Agents' World" (ESAW)*, 2009.