

Instructing BDI Agents with Natural Language

January 4, 2026

Research Question

Can LLMs be integrated with the Beliefs-Desires-Intentions paradigm to provide an alternative to programmer-crafted plans?

Background

In [4] Ichida et al create a Beliefs-Desires-Intentions (BDI) framework (NatBDI) that can 1) be instructed on how to act given plans in a structured form of natural language leveraging a pre-trained large language model and 2) used a fallback policy based on deep reinforcement learning in the event that no human instruction exists. This was evaluated on the ScienceWorld benchmark set [5] on the *melt* and *find-non-living-object* tasks.

Approach

The object of this project would be to reproduce this work but instead of developing a BDI framework from scratch (as in the paper), to integrate the proposal within an existing framework such as Jason [1], Gwendolen [2] or SimpleBDI (described in [3]). The system should be evaluated against additional tasks in the ScienceWorld Benchmark set. If time is available at this point, the work can be extended to allow an additional fallback policy of questioning a user for more input to implement a form of one shot learning.

Milestones

1. Assessment of candidate BDI frameworks to identify the one most amenable to extension in the style of NatBDI.

2. Extension of this framework to enable the representation of plans and beliefs in natural language and reasoning about those representations.
3. Evaluation of this system against tasks in ScienceWorld.
4. Design and implementation of extensions to this work, such as the ability to question a human about what to do as part of the fallback procedure and then update the plan database accordingly in order to implement a form of one-shot learning.

References

- [1] BORDINI, R. H., HÜBNER, J. F., AND WOOLDRIDGE, M. *Programming Multi-Agent Systems in AgentSpeak using Jason*. John Wiley & Sons, 2007.
- [2] DENNIS, L. A. Gwendolen semantics: 2017. Tech. Rep. ULCS-17-001, University of Liverpool, Department of Computer Science, 2017.
- [3] DENNIS, L. A., AND OREN, N. Explaining BDI agent behaviour through dialogue. *Journal of Autonomous Agents and Multi-Agent Systems* (2022).
- [4] ICHIDA, A. Y., MENEGUZZI, F., AND CARDOSO, R. C. BDI agents in natural language environments. In *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2024, Auckland, New Zealand, May 6-10, 2024* (2024), M. Dastani, J. S. Sichman, N. Alechina, and V. Dignum, Eds., International Foundation for Autonomous Agents and Multiagent Systems / ACM, pp. 880–888.
- [5] WANG, R., JANSEN, P., CÔTÉ, M.-A., AND AMMANABROLU, P. Scienceworld: Is your agent smarter than a 5th grader?, 2022.