BDI Agents in Natural Language Environments

Alexandre Yukio Ichida, Felipe Meneguzzi, Rafael C. Cardoso

✉ alexandre.yukio@edu.pucrs.br

✉ felipe.meneguzzi@abdn.ac.uk

✉ rafael.cardoso@abdn.ac.uk

Contribution

- We develop the seminal approach for an entire class of BDI-based agent architectures that use machine learning components to deal with natural language information named NatBDI.
- Combining a natural language interface and reasoning capabilities with the folk psychology abstraction of mental states in the BDI model [1] improves human understanding of the agent's handling of noisy information.
- We include a natural language plan library to allow humans to create plans to customise the agent's behaviour and helps avoid unintended conduct.
- We leverage Natural Language Inference (NLI) models to support logical inferences over belief base and plan contexts encoded in natural language.

Text-based Environment - ScienceWorld

- ScienceWorld is an interactive textual environment that simulates engines for thermodynamics, electrical circuits, matter and chemistry reactions, and biological processes at the level of a standard elementary school science curriculum [5].
- The ScienceWorld environment evaluates the agent's capacity to use declarative scientific knowledge to act or plan in order to solve tasks that humans can perform with ease (e.g., melting ice). These tasks cover topics such as the change of state (boiling, melting, freezing), taking measurements (thermometer, boiling point), classification (find a non-living thing, find a plant), etc.

Agent Architecture and Components

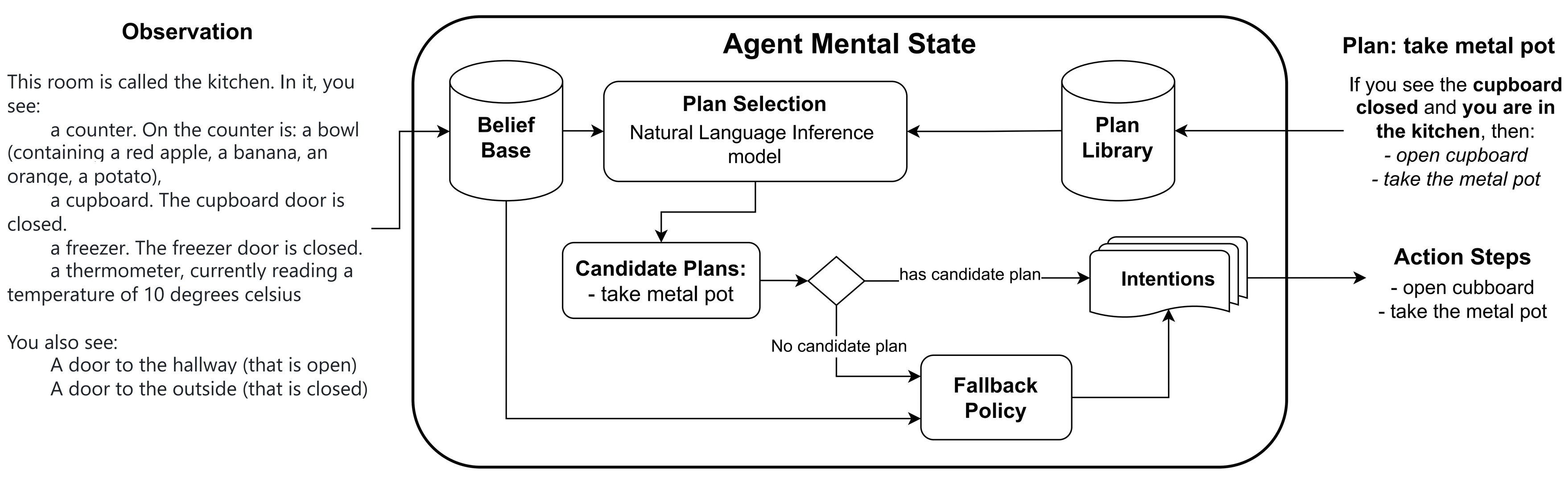

Figure 1: Diagram illustrating NatBDI to handle and actuate over natural language environments.

Natural Language Plans

- Plan rules inspired by AgentSpeak language [3] consisting in: Triggering event, plan context and body encoded as natural language sentences.

```
IF your task is to get the metal pot
CONSIDERING you are in the kitchen AND you see the cupboard closed
THEN:
open the cupboard,
take the metal pot

IF your task is to melt water
THEN:
PLAN TO get the metal pot
pick up thermometer
...

```

Listing 1: Plans in natural language to pick the metal pot and melt water in ScienceWorld.

Plan Inference Steps

1. Create a matrix M to evaluate all belief, context sentence pairs

$$M_{i,j} = \mathcal{C} \times \mathcal{B} = \{(c_i, b_j) \mid c \in \mathcal{C} \land b \in \mathcal{B}\} \quad (1)$$

2. Create an entailment matrix E with all NLI result obtained from M

$$E_{i,j} = \{nli(c_i, b_j) \mid (c_i, b_j) \in M_{i,j}\} \quad (2)$$

3. Check whether the plan context are entailed by the belief base ($\mathcal{B} \models \mathcal{C}$) using matrix E as follows:

$$\mathcal{B} \models \mathcal{C} \doteq \bigwedge_{c_i \in \mathcal{C}} c_i \bigvee_{b_j \in \mathcal{B}} b_j \quad (3)$$

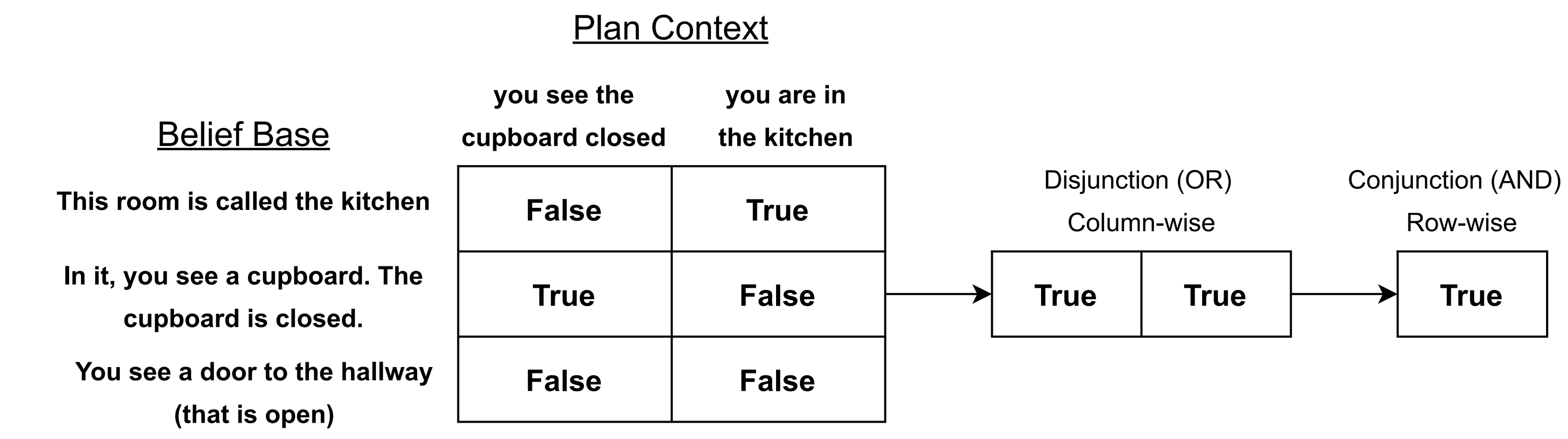

Figure 2: Diagram illustrating an example of inference steps.

Fallback Policy

- If there is no plan candidate, the agent triggers the fallback policy to select an action without human intervention.
- The fallback policy is a function $f : \mathbb{R}^{|\mathcal{B}|} \to \mathbb{R}^{a}$ that maps the belief base and results into a plan. We model the fallback policy as a POMDP [4].
- We use the Deep Relevance Reinforcement Network (DRRN) [2], which shows best results according to [5].

Results

- We perform experiments of the NatBDI in two tasks in the ScienceWorld environment with different plan library sizes, described in Table 1, and compare the use of different NLI model types in Table 2.
- In Figure 3, we show the trade-off between including plan-rules and increasing the number of episodes in fallback policy training phase.

Table 1: Average scores and number of actions performed in each phase out of all task variations.

Task	Variations	Episodes	Number plan-rules	Score (Total)	Score (BDI)	Score (DRRN)	Number BDI actions	Number RL actions
find-non-living-thing	75	242	0	0.66	0.00	**0.66**	0.00	50.00
			8	0.75	0.30	**0.45**	3.33	38.00
			15	0.84	**0.58**	0.26	6.25	24.00
			23	0.91	**0.79**	0.12	7.64	13.33
			30	0.98	**0.98**	0.00	9.19	4.00
melt	9	457	0	0.03	0.00	**0.03**	0.00	50.00
			4	0.14	**0.11**	0.03	5.11	44.44
			7	0.36	**0.34**	0.02	10.89	33.33
			10	0.57	**0.56**	0.01	17.11	22.22
			13	0.67	**0.67**	0.00	20.89	16.67

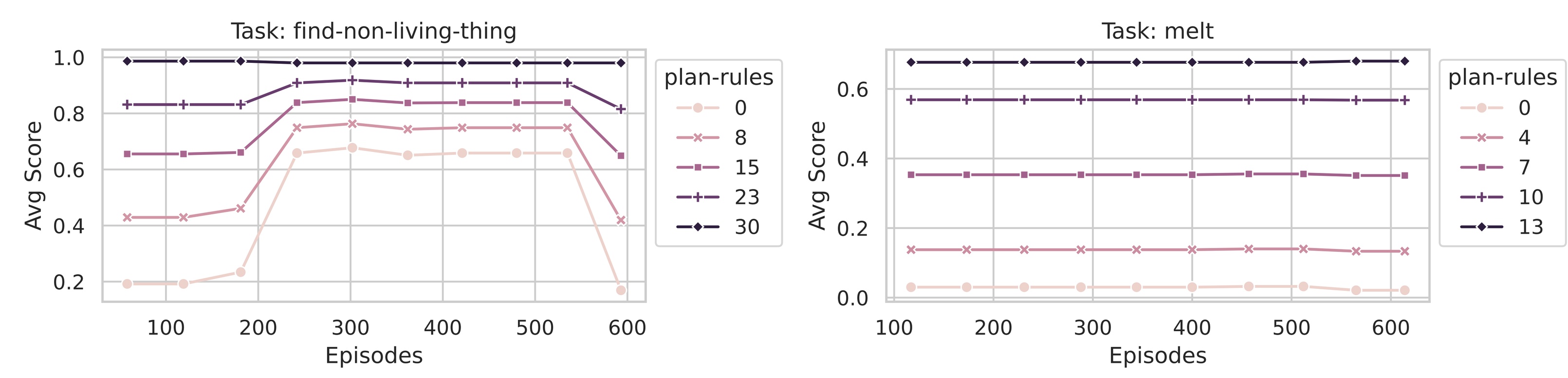

Figure 3: Scores per episodes when scaling the number of plan-rules.

Table 2: Results of using different LLMs for NLI. The following columns describe them: model size (*Params*); accuracy on *MNLI*-m test set; score obtained using NatBDI; average number of *actions* performed, *errors* raised and plan-rules (*Plans*) used; lexical overlap computed on entailment pairs (*LO(E)*).

Model	Params	MNLI-m	Task	Score	Actions	Errors	Plans	LO(E)
MiniLM (L6)	22M	82.2	find-non-living-thing	0.69	7.65	0.37	2.57	0.64
			melt	0.23	8.78	1.00	3.59	1.34
Bert (base)	110M	84.6	find-non-living-thing	0.84	9.61	0.20	2.72	0.60
			melt	0.33	11.44	0.67	3.56	1.16
Roberta (large)	355M	90.8	find-non-living-thing	**0.98**	9.19	0.08	2.84	0.40
			melt	**0.67**	20.89	0.33	5.67	1.21

Conclusion and Future Work

- This work introduces a new class of agent architecture that uses the BDI reasoning cycle with components driven by natural language processing.
- We show that manually designed plan-rules in natural language can substantially improve performance of NatBDI in ScienceWorld.
- Future work lies in learning plan-rules from data to allow humans to co-design an agent's plan library in an efficient and transparent way.

References

[1] Michael E Bratman. Two faces of intention. *Philosophical Review*, 1984.

[2] Ji He, Jianshu Chen, Xiaodong He, Jianfeng Gao, Lihong Li, Li Deng, and Mari Ostendorf. Deep reinforcement learning with a natural language action space. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*.

[3] Anand S Rao. AgentSpeak(L): BDI agents speak out in a logical computable language. In *Proceedings of the 7th European Workshop on Modelling Autonomous Agents in a Multi-agent World*, 1996.

[4] Martijn Schut, Michael Wooldridge, and Simon Parsons. On partially observable mdps and bdi models. In *Foundations and Applications of Multi-Agent Systems*. Springer, 2002.

[5] Ruoyao Wang, Peter Jansen, Marc-Alexandre Côté, and Prithviraj Ammanabrolu. Scienceworld: Is your agent smarter than a 5th grader? *arXiv*, (arXiv:2203.07540 [cs.CL]), 2022.