Engineering Norm-aware BDI Agents

Extended Abstract

Michael Winikoff
Victoria University of Wellington
Wellington, New Zealand
michael.winikoff@vuw.ac.nz

Frank Dignum
Umeå University
Umeå, Sweden
dignum@cs.umu.se

Sebastian Rodriguez
RMIT University
Melbourne, Australia
sebastian.rodriguez@rmit.edu.au

John Thangarajah
RMIT University
Melbourne, Australia
john.thangarajah@rmit.edu.au

ABSTRACT

Norms are an important abstraction for multi-agent systems (MAS) because they specify *flexible constraints*. As constraints they help bring order to agent behaviour. However, being flexible, they can be violated, and thus do not over-constrain autonomous entities. In order to build MAS with cognitive agents that use norms we need to be able to engineer these agents to be *norm-aware*, i.e. being able to take norms into account in their reasoning. This paper defines an approach to engineering norm-aware BDI (Belief-Desire-Intention) agents that can deal with cases including adding steps in order to comply with norms and that can be used with off-the-shelf BDI agent languages.

KEYWORDS

Norms; Norm-aware agents; Goal Plan Tree; Belief-Desire-Intention

ACM Reference Format:
Michael Winikoff, Frank Dignum, Sebastian Rodriguez, and John Thangarajah. 2026. Engineering Norm-aware BDI Agents: Extended Abstract. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 3 pages. https://doi.org/10.65109/SEJS3352

1 INTRODUCTION

Norms are important for MAS [1, 3, 6, 10, 12, 13] because they specify *flexible constraints* that are usually obeyed, simplifying interactions, but that can also be violated. Agents therefore need to be able to reason about whether to violate a norm and when they should adapt their behaviour to comply with it.

Running example: we have a goal to go to work, and can do this in four different ways: walking (if the distance is small), using a ride sharing service, driving ourselves (if we have access to a car), or using public transport (“PT”, which involves the steps: go to the station, board the train, wait until one arrives at the destination, and then disembark). The use of public transport is subject to a norm that passengers must have a valid ticket when using public transport, or each be subject to a fine. In the ADICO format [5] this could be written as: **A**ttributes: a person who is a passenger of public transport, **D**eontic: must, a**I**m: have a valid ticket, **C**ondition: when riding public transport, **O**r else: pay a fine.

In this context the sorts of reasoning that we want our norm-aware BDI agent to be able to do, if it chooses the PT option, include: (1) adapting its plan to add buying a ticket (if it decides to comply with the norm); (2) deciding to violate the norm *knowingly* by taking account of the consequences of doing so (the possible fine); and (3) when reasoning about what course of action to choose, take account of the consequences of violating the norm and of the additional action(s) required to conform with the norm.

This paper proposes an approach for engineering norm-aware BDI agents (for full details see [15]). The core idea of our approach is to transform each norm into a goal-plan tree in a way that allows the agent to be norm-aware without requiring the BDI platform to be extended. Although there has been prior work on norm-aware agents, none of the approaches (see Table 1) cover the three cases and are usable with existing (unmodified) BDI platforms.

Paper(s)	Add Steps	Violate Norm	Select Plan	off-the-shelf BDI?
[12]	✘	✘	~	✘
N-2APL [1]	✘	✔	✘	✘
N-Jason [10]	✘	✔	✔	✘
NBDI [8, 9]	✘	✔	✔	✘
ν-BDI [11]	✘	✔	✔	✘
JaCaMo [2]	✘	✔	✘	✘
Soc. Comm. [4, 14]	✘	✘	✔	✔
This paper	✔	✔	✔	✔

Table 1: Prior work: support for the three cases and platform.

2 ADDING NORMS TO BDI AGENTS

Our overall process is to:

(1) translate each norm N_i into a corresponding goal with Goal-Plan Tree (GPT) G_{N_i}, following a pattern (Figure 2);
(2) augment the original GPT G with links to a norm from the point where additional steps to comply with the norm might be required;
(3) augment G and G_{N_i} with *valuings*, which capture key attributes of the various options, including the consequences of violating norms, and of any additional steps required to conform with norms.

Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), C. Amato, L. Dennis, V. Mascardi, J. Thangarajah (eds.), May 25 – 29, 2026, Paphos, Cyprus. © 2026 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org). https://doi.org/10.65109/SEJS3352

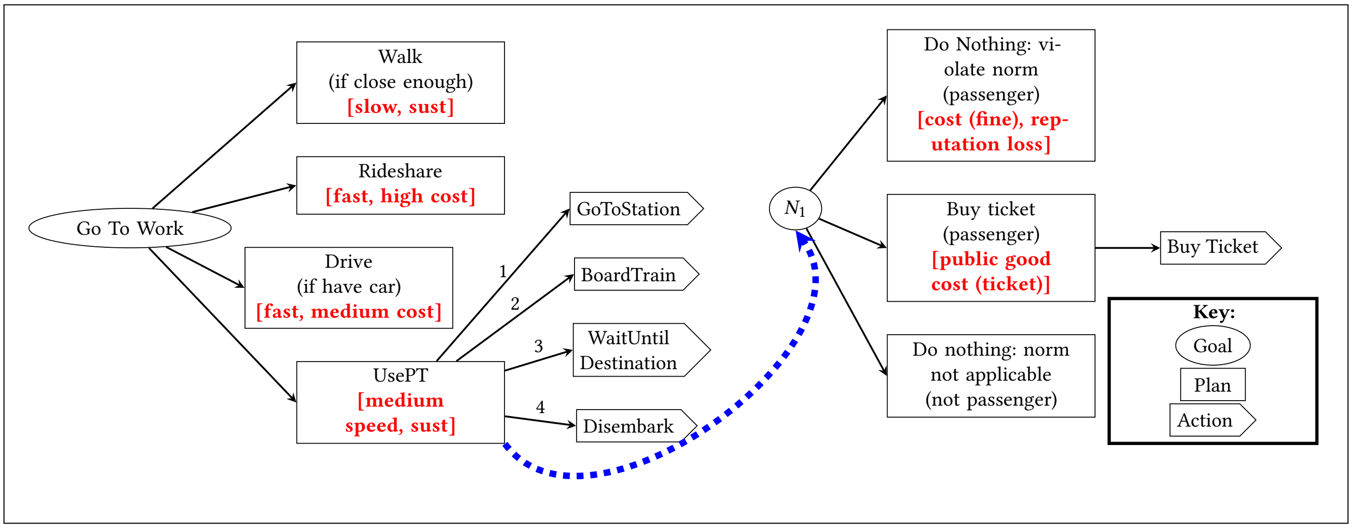

Figure 1: GPT for the example: G (left, black only), the link (blue dashed), G^V (include red bold text), G_{N_1} (right, black only), $G^V_{N_1}$ (include red bold text). Notation: each goal has multiple plans that achieve it ("or" decomposition), numbers on edges denote order, "PT" = "Public Transport", conditions are in brackets, red bold text in square brackets = valuing.

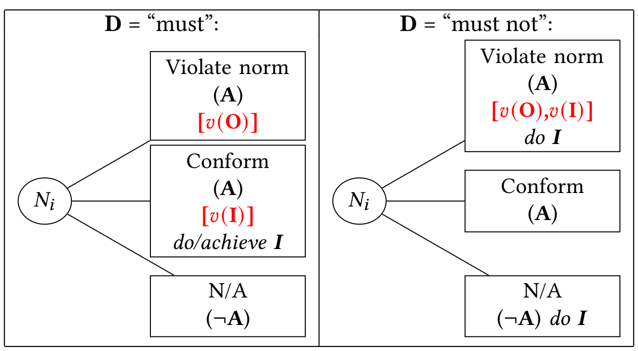

Figure 2: Mapping Norms to GPTs. Notation: letters A,D,I,C,O correspond to those components of the norm, plan context condition is in brackets, valuings are red bold in square brackets, plan body is italic, "do/achieve" = "do" if I is an action, "achieve" if it is a condition, $v(\mathbf{I})$ is the valuing of doing/achieving I, and $v(\mathbf{O})$ is the valuing of O.

Figure 1 shows the result for our running example. The agent then uses meta-reasoning—taking into account the valuings—to decide which plan to select.

3 IMPLEMENTATION & EVALUATION

Our running example has been fully implemented in Jason. This involves: using Jason meta-events to implement the link between G and G_{N_i}, translating the valuings into beliefs, defining the agent's preferences over options, and adding meta-reasoning to decide between options taking into account its preferences over the consequences. Where we have used features that are particular to Jason, we also have an alternative approach that is generic (for full details see [15]).

We evaluated our work by demonstrating, using our example:

(1) that our approach handles the three cases highlighted in the introduction: choosing to conform with a norm, requiring plan modification; knowingly violating a norm; and changing the choice of plan.
(2) how our framework operates at the right level of abstraction by showing that it is easy to accommodate changes, such as: adjusting norm applicability (e.g. young children do not require a ticket), adding new norms, providing different ways of satisfying the norm condition, and dealing with plan failure.

4 DISCUSSION

We presented an approach for engineering norm-aware BDI agents that can be used with existing AOPLs without requiring them to be modified. Our approach includes a simple process for translating norms into goal-plan trees, and a method for implementing these in Jason.

Future work includes: automating the generation of code; automating the identification of where links are required; developing a larger norm-aware BDI system to provide a basis for assessing the run-time overhead, possibly building on the agent-based social simulation of the COVID crisis [7]; adding dynamic norms using runtime plan addition; and exploring using norms in explanations (e.g. "I chose to walk in order to avoid a fine for using public transport without a ticket").

ACKNOWLEDGMENTS

Michael Winikoff would like to thank the University of Ljubljana for hosting him on sabbatical while this paper was written.

REFERENCES

[1] Natasha Alechina, Mehdi Dastani, and Brian Logan. 2012. Programming norm-aware agents. In *International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2012, Valencia, Spain, June 4-8, 2012 (3 Volumes)*, Wiebe van der Hoek, Lin Padgham, Vincent Conitzer, and Michael Winikoff (Eds.). IFAAMAS, 1057–1064. http://dl.acm.org/citation.cfm?id=2343848

[2] Olivier Boissier, Rafael H. Bordini, Jomi Fred Hübner, Alessandro Ricci, and Andrea Santi. 2013. Multi-agent oriented programming with JaCaMo. *Sci. Comput. Program.* 78, 6 (2013), 747–761. https://doi.org/10.1016/J.SCICO.2011.10.004

[3] Cristiano Castelfranchi, Frank Dignum, Catholijn M. Jonker, and Jan Treur. 2000. Deliberative Normative Agents: Principles and Architecture. In *Intelligent Agents VI LNAI 1757*, Nicholas R. Jennings and Yves Lespérance (Eds.). Springer, 364–378.

[4] Amit K. Chopra, Matteo Baldoni, Samuel Christie, and Munindar P. Singh. 2025. Azorus: Commitments over Protocols for BDI Agents. In *Proceedings of the 24th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2025, Detroit, MI, USA, May 19-23, 2025*, Sanmay Das, Ann Nowé, and Yevgeniy Vorobeychik (Eds.). International Foundation for Autonomous Agents and Multiagent Systems / ACM, 490–499.

[5] Sue ES Crawford and Elinor Ostrom. 1995. A grammar of institutions. *American political science review* (1995), 582–600.

[6] Frank Dignum. 1999. Autonomous agents with norms. *Artificial Intelligence and Law* 7, 1 (March 1999), 69–79.

[7] Frank Dignum (Ed.). 2021. *Social Simulation for a Crisis: Results and Lessons from Simulating the COVID-19 Crisis.* Springer. https://doi.org/10.1007/978-3-030-76397-8

[8] Baldoino Fonseca dos Santos Neto, Viviane Torres da Silva, and Carlos José Pereira de Lucena. 2010. Using Jason to Develop Normative Agents. In *Advances in Artificial Intelligence - SBIA 2010 - 20th Brazilian Symposium on Artificial Intelligence, São Bernardo do Campo, Brazil, October 23-28, 2010. Proceedings (Lecture Notes in Computer Science, Vol. 6404)*, Antônio Carlos da Rocha Costa, Rosa Maria Vicari, and Flavio Tonidandel (Eds.). Springer, 143–152. https://doi.org/10.1007/978-3-642-16138-4_15

[9] Baldoino Fonseca dos Santos Neto, Viviane Torres da Silva, and Carlos José Pereira de Lucena. 2011. Developing Goal-Oriented Normative Agents: The NBDI Architecture. In *Agents and Artificial Intelligence - Third International Conference, ICAART 2011, Rome, Italy, January, 28-30, 2011. Revised Selected Papers (Communications in Computer and Information Science, Vol. 271)*, Joaquim Filipe and Ana L. N. Fred (Eds.). Springer, 176–191. https://doi.org/10.1007/978-3-642-29966-7_12

[10] Jeehang Lee, Julian A. Padget, Brian Logan, Daniela Dybalova, and Natasha Alechina. 2014. N-Jason: Run-Time Norm Compliance in AgentSpeak(L). In *Engineering Multi-Agent Systems - Second International Workshop, EMAS 2014, Paris, France, May 5-6, 2014, Revised Selected Papers (Lecture Notes in Computer Science, Vol. 8758)*, Fabiano Dalpiaz, Jürgen Dix, and M. Birna van Riemsdijk (Eds.). Springer, 367–387. https://doi.org/10.1007/978-3-319-14484-9_19

[11] Felipe Meneguzzi, Odinaldo Rodrigues, Nir Oren, Wamberto Weber Vasconcelos, and Michael Luck. 2015. BDI reasoning with normative considerations. *Eng. Appl. Artif. Intell.* 43 (2015), 127–146. https://doi.org/10.1016/J.ENGAPPAI.2015.04.011

[12] Felipe Rech Meneguzzi and Michael Luck. 2009. Norm-based behaviour modification in BDI agents. In *8th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2009), Budapest, Hungary, May 10-15, 2009, Volume 1*, Carles Sierra, Cristiano Castelfranchi, Keith S. Decker, and Jaime Simão Sichman (Eds.). IFAAMAS, 177–184. https://dl.acm.org/citation.cfm?id=1558037

[13] Sofia Panagiotidi, Javier Vázquez-Salceda, and Frank Dignum. 2012. Reasoning over Norm Compliance via Planning. In *Coordination, Organizations, Institutions, and Norms in Agent Systems VIII - 14th International Workshop, COIN 2012, Held Co-located with AAMAS 2012, Valencia, Spain, June 5, 2012, Revised Selected Papers (Lecture Notes in Computer Science, Vol. 7756)*, Huib Aldewereld and Jaime Simão Sichman (Eds.). Springer, 35–52. https://doi.org/10.1007/978-3-642-37756-3_3

[14] Michael Winikoff. 2007. Implementing commitment-based interactions. In *6th International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS 2007), Honolulu, Hawaii, USA, May 14-18, 2007*, Edmund H. Durfee, Makoto Yokoo, Michael N. Huhns, and Onn Shehory (Eds.). IFAAMAS, 128. https://doi.org/10.1145/1329125.1329283

[15] Michael Winikoff, Frank Dignum, Sebastian Rodriguez, and John Thangarajah. 2026. Engineering norm-aware BDI agents. https://doi.org/10.5281/zenodo.18603467