Raise BDI Agents, First Steps

Doctoral Consortium

Andrea Gatti
Università degli Studi di Genova
Genova, Italy
andrea.gatti@edu.unige.it

ABSTRACT

Recent developments in game engines have opened the door to three-dimensional worlds for agents. VEsNA (Virtual Environments via Natural language Agents) is a framework for building embodied BDI agents extending the Jason framework. It allows agents to have a situated body, communicate with other agents and users, interact with objects, and reason logically about space. VEsNA agents also interact with humans in natural language through ChatBDI and act according to a simulated personality and mood. The framework has been validated through several experiments: an evaluation of the *nl2kqml* natural language pipeline, an assessment of temper-driven plan selection, a complex multi-agent scenario developed in collaboration with the game company Untold Games, and a series of theoretical contributions extending the syntax and semantics of BDI agents. In this paper, I summarise the work carried out during the first two years of my PhD, outline the trajectory for the final year, and critically assess the strengths, limitations, and practical potential of VEsNA.

KEYWORDS

BDI Agents; Virtual Environments; NLP; Logics;

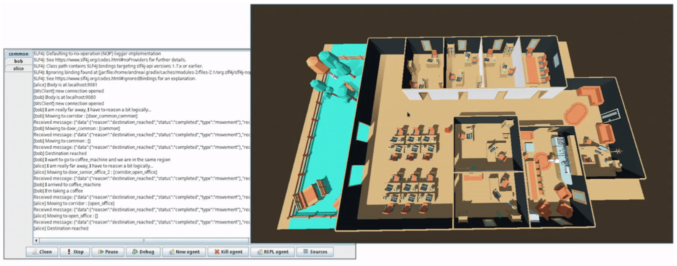
ACM Reference Format:
Andrea Gatti. 2026. Raise BDI Agents, First Steps: Doctoral Consortium. In *Proc. of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2026), Paphos, Cyprus, May 25 – 29, 2026*, IFAAMAS, 3 pages. https://doi.org/10.65109/JYPA1233

1 INTRODUCTION

Belief-Desire-Intention (BDI) agents are grounded in psychological concepts and formalized by logicians. While modern neural networks attempt to emulate human brains from a microscopic perspective, the BDI architecture aims to emulate cognition at a macro level. Existing implementations of cognitive agents, however, typically operate in simplified 2D grid-worlds or abstract, non-spatial environments that offer little opportunity for genuine embodied reasoning. With the advancements in game engines over the past few years, it is now possible to bring these agents into fully three-dimensional worlds with all their associated challenges.

VEsNA (Virtual Environment via Natural Language Agents) [9][6][10] addresses this scenario by tackling many of the challenges and features associated with embodiment for BDI agents. It is built as an extension of Jason [2], and leverages CArtAgO, the artifact-based environment infrastructure provided by the JaCaMo platform [1], for workspace and object management. The main goal of this framework, and of my PhD research, is to create a system that natively supports embodied reasoning and interaction, enabling BDI agents to engage with users.

The current implementation uses Jason on the agent side and Godot[13] for virtual environment development. All the code is available on GitHub. A screenshot of VEsNA working is visible in Fig 1.



```
SLF4J: Defaulting to no-operation (NOP) logger implementation
SLF4J: See https://www.slf4j.org/codes.html#noProviders for further details.
SLF4J: Class path contains SLF4J bindings targeting slf4j-api versions 1.7.x or earlier.
SLF4J: Ignoring binding found at [jar:file:/home/andrea/.gradle/caches/modules-2/files-2.1/org.slf4j/slf4j-no
SLF4J: See https://www.slf4j.org/codes.html#ignoredBindings for an explanation.
[alice] Body is at localhost:9081
[WsClient] new connection opened
[bob] Body is at localhost:9080
[WsClient] new connection opened
[bob] I am really far away, I have to reason a bit logically...
[bob] Moving to corridor : [door_common,common]
Received message: {"data":{"reason":"destination_reached","status":"completed","type":"movement"},"rec
[bob] Moving to door_common : [common]
Received message: {"data":{"reason":"destination_reached","status":"completed","type":"movement"},"rec
[bob] Moving to common : []
Received message: {"data":{"reason":"destination_reached","status":"completed","type":"movement"},"rec
[bob] Destination reached
[bob] I want to go to coffee_machine and we are in the same region
[alice] I am really far away, I have to reason a bit logically...
[alice] Moving to door_senior_office_2 : [corridor,open_office]
Received message: {"data":{"reason":"destination_reached","status":"completed","type":"movement"},"rec
[bob] I arrived to coffee_machine
[bob] I'm taking a coffee
Received message: {"data":{"reason":"destination_reached","status":"completed","type":"movement"},"rec
[alice] Moving to corridor : [open_office]
Received message: {"data":{"reason":"destination_reached","status":"completed","type":"movement"},"rec
[alice] Moving to open_office : []
Received message: {"data":{"reason":"destination_reached","status":"completed","type":"movement"},"rec
[alice] Destination reached
```

Figure 1: A VEsNA Screenshot: the Office scenario is the default used of tests.

2 ARCHITECTURE AND COMPONENTS

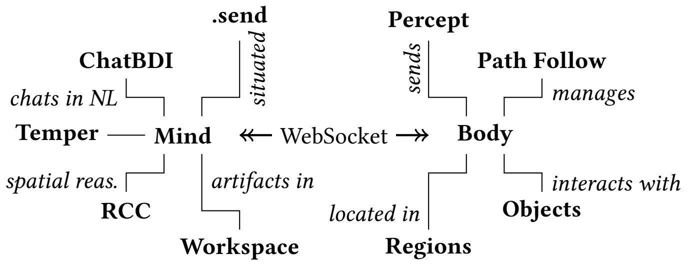


Figure 2: VEsNA Components

Fig. 2 illustrates the current architecture of VEsNA. Architecturally, a VEsNA agent is composed of a *mind* and a *body*. In the current implementation, these two components communicate via WebSocket using JSON messages. Some components have a digital twin (*RCC–Regions* and *Workspace–Objects*), while others exist on only one side.

The mind manages goals, determines strategies, interprets environmental signals, reacts to events, and maintains awareness of

---

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/JYPA1233

available artifacts. The body executes the mind's directives, reports environmental events, and provides perceptual feedback about action progress and failures.

Spatial Reasoning. The agent performs spatial reasoning[3][7] using Region Connection Calculus[14] (RCC). RCC provides a set of logical predicates that define space through relationships between regions. Using RCC, the agent's mind reasons about the semantic structure of space rather than relying on coordinate-based representations. Movement planning and execution are handled by the body, which maintains a physical representation of regions and navigates to locations specified by the mind.

Artifacts. The environment may contain objects with enhanced features. These special objects, namely *artifacts*, are known to agents and can be utilized by them [8]. VEsNA artifacts are implemented as CArtAgO artifacts: they inherit their observable properties and operations interface, while adding spatial and physical characteristics (position, occupancy, and portability). In VEsNA, there are two types of artifacts:

- *situated*: these objects are fixed in place and support a maximum number of simultaneous users;
- *grabbable*: these objects are movable and can be held by only one agent at a time.

We evaluated this feature in a scenario where agents need to prepare coffee. To accomplish this task, agents must locate a cup (*grabbable artifact*) and a coffee machine (*situated artifact*) and use them together.

Situated Communication. VEsNA agents can communicate only when they are in the same room [4]. This constraint requires agents to move close to each other to exchange information, making communication a situated action. VEsNA extends the *.send* internal action that normally makes the agents communicate modifying the message delivery depending on the conditions. VEsNA supports two communication modalities: (1) *private*: the message is sent exclusively to a specific agent; (2) *public*: the message is addressed to target agents but is also overheard by the others in the same room, if any.

When an agent overhears something, it can decide how to respond based on its design. Notably, if the intended receiver is not in the room, other agents who overhear the message may handle the request or information instead.

Chat with Agents. Agents can engage in natural language conversations with users [11][12] through *chatBDI*, a framework that implements two main methods: *nl2kqml* and *kqml2nl*.

The *nl2kqml* method enables Jason agents to interpret natural language messages based on their belief base and plan library. The incoming message is classified by KQML illocutionary force, embedded via a pre-trained language model, and compared with the agent's beliefs and/or plans to find the best matching template. This template is then used by another component (currently an LLM) to generate the final answer, which is delivered to the receivers as a standard KQML message.

Conversely, when an agent needs to communicate with a user, it uses *kqml2nl*. This method leverages an LLM to generate a natural language message from the agent's KQML message.

Temper. The framework extends the Jason reasoning cycle to consider each agent's temper during the decision-making process. The temper is subdivided into an *immutable* part and a *mutable* part (*mood*). Plans are annotated with a temper and with effects on the temper itself. Agents choose how to act by selecting either the plan with the most similar temper to their own, or through weighted randomness. The framework is model-agnostic with respect to the temper representation: any consistent set of personality dimensions can used. In the current implementation and experiments, the *Big Five* personality model was adopted to define the temper space. This work has been tested during a master thesis in collaboration with Untold Games, a game company located in Genoa. It is better presented in [5].

3 FUTURE WORK

While significant progress has been made, many challenges remain and several directions merit further exploration.

Collaborative Plans. This direction is currently under active development. The goal is to allow plans to be annotated as collaborative, specifying requirements for cooperation among agents. With this feature, VEsNA would autonomously manage actions that require the involvement of multiple agents, coordinating their efforts and handling potential failures gracefully.

Exploration. One of the main current limitations of VEsNA is that the map must be known in advance. It would be far more interesting to allow agents to build their spatial representation at runtime, but this problem is far from trivial. Some feasibility studies have been conducted and initial steps have been taken, but without significant results so far.

Sight. Another aspect that deserves investigation, closely related to the exploration problem, is agent vision. This poses significant challenges, such as: how can an agent notice that an object is no longer present? How can it perceive that an object is being held by another agent?

Learning. VEsNA agents could attempt to emulate, in logical form, well-known learning algorithms to learn how to generate higher-level plans starting from a predefined set of actions and plans. This requires an in-depth study of the system to fully understand the concept of "state" within an agent's lifecycle and its evolution over time. A first step in this direction consists of generating execution traces during runtime.

Interaction with the User. Another open challenge is physically placing the user within the same environment as the agents, either by controlling an avatar on screen or by being fully immersed in virtual reality. This development poses many challenges, particularly in handling all possible user inputs and actions, and in enabling agents to interpret them appropriately.

Stabilization. Another avenue, more implementation-oriented than research-focused, would be to develop a C-based version of Jason for game engines such as Unity and Unreal. Finally, it is important to dedicate time to consolidating all the components and stabilizing the framework, ensuring that it becomes usable by third-party users in a practical and reliable manner.

REFERENCES

[1] Olivier Boissier, Rafael H Bordini, Jomi Hubner, and Alessandro Ricci. 2020. *Multi-agent oriented programming: programming multi-agent systems using JaCaMo*. MIT Press.

[2] Rafael H Bordini, Jomi Fred Hübner, and Michael Wooldridge. 2007. *Programming multi-agent systems in AgentSpeak using Jason*. John Wiley & Sons.

[3] Angelo Ferrando, Andrea Gatti, and Viviana Mascardi. 2024. Geometric and Spatial Reasoning in BDI Agents: A Survey. In *Proceedings of the 39th Italian Conference on Computational Logic, Rome, Italy, June 26-28, 2024 (CEUR Workshop Proceedings, Vol. 3733)*, Emanuele De Angelis and Maurizio Proietti (Eds.). CEUR-WS.org. https://ceur-ws.org/Vol-3733/paper1.pdf

[4] Angelo Ferrando, Andrea Gatti, and Viviana Mascardi. 2025. Oops, I Heard That! Situated Communication with Locality-Aware KQML. In *Informal Proceedings of the 13th International Workshop on Engineering Multi-Agent Systems, EMAS 2025, Co-located with AAMAS 2025, Detroit, MI, USA, May 19-23, 2025*.

[5] Angelo Ferrando, Andrea Gatti, Viviana Mascardi, Andrea Stucchi, and Fabio Casale. [n.d.]. VEsNA-Pro: Exploiting BDI Agents with Propensities for Emergent Narrative. In *The 25th International Conference on Autonomous Agents and Multi-Agent Systems*.

[6] Andrea Gatti. 2023. Virtual Environments via Natural Language Agents. In *Multi-Agent Systems - 20th European Conference, EUMAS 2023, Naples, Italy, September 14-15, 2023, Proceedings (Lecture Notes in Computer Science, Vol. 14282)*, Vadim Malvone and Aniello Murano (Eds.). Springer, 486–492. https://doi.org/10.1007/978-3-031-43264-4_36

[7] Andrea Gatti. 2024. Reason Logically, Move Continuously. In *Agents and Robots for reliable Engineered Autonomy - 4th Workshop, AREA 2024, Santiago de Compostela, Spain, October 19, 2024, Proceedings (Communications in Computer and Information Science, Vol. 2230)*, Angelo Ferrando and Rafael C. Cardoso (Eds.). Springer, 115–127. https://doi.org/10.1007/978-3-031-73180-8_8

[8] Andrea Gatti, Angelo Ferrando, and Viviana Mascardi. 2025. Situated Agents in Action: Extending VEsNA with Spatial and Grabbable Artifacts. In *Advances in Practical Applications of Agents, Multi-Agent Systems, and Computational Social Science: The PAAMS Collection - 23rd International Conference, PAAMS 2025, Lille, France, June 25-27, 2025, Proceedings (Lecture Notes in Computer Science, Vol. 16031)*, Philippe Mathieu and Fernando De la Prieta (Eds.). Springer, 348–353. https://doi.org/10.1007/978-3-032-07638-0_29

[9] Andrea Gatti and Viviana Mascardi. 2022. Towards VEsNA, a Framework for Managing Virtual Environments via Natural Language Agents. In *Proceedings of the Second Workshop on Agents and Robots for reliable Engineered Autonomy, AREA@IJCAI-ECAI 2022, Vienna, Austria, 24th July 2022 (EPTCS, Vol. 362)*, Rafael C. Cardoso, Angelo Ferrando, Fabio Papacchini, Mehrnoosh Askarpour, and Louise A. Dennis (Eds.). 65–80. https://doi.org/10.4204/EPTCS.362.8

[10] Andrea Gatti and Viviana Mascardi. 2023. VEsNA, a Framework for Virtual Environments via Natural Language Agents and Its Application to Factory Automation. *Robotics* 12, 2 (2023), 46. https://doi.org/10.3390/ROBOTICS12020046

[11] Andrea Gatti, Viviana Mascardi, and Angelo Ferrando. 2025. ChatBDI: Think BDI, Talk LLM. 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, 2541–2543. https://doi.org/10.5555/3709347.3743930

[12] Andrea Gatti, Viviana Mascardi, Angelo Ferrando, et al. 2025. Let Me Talk to You! Natural Language Interaction Between Humans and BDI Agents via ChatBDI. In *28th European Conference on Artificial Intelligence, 25-30 October 2025, Bologna, Italy–Including 14th Conference on Prestigious Applications of Intelligent Systems (PAIS 2025)*.

[13] Juan Linietsky, Ariel Manzur, and The Godot Contributors. 2014. Godot Engine. https://godotengine.org/. Open-source game engine, MIT License.

[14] David A Randell, Zhan Cui, Anthony G Cohn, et al. 1992. A spatial logic based on regions and connection. *KR* 92, 165-176 (1992), 40–40.