Ontology-Based Reflective Communication for Shared Human-AI Recognition of Emergent Collaboration Patterns

van Zoelen, Emma M.; van den Bosch, Karel; Abbink, David; Neerincx, Mark

DOI
10.1007/978-3-031-21203-1_40

Publication date
2023

Document Version
Final published version

Published in
PRIMA 2022

Citation (APA)
van Zoelen, E. M., van den Bosch, K., Abbink, D., & Neerincx, M. (2023). Ontology-Based Reflective Communication for Shared Human-AI Recognition of Emergent Collaboration Patterns. In R. Aydoğan, N. Criado, V. Sanchez-Anguix, J. Lang, & M. Serramia (Eds.), *PRIMA 2022: Principles and Practice of Multi-Agent Systems - 24th International Conference, Proceedings* (pp. 621-629). (Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Vol. 13753 LNAI). Springer. https://doi.org/10.1007/978-3-031-21203-1_40

Ontology-Based Reflective Communication for Shared Human-AI Recognition of Emergent Collaboration Patterns

Emma M. van Zoelen[1,2](✉), Karel van den Bosch[2], David Abbink[1], and Mark Neerincx[1,2]

[1] Delft University of Technology, Mekelweg 5, 2628 CD Delft, The Netherlands
e.m.vanzoelen@tudelft.nl
[2] TNO, Kampweg 55, 3769 DE Soesterberg, The Netherlands

Abstract. When humans and AI-agents collaborate, they need to continuously learn about each other and the task. We propose a Team Design Pattern that utilizes adaptivity in the behavior of human and agent team partners, causing new Collaboration Patterns to emerge. Human-AI Co-Learning takes place when partners can formalize recognized patterns of collaboration in a commonly shared language, and can communicate with each other about these patterns. For this, we developed an ontology of Collaboration Patterns. An accompanying Graphical User Interface (GUI) enables partners to formalize and refine Collaboration Patterns, which can then be communicated to the partner. The ontology was evaluated empirically with human participants who viewed video recordings of joint human-agent activities. Participants were requested to identify Collaboration Patterns in the footage, and to formalize patterns by using the ontology's GUI. Results show that the ontology supports humans to recognize and define Collaboration Patterns successfully. To improve the ontology, it is suggested to include pre- and post-conditions of tasks, as well as parallel actions of team members.

Keywords: Human-agent team · Ontology · Collaboration Patterns · Co-learning

1 Introduction

A growing body of research on human-agent teaming [1, 2] and human-robot collaboration [3] studies how to make optimal use of the qualities of both humans and AI agents by making them team partners. An important aspect of becoming successful team members is to continuously learn about each other and the task, to make sure the team becomes a fluently functioning unit; a process called co-learning [4].

To us humans, adapting to and learning with our fellow human team members often comes natural. In a hybrid human-agent team, successful adaptation and learning is not self-evident, as humans and AI agents differ in the way in which they learn and adapt. Still, implicit co-adaptation is bound to occur as the learning processes of human and agent will influence each other while they collaborate on a task. As a result, new team behaviors will emerge [5]; successful emergent team behaviors (coordination and

R. Aydoğan et al. (Eds.): PRIMA 2022, LNAI 13753, pp. 621–629, 2023.
https://doi.org/10.1007/978-3-031-21203-1_40

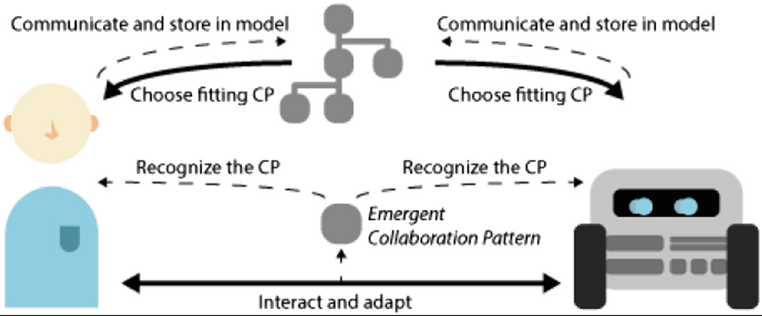
cooperation as defined by [6]) can be specified as 'Collaboration Patterns[1]'. To make sure that the team can successfully co-learn by achieving reflective communication (the highest quality collaboration [6]), it is necessary for team members to consolidate and organize their collaborative efforts by developing a shared model of the collaboration.

Based on exploratory work done in [5], this process can be expressed as a Team Design Pattern (Table 1) that describes how human and agent team members co-learn by communicating about emergent Collaboration Patterns. In this paper we address the following research question:

What kind of model and communication interface enables a human-agent team to establish shared recognitions of emergent Collaboration Patterns?

Existing frameworks of collaboration usually predefine Collaboration Patterns (e.g. as Plays [7] or Social Practices [8]), and do not study how Patterns can be created or updated through communication during human-agent collaboration. We propose a model in the form of an ontology, along with an accompanying communication interface that can be used for communication about and formalization of emergent Collaboration Patterns. An advantage of ontologies is that they provide shared univocal conceptualizations that can be used for communication and reasoning [9]. The concepts and structure of the ontology and its communication interface were evaluated empirically, with human participants. With this ontology-based reflective communication we aim to enable human-agent co-learning of successful coordination and collaboration behaviors.

Table 1. Team design pattern for co-learning in a human-agent team. The pattern supports definition of emergent Collaboration Patterns in a shared ontology, enabling partners to communicate about them. In this paper we focus on the dashed arrows on the human side.

Name	Human-AI Co-Learning
Description	When human and an adaptive AI agent collaborate as team partners, they both adapt their behavior constantly to dynamically changing requirements of the task. When doing so, Collaboration Patterns emerge. A team member recognizes a CP as valuable to the task, and communicates this to its team member. By jointly reflecting on the CP they can refine or adjust the CP until they agree on its value and use. The CP is defined and stored in the shared ontology; now both team partners are explicitly aware of this CP, and can use it when relevant.
Structure	

[1] Previously called 'Interaction Patterns' in [5].

2 Ontology: Requirements and Background

2.1 Requirements for Collaboration Pattern Ontology

To enable an agent to reason and communicate about patterns of collaboration, the agent needs a model representing the relevant concepts that underlie the collaboration. Relevant concepts are, for example, the entities (e.g. actors, objects) that take part in the collaboration, details about the actions that should be executed (e.g. which actions, when, and in what order) and the context in which the Collaboration Pattern takes place (e.g. when does it start and end). Using these concepts, an agent can connect a particular pattern of actions to a particular instance of a context. In addition, the model should allow the agent to define the success of a Collaboration Pattern, in terms of its contribution (or harm) to the team's task.

As the model will be used for storing and updating newly emergent CPs, and for communication during collaboration, it is necessary that it can be dynamically updated. Given that the ontology should function as a shared model between team members, the structure and concepts of the ontology should be fixed, while instances of specific Collaboration Patterns can be updated. This is analogous to a frame-based approach, in which an *Upper Ontology* describes concepts and relations in a generic manner, while a *Lower Ontology* describes unique instances of the concepts and relations [9].

To summarize, the requirements for the ontology are as follows: (1) it should store and specify the structure of patterns of collaboration; (2) it should contain a model of context at a level that is understandable by humans; (3) it should support the agent to reason about the appropriateness of patterns in specific contexts; (4) it should allow live updating; (5) there should be a distinction between high level concepts (Upper Ontology) that provide the structure of the model, and low level instances (Lower Ontology) that can be used directly in a task by the team members.

2.2 Ontologies in Human-Agent Teaming

There are several existing ontologies in the areas of human-agent teaming and human-robot collaboration. Some of these address team configuration [10, 11], but do not cover collaborative actions, hence are not directly useful for our purposes. Ontologies that do represent coordinated actions focus on high-level tasks and goals (e.g. [12]). Some more recent papers touch upon complex aspects of team behavior, such as the integration of intent (e.g. [2]), or a combination of tasks, goals and intent by introducing the concept of 'Plays' in their ontology (e.g. [1, 13]). 'Plays' were introduced and are used mostly in human-agent teaming applications in the military [7]. A play is a set of instructions that tells actors in a team how to act in a particular situation. This concept is similar to our notion of Collaboration Patterns, although a play is a predefined set of instructions, whereas CPs emerge and develop during collaboration. While the concept of plays is relevant for our work, the reported ontologies that represent plays do not provide information on the structure of a play and therefore do not meet requirement 1.

For human-robot collaboration, many ontologies exist that enable robots to behave autonomously in a certain practical task [14]. They contain concepts such as object, task, actor, etc., and have a large overlap with task models (such as [15]), but add aspects that are required for robots, such as hardware knowledge (e.g. about sensors and actuators that the robot is equipped with). Some of these ontologies support communication to humans, but with a focus on context-dependent information sharing. The task model related aspects of these ontologies are reusable for the description of context in our ontology (requirement 2).

In conclusion, existing ontologies that use task models as described in [14, 15] can provide a basis for formalizing context within an ontology. We used these as a starting point for designing the context part of our ontology (requirement 2).

2.3 Frameworks for Describing Patterns of Collaboration

We have also looked at frameworks that describe patterns of collaboration between partners, for example from sociology, that have not been formalized in an ontology. As we aim to represent CPs that are defined while collaborating, we are looking for a structure of the concepts and relations that build up a CP, to incorporate that in the ontology.

One such framework is Social Practices, which originates from sociology, although it has been formalized for agent reasoning [8]. Social Practices are described as ways of doing things that are shared between actors. They contain patterns of behavior that are strongly tied to the specific contexts in which they are executed. Social Practices emerge as humans interact with each other in a particular environment. Although Social practices are predefined (and not emergent) in the formalization in [8], the formalization is detailed and contains the elements that build up a Social Practice. The formalization in [8] uses concepts such as Actors, Roles, and Places to describe context, and also includes concepts such as Possible Actions and Strategies to describe expected sequences of behavior.

We used the Social Practice formalization as a starting point for designing the part of the ontology that represents Collaboration Patterns (requirement 1).

3 Collaboration Pattern Ontology

Using the aforementioned frameworks, we chose a minimal set of concepts and relations, that can be extended for use in specific domains. This minimal set of concepts and relations serves as the Upper Ontology. In choosing the concepts and relations, we have based ourselves on previous work; we wanted to ensure that our ontology could at least describe CPs found in [5]. An example of such a CP is ‘Alternating actively working on the task and waiting for a team member’; in this CP, a human team member clears away some small rocks from a pile, then waits for their agent team member to clear away large rocks from that pile. This CP would take place when there is a rock pile that contains both large and small rocks. A definition of a CP based on this example needs tasks (clear away, wait), actors (human, robot) and objects (small and large rocks), as well as a way to describe the order in which tasks happen, and a condition to choose the CP. Figure 1 presents a graphical overview of the ontology.

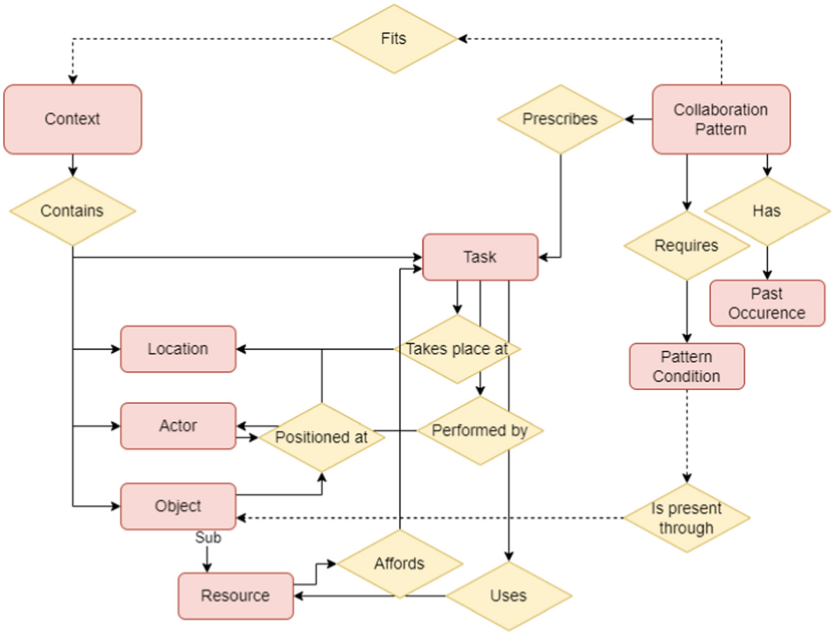


Fig. 1. An overview of the Collaboration Pattern ontology. The red items are the concepts, the yellow items are the relations. Dashed lines represent relations that enable an agent to reason about whether CPs fit a certain context. (Color figure online)

4 Ontology in Practice: Translation to Context

To evaluate whether the chosen approach supports humans in expressing Collaboration Patterns to agents, we chose to use a USAR task in which a human-agent team collaborated in saving an earthquake victim from underneath a pile of rocks (see [5]). We created a context-dependent set of specifications for several concepts, as well as a communication interface, based on use of the ontology within this task. The context-dependent set of specifications are translations of the concepts specified in Fig. 1.

To support a human team partner in defining and communicating a Collaboration Pattern and the contextual conditions for its application to their agent team partner, a drag-and-drop graphical user interface (GUI) was developed (see Fig. 2). The GUI consisted of predefined 'building blocks' based on the context-dependent concepts.

4.1 Evaluation Goals and Method

The evaluation addressed whether human participants were able to identify emergent Collaboration Patterns in video footage. We verified whether the detected patterns matched those previously identified by the researchers, as well as what the differences and similarities were. Moreover, we investigated whether participants were able to describe these CPs with the ontology and the designed GUI, as well as what combinations of the concepts they used to create their descriptions.

A series of 14 short video clips of human-agent collaboration were presented to participants. These clips were taken from a previous experiment [5]. All clips contained a fragment of human-agent collaboration in which a Collaboration Pattern was previously identified by the experimenter. The experiment was conducted by video call with each individual participant. It consisted of two parts: (1) participants practiced with the task of the experiment described in [5], and (2) participants watched the videos and defined Collaboration Patterns they distinguished using the GUI. They were asked to clarify their actions and thinking verbally to the experiment leader. Ten students participated (7 M, 3 F). All participants were in the final stages of AI-related Master programs. The procedure was approved by the Human Research Ethics Committee at Delft University of Technology on November 12th, 2021.

The GUI descriptions of the Collaboration Patterns were coded in two different ways: (1) correct, incorrect or semi-correct when compared to CP descriptions made by the researchers prior to the experiment, and (2) open coded based on how the descriptions were built up. The verbal clarifications were used to understand what participants meant in case their CP descriptions made with the GUI were unclear. They were also used to explore what aspects of the ontology, GUI and approach can be improved.

Fig. 2. An example Collaboration Pattern and accompanying Pattern Condition in the GUI. The user selects context-factors from colored blocks and drags them into the left grey area (“Situation”) to define the contextual conditions of the CP, and into the right grey area (“What we do”) to define the actions constituting the CP. The description portrayed is the example mentioned in Sect. 3, ‘Alternating actively working on the task and waiting for a team member’.

4.2 Results: Correctness of Collaboration Patterns Recognized

One participant (participant 13) did not follow the syntax of the GUI as instructed, and described Collaboration Patterns in an extremely minimal manner. As no feedback was given on the descriptions during the experiment, there was limited opportunity for them to learn. All participants had difficulty understanding what was happening in video 5, which was understandable, as the course of actions was dependent on the implicit intentions of the human, and not very apparent in the behavior. We observed that participants became better at recognizing and describing that the human was directing the robot as the experiment progressed. Overall, they were able to create descriptions that made sense and at least partly fit with the descriptions made by the researchers.

Interestingly, anything that required the human to wait for the robot was ignored in the descriptions. On the other hand, waiting behavior by the robot was expressed often. Moreover, some participants decided to decouple the human and robot behavior completely, describing both separately and therefore ignoring any interaction or coordination between the two team members. Most participants did this for some descriptions.

4.3 Results: Patterns of Descriptions

Many participants noticed that the human team member directed the robot in several of the videos. However, they differed in the way they accounted for this in their CPs. They would for example describe the human standing still, or moving back and forth, they sometimes described ‘Robot move to Human’, and some by put the human actions in the situation description. Several participants also indicated that they had difficulty expressing this kind of behavior. Participants indicated that they felt the need to describe causality. They sometimes attempted to create a pre-condition or trigger for an action, but sometimes also a post-condition or consequence of an action.

Contextual information (such as locations, objects or agents) was often left out of action descriptions, possibly because participants deemed the information self-evident (e.g. the information is already in the situation description, or two consecutive actions that are done by the same agent). Sometimes, participants added extra information to make their descriptions more specific, for example by using double location specifications to describe a more specific location.

About half of the participants took a modular approach in describing the Collaboration Patterns. They used several small CPs to describe behavior observed in one video. Some participants created separate CPs for the human and the robot. These participants often attempted to create a complete set of CPs that would describe all possible behavior sequences observed; therefore, many of these small CPs were reused in several videos.

5 Discussion

Our work presents an approach for enabling a human-agent team to co-learn, and to identify and share emergent Collaboration Patterns. The ontology has been designed to be generic; it can be used in or easily adapted to other contexts. It should be seen as a first iteration for creating a model to formalize emergent CPs, that can be built upon.

To enable a human-agent team to use the ontology of CPs, we designed an interface. We chose to design a drag-and-drop GUI (rather than, e.g., a natural language interface), to restrict formalization of and communication about the CPs to concepts and relations present in the ontology. The GUI allowed us to perform an experiment with human participants. The results show that it supports people in formalizing CPs within the boundaries of the ontology. Research into what interfaces are suitable for communicating about CPs in different task contexts will be valuable.

Results of the evaluation can be used to improve the design of the ontology and the GUI. Participants tended to not explicitly formalize information that was described earlier in the same CP. To ensure that this does not hamper the process of developing shared representations of CPs, we might want to equip the agent with inferencing capabilities, or expand the interface to guide the human in checking whether assumptions are met. Several participants also formalized human and agent behavior separately. The GUI should support the user more extensively in formalizing the interactions as a CP, rather than a sequential series of actions. This requires elaboration of the ontology, by allowing parallel tasks, and/or the coupling of actions through pre- and post-conditions.

The evaluation was done offline, through video recordings, instead of on-task, based on experienced collaboration. Therefore, it does not address how descriptions might change over time due to behavior portrayed by the agent team partner. A next challenge is to implement the additional design requirements obtained in this study, and to investigate whether they support human-agent co-learning on the job. This will improve our understanding of how participants' perception of task, agent and CPs evolves over time.

6 Conclusion

The ontology in this paper enables a human-agent team to represent emergent Collaboration Patterns explicitly, thus making them available for use in future task situations. The formal representation in the ontology enables partners to correct and refine CPs, based upon new experiences or reflections. This way, the ontology system supports co-learning within the team. The drag-and-drop graphical user interface that we presented provides a common language for the team members by translating high-level concepts from the ontology to more contextualized concepts. Evaluation with human participants showed that people are able to identify relevant CPs from videos, and that they were able to formalize them using the GUI. CPs in which the human directs the agent proved difficult to describe. It is therefore considered necessary to expand the ontology system with a function to represent behaviors conducted in parallel by human and agent, as well as to explicitly provide ways of representing pre- and post-conditions.

Our work contributes to developing representations of emergent Collaboration Patterns that support co-learning by humans and agents. Further research is needed to evaluate, expand and refine the proposed ontology system.

References

1. van Diggelen, J., Barnhoorn, J., Post, R., Sijs, J., van der Stap, N., van der Waa, J.: Delegation in human-machine teaming: progress, challenges and prospects. In: Russo, D., Ahram, T., Karwowski, W., Di Bucchianico, G., Taiar, R. (eds.) IHSI 2021. AISC, vol. 1322, pp. 10–16. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-68017-6_2
2. Schneider, M.F., Miller, M.E., Ford, T.C., Peterson, G., Jacques, D.: Intent integration for human-agent teaming. Syst. Eng. 1–13 (2022). https://doi.org/10.1002/sys.21616
3. Ajoudani, A., Zanchettin, A.M., Ivaldi, S., Albu-Schäffer, A., Kosuge, K., Khatib, O.: Progress and prospects of the human–robot collaboration. Auton. Robot. **42**(5), 957–975 (2017). https://doi.org/10.1007/s10514-017-9677-2

4. Schoonderwoerd, T.A.J., van Zoelen, E.M., van den Bosch, K., Neerincx, M.A.: Design patterns for human-AI co-learning: a wizard-of-Oz evaluation in an urban-search-and-rescue task. Int. J. Hum.-Comput. Stud. **164**, 102831 (2022). https://doi.org/10.1016/j.ijhcs.2022.102831
5. van Zoelen, E.M., van den Bosch, K., Neerincx, M.: Becoming team members: identifying interaction patterns of mutual adaptation for human-robot co-learning. Front. Robot. AI. **8**, 200 (2021). https://doi.org/10.3389/frobt.2021.692811
6. Engeström, Y.: Interactive expertise: studies in distributed working intelligence. Research Bulletin 83. Department of Education, University of Helsinki, Bulevardi 18, SF-00120 Helsinki, Finland (1992)
7. Miller, C.A.: Delegation architectures: playbooks and policy for keeping operators in charge. Presented at the Workshop on Mixed-Initiative Planning and Scheduling, Monterey, California, USA (2005)
8. Dignum, F.: Interactions as social practices: towards a formalization. ArXiv180908751 Cs (2018)
9. Fensel, D.: Ontologies: a silver bullet for knowledge management and electronic commerce (2001)
10. van Diggelen, J., Neerincx, M., Peeters, M., Schraagen, J.M.: Developing effective and resilient human-agent teamwork using team design patterns. IEEE Intell. Syst. **34**, 15–24 (2019). https://doi.org/10.1109/MIS.2018.2886671
11. Pico-Valencia, P., Holgado-Terriza, J.A., Sierra Martínez, L.M.: A preliminary ontology for human-agent collectives. In: De La Prieta, F., et al. (eds.) PAAMS 2019. CCIS, vol. 1047, pp. 176–187. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-24299-2_15
12. Madni, A.M., Madni, C.C.: Architectural framework for exploring adaptive human-machine teaming options in simulated dynamic environments. Systems **6**, 44 (2018). https://doi.org/10.3390/systems6040044
13. Kasmier, D., et al.: Ontology of plays for autonomous teaming and collaboration. In: Proceedings of the XIV Seminar on Ontology Research in Brazil, ONTOBRAS, pp. 9–22 (2021)
14. Olivares-Alarcos, A., et al.: A review and comparison of ontology-based approaches to robot autonomy. Knowl. Eng. Rev. **34**, 1–29 (2019). https://doi.org/10.1017/S0269888919000237
15. van Welie, M., van der Veer, G.C., Eliëns, A.: An Ontology for task world models. In: Markopoulos, P., Johnson, P. (eds.) Design, Specification and Verification of Interactive Systems 1998. Eurographics, pp. 57–70. Springer, Vienna (1998). https://doi.org/10.1007/978-3-7091-3693-5_5