Human-Aware Planning: A Survey related to Joint Human-Agent Activities

Sebastian Ahrndt, Johannes Fähndrich, and Sahin Albayrak

DAI-Laboratory, Technische Universität Berlin,
Faculty of Electrical Engineering and Computer Science,
Ernst-Reuter-Platz 7, 10587 Berlin, Germany
sebastian.ahrndt@dai-labor.de (Corresponding author)

Abstract. To become a part of a joint human-agent team, artificial agent are required to achieve joint goals with humans not only performing task for humans. This includes the ability to coordinate actions between team-members, which is e.g. addressed by Human-Aware Planning approaches. This work surveys available solutions regarding the special requirements identified for joint human-agent activities. In particular, the work concentrates on the requirement of interpredictability, which requires to include the course of actions of other team-members into the planning process of one's own course of action.

1 Introduction

Focus Human-sided self-explanation, aims to integrate users into existing systems not only as supervisors but also as a regular component and to offer the opportunity to interact with human beings, by means of the available communication channels [11]. Although the addressed systems are goal-driven and act (as much as possible) autonomously, humans have to be able to define goals, to restrict the system by means of constraints and to inspect the results of the self-organisation process [15, 21]. Furthermore, humans must not only be seen as supervisors but also as agents that can be asked for help to reach specific goals [21]. These mechanisms rises significant engineering challenges for both, the decision making process of agents and the interaction between agents and human users. Concentrating on the decision making process, this work surveys requirements and available state-of-the-art solutions. In particular, we focus on the challenge to decide at which moment an interaction with humans is useful and at which situation and in which way the interventions of human users should be included into the course of actions of agents.

Motivation Autonomous agents have to decide what actions to perform, when they should be performed and—in some cases—determine which entity should be asked for help to fulfil a specific task [28]. Once the available entities do not only comprise artificial agents but also natural agents or in other words human beings, joint human-agent activities emerge. Decision making procedures

for joint human-agent activities are done by Human-Aware Planning [8] (HAP) components, that encourage the basic idea of human-agent activities, which is that agents and humans have different capabilities, advantages and disadvantages that can complement each other. One example of a human advantage is the selection of relevant information while one example of the strengths of computers is the management of vast amounts of data. Furthermore, humans excel at finding solutions to new or evolving problems, building knowledge and learning, but are less efficient and not suited for the execution of multiple operations at the same time.
To build effective team-players agents planning for such activities have to tackle the dynamic nature of humans as the human behaviour features several aspects of uncertainty for the planning process [1, 16, 19]. As an example, consider the fact that humans select merely good and feasible actions to fulfil a task rather then selecting the optimal ones [25]. Furthermore, after committing to a specific task humans may change their goals from one moment to another without a (for computers) comprehensible reason. These aspects of uncertainty imply a dynamical behaviour that can be seen as some kind of 'Quality of Behaviour' a human is able to provide and influences the planning process in different ways. For example, the non-optimal execution of an action influences the execution time whereas the sudden interruption of a task endangers the whole plan and might imply a replanning to reach a given goal.
One assumption of currently available human-aware planning components is that whenever a task is assigned to a human that possesses the ability to fulfil it, the human will provide results in a timely fashion. This assumption is questionable since the 'Quality of Behaviour' a human is able to provide depends on the current context of the human. For example, a human might not provide required information if he/she is currently occupied with performing another task. The primary objective of our work is to relax this assumption to a more general one. That is, whenever a task is assigned to a human with the ability to fulfil it, the human may accept or decline such task and provide results either in time or delayed [1]. This form of 'context-dependent' behaviour is essential for planning processes that account for the abilities of human beings.

Structure The remainder of this work is structured as follows. In Section 2 we give some background information about joint human-agent activities and human-aware planning components. Section 3 introduces requirements for HAP. Section 4 presents state-of-the-art HAP projects and approaches. Afterwards, in Section 5 we discuss the results and implications of this survey. Finally, Section 6 concludes our study and gives an outlook on future work.

2 Terminology

Joint Human-Agent Activities Following the definition for joint-activities given by *H.H. Clark* [10, p. 3], joint human-agent activities can be defined as an extended set of actions that is executed by an ensemble of natural and artificial agents who are coordinating with each other [10, 19]. Agents cooperate with each other to overcome some certain kind of inherent limitations. Nevertheless, cooperation would be avoided if no other stimuli exist, as it adds additional cost to an activity, e.g. in terms of a communication overhead. This can be an external stimulus like a goal that is not reachable without cooperation either caused by resource or capability constraints or an internal stimulus like an extroverted character that is forced to cooperate with others [24]. These limitations can be found either on the perception level, the cognition level and/or the execution level [23, 26]. As examples, consider agents with sensory malfunction (perception level), humans with a disease like dementia (cognition level) or robots that are not able to overcome obstacles like stairs (execution level). In consequence they form a symbiotic relationship in which agents perform tasks for humans and humans in return help agents [23].

Human-Aware Planning Planning procedures that account for joint human-agent activities are computed by human-aware planning components. HAP is an evolving branch of AI planning systems for collaborative settings where agents coexists with humans [8] (e.g., socially assistive robots in household environments). Following *Cirillo et al.* [9, p. 17], human-aware planning can be applied in situations in which there are artificial agents whose actions can be planned and natural agents whose action can be only predicted, sharing the same environment. Here, to build effective team-players the agents are required to include the state of the human into their planning process to anticipate the actions of the human [14]. This information then can be used to generate plans including the human as actor and respecting a set of e.g., social or interaction constraints [8, 20].

3 Requirements

Advances in AI stimulate the development of more complex teamwork scenarios then those of these days, in which artificial agents become part of the team itself [22]. *Klein et al.* [18, 19] emphasises that such team-work requires the participants at least to enter into an agreement to work together (named Basic Compact), to be mutually predictable and directable and to maintain a common ground. Furthermore, the authors formulate several other challenges like observability and cost control, most of them beyond the scope of this work. However, two of the challenges—namely collaboration and interpredictability—are of particular interest. Collaboration addresses the decision making process of agents—the actual planning. It includes the ability to understand and solve a problem in an incremental fashion and negotiate the course of actions with other team-members. Requirements for such planning components include the abilities to [5, 6, 19]:

- reveal the current status of the overall plan,
- detect possible failures during plan execution and to replan if failure occur,

– evaluate the viability of plan changes,
– replan in situations where an individual agent's capabilities are outperformed,
– recruit more capable agents to perform the replanning,
– manage retasking when plan changes occur, and
– adjust the communication capability to the agents' capabilities (human agents require UIs where software agents expect procedural calls).

Additionally, we can identify some real time constraints. If the agent application takes too much time to act on behalf of a user (e.g., for plan generation), the human might feel misunderstood or might think that the application has malfunction [20]. These stop-and-go like interaction avoids the fluent meshing of actions, which is typically for good teamwork [13].
Furthermore, we have to consider the context-dependent behaviour of humans. That means, that whenever a human is predicted to fulfil a task, the human may perform this task or not and provides results either in time or delayed [1]. This is also called interpredictability and requires to plan actions considering the actions of others [6, 18]. For this, human-behaviour models either hand-crafted or derived from psychological studies can be used [16, 17]. Such models define the possible behaviour, capabilities and habits of humans and should be updated during runtime to adapt to the individuals that interact with the system.

4 Related Work

Several of the presented requirements can be imported 'out of the box' from state-of-the-art approaches either from the field of human-aware planning or from adjacent research fields [27]. For example, the ability to reveal the current status of the overall plan and to detect possible failures during the plan execution is given by the 3-layer architecture of current dynamic planning components [12, p. 9]. Here, a closed loop between the planning level, the monitoring level and the execution/controller level enables interleaved planning and execution, which is utilised for plan supervision, plan revision and replanning.
This multi-layered structure is for example used in the HAP framework presented by *Cirillo et al.* [8]. The approach describes the use of intention models to decide whether an agent is allowed to perform its task or if the agent would disturb the human user and should not perform the task now. This information is used to postpone agent tasks to a more acceptable time frame. This use case is interesting but beyond the scope of our work as it is our intention to consider the abilities of humans when planning and to assign tasks to them.
Another HAP approach—the Human Aware Task Planner (HATP)—is presented by *Montreuil et al.* [20]. HATP is able to estimate the viability of a plan according to several social constraints, reaching from undesirable states to effort balancing and abstraction legibility. An extension of HATP is presented by *Alili et al.* [4]. The work introduces a supervision layer that refines tasks based on the current context in an incremental fashion and monitors human behaviour,

i.e. whether a human user commits to a task assignment or not. This allows the system to recognise in which context a human user accepts or declines a specific task. Here, the work lacks detail on the usage of such information.
Alami et al. [3] present a way to adjust the planning procedures to different types of humans using so called InterActionAgents (IAA) as a representation of humans interacting with the system similar to proxies. The author further discuss a concept for a framework using the information provided by the IAAs to produce legible behaviour.[1]
Rosenthal et al. [23] emphasise how human-agent cooperation enables a team to accomplish tasks that the team members cannot fulfil on their own. The work makes extensive use of plan changes and retasking and also introduces ways to replan in situations where individual agent capabilities are outperformed. Models of human behaviour are not used.
The work of *Kirsch et al.* [16] uses models of human abilities to predict human behaviour and reactions and is in consequence able to produce plans that are socially acceptable. Furthermore, the authors state that these models are adjusted through a learning cycle. Nevertheless, the strengths of the system are not the planning techniques or the learning algorithms but the concept of combining two frameworks to facilitate joint human-agent activities.

5 Discussion

As mentioned above several of the presented requirements can be imported 'out of the box' from related work. Nevertheless, we are able to identify subjects that are merely covered by state-of-the-art solutions. Table 1 classifies the introduced approaches in comparison to our own study named the Hplan [2] (Human-Plan) project. The table includes the ability of the solutions to monitor the execution phase and to evaluate the viability of plan changes, which is subsumed using the term manage state. Furthermore, we evaluated the approaches abilities to detect failures during execution and to replan if necessary including retasking of the planned actions. The ability to recruit more capable agents to perform the replanning was not found in any of the approaches. One reason might be a missing exchange between research in e.g., mixed-initiative planning [7] and research in planning for joint human-agent activities. As the Hplan projects focuses interpredictability as research goal, we further evaluated whether the approaches take representation of humans into account. Here, we distinguish social constraints, behavioural models, intention models and whether these models are static or introduce some certain kind of learning.
Table 1 emphasises that technical requirements such as manage the current state, replan and the detection of failures are satisfied by most of the approaches. This is typically done by implementing the 3-layer architecture of dynamic planning components and establishing a life-cycle of planning, executing and monitoring. Additionally, we can conclude that there exist some work related to the

[1] *Kirsch et al.* [17] gives an introduction to the term 'legible behaviour' in the context of joint human-agent activities.

Table 1: A classification of human-aware planning approaches. The table entries reads as follow: + the approach supports this feature, o the approach does not fully support this feature but it supports it in some (weak) way (e.g. extension), – the approach does not support this feature.

	[20]	[4]	[16]	[8]	[23]	[3]	HPLAN
Manage State	+	+	o	+	+	o	+
Replanning	+	+	o	+	+	o	+
Failure Detection	+	o	o	o	+	o	+
Social Constraints	+	+	o	o	−	+	−
Behavioural Models	o	o	+	o	o	−	+
Intention Models	o	o	o	+	−	−	−
Learning	−	−	o	−	−	−	+

integration of specific information about humans. For example, *Montreuil et al.* [20] presents a solution that integrates social constraints into the actual planning process by assigning a social score to evaluate plans. This social score can comprise several dimensions. In contrast to social constraints, the integration of information about behaviour (e.g., personality traits) and abilities are not as well advanced. Even so, the different authors involved highlight the need for a novel representation of humans [8, 16]. The conceptual work of *Kirsch et al.* [16] presents a first approach to integrate such representations, which can be also adapted during runtime using learning techniques. The concept of InterActionAgents introduced by *Alami et al.* [3] presents a way to encapsulate different information about individuals and is a natural way when developing agent-oriented.

6 Conclusion

This article provided an overview about requirements for human-aware planning and available human-aware planning approaches. In particular, we focussed on the challenges associated with joint human-agent activities. It was argued that the decision making process of artificial agents can benefit when anticipating impending actions of humans. This requirement, which is also called interpredictability, implies shared knowledge between team-members developed through experience gained during the actual teamwork [6]. Here, a novel representation about human behaviour, abilities, habits, social rules and intentions is required. These representations can be derived, e.g., from social and psychological studies and should be adapted during the interaction with humans. Although the related work fulfils several of the presented requirements, the integration of such representations is mainly missing. It was shown that major advancement were made integrating social constraints into the planning process, that the integration of behavioural and intention models is a barely sufficient explored and that the online adaptation during the interaction is nearly non-existent nowadays. In future work, we want address these issues in the HPLAN project concatenating contemporary planning techniques with a human behavioural model and reinforcement learning techniques.

References

1. Ahrndt, S.: Improving human-aware planning. In: Klusch, M., Thimm, M., Paprzycki, M. (eds.) Multiagent System Technologies, pp. 400–403. No. 8076 in Lecture Notes in Artificial Intelligence, Springer (2013)
2. Ahrndt, S., Ebert, P., Fähndrich, J., Albayrak, S.: Hplan: Facilitating the implementation of joint human-agent activities. In: Demazeau, Y., Zambonelli, F., Corchado, J.M., Bajo, J. (eds.) Advances in Practical Applications of Heterogeneous Multi-Agent Systems. The PAAMS Collection. pp. 1–12. Springer (2014), to appear in June
3. Alami, R., Clodic, A., Montreuil, V., Sisbot, E.A., Chatila, R.: Task planning for human-robot interaction. In: Bailly, G., Crowley, J.L., Privat, G. (eds.) Proc. of the sOc-EUSAI2005. pp. 81–85. ACM Press (2005)
4. Alili, S., Warnier, M., Ali, M., Alami, R.: Planning and plan-execution for human-robot cooperative task achievement. In: Proc. of the 19^{th} ICAPS. pp. 1–6 (2009)
5. Allen, J., Ferguson, G.: Human-machine collaborative planning. In: Proceedings of the Third International NASA Workshop on Planning and Scheduling in Space. pp. 1–10 (2002)
6. Bradshaw, J.M., Feltovich, P., Johnson, M., Breedy, M., Bunch, L., Eskridge, T., Jung, H., Lott, J., Uszok, A., Diggelen, J.: From tools to teammates: Joint activity in human-agent-robot teams. In: Kurosu, M. (ed.) Human Centered Design, Lecture Notes in Computer Science, vol. 5619, pp. 935–944. Springer Berlin Heidelberg (2009), http://dx.doi.org/10.1007/978-3-642-02806-9_107
7. Burstein, M.H., McDermott, D.V.: Issues in the development of human-computer mixed-initiative planning systems. In: Gorayska, B., Meyer, J. (eds.) Cognitive Technology: In Search of a Humane Interface, chap. 17, pp. 285–304. Elsever Science B.V. (July 1996)
8. Cirillo, M., Karlsson, L., Saffiotti, A.: Human-aware task planning: An application to mobile robots. ACM Trans. Intell. Syst. Technol. 1(2), 1–26 (2010)
9. Cirillo, M.: Planning in Inhabited Environments: Human-Aware Task Planning and Activity Recognition. Ph.D. thesis, Örebro University, http://www.publications.oru.se/4DCGI/dropacta?r=23850&t=t&id=510 (November 2010)
10. Clark, H.H.: Using Language. Cambridge Univ. Press (1996)
11. Fähndrich, J., Ahrndt, S., Albayrak, S.: Self-explaining agents. Jurnal Teknologi (Science & Engineering) 63(3), 53–64 (2013)
12. Ghallab, M., Nau, D., Traverso, P.: Automated Planning: Theory & Practice. Morgan Kaufmann (2004)
13. Hoffman, G., Breazeal, C.: Effects of anticipatory action on human-robot teamwork – efficiency, fluency, and perception of team. In: Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction. pp. 1–8. HRI '07, ACM, New York, NY, USA (2007), http://doi.acm.org/10.1145/1228716.1228718
14. Jenkins, O.C., González, G., Loper, M.M.: Tracking human motion and actions for interactive robots. In: Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction. pp. 365–372. HRI '07, ACM, New York, NY, USA (2007), http://doi.acm.org/10.1145/1228716.1228765

15. Kephart, J.O., Chess, D.M.: The vision of autonomic computing. Computer 36(1), 41–50 (January 2003)
16. Kirsch, A., Kruse, T., Mösenlechner, L.: An integrated planning and learning framework for human-robot interaction. In: Proc. of the 19th ICAPS. pp. 1–6 (2009)
17. Kirsch, A., Kruse, T., Sisbot, E.A., et al.: Plan-based control of joint human-robot activities. KI – Künstliche Intelligenz 24(3), 223–231 (2010)
18. Klein, G., Feltovich, P.J., Bradshaw, J.M., Woods, D.D.: Common ground and coordination in joint activitiy. In: Rouse, W.B., Boff, K.R. (eds.) Organizational Simmulation, chap. 6, pp. 139–184. John Wiley & Sons (2005)
19. Klein, G., Woods, D.D., Bradshaw, J.M., Hoffmann, R.R., Feltovich, P.J.: Ten challenges for making automation a 'team player' in joint human-agent activity. Human-Centered Computing 19(6), 91–95 (2004)
20. Montreuil, V., Clodic, A., Alami, R.: Planning human centered robot activities. In: IEEE SMC. pp. 2618 –2623 (2007)
21. Müller-Schloer, C., Schmeck, H., Ungerer, T. (eds.): Organic Computing – A Paradigm Shift for Complex Systems, vol. 1. Springer (2011)
22. Neef, M., van Rijn, M., Keus, D., Mark, J.W.: Organizing smart networks and human into augmented teams. In: Harris, D. (ed.) Engineering Psychology and Cognitive Ergonomics, Lecture Notes in Computer Science, vol. 5639, pp. 213–222. Springer Berlin Heidelberg (2009)
23. Rosenthal, S., Biswas, J., Veloso, M.: An effective personal mobile robot agent through symbiotic human-robot interaction. In: Proceedings 9th Autonomous Agents and MultiAgent Systems (AAMAS). vol. 1, pp. 915–922 (May 2010)
24. Salvit, J., Sklar, E.: Modulating agent behavior using human personality type. In: Proceedings of the Workshop on Human-Agent Interaction Design and Models (HAIDM) at Autonomous Agents and MultiAgent Systems (AAMAS). pp. 145–160 (2012)
25. Simon, H.A.: The Sciences of the Artifical. MIT Press, 3 edn. (October 1996)
26. Sycara, K.P.: Multiagent systems. AI Magazine 19(2), 79–92 (1998), `http://www.aaai.org/ojs/index.php/aimagazine/article/view/1370`
27. Talamadupula, K., Kambhampti, S., Schermerhorn, P., Benton, J., Scheutz, M.: Planning in human-robot teaming. In: Proceedings of the 21th International Conference on Automated Planning and Scheduling (ICAPS'2011). pp. 82–89 (2011)
28. Wooldridge, M.: An Introduction to Multi-Agent Systems. Wiley Publishing, 2nd edn. (2009)