Challenges in Planning for Human Robot Teams

***Abstract*—Among the many anticipated roles for future robots is that of being a human teammate. Aside from all the technological hurdles that have to be overcome on the hardware and control sides to make robots fit for work with humans, the added complication here is that humans have many conscious and subconscious expectations of their teammates – teaming is mostly a *cognitive* rather than *physical* coordination activity. This focus on cognitive coordination, however, introduces new challenges for robotics that require fundamental changes to the traditional view of autonomous agents.**

In this paper, we provide an analysis of the differences between traditional autonomous robots and robots that team with humans, identifying the necessary teaming capabilities that are largely missing from current robotic systems. We then focus on the important challenges that are unique and of particular importance to human-robot teaming, particularly from the point of view of the deliberative process of the autonomous agent, and sketch potential ways to address them.

I. INTRODUCTION

An increasing number of applications demand that humans and robots work together. Although a few of these applications can be handled through "teleoperation", technologies that act in concert with the humans in a teaming relationship with increasing levels of autonomy are often desirable if not required. Even with a robust human-robot interface, robots will need to be good team players and exhibit characteristics common in human-human teams, such as the ability to recognize the intentions of the human teammates, and to interact in a way that is comprehensible to them. At some level, autonomous robots need to understand human behavior and adapt to that behavior in an efficient manner, much like humans adapt to the behavior of other humans.

Take, for example, a search and rescue team where one searcher is in need of a tool to open a blocked passage way and another searcher upon watching the first immediately dashes over to provide help with lifting a heavy pole that could be used to pry open the broken door. Or take the surgery assistant that proactively picks a gauze to press down on the bleeding tissue after the surgeon finished making the incision. These are two simple, but typical examples of how humans coordinate their activities in teams, especially experienced teams, often without the need of explicit communication during coordination. It is this type of coordination that contributes to the highly efficient teaming between humans.

The reason why humans are able to perform these proactive steps and support others in their goals is the human ability to quickly (1) recognize teaming context in terms of the current status of the team (i.e., status of the team task and states of the teammates), (2) anticipate next team behavior under the current context to decide their individual subgoals to be achieved for the team, and (3) take proper actions to support the advancement of those subgoals with the consideration of the other teammates. Within these steps, explicit communication may be used to facilitate the collaboration when necessary.

The three steps above form an integrated loop such that they are intermixed with each other during the coordination process. This loop allows humans to adapt to changing circumstances to update their understanding about the current status of the team, infer about teammate intentions, goals and preferences to predict their behavior in the context of the team task, and act in ways that are consistent with the expected teaming behavior. Furthermore, this process is constantly evolving for the team during teaming experience. Critically, humans will likely expect all of the above capabilities from a robotic teammate, as otherwise team dynamics will suffer.

As such, the challenge in human-robot teaming is primarily *cognitive*, rather than physical. Cognitive teaming allows the robots to adapt more proactively to the many conscious and subconscious expectations of their human teammates. At the same time, improper design of such robot autonomy could increase the human's cognitive load, leading to the loss of teaming situation awareness, misaligned coordination, poorly calibrated trust, and ultimately slow decision making, deteriorated teaming performance, and even safety risks to the humans. As designers of robotic control architectures, we thus have to first isolate the necessary functional capabilities that are common to realize such autonomy for teaming robots. The aim of the article is to do just that and thus provide a framework that can serve as the basis for the development of cognitive robotic teammates.

II. RELATED WORK

In human-human teams, it is well understood that every team member maintains a cognitive model of the other teammates they interact with [13]. These models not only captures their physical states, but also mental states such as the teammate intentions and preferences, which can significantly influence how an agent interacts with the other agents in the team. Although such modeling has been identified as an important characteristic of effective teaming [12], [14], [25], it is less clear how it is maintained at the individual level. Furthermore, the relative importance of different aspects of such models cannot be easily isolated in experiments with human teammates, but must be separately considered for robots since they often require very different modeling technologies.

Such forms of modeling is critical for the design of robotic teammates that work alongside humans. It allows the robots to understand about their human partners, and in turn use this knowledge to plan their coordination to improve teaming experience. However, although there exists work that has investigated the various aspects of this modeling [16], [54], [64], a systematic summary of the important challenges is still missing.

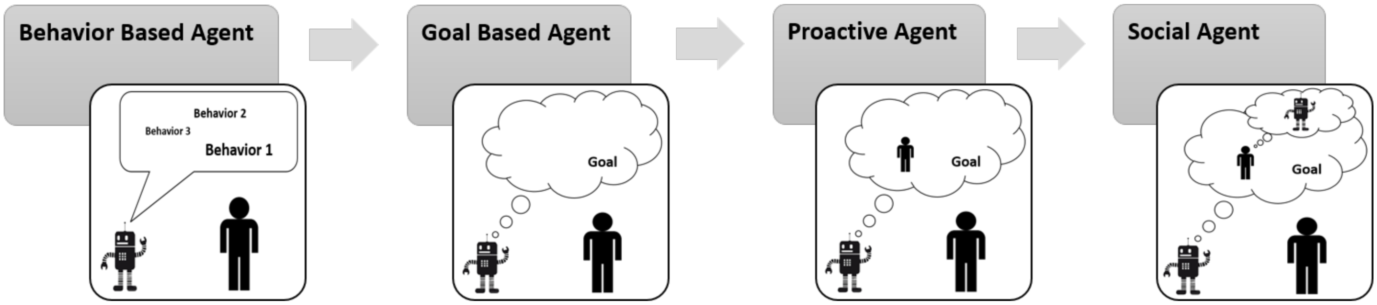

Fig. 1. An categorical view of different types of agents. Each type is deeper than the previous types in terms of modeling complexity (from left to right).

Next, we provide a review of the related work in terms of agent types (Figure 1) that can be used to implement robotic teammates, following the categorization in [59]. The first two types correspond to classical agent architectures in robotics and artificial intelligence. We list them to facilitate the comparison of earlier teaming agents and cognitive teaming agents. We show that all of them fall within a spectrum of agents that differ based on the extent to which the agent interaction with the external world and other agents is modeled.

A. Behavior-based agent

Behavior-based agents [5] have been an important design paradigm for embodied agents (especially robots), in which complex behaviors result from a collection of basic behaviors interacting with each other. These basic behaviors often operate in parallel via cooperative or competitive arbitration schemes [52]. Behavior-based agents have been applied to various tasks such as formation control [2], [3], box pushing [42], [22], navigation [43], and surveillance [60]. One issue with behavior-based agent is that the interactions between basic behaviors often have to be provided manually. This can quickly become impractical when complex interactions are desired. Furthermore, since this type of agent does not maintain a model of the world, it cannot reason about its dynamics and hence is often reactive. Hence, mere behavior-based agents are usually restricted to simple tasks.

B. Goal-based agent

In contrast to behavior-based agents, a goal-based agent maintains a model of the world (Fig. 2) and how its actions can change the state of the world. As a result, it can predict how the world will respond before it executes any action. The earliest agent of this type is Shakey [40]. The model that is maintained is often specified at a factored level using planning languages, such as STRIPS [17], PDDL [18] or its probabilistic extensions [49], or at the atomic level using MDP specification [45], [16]. A goal-based agent can also maintain its own epistemic state [19], such as beliefs and desires [47], [21]. Goal-based agents typically assume that the given model is complete, which may not be realistic in open world domains [55].

Remark: Both behavior and goal based agent can handle multi-agent coordination [60], [43], [2], [41], [26]. However, it is often assumed that the team is given a specific goal, and the team members either are provided information about each other a priori, or can explicitly exchange such information. As a result, agents can readily maintain a model of the others in teaming. While this assumption may be true for robots teaming with robots, we would definitely not observe such convenience in human-robot teams (e.g., requiring humans to provide the information constantly can significantly increase their cognitive load); furthermore, the goal is often spontaneous rather than given. As a result, a robotic teammate that is solely behavior or goal based can only handle specific tasks and will rely on human inputs for task assignments.

C. Proactive agent

A proactive agent, on the other hand, is supposed to maintain a model of the others through both observations and communications (if available and necessary). This model is not only about the others' physical state (e.g., location), but also mental state which includes their goals [56], capabilities [61] (which include the consideration of physical capabilities), preferences [38], and knowledge [4].

Given that none of these are directly available, they must be inferred [46] or learned [61], [1], [6] from observations. As a result, the model of the other agents is often subject to a high-level of incompleteness and uncertainty. This is especially true when human teammates are involved. Nevertheless, even such an *approximate* model of other agents can be important for efficient teaming when used properly. For example, it can be used by the agents to plan for coordination to exploit opportunities to help the humans in a proactive way [8], [16] while avoiding conflicts [10], [9].

Remark: In addition to using the model of the others to plan for coordination, a proactive agent can also act proactively to change the others' modeling of itself when necessary [4]. For example, a robot can explicitly convey its intention through natural languages [57] or gestures [44] to let the human understand its intention to help or request for help.

D. Social agent

A deeper level of modeling is not only about the other agents, but also the other agents' modeling of the agent itself

[64]. This includes, for example, the others' expectation and trust of the agent itself. Such modeling allows the robot, for example, to infer the human expectation of its own behavior and in turn choose behaviors that are consistent with this expectation. Expectation and trust, in particular, represent the social aspects of agent interactions since they are particularly relevant when agents form groups or teams together. An agent that behaves socially [33], [15], [64] allows the other agents to better understand and anticipate its behavior, thus contributing to the maintenance of teaming situation awareness [12]. In human-human teams, social behaviors contribute significantly to fluent teaming [51].

Remark: Similar to a proactive agent, a social agent often has to learn and maintain a model about the various social aspects from observations [64]. In addition to using these social aspects to guide its behavior generation, a social agent can also act to change these aspects (via informing the others about the discrepancies in their modeling about itself *AND* updating the same in its model of the others). For example, a robot can maintain its trust from the human by constructing excuses when a task cannot be achieved [24] to provide explanations to the human from the robot's own perspective, while taking into account the human's understanding of itself.

Section Remark: Although various types of agents can be used to realize a robotic teammate, based on the above discussion, the challenges that are introduced by the humans in the loop lie in particular in the implementation of a proactive and social agents. A common characteristics among these two types of agents is that both require the notion of ***mental modeling*** of the other teammates, which cannot be directly observed and must be inferred cognitively. This is the key requirement of a cognitive teaming capability.

III. Exemplary Human-Robot Teaming Scenario

To better illustrate how mental modeling of teammates can contribute to the different capabilities needed for cognitive teaming agents, we will now consider scenarios from a mixed human-robot team performing an Urban Search and Rescue (USAR) task where each subteam i consists of one human H_i and one robot R_i.

For subteam 1: Based on the floor plan of the building in its search area, R_1 realizes that the team needs to use an entrance to a hallway to start the exploration. R_1 notices that a heavy object blocks the entrance to the hallway. Based on its capability model of H_1 (i.e., what H_1 can and cannot lift) and H_1's goal, R_1 decides to interrupt its current activity and move the block out of the way. H_1 and R_1 then continue exploring different parts of the area independently when H_1 discovered a victim and informs R_1. R_1 understands that H_1 needs to get a medical kit to be able to triage this victim as soon as possible but knows that H_1 does not know where a medical kit is located. Since R_1 has a medical kit already, but cannot deliver it due to other commitments, it places its medical kit along the hallway that it expects H_1 to go through, and informs H_1 of the presence of the kit.

For subteam 2: Based on the floor plan of the building in its search area, R_2 finds that all the entrances are automatic doors that are controlled from the inside. Since the connection cannot be established due to power lost, the team needs to break a door open first. R_2 infers that H_2 is about to break a door open based on the teaming context and its observations. Since it knows that breaking the door open may cause a board to fall on H_2, R_2 moves to catch the board preventatively. Once H_2 and R_2 are inside, however, H_2 is uncertain about the structural integrity and has no information on which parts may easily collapse. R_2 has access to the building structure information and proposes a plan to split the search in a way that minimizes human risk.

For both subteams: As both teams are searching their areas, they receive information about a third area to be explored. Since neither H_1 nor H_2 are finished with their current search task, they assume that the other will take care of the third area. Since R_1 understands H_1 and H_2's current situation, and expects itself to be done with its part of the task soon, R_1 decides to work on the third area since it does not expect H_1 to need any help. R_1 informs H_1. H_1 is OK with it and informs H_2 that team 1 is working on the third area. When R_1 arrives at the third area, it notices new situations which require certain equipment from team 2. R_1 communicates with R_2 about the availability of the missing items. R_2 quickly predicts equipment needs and anticipates that those items are not needed for a while. After quickly getting the OK from H_2 to lend the equipment to R_1, R_2 drives off to meet R_1 half-way, hand over the equipment, and R_1 returns to the third area with the newly acquired equipment. H_1 was not informed during this process since R_1 understands that H_1 has a high workload. Once the equipment is no longer needed, R_1 meets up with R_2 again, returning the equipment in time for H_2 to have it available.

Based on the above scenario, we can see that the mental modeling of the others on a cognitive robotic teammate contributes to the fluent operation of the team:

Modeling the human mental state: For example, R_1 needs to understand the capabilities of H_1 (i.e., what H_1 can and cannot lift); both R_1 and R_2 need to be able to infer about the intention of the human teammates. The modeling may also include the human's knowledge, belief, mental workload, trust and etc.

Furthermore, we connect this human mental modeling with the three capabilities that are critical in human-human teams (see introduction) by showing that they are enabled by such modeling on a cognitive robotic teammate:

C1. Recognizing teaming context to identify the status of the team task and states of the teammates: For example, based on the floor plan of the building, R_1 realizes that the team needs to use an entrance to a hallway to start the exploration. R_2 finds that all the entrances are automatic doors that are controlled from the inside. Consequently, it infers that the team needs to break a door open first. This inference process takes into account the modeling of the teammate's state (e.g., the intention to enter the building).

C2. Anticipate team behavior under the current context: For example, given that a heavy object blocks the entrance to the hallway, R_1 infers that the human will be finding a way to clear the object. R_2 infers that H_2 is going to break a door

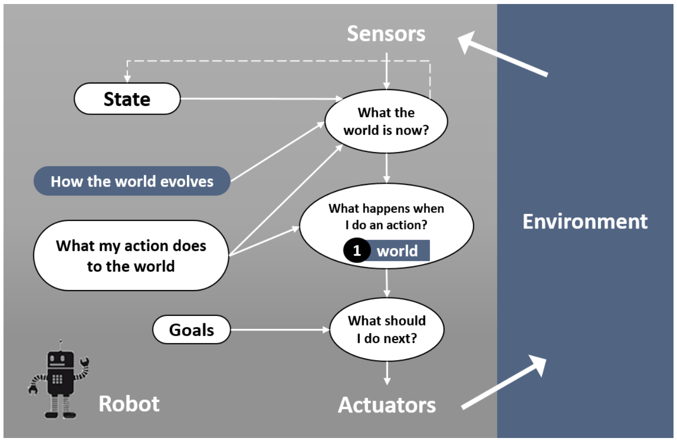

Fig. 2. Traditional view of the goal-based intelligent agent architecture [48] that describes how the agent models the world, senses changes in the environment, plans to achieve goals and acts to execute the plans.

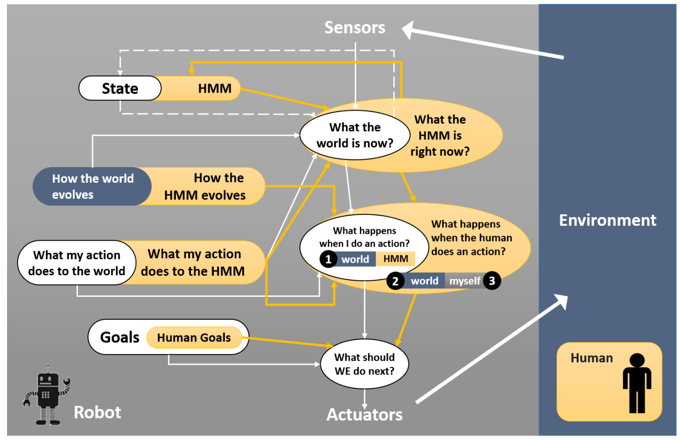

Fig. 3. An updated view of the architecture of a cognitive teaming agent acknowledging the need to account for the human's mental state by means of what we refer to as **Hu**man **M**ental **M**odeling or HuMM.

open based on the teaming context and its observations. This prediction takes into account of the modeling of the human's capabilities and knowledge about the teaming context.

C3. Take proper actions to advance the team goal while taking into account the teammates: For example, after anticipating the human's plan, the robots should proactively help the humans (e.g., R_1 helps H_1 move the block away and R_2 catches the board preventively that can potentially hurt the human), while taking the account the modeling of the human's capabilities, mental workload, and expectation.

Remark: C3 above not only includes actions that contribute to the team goal, but also actions for maintaining teaming situation awareness (e.g., making explanations). As such, C3 feeds back to C1 and the three capabilities in turn form a loop that should be constantly exercised to achieve fluent teaming. Furthermore, although we have been focusing on implicit communication (e.g., through observing behaviors) to emphasize the importance of mental modeling, explicit communication (e.g., using natural language) is also an important part of the loop. Another note is that since both implicit and explicit communication can update the modeling of the other teammates' mental states as discussed, they are anticipated to evolve the teaming process in the long term.

IV. Transitioning to Cognitive Teaming Agent

The mental modeling of the other agents (in particular humans in this case) changes the view of an agent. In this section, we start with a typical view of a goal-based agent as depicted in Fig. 2 and consider the extensions that have to be made to this generic agent view to enable cognitive teaming agent to support the various capabilities we described earlier. In particular, our objective is to characterize how the "*Sense-Model-Plan-Act*" (SMPA) cycle of a typical goal-based robotic agent will change for teaming agents. Next, we provide an overview of the new aspects of this "*Sense-Model-Plan-Act*" cycle, focusing on the changes that must be brought into each functionality:

Sense: The robotic agent can no longer sense passively to check that the preconditions of an action are satisfied, or after it applies an action to the world to confirm that it is updated accordingly ("what the world is like now" in Fig. 2). In teaming scenarios, the robotic agent needs to proactively make complex sensing plans that interact closely with other functionalities (i.e., *Model* and *Plan*) to maintain a proper mental state (such as intentions, knowledge and beliefs) of its human teammates in order to infer about their needs. For example, how the robot should behave is dependent on how much and what type of help the human requires, which in turn depends on the observations about the human teammates such as their behavior and workload. Furthermore, the inference about the human mental state should be informed by the human model that the robot maintains about the human's capabilities and preferences. For example, if a human teammate does not drink coffee, the robot should probably consider tea instead when asked to fetch a drink. Note that directly asking humans (i.e., explicit communication) should also be considered as a specific form of sensing.

Model: Correspondingly, "what the world is like now" needs to include not only environmental states, but also mental states of the team members (including humans and maybe even other robots under certain situations). Hence, we need an augmented state model that can also represent the mental, in addition to the physiological (e.g., human temperature, levels of exhaustion, bodily stress, and remaining energy) and physical states of human teammates (e.g., maximum running speed, and lifting strength). Here, mental states may not only include cognitive and affective states such as the human's task-relevant beliefs, goals, preferences, and intentions, but also, more generally, emotions, workload, expectations, trust and etc. "What my actions do to the world" then needs to include the effects of the robot's actions on the team member's mental state, in addition to the effects on their physiological and physical states and the observable environment; "How the world evolves" now also requires rules that govern the evolution of agent mental states based on their interactions with the world (including information exchange through communication); "What it will be like" will thus be an updated state presentation that not only captures the world state, agent physiological and physical state changes based on their actions and current states, but also those mental state changes caused by the agent itself and other team members.

Plan: "What action I should do" now involves more complex decision-making that must also consider human mental state. Furthermore, since the robot actions now can influence not only the state of the world but also the mental state of the humans, the planning process must also consider how the actions may influence their mental state and even how to manipulate such mental state. For example, in teaming scenarios, it is important to maintain a shared mental state between the teammates. This requires the robots to generate behavior that is expected and predictable to the human teammates such that the humans would be able to understand the robot intention. On the other hand, a shared mental state does not necessarily mean that every piece of information needs to be synchronized. Given the limitation on human cognitive load, sharing only necessary information is more practical between different teammates working on different parts of the team task. A properly maintained shared mental state between the teammates can contribute significantly to the efficiency of teaming since it can reduce the necessity of explicit communication in many situations.

Act: In addition to physical actions, we now also have communicative actions that can change the mental state of the humans by changing their beliefs, intents, etc. As we mentioned above, actions to affect the human's do not have to be linguistic (direct); stigmergic actions to instrument the environment, for example, leaving a tool for a person, can also inform the humans such that their mental states can be changed. Given that an action plan is eventually realized via the activation of effectors by providing motor commands, *Act* must be tightly integrated with *Plan*. While *Plan* generates the sequence of actions to be realized, motor commands can create different motion trajectories to implement each action and can in turn impact how the plan would be interpreted since different realizations can exert different influences on the human's mental states based on the context. Hence, it is important to carefully choose the motor commands so that the trajectories are "legible" to the humans for them to understand the behaviors of the robot at the motion level.

Remark: Changes to "model" in Figure 3 is a direct result of the requirement of human mental modeling. Coarsely speaking, changes to "sense" contribute to the recognition of teaming context, changes to "plan" contribute to the anticipation of team behavior and "act", and changes to "act" contribute to the determination of proper actions at both the action and motion levels. In practice, these four functionalities are tightly integrated in the behavior loop.

V. CHALLENGES

The desired capabilities for teaming robots that are reflected in this new view of agent present several challenges for robotic teammates which have not been addressed yet:

A. Learning Human Mental Model

The first challenge in cognitive teaming is for robots to create and maintain models of the human teammates that capture their capabilities, preferences, intentions and etc. This naturally occurs between human teammates as they work collaboratively in the environment. These models are inherently partial from the robot's perspective and must be learnable. As a result, they are often subject to a high level of incompleteness. However, most work on planning has hitherto focused on complete world models [23], [20], [27], while most real-world scenarios, especially when they involve humans, have a critical defining property that eliminates these standard planning and problem-solving methods: they are *open-ended* in that planning agents typically do not have sufficient knowledge about all task-relevant information (e.g., human models) at planning time – in other words, the planning models would be incomplete [28]. Despite being incomplete, such models must support reasoning as well as being improvable from sensing (which includes explicitly asking the human). Hence, an important challenge is to develop representations of approximate and incomplete models that are easy to learn (for human mental modeling) and can support planning/decision-making (for anticipating human behavior).

There exists prior work that introduced various incomplete models (Fig. 4). They differ in the information that is available for model learning, as well as how planning is performed. Some of them start with complete action models, such as STRIPS, and modified them with possible precondition/effect annotations to support incompleteness [36], [37], [66], [65], [67]. Although these models support principled approaches for robust planning, they are still quite difficult to learn. On the other end of the spectrum, we can have very shallow models [58] that assume no structured information at all which are used mainly in short-term planning support. There are also partial models that are somewhere in between [61], which has more structured information while is still easy to learn. However, planning under this model is incomplete. These prior models only form a starting point of work in this direction. Research effort still remains to be invested to introduce other incomplete models, analyze their applicability, and apply them to various domains.

B. Human-aware Planning

Most traditional approaches to planning focus on one-shot planning in closed worlds given complete domain models. While even this problem is quite challenging, and significant strides have been made in taming its combinatorics, planners for robots in human-robot teaming scenarios require the ability to be *human-aware*. Human-aware planning is especially challenging since it must not only consider human physical state, but also their mental state as we discussed, which are inherently incomplete. Furthermore, different human models capture different aspects of the human (e.g., capabilities [61], intentions [54] and emotions [50]) which are closely interrelated, and it is not clear how they can be combined. In the robotics and automated planning communities, there has been some work on human-aware planning, from the aspect of both path planning [53], [31], [8], [9] and task planning [29], [11], with the intention of making robot plans more socially acceptable (e.g., resolving conflicts with the plans of

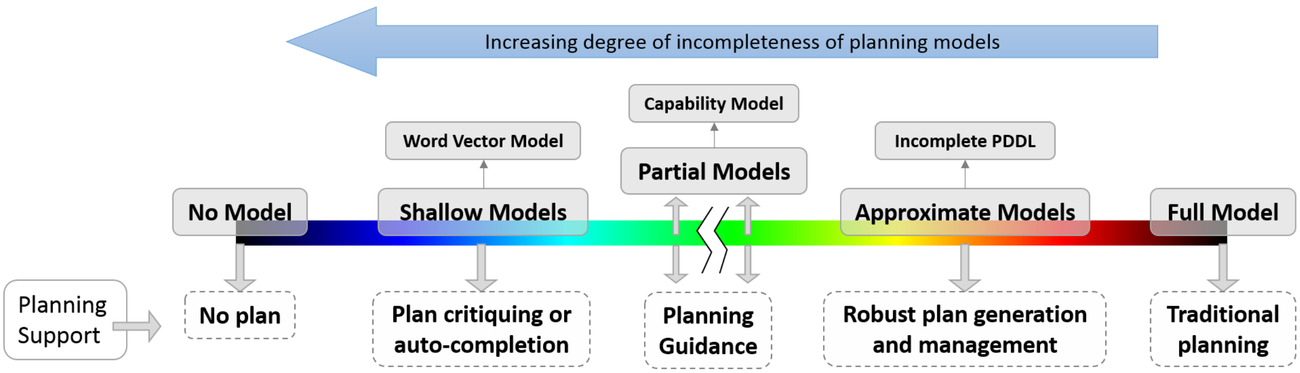

Fig. 4. A schematic view of the different classes of incomplete models and the relationships between them in the spectrum of incompleteness [58].

fellow humans). However, most of these approaches assume that certain aspects of the human models are provided, and they are accurate and complete.

Explainable Planning: There exists some recent work on planning while considering the human expectation [64], [63], [32] with the human model learned approximately from labeled plan traces. This work introduced a plan explicable measure [64] that captures the human expectation of the robot, which can be used by agents to proactively choose, or directly incorporated into the planning process to generate plans that are more comprehensible without affecting the quality too much. This is, however, only a start in this research direction. Performing human-aware planning with incomplete models remains an important challenge in human-aware planning, especially given that we do not yet understand how these different human models interact.

Plan Explanations:

C. Long-term Teaming

We have so far only discussed how a robot can maintain the human mental models. However, in teaming, this modeling is bi-directional. When robots do not have certain information, the robot can plan to sense and communicate with the human. You cannot always expect such behavior from humans when inconvenience can be caused. Instead, in many cases when the human is expected to have insufficient information about the robot, the robot needs to proactively communicate its model to the human. This communication can be, for example, about the intention, plan, explanations and excuses of behavior [24] for the robot; communication can involve different modalities such as visual projection, natural languages, gesture signaling, and a mixture of them. This capability is important for the human and robot to evolve their mental models to improve teaming in the long term. Note that the models (i.e., the human actual model and its representation on the robot) are not required to be aligned, which is often only applicable for repetitive tasks [39]. Much existing work on robots communicating with humans using different modalities can be utilized [30], [57]. However, a more critical challenge for robotic teammate is to compute when, what, and how to communicate for model adaptation for long-term teaming. Communicating too much information can increase the cognitive load of the human teammates while communicating too little can decrease the teaming situation awareness.

When designing robotic teammate, it must also be realized that many human-robot teaming tasks are not only complex, but can also span multiple episodes for an extended period of time. In such scenarios, the system's performance is dependent on how the teams perform in the current task, as well as how they perform in future tasks. A prerequisite to consider long-term teaming is to maintain mental states of the agents (e.g., trust) that influence their interactions, and analyze how these states dynamically affect the teaming performance and how they evolve over time.

D. Human-factors Analysis and Evaluation

Another challenge is the analysis of human-human and human-animal teams, as well as the evaluation of developed robotic systems to verify their performance. While this may seem less obvious, it is an important part of the equation to develop cognitive robotic teammate. Although the teamwork literature has identified characteristics of effective and ineffective teams, it is less clear what characteristics are essential at the individual level to be a good team player. What are the most important behaviors that a teammate must exhibit for the team to be effective? For instance, how should we tradeoff closed-loop-communication with the development of a good understanding of team members' intentions and capabilities? The challenge here is to understand the relative importance of these aspects under different contexts. Such analysis would undoubtedly help with the development of robotic teammates.

In terms of evaluation, it is well known that automation can have both positive and negative effects on human performance. Fortunately, there are many theories and findings from the human teaming literature that are relevant to human-robot teaming. For instance, we know that team effectiveness is tied to shared mental models [7], [34], team situational awareness [25], and interaction [13]. Although it seems only naturally to assume that effective robotic team players will need to have similar capabilities for understanding their teammates and the situation and for communicating their actions to teammates, and there exists work that tried to partly verify this claim [35], [62], it still remains to be evaluated more extensively.

VI. Conclusion

In this paper, we discussed the challenges in design of autonomous robots that are cognizant of the cognitive aspects of working with human teammates. We argued that traditional goal-based and behavior-based agent architectures are insufficient for building robotic teammates. Starting with the traditional view of a goal-based agent, we expand it to include a critical missing component: human mental modeling. We discussed the various tasks that are involved when such models are present, along with the challenges that need to be addressed to achieve these tasks. We hope that this paper can serve as guidance for the development of robotic systems that can enable more natural teaming with humans.

References

[1] Pieter Abbeel and Andrew Y. Ng. Apprenticeship learning via inverse reinforcement learning. In *Proceedings of the Twenty-first International Conference on Machine Learning*, ICML '04, pages 1–, New York, NY, USA, 2004. ACM.

[2] R.C. Arkin. Motor Schema – Based Mobile Robot Navigation. *International Journal of Robotics Research*, 8(4):92–112, Aug. 1989.

[3] T. Balch and R.C. Arkin. Behavior-based formation control for multirobot teams. *IEEE Transactions on Robotics and Automation*, 14(6):926–939, Dec. 1998.

[4] Chitta Baral, Gregory Gelfond, Enrico Pontelli, and Tran Cao Son. An action language for multi-agent domains: Foundations. *under submission to Artificial Intelligence*, 2015.

[5] Rodney Brooks. Intelligence without representation. *Artificial Intelligence*, 47:139–159, 1991.

[6] Benjamin Burchfiel, Carlo Tomasi, and Ronald Parr. Distance minimization for reward learning from scored trajectories. In *National Conference on Artificial Intelligence*.

[7] J.A. Cannon-Bowers, E. Salas, and S. Converse. Shared mental models in expert team decision making. *Current issues in individual and group decision making*.

[8] T. Chakraborti, G. Briggs, K. Talamadupula, Yu Zhang, M. Scheutz, D. Smith, and S. Kambhampati. Planning for serendipity. In *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*, pages 5300–5306, 2015.

[9] Tathagata Chakraborti, Yu Zhang, David Smith, and Subbarao Kambhampati. Planning with resource conflicts in human-robot cohabitation. In *AAMAS*, 2016.

[10] M. Cirillo, L. Karlsson, and A. Saffiotti. Human-aware task planning for mobile robots. In *Advanced Robotics, 2009. ICAR 2009. International Conference on*, pages 1–7, June 2009.

[11] Marcello Cirillo, Lars Karlsson, and Alessandro Saffiotti. Human-aware task planning: An application to mobile robots. *ACM Trans. Intell. Syst. Technol.*, 1(2):15:1–15:26, December 2010.

[12] N. J. Cooke. Team cognition as interaction. *Current Directions in Psychological Science*, 2015.

[13] N. J. Cooke, J. C. Gorman, C. W. Myers, and J.L. Duran. Interactive team cognition. *Cognitive Science*, 2013.

[14] N. J. Cooke and M. L. Hilton. Enhancing the effectiveness of team science. *National Academies Press*, 2015.

[15] Anca Dragan and Siddhartha Srinivasa. Generating legible motion. In *Proceedings of Robotics: Science and Systems*, 2013.

[16] Alan Fern, Sriraam Natarajan, Kshitij Judah, and Prasad Tadepalli. A decision-theoretic model of assistance. *J. Artif. Int. Res.*, 50(1):71–104, May 2014.

[17] Richard E. Fikes and Nils J. Nilsson. Strips: A new approach to the application of theorem proving to problem solving. In *Proceedings of the 2Nd International Joint Conference on Artificial Intelligence*, IJCAI'71, pages 608–620, San Francisco, CA, USA, 1971. Morgan Kaufmann Publishers Inc.

[18] M. Fox and D. Long. PDDL2. 1: An extension to PDDL for expressing temporal planning domains. *Journal of Artificial Intelligence Research*, 20(2003):61–124, 2003.

[19] Andr Fuhrmann. *The Journal of Symbolic Logic*, 57(4):1479–1481, 1992.

[20] Hector Geffner and Blai Bonet. *A Concise Introduction to Models and Methods for Automated Planning*. Synthesis Lectures on Artificial Intelligence and Machine Learning. Morgan & Claypool, 2013.

[21] Michael Georgeff, Barney Pell, Martha Pollack, Milind Tambe, and Michael Wooldridge. *The Belief-Desire-Intention Model of Agency*, pages 1–10. Springer Berlin Heidelberg, Berlin, Heidelberg, 1999.

[22] B.P. Gerkey and M.J. Mataric. Sold!: Auction methods for multi-robot coordination. *IEEE Transactions on Robotics and Automation, Special Issue on Multi-robot Systems*, 18(5):758–768, 2001.

[23] Malik Ghallab, Dana Nau, and Paolo Traverso. *Automated planning: theory & practice*. Access Online via Elsevier, 2004.

[24] Moritz Göbelbecker, Thomas Keller, Patrick Eyerich, Michael Brenner, and Bernhard Nebel. Coming up with good excuses: What to do when no plan can be found. In *ICAPS*, pages 81–88, 2010.

[25] J.C. Gorman, N.J. Cooke, and J.L. Winner. Measuring team situation awareness in decentralized command and control environments. *Ergonomics*, 49:1312–1325, 2006.

[26] Barbara J. Grosz and Sarit Kraus. Collaborative plans for complex group action. *Artif. Intell.*, 86(2):269–357, October 1996.

[27] Subbarao Kambhampati. Refinement planning as a unifying framework for plan synthesis. *AI Magazine*, 18(2):67–97, 1997.

[28] Subbarao Kambhampati. Model-lite planning for the web age masses: The challenges of planning with incomplete and evolving domain models. In *AAAI*, 2007.

[29] Uwe Koeckemann, Federico Pecora, and Lars Karlsson. Grandpa hates robots - interaction constraints for planning in inhabited environments. In *Proc. AAAI-2010*, 2014.

[30] Thomas Kollar, Stefanie Tellex, Deb Roy, and Nick Roy. Toward understanding natural language directions. In *International IEEE/ACM Conference on Human-Robot Interaction*, 2010.

[31] Markus Kuderer, Henrik Kretzschmar, Christoph Sprunk, and Wolfram Burgard. Feature-based prediction of trajectories for socially compliant navigation. In *Proceedings of Robotics: Science and Systems*, Sydney, Australia, July 2012.

[32] Anagha Kulkarni, Tathagata Chakraborti, Yantian Zha, Satya Gautam Vadlamudi, Yu Zhang, and Subbarao Kambhampati. Explicable robot planning as minimizing distance from expected behavior. *CoRR*, abs/1611.05497, 2016.

[33] Jim Mainprice, E Akin Sisbot, Thierry Siméon, and Rachid Alami. Planning safe and legible hand-over motions for human-robot interaction. *IARP workshop on technical challenges for dependable robots in human environments*, 2(6):7, 2010.

[34] J. E. Mathieu, T. S. Heffner, G. F. Goodwin, E. Salas, and J. A. Cannon-Bowers. The influence of shared mental models on team process and performance. *Journal of Applied Psychology*, 2000.

[35] Vignesh Narayanan, Yu Zhang, Nathaniel Mendoza, and Subbarao Kambhampati. Automated planning for peer-to-peer teaming and its evaluation in remote human-robot interaction. In *ACM/IEEE International Conference on Human Robot Interaction (HRI)*, 2015.

[36] Tuan Nguyen and Subbarao Kambhampati. A heuristic approach to planning with incomplete STRIPS action models. In *International Conference on Automated Planning and Scheduling*, 2014.

[37] Tuan Nguyen, Subbarao Kambhampati, and Minh B. Do. Synthesizing robust plans under incomplete domain modesl. In *Proc. Neural Information Processing Systems (NIPS)*, 2013.

[38] Tuan Anh Nguyen, Minh Do, Alfonso Emilio Gerevini, Ivan Serina, Biplav Srivastava, and Subbarao Kambhampati. Generating diverse plans to handle unknown and partially known user preferences. *Artificial Intelligence*, 190(0):1 – 31, 2012.

[39] Stefanos Nikolaidis and Julie Shah. Human-robot cross-training: Computational formulation, modeling and evaluation of a human team training strategy. In *Proceedings of the 8th ACM/IEEE International Conference on Human-robot Interaction*, HRI '13, pages 33–40, Piscataway, NJ, USA, 2013. IEEE Press.

[40] Nils J. Nilsson. Shakey the robot. Technical report.

[41] Raz Nissim, Ronen I. Brafman, and Carmel Domshlak. A general, fully distributed multi-agent planning algorithm. In *AAMAS*, pages 1323–1330, Richland, SC, 2010. International Foundation for Autonomous Agents and Multiagent Systems.

[42] L.E. Parker. ALLIANCE: an architecture for fault tolerant multirobot cooperation. *IEEE Transactions on Robotics and Automation*, 14(2):220–240, 1998.

[43] L.E. Parker and F. Tang. Building multirobot coalitions through automated task solution synthesis. *Proceedings of the IEEE*, 94(7):1289–1305, Jul. 2006.

[44] Dennis Perzanowski, Alan C. Schultz, and William Adams. Integrating natural language and gesture in a robotics domain. In *Proceedings of the 1998 IEEE International Symposium on Intelligent Control*, 1998.

[45] Martin L. Puterman. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc., New York, NY, USA, 1st edition, 1994.

[46] Miquel Ramírez and Hector Geffner. Plan recognition as planning. In *IJCAI*, pages 1778–1783, 2009.

[47] Anand S. Rao and Michael P. Georgeff. BDI Agents: From Theory to Practice. In *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95*, pages 312–319, 1995.

[48] Stuart J Russell, Peter Norvig, and Ernest Davis. *Artificial intelligence: a modern approach*. Prentice Hall, 3 edition, 2010.

[49] Scott Sanner. Relational dynamic influence diagram language (rddl): Language description, 2011.

[50] M. Scheutz and P. Schermerhorn. Affective goal and task selection for social robots. In J. Vallverdú and D. Casacuberta, editors, *Handbook of Research on Synthetic Emotions and Sociable Robotics: New Applications in Affective Computing and Artificial Intelligence*. Idea Group Inc., 2009.

[51] Matthias Scheutz, Julie Adams, and Scott DeLoach. A framework for developing and using shared mental models in human-agent teams. *JCEDM*, forthcoming.

[52] Matthias Scheutz and Virgil Andronache. Architectural mechanisms for dynamic changes of behavior selection strategies in behavior-based systems. *IEEE Transactions of System, Man, and Cybernetics Part B: Cybernetics*, 34(6):2377–2395, 2004.

[53] E.A. Sisbot, L.F. Marin-Urias, R. Alami, and T. Simeon. A human aware mobile robot motion planner. *Robotics, IEEE Transactions on*, 23(5):874–883, Oct 2007.

[54] K. Talamadupula, G. Briggs, T. Chakraborti, M. Scheutz, and S. Kambhampati. Coordination in human-robot teams using mental modeling and plan recognition. In *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, pages 2957–2962, Sept 2014.

[55] Kartik Talamadupula, J. Benton, Subbarao Kambhampati, Paul Schermerhorn, and Matthias Scheutz. Planning for human-robot teaming in open worlds. *ACM Transactions on Intelligent Systems and Technology. (Special Issue on Applications of Automated Planning)*, 1(2), 2010.

[56] Kartik Talamadupula, Gordon Briggs, Tathagata Chakraborti, Matthias Scheutz, and Subbarao Kambhampati. Coordination in human-robot teams using mental modeling and plan recognition. In *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*, pages 2957–2962, Sept 2014.

[57] Stefanie Tellex, Adrian Li, Daniela Rus, and Nicholas Roy. Asking for help using inverse semantics. In *In RSS*, 2014.

[58] Xin Tian, Hankz Hankui Zhuo, and Subbarao kambhampati. Discovering Underlying Plans Based on Distributed Representations of Actions. In *AAMAS*, 2016.

[59] Michael Wooldridge. *An Introduction to MultiAgent Systems*. Wiley Publishing, 2nd edition, 2009.

[60] Y. Zhang and Parker, L.E. IQ-ASyMTRe: Forming executable coalitions for tightly coupled multirobot tasks. *IEEE Transactions on Robotics*, 29(2):400–416, 2013.

[61] Y. Zhang, S. Sreedharan, and S. Kambhampati. Capability models and their applications in planning. In *International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*, 2015.

[62] Yu Zhang, Vignesh Narayanan, Tathagata Chakraborty, and Subbarao Kambhampati. A human factors analysis of proactive assistance in human-robot teaming. In *IROS*, 2015.

[63] Yu Zhang, Sarath Sreedharan, Anagha Kulkarni, Tathagata Chakraborti, and Subbarao Kambhampati Hankz Hankui Zhuo. Plan explainability and predictability for cobots. *CoRR*, abs/1511.08158, 2015.

[64] Yu Zhang, Sarath Sreedharan, Anagha Kulkarni, Tathagata Chakraborti, Hankz Hankui Zhuo, and Subbarao Kambhampati. Plan explicability for robot task planning. In *Proceedings of the RSS Workshop on Planning for Human-Robot Interaction: Shared Autonomy and Collaborative Robotics*, 2016.

[65] Hankz Hankui Zhuo and Subbarao Kambhampati. Action-model acquisition from noisy plan traces. In *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, IJCAI'13, pages 2444–2450. AAAI Press, 2013.

[66] Hankz Hankui Zhuo, Tuan Anh Nguyen, and Subbarao Kambhampati. Model-lite case-based planning. In *AAAI*, 2013.

[67] Hankz Hankui Zhuo, Tuan Anh Nguyen, and Subbarao Kambhampati. Refining incomplete planning domain models through plan traces. In *IJCAI*, 2013.