An Agent Framework for Manipulation Games

Javier M. Torres
Brainific S.L.
javier.m.torres@brainific.com

Abstract

Current agents use communication in a collaborative setting, exchanging truthful information to achieve a common plan. This paper defines games where agents may exchange information about the physical situation (both fluents and action events), arbitrarily nested beliefs, and action consequences, to manipulate other agents for their own goals, i.e. guide the other agents' own reasoning and planning. We propose a model for an "agent mind" that can cater for all these aspects through revisable, prioritized belief bases; goal recognition including epistemic situations; or planning including speech acts with structured content. We also discuss recent algorithms to address each one of them, and propose a concrete implementation for a future stage.

Introduction

We define *manipulation games* as games in which players not only affect some shared physical state, but also exchange information in the hope of influencing the other players so their goals can be achieved. Goals are often hidden from other players, and may or may not be conflicting. Observability is often restricted, and information can only be gained through third-party accounts. In this paper, we propose an action language specification and a doxastic model to reason about and share both actions and beliefs that enable manipulation of agents in games. Not all aspects of reasoning (e.g. fully reasoning about actions, or the planning of actions itself) have been implemented in source code, but suitable alternatives have been identified for each of the processes in the model.

Manipulation Games

Interactions with NPCs in RPG-like games can be modelled as non-zero sum games with hidden goals. Although there are sources of knowledge, stemming from observability (either an agent observes a situation or another agent directly), most information is more or less grounded belief, specially about the beliefs and motivations of other agents. Since only actions can be truthfully observed, intent and plan recognition are the only ways to estimate what another agent wants and believes, so the agent can adjust their own plans. After these models are built, one agent can plan to provide such information that other agents conclude beliefs or take actions in a way that benefits the planning agent. We need to consider such capabilities as:

- Reasoning and representation: reason about predicates with open world assumptions with a base of prioritized beliefs, and represent communicative actions
- Goals, goal recognition and goal recognition design: find out about other agents' plans to guide the planning
- Discourse and action planning: plan actions and predicates to communicate to indirectly guide other agents' actions

Agents, in fact, assume that all of them perform the same loop when facing a change in the environment: first non-obvious predicates are deduced from existing and new information; then the goals of the agents involved in the new situation are re-assessed; and then current actions are re-planned, or new actions are planned. The classic game Diplomacy, in fact, restricts its mechanics in such a degree that these mechanisms are the basis for the game.

A related field, from which we borrow, is that of persuasion and argumentation theory. A description of how automated planning techniques can be used to promote arguments can be found in (Black, Coles, and Bernardini 2014) and (Black, Coles, and Hampson 2017). We argue that manipulation expands the set of actions available in an argumentation setting by considering false or incomplete predicates, and relying on the other agents' internal processing like its own goal recognition or higher order reasoning.

Small treatise about manipulation for honest people[1]

Let us consider a sample fantasy RPG scenario with three actors: Aisha, Bo Yang and Chinira. The three of them are officers in the same army, and Chinira is the common boss of Aisha and Bo Yang.

Aisha has currently a *goal* of recovering the McGuffin of Diabolic Wisdom, which she hid in the common room of

[1] A humble tribute to (Joule, Beauvois, and Deschamps 1987)

Bo Yang's squad instead of her own, so nobody could think she was in its possession. The issue is not trespassing, she can freely enter this room, but rather concealment; she cannot risk Bo Yang's soldiers and Bo Yang himself seeing the artifact and learning about it. At least one person from Bo Yang's squad is always present there, and they would question Aisha if she searched the room. Additionally, Bo Yang despises Chinira, *believing* her to act in a purely selfish manner, attributing her a *goal* of self promotion.

During the paper we will show a formal description of the messages exchanged, and the beliefs and plans generated by the agents within bounding boxes.

Previous Work

This framework is very similar to a BDI architecture (Rao, Georgeff, and others 1995), as it endows agents with *beliefs* and *desires*, or goals. We believe, however, that few, if any, BDI implementations use the kind of techniques that we propose support manipulation games, like higher order reasoning, goal recognition or epistemic planning.

Horswill's MKULTRA (Horswill 2018) is a superb implementation of a manipulation game and an inspiration for the current work. The player can insert beliefs into other agents' minds to solve various puzzles. The areas where this work aims at improving MKULTRA are the use of full-fledged logical reasoning, instead of logic programming; higher-order beliefs; and a more oblique manipulation through planning/goal recognition and the re-evaluation of source reliabilities.

Ryan's Talk of the Town (Ryan et al. 2015) presents a system where bounded rationality and memory in agents creates a compelling narrative. Talk of the Town does not implement a complex model for agent reasoning, but on the other hand, agents follow complex schedules through which they acquire first- and second information about other agents.

(Ware and Siler 2021) describe a narrative planner that takes into consideration intentions and beliefs. However, characters themselves seem to use utility functions to choose actions, and do not reason about their own beliefs and those of other agents.

Within the automated planning community, epistemic planning (taking the epistemic state of other agents into consideration) has become so important as to have a dedicated workshop in ICAPS 2020 (https://icaps20subpages.icaps-conference.org/workshops/epip/). The work in (Shvo et al. 2020) includes epistemic plan recognition (which includes epistemic planning itself) leveraging the planners in (Le et al. 2018), (Wan, Fang, and Liu 2021) and (Muise et al. 2015). The authors are unsure whether belief (KD45) or knowledge (S5) is considered in these planners, and whether communication extends beyond the truth value of single fluents.

Multi agent systems such as (Panisson et al. 2018) provide a good foundation when it comes to theory of mind and speech acts, but deal with agent collaboration, whether implicit or explicit. This work instead focuses on taking advantage of the agents' reasoning strategies to obtain the desired result, regardless of whether this result is beneficial to

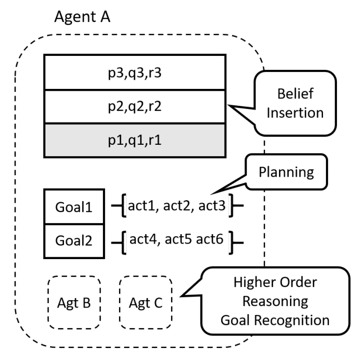

Figure 1: Internal Agent Model

the other agent. A noteworthy exception is (Black, Coles, and Bernardini 2014), which studies persuasion (a component of manipulation). The work on prioritized belief bases with non-idempotent operations presented in (Velázquez-Quesada 2017) models manipulation of human agents in our opinion.

Reasoning and Knowledge Representation

As we have mentioned before, each agent performs a loop of goal recognition and sensing followed by action planning, supported by doxastic reasoning, in a variation of the traditional Sense-Plan-Act loop.

The internal model of each agent, as illustrated in Figure 1, consists of a prioritized and time-versioned belief base including certain knowledge (sensing actions), a list of internal goals and associated plans, and a set of agent models with the same structure, rebuilt whenever new information is added. The state is updated by processes of belief insertion, planning, goal recognition and higher order epistemic reasoning, as described in the rest of the paper.

Dynamic Epistemic Logic has often been used to model reasoning of higher order beliefs such as those that can be expressed in this model. It is a formalism describing *epistemic states* and their changes after executing *actions*, which has experienced a substantial growth in the last 15 years. A very complete account of its evolution from public announcement logics to its current form can be found in (Van Ditmarsch, van Der Hoek, and Kooi 2007). It has been applied to areas such as cryptography or logic puzzles, and extended to different areas like modelling questions or epistemic planning.

However, we have found two important shortcomings in this family of logics:

- it puts the burden on the problem to fully specify the initial model;
- it models actions as semantic changes that directly modify this model, without few guidelines about what conditions should the actions fulfill to preserve model properties (e.g. KD45 for belief or S5 for knowledge) across updates, as Herzig has pointed out in (Herzig 2017).

Hence, we have decided to focus on actions with syntactic effects as much as possible (e.g. forcing as a result that

A believes in p after an action), so the user of the model needs to build partial models from the acting agent's point of view using techniques like tableaux whenever it is necessary, adding computational burden at the expense of flexibility. We have nonetheless studied the formalization of prioritized belief bases from a DEL perspective as described in (Baltag and Smets 2008). Observability and sensing primitive actions also allow us to derive knowledge (S5) modal formulas, in a way similar to that described in (Baral et al. 2015). A separate language for action specifications describes what an agent plans to do or is doing, within the syntax of the belief logic.

Doxastic Resoning

The doxastic model proposed consists of *prioritized belief bases* as described in (Rott 2009): an ordered list of sets of sentences, with most plausible sentences placed closer to the head of the list, followed by first hand, present-time, direct knowledge. We keep an open world assumption in our framework: having a p or ¬p explicitly in an agent's belief base means that they actually B(p) or B(¬p). The lack of belief about p means that they will not commit on any valuation for p: complete uncertainties (formulas for which agents do not have any preference) will not be represented in the base. Prioritized belief bases generate a corresponding system of spheres model, where possible world sets are filtered by each layer in the base. Since sentences can be removed due to their origin during an agent's lifetime, conflicting sentences may be kept in different levels; the actual belief of the agent will depend on the relative position of each sentence. An example base is presented in Figure 2 with annotations about the source of the beliefs.

These structures are more succinct than, for example, POMDP models, since they use logic sentences to express sets of worlds, and human agents tend to use vaguely defined confidence or plausibility levels instead of exact probabilities.

Sentences in such a model keep track of their origin, such as:

- Past direct observations. A belief could be implicitly formed about the current situation depending on a state that was observable *in the past, but not anymore*, with its plausibility degrading with time up to complete uncertainty being removed from the belief base.
- Accounts from other agents, accepted according to observed certainties and the perceived "honesty record" of other agents.
- Abduction, mainly targeted at action reasoning, so causes will be ordered according to their plausibility depending on the simplicity of their attributions to effects.
- Induction, for agents that perform some kind of statistical analysis of observed facts.

Note that the current paper does not propose explicit mechanisms for the inclusion of a belief in the base, apart from these suggestions.

A version timeline of the belief base is kept, so past and point temporal modalities like $\textsc{At}_{(t=3)}(p)$ can be used to refer to any (including the acting) agent's beliefs. Deduced propositions are indirectly referenced whenever a check using a tableau starts.

We allow the following types of predicates in the belief base:

- first order predicates; e.g. $has(B, knife)$
- visibility statements of first order predicates; e.g. $see(A, has(B, knife))$
- temporal statements of any other item; e.g. $\textsc{At}_{t=3}(has(B, knife))$
- goal statements about agents; e.g. $\textsc{Goal}_A(catch\_killer)$
- statements about actions with preconditions and postconditions, with probabilistic outcomes; e.g. $search\_room\{pre : empty(room\_123);\ post : \{t := t + 3\ with\ p = 1;\ in(knife, room\_123)\ with\ p = 0.5; \neg in(knife, room\_123)\ with\ p = 0.5\}\}$
- predicates that express that an action has just been performed; e.g. $done(\mathbf{fired}(A, B))$
- beliefs from other agents; e.g. $B_A(is\_killer(B))$

Special predicates like those described in (Marques and Rovatsos 2015) can be expressed using modalities about goals, actions and beliefs. The $\textsc{Goal}_A(p)$ modality expresses that agent A will take actions that make it more probable for proposition p to become true. One can express the preferred next action for agent A as $\textsc{Goal}_A(done(\mathbf{act}()))$. Predicates about knowledge like *unknown(a; que)* (the answer to question *que* is unknown to agent *a*) can be expressed as $\forall X(\neg B_A(que(X)))$.

To allow some approximation to probabilistic reasoning, plausibilities are related to discrete probabilities. Values from 0 to 2^n, expressing probabilities from 0 to 1, are used. Operations that would result in intermediate values are rounded towards $2^{(n-1)}$, a probability of 0.5. This value is important since it represents uncertainty, and as such can be removed from the belief base. We expect long term reasoning to be "diluted" in this way to control state explosion, since the further a result is in terms of operations (e.g. a situation several steps ahead in a plan), the more probable it is to turn into an uncertainty. In no way are complex probabilistic logic frameworks (e.g. Markov logic networks) involved in these estimations: derived statements always inherit the least plausible value from those among all the input statements.

In Figure 2 we can see a prioritized belief base consisting of three layers, each with a certain plausibility. In this example, we will use a value of 0 to 16 levels of probability, with a plausibility of 0 corresponding to a probability of 1, and 8 corresponding to complete uncertainty (50/50 estimation), and therefore not represented. Note that believing p with plausibility $plaus$ higher than 8 is equivalent to believing $\neg p$ with plausibility $16 - plaus$. In the example, we see levels of plausibility from 1 to 6, that would correspond to probabilities $\frac{15}{16}$ (almost certain) to $\frac{9}{16}$ (more likely than not). Note that certain knowledge is assumed to come only from direct observation in the current moment, so it is tracked separately. This structure induces a system of spheres, where each sphere includes layers from the base incrementally, as

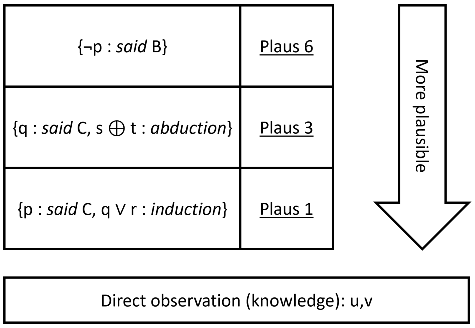

Figure 2: Prioritized Belief Base

POSTCONDITIONS

ontic	open(door)	Directly applicable
epistemic	see(B, open(door))	
doxastic	B_B(GOAL$_A$(out(A, room)))	Estimated through: • Doxastic reasoning • Goal recognition

Figure 3: Action Postconditions

long as the sentences from a less plausible layer do not contradict those from a more plausible one. Let us see what spheres would be induced by this base:

- Plausibility 1 / Probability $\frac{15}{16}$: all worlds complying with $p, q \vee r$ (e.g. $pqrst, p\bar{q}rst, pq\bar{r}s\bar{t}$)
- Plausibility 3 / Probability $\frac{12}{16}$: all worlds complying with $p, q, s \oplus t$ (e.g. $pqrst, pq\bar{r}\bar{s}\bar{t}$), within the previous sphere

The last layer contradicts previous, more plausible beliefs, and therefore does not induce any layer.

Keeping past states of this base and referring to them using dedicated past and point modalities allows us non-monotonic reasoning, since only new beliefs are added to the base, as all beliefs are implicitly tagged with the moment in which they were believed. A belief in p is not replaced, but rather it is asserted that $\text{AT}_{t=3}\neg p$ but also $\text{AT}_{t=6}p$.

First-hand, certain knowledge is handled apart from belief. This may exclude certain scenarios of misunderstandings: the step where an agent creates a false belief from a truthful observation. In the film "Knives Out", Great Nana mistakes character Marta for another, even though Marta is standing at plain sight in front of her. Marta then derives the incorrect belief that Great Nana has recognized her, since this model takes a "sensing" action by another agent as something about which we can have certain knowledge (and we see how this may not always be the case).

Action Representation

Actions are first class objects in the language. Preconditions and postconditions can be communicated and they are certainly used in planning, but in no way are they considered immutable or fixed. This has already been described in (Steedman and Petrick 2007), which uses a special purpose database. This is specially the case for communicative actions, which have few if any preconditions (e.g. $\neg B_A(p) \wedge \neg B_A(\neg p)$ for $\textbf{ask}(A,\ "p",\ \{\},\ \{B\})$) and can be easily extended as we will see later in .

The postconditions of actions can have three different natures, as summarized in 3:

- ontic, e.g. $open(door)$. The specification for ontic (physical) actions is very similar to traditional specification in STRIPS planning: a list of set/unset atomic predicates.
- epistemic, e.g. $see(B, open(door))$. Epistemic effects are tracked through observability predicates, as opposed to epistemic model modification as in logics in the DEL family, due to the issues of semantic action models as explained before. Note that observability itself is directly observable and applicable (we know for sure whether agent A sees something if we see them), and hence is an epistemic, not doxastic, effect.
- doxastic, e.g. $B_B(\text{GOAL}_A(out(A, room)))$. These effects can be computed using *doxastic logic* and *goal recognition*, as will be detailed later, and therefore depend on what the acting agent believes about the other agents; these complex effects will need to be evaluated again in every individual planning step, and may of course be incorrect if the higher order beliefs are themselves incorrect.

An agent may communicate action specifications (preconditions and effects). Both linear and contingent action plans can be communicated as *composite actions* using sequence (;), nondeterministic ($\cup$) and test (?) operators as in dynamic logics (e.g. as described in (Bolander et al. 2019)). We have decided not to cover unbounded iteration, since finite plans will be easier to check.

We consider the following basic actions in our framework:

- Perform an action with pure ontic or epistemic effects; e.g. $\textbf{fire}(E, D)$
- Say a proposition to a set of agents; e.g. $\textbf{say}(A,\ "\exists X\, \text{AT}_{t=3}(done(\textbf{fire}(X, D)))",\ \{B, C\})$, which means that A says to B and C that someone fired upon D at t=3.
- Ask an agent about something to a set of agents, that is, check the validity of a statement or request a value for the free variables in a statement that makes it true according to their beliefs, e.g. $\textbf{ask}(A,\ "\text{AT}_{t=3}(done(\textbf{fire}(E, D)))",\ \{\},\ \{B, C\}$ or $\textbf{ask}(A,\ "\text{AT}_{t=3}(done(\textbf{fire}(X, D)))",\ \{X\},\ \{B, C\}$
- Request an agent to do someghing, e.g. $\textbf{request}(A,\ "\textbf{ask}(B,\ "\text{BEL}_C(\text{AT}_{t=3}(done(\textbf{fire}(X, D)))",\ \{X\},\ \{C\}",\ \{B\})$

Internal actions are considered in planning, but they are not modelled as communicative acts. The reason is that they are clear enough to be modelled by each agent, and also known to happen whenever enough information is provided (a goal may be recognized as soon as there is evidence for it). There is no need to communicate anything about internal actions because all agents have enough stable knowledge to reason about them, even though the specific treatment of each agent of course depends on their beliefs. Agents however need to have a bounded rationality, so we cannot rely on other agents to come to conclusions, even though they may be logically valid.

First Act: Aisha Talks to Chinira

Aisha provides Chinira with the information that an incoming squad includes a spirit, vulnerable to a ritual from a certain book. This information comes from Deepak, an unreliable source, but Aisha hides this uncertainty from the message to let Chinira draw her own conclusions.

A

$$\textbf{say}(\forall X(in(patrol, X) \Rightarrow B_X(in(squad, spirit))))$$

Chinira incorporates Aisha's account with high plausibility in their *prioritized belief base* based on a previous track record of complete and accurate information. A lower plausibility could have been assigned if e.g. induction had shown that information from source Aisha is not reliable when checked against facts. Also, a higher, conflicting evidence already present (e.g. the report from an inside informant) would also have invalidated Aisha's information due to the construction of the spheres from the base. Using a simple breadth-first search, Chinira decides to request Bo Yang to stay in the garrison while she goes with her own squad to the ruins where this book is located, as other actions (sending Aisha or Bo Yang, perceived as inferiors; using some other strategy against the spirit; not doing anything in the hope of the spirit posing a lesser threat) would pay a lower performance/cost balance, always according to her current beliefs.

C

$$\boxed{in(squad, spirit) \ : \ \textbf{plaus 1}}$$

Plan: $(\textbf{go}(ruins); \textbf{get}(book); \textbf{fight}(squad))$

$\textbf{request}(C, \textbf{request}(A, \textbf{stay}()), B), \{A\})$

Goals, Goal Recognition and Goal Recognition Design

Beliefs and goals are not directly observable: an agent can only infer them in another agent through observation of their behaviors. Goal recognition is, thus, a very important piece of manipulation games. Whether an action is taken or not depends on the beliefs about preconditions and effects, and whether the effects lead to a goal. Goal recognition is a kind of abduction process, where agents' goals are deemed the most probable or concise explanation for those agents' actions. To illustrate the importance of goals and goal recognition, let us take the muddy children puzzle, a staple in dynamic epistemic logic. Agents need to assume that everyone's goal includes a truthful account of their observations. Without this assumption, for example with a lying agent, the puzzle cannot proceed.

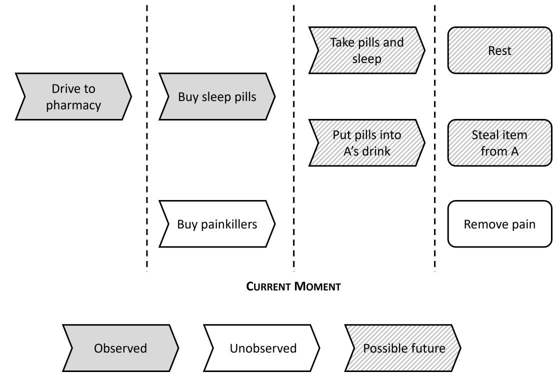

Figure 4: Plan and Goal Recognition

The current section only describes the open issue of the need for goal recognition mechanisms to compute doxastic preconditions of actions in a setting where these include speech acts like saying, asking or requesting, and beliefs of other agents may not match the contents of such statements. We do identify some algorithms as potential candidates for implementation.

We approach the concepts of *desire* and *intention* from BDI logics (Rao, Georgeff, and others 1995) deriving it from goal recognition in automated planning and dynamic logics, rather than the usual Computational Tree Logic. The intuition is that an agent *desires* a formula ϕ if, when allowed a choice, that agent takes an action that maximizes the estimated probability of reaching a state where ϕ holds, when compared with any other action to take, *according to that agent's beliefs*. A Bayesian formulation of goal recognition can be found in (Baker, Tenenbaum, and Saxe 2006). Goal recognition can be performed using the same planning algorithms as the agent uses for its own plans, as described in (Ramírez and Geffner 2009) and (Ramírez and Geffner 2010), instead of relying on a plan library. As described in the latter, a prior distribution $P(G)$ over the goals G (priors for goal preference can be incorporated as explained in (Gusmão, Pereira, and Meneguzzi 2021)) is used to obtain the likelihoods $P(O|G)$ of the observation O given the goals G using the cost differences as obtained by a classical planner. In (Sohrabi, Riabov, and Udrea 2016) additional features like unreliable observations and plan recognition, in addition to goal recognition, are introduced.

In Figure 4 we can see two observed steps, with two possible plan completions to different goals. We can rule out the goal in white, due to the second event observed. However, without some evidence pointing towards one of the grey-patterned goals, we cannot predict further actions.

It is worth mentioning Geib's PHATT algorithm as described in (Geib and Goldman 2009) for its use of probabilistic actions and AND-OR trees (suitable for contingent planning and very similar to behaviour trees in game agent logic). It relies, however, on the description of the structure of tasks and a preexisting plan library. However, further research into the induction of task structure and hierarchies, as well as using planning itself as a plan library generation would be necessary to consider this algorithm.

A further concept in automated planning is *goal recognition design*, whereby some action is designed to make an agent's goals as easy to discover as possible, as described in (Keren, Gal, and Karpas 2014). By making explicit statements about goals and beliefs in our doxastic model, planning algorithms as described in the following section can include actions that reduce uncertainty, like sensing actions and goal recognition design.

Action and Discourse Planning

Ontic actions can be planned using planners that can handle a probabilistic outcome using the action specifications stored in the agent's belief base, like probabilistic planners or a deterministic planner with replanning like (Yoon, Fern, and Givan 2007). Also, other multiagent frameworks and proposals specify preconditions and postconditions for communicative actions, e.g. the FIPA standards (FIPA 2008). FIPA defines epistemic and doxastic preconditions (*feasibility preconditions, FP*) and postconditions (*rational effects, RE*). These preconditions and postconditions are asserted in the belief base if an agent detects that action. We could consider these as "social protocols" that state clearly the goals of the speaker.

However, a complication for communicative actions in our setting is that traditional agent frameworks are oriented toward *collaborative* agents. Feasibility preconditions express a socially agreed reason to why the action is performed, but in fact nothing prevents an agent to say whatever it wants. Furthermore, when seen from the perspective of the "sender", a communicative act may be issued precisely to guide the goal recognition process in the "receiver" towards a certain goal or plan. The possession or not of a certain knowledge does not enable us to ask a question; rather, our goal of reducing uncertainty or of making another agent believe that one does not know something is what will compel us toward that action. In a similar way, evaluating the outcome of a communicative action needs to take goals into account.

Postconditions in communicative actions become complicated to compute: the sender has to try and replicate a goal recognition step, using the receiver's beliefs about the sender and goals *to the extent to the sender itself's belief*, and then try to predict what will the receiver belief about the sender's intentions. Note that even ontic actions may carry a doxastic effect, in the sense that *any* action is framed within an estimation of goals. Opening a door is evidence for the other agent to have run through it, but it may have been left open on purpose to lure the observer into that conclusion. We believe that the increase in memory and computation power of user equipment justifies exploring this modelling. As mentioned before, re-planning or MCTS techniques in automated planning have yielded satisfactory results.

Selecting the content to present in a speech act can be guided by bulding a model for the receiver, including goals and beliefs, so candidates for items in sentences can be proposed from incomplete proofs (e.g. whether they can close open branches in a tableaux) or plans (e.g. communicating action preconditions to the receiver, *regardless of their actual truth*) that are related to the current goals. The whole loop of goal recognition, reasoning and planning is performed in the simulated model for the other agent, to the extent to which resources can be dedicated. If we belief that $\mathrm{GOAL}_A(catch\_killer)$, informing A of $has(B, knife)$ allows it to plan further actions, like asking B for the knife. If we believe that $B_A(\forall X(has(X, knife) \Rightarrow is\_killer(X)))$ this information item is a particularly powerful lever to guide A's actions.

Some authors use existing planners adapted with epistemic predicates. For example (Marques and Rovatsos 2016) modifies the Contingent-FF planner to include requests and yes/no questions. Also (Muise et al. 2015) apply a classical planner, the Fast Downward planner (Helmert 2006), to a multiagent epistemic setting with higher order beliefs using additional fluents derived from epistemic logic axioms. However, such an approach must accomodate the expensive computations for doxastic postconditions.

Second Act: Aisha Talks to Bo Yang

Aisha then takes Chinira's request to Bo Yang. However, she mentions Deepak when mentioning the patrol report to Bo Yang. When issuing Chinira's request, Aisha does not reveal how Chinira has come to this conclusion, but she makes it clear that it comes from Chinira.

A

$$\textbf{say}(B_D(in(squad, spirit)))$$
$$\textbf{say}(done(\textbf{request}(C, \textbf{request}(A, \textbf{stay}(), \{B\}), \{A\})))$$
$$\textbf{request}(A, \textbf{stay}(), \{B\})$$

The specification language for higher order epistemic actions would allow Bo Yang to examine a possible plan where Aisha's goal is represented and this observation is matched, but goal probability priors would rank the corresponding goal fairly lower than alternative goals from other agents, or would have too many unknown factors. Bo Yang may suspect that $\exists X(\textsc{Goal}_X(\neg in(B, city)))$, but he would not be able to progress the reasoning much further without a lot of time consuming sensing actions. Bo Yang believes Aisha to be a loyal individual due again to induction from past observations. A more plausible goal at play is assuming that Bo Yang has the goal of impressing higher officers (a wrong belief that has nonetheless crept up high in Bo Yang's prioritized belief base due to previous interactions with Aisha) by recovering the book and keeping him in the garrison. Bo Yang decides instead to face the incoming army (as he cannot go for the book himself and clash with Chinira, staying would result in a loss of face, and no sentence could possibly land higher in Chinira's belief base to change their mind in Bo Yang's belief). This results in Bo Yang's soldiers leaving their barracks for a few days.

B

$$\boxed{\neg in(squad, spirit)\ :\ \textbf{plaus 4}}$$
$$\boxed{\textsc{Goal}_C(honour(C) > honour(B))\ :\ \textbf{plaus 1}}$$

Plan: (**fight**$(squad)$)

Conclusion

We have presented a definition for games of manipulation, as games with open communication and unknown goals, whose players use models of other agents to guide their actions. We have pointed out at three aspects that enable these games: doxastic higher order reasoning, goal recognition, and epistemic planning. For each of these areas we have identified alternatives for data structures and algorithms that can support these aspects. In future communications we plan to present prototypes for each of them.

References

Baker, C. L.; Tenenbaum, J. B.; and Saxe, R. 2006. Bayesian models of human action understanding. *Advances in neural information processing systems* 18:99.

Baltag, A., and Smets, S. 2008. A qualitative theory of dynamic interactive belief revision. *Logic and the foundations of game and decision theory (LOFT 7)* 3:9–58.

Baral, C.; Gelfond, G.; Pontelli, E.; and Son, T. C. 2015. An action language for multi-agent domains: Foundations. *arXiv preprint arXiv:1511.01960*.

Black, E.; Coles, A.; and Bernardini, S. 2014. Automated planning of simple persuasion dialogues. In *International Workshop on Computational Logic and Multi-Agent Systems*, 87–104. Springer.

Black, E.; Coles, A. J.; and Hampson, C. 2017. Planning for persuasion. In *16th International Conference on Autonomous Agents and Multiagent Systems, AAMAS 2017*, 933–942. International Foundation for Autonomous Agents and Multiagent Systems (IFAAMAS).

Bolander, T.; Engesser, T.; Herzig, A.; Mattmüller, R.; and Nebel, B. 2019. The dynamic logic of policies and contingent planning. In *European Conference on Logics in Artificial Intelligence*, 659–674. Springer.

FIPA, T. 2008. Fipa communicative act library specification. *Avaliable online: https://www. fipa. org/specs/fipa00037/SC00037J. html (accessed on 19 July 2021)*.

Geib, C. W., and Goldman, R. P. 2009. A probabilistic plan recognition algorithm based on plan tree grammars. *Artificial Intelligence* 173(11):1101–1132.

Gusmão, K. M.; Pereira, R. F.; and Meneguzzi, F. 2021. Inferring agents preferences as priors for probabilistic goal recognition. *arXiv preprint arXiv:2102.11791*.

Helmert, M. 2006. The fast downward planning system. *Journal of Artificial Intelligence Research* 26:191–246.

Herzig, A. 2017. Dynamic epistemic logics: promises, problems, shortcomings, and perspectives. *Journal of Applied Non-Classical Logics* 27(3-4):328–341.

Horswill, I. 2018. Postmortem: Mkultra, an experimental ai-based game. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, volume 14.

Joule, R.-V.; Beauvois, J.-L.; and Deschamps, J. C. 1987. *Petit traité de manipulation à l'usage des honnêtes gens*. Presses universitaires de Grenoble Grenoble.

Keren, S.; Gal, A.; and Karpas, E. 2014. Goal recognition design. In *Proceedings of the International Conference on Automated Planning and Scheduling*, volume 24.

Le, T.; Fabiano, F.; Son, T. C.; and Pontelli, E. 2018. Efp and pg-efp: Epistemic forward search planners in multi-agent domains. In *Twenty-Eighth International Conference on Automated Planning and Scheduling*.

Marques, T., and Rovatsos, M. 2015. Toward domain-independent dialogue planning. In *Fourth International Workshop on Human-Agent Interaction Design and Models (HAIDM 2015)*.

Marques, T., and Rovatsos, M. 2016. Classical planning with communicative actions. In *Proceedings of the Twenty-second European Conference on Artificial Intelligence*, 1744–1745.

Muise, C.; Belle, V.; Felli, P.; McIlraith, S.; Miller, T.; Pearce, A. R.; and Sonenberg, L. 2015. Planning over multi-agent epistemic states: A classical planning approach. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*.

Panisson, A. R.; Sarkadi, S.; McBurney, P.; Parsons, S.; and Bordini, R. H. 2018. On the formal semantics of theory of mind in agent communication. In *International Conference on Agreement Technologies*, 18–32. Springer.

Ramírez, M., and Geffner, H. 2009. Plan recognition as planning. In *Twenty-First International Joint Conference on Artificial Intelligence*.

Ramírez, M., and Geffner, H. 2010. Probabilistic plan recognition using off-the-shelf classical planners. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 24.

Rao, A. S.; Georgeff, M. P.; et al. 1995. Bdi agents: From theory to practice. In *ICMAS*, volume 95, 312–319.

Rott, H. 2009. Shifting priorities: Simple representations for twenty-seven iterated theory change operators. In *Towards mathematical philosophy*. Springer. 269–296.

Ryan, J.; Summerville, A.; Mateas, M.; and Wardrip-Fruin, N. 2015. Toward characters who observe, tell, misremember, and lie. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, volume 11.

Shvo, M.; Klassen, T. Q.; Sohrabi, S.; and McIlraith, S. A. 2020. Epistemic plan recognition. In *Proceedings of the 19th International Conference on Autonomous Agents and Multi-Agent Systems*, 1251–1259.

Sohrabi, S.; Riabov, A. V.; and Udrea, O. 2016. Plan recognition as planning revisited. In *IJCAI*, 3258–3264.

Steedman, M., and Petrick, R. 2007. Planning dialog actions. In *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue*, 265–272.

Van Ditmarsch, H.; van Der Hoek, W.; and Kooi, B. 2007. *Dynamic epistemic logic*, volume 337. Springer Science & Business Media.

Velázquez-Quesada, F. R. 2017. On subtler belief revision policies. In *International Workshop on Logic, Rationality and Interaction*, 314–329. Springer.

Wan, H.; Fang, B.; and Liu, Y. 2021. A general multi-agent epistemic planner based on higher-order belief change. *Artificial Intelligence* 301:103562.

Ware, S. G., and Siler, C. 2021. The sabre narrative planner: multi-agent coordination with intentions and beliefs. In *Proceedings of the 20th International Conference on Autonomous Agents and Multiagent Systems*, 1698–1700.

Yoon, S. W.; Fern, A.; and Givan, R. 2007. Ff-replan: A baseline for probabilistic planning. In *ICAPS*, volume 7, 352–359.