Extended Goal Recognition: A Planning-Based Model for Strategic Deception

Peta Masters
The University of Melbourne
Victoria, Australia
peta.masters@unimelb.edu.au

Michael Kirley
The University of Melbourne
Victoria, Australia
mkirley@unimelb.edu.au

Wally Smith
The University of Melbourne
Victoria, Australia
wsmith@unimelb.edu.au

ABSTRACT

Goal recognition is the problem of determining an agent's intent by observing its actions. In the context of AI research, the problem is tackled for two quite different purposes: to determine an agent's most probable goal or, for human-aware planning including planned—or strategic—deception, to determine an observer's most likely belief about that goal. Making no distinction, contemporary models tend to assume an infallible observer, deceived only while it has limited access to information or if the environment itself is only partially observable. Focusing on the second purpose, we propose an extended framework that incorporates formal definitions of confirmation bias, selective attention and memory decay. In contrast to pre-existing models, our approach combines explicit consideration of prior probabilities with a principled representation of observer confidence and distinguishes between potential observations—i.e., every observable event within the observer's frame of reference—and recalled observations which we model as a function of attention and memory. We show that when these factors are taken into consideration, false beliefs may arise and can be made to persist, even in a fully observable environment—thus providing a perceptual model readily incorporated into the "thinking" of an adversarial agent for the purpose of strategic deception.

KEYWORDS

goal recognition; deception; planning; probabilistic reasoning

ACM Reference Format:
Peta Masters, Michael Kirley, and Wally Smith. 2021. Extended Goal Recognition: A Planning-Based Model for Strategic Deception. In *Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), Online, May 3–7, 2021*, IFAAMAS, 9 pages.

1 INTRODUCTION

Strategic deception explicitly sets out to manipulate an opponent's perceptual reality [4, 12]. Typically, this demands a model of that reality and the better the model, the more successful the deceptions based upon it are likely to be. One such model comes from goal recognition (GR), the problem of inferring an agent's intent from its behaviour. GR is studied for two purposes: literally, to determine an agent's most probable goal [e.g., 30, 34] but also—particularly in the context of human-aware planning—to determine an observing agent's most likely *belief* about the observed agent's goal [e.g., 9, 24], that is, as a proxy for the observer's perceptual reality. This is the problem that extended GR sets out to solve. It seeks not necessarily the correct goal but the goal that a human-like observer is most likely to believe to be correct, thus providing a planning-based model readily incorporated into the "thinking" of an adversarial agent for the purposes of strategic deception.

In computer science, perceptual reality begins with observation. When we declare an environment "fully observable", we assume not only that all observable phenomena will be observed but that they will be believed and remembered; i.e., encoded and available for future decision-making. Under these constraints, deception seems to depend on a degree of partial observability such that the observer is literally incapable of perceiving one or more aspects of the ground truth (owing to faulty sensors, for example [e.g., 20] or because essential data has been withheld or has gone missing [31]). But this simplistic representation fails to account for many aspects of perception and deception as they occur in the human realm, where observation and belief do not necessarily go hand-in-hand. Effectively, an environment that ought to be fully observable—in that everything in it has at some point been available to the sensors of the observing agent—may become partially observable owing to bounded rationality and the predictable cognitive limits of the observer [12] so that, by the time the agent makes its decision, data may indeed be missing and stored sensory data unreliable.

We focus on three notions in particular. Firstly, not everything observable is necessarily encoded or available for recall. Secondly, not everything initially available is remembered indefinitely. And thirdly, humans sometimes fail to believe the evidence of their own senses. We show how misperceptions such as these can be incorporated into online GR to facilitate human-aware reasoning. While we demonstrate our results in the context of deception, they are equally relevant to cooperative planning where they can help to achieve transparency and avoid the need for explanation.

Our framework accepts a core premise of Masters and Sardina [24]'s model for path-planning that deception occurs when the observing agent—assumed to be performing GR—estimates the probability of the real goal to be less than (or equal to) the probability of at least one other possible goal. Concretely, we incorporate into a contemporary model of GR predictable aspects of memory and attention known to apply to human observers [3, 17, 21]: (a) improved consideration of ***prior probabilities***, somewhat neglected in recent work, as we discuss; (b) self-modulation of a confidence parameter to model ***confirmation bias***; (c) an implementation of ***selective attention*** which extends the usual definition of observations to distinguish between *potential* observations—which include every observable event within the observer's frame of reference—and *recalled* observations, which are the only observations ultimately available to the agent for decision-making or prediction; and (d) recalled observations are subject to ***decay***.

Proc. of the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2021), U. Endriss, A. Nowé, F. Dignum, A. Lomuscio (eds.), May 3–7, 2021, Online.

We show that, when these factors are taken into account, deception can occur—as it does in the real world—even in a fully observable domain. Furthermore, belief in a false goal can persist beyond limits that contemporary scholarship has seemed to show are fixed constraints imposed by the domain [e.g., 14, 25].

In the following sections, we set out the technical background before presenting this paper's main contributions: examination of prior probability in the context of contemporary GR, showing how false belief can arise in a fully observable domain; and an extended GR framework, which incorporates key aspects of perception and memory that make human agents vulnerable to deception. Finally, we examine related work and conclude with discussion.

2 BACKGROUND

In this section, we present the technical background that we rely on for the remainder of the paper.

"Plan recognition as planning", also known as "cost-based" GR, has been modelled in various domains including task-planning [30], discrete path-planning [23] and continuous motion-planning [34]. Our extended model builds on a generic interpretation of the problem from [26], which can be applied to any of the above, requiring only that the domain supports the notion of state-to-state transitions that can be costed and that the observable phenomena within the domain (whether actions, states, fluents, trajectories, etc.) *are* or can be *associated with* the transitions that give rise to them.

Definition 1. *A **generic cost-based GR problem** is a tuple $\mathcal{P} = \langle \mathcal{D}, \Omega, \vec{o}, G, s, Prob \rangle$ where:*

- *$\mathcal{D}$ is a model of the GR domain (which defines states, transitions between states and their cost);*
- *Ω is the set of all the observable phenomena in $\mathcal{D}$;*
- *$\vec{o} = o_1, o_2, .., o_n$ is an observation sequence, $o_i \in \Omega$;*
- *G is the set of candidate goals;*
- *s is the initial state, which is fully observable; and*
- *$Prob$ is the prior probability distribution across G.*

Generally, a ***plan*** π in $\mathcal{D}$ is a sequence of elements or events that imply transitions from state to state. Given a set of all such elements E, each element has a cost $c : E \mapsto \mathbb{R}$ and the ***cost of a plan*** $cost(\pi) = \sum_{i=1}^{m} c(e_i)$. A plan $\pi = e_1, \ldots, e_m$ is said to ***satisfy observations*** $\vec{o} = o_1, \cdots, o_n$, if there exists a monotonic function $f : \{1, \cdots, n\} \mapsto \{1, \ldots, m\}$ such that $e_{f(i)} = o_i$ for all $i \in \{1, \cdots, n\}$. That is, the ordering (in both the plan and the observation sequence) is preserved. The ***optimal (lowest) cost*** of a plan from s to a goal $g \in G$ is denoted by $optc(s, g)$ and the lowest cost plan from s to g that satisfies observations $\vec{o}$ is denoted $optc(s, \vec{o}, g)$.

The solution to $\mathcal{P}$ is a probability distribution which prefers those goals that best satisfy the observations. In seminal work, Ramirez and Geffner [29, 30] introduce the notion of *cost difference* as a basis on which to make that distinction, being the difference between the cost of a plan that satisfies observations and the cost of a plan that does not. The power of the formula lies in the fact that both terms can be calculated by a classical planner while one of the key insights is that the lower a goal's cost difference, the higher its probability (relative to other goals in the distribution).

The cost difference formula has since been analysed by others [10, 23, 25] and we adopt a less computationally demanding construction than the original, proved to return identical results in all but one corner case:[1]

$$costdif_{MS}(s, \vec{o}, g) = optc(s, \vec{o}, g) - optc(s, g). \tag{1}$$

The solution to a problem $\mathcal{P}$, adapted from [26], is given by the probability distribution at (2) below. In words, the likelihood of a goal is inversely proportional to the cost difference. The results are then multiplied by priors (*Prob*) and normalised.

$$Pr_{MS}(G \mid \vec{o}) = \alpha \cdot \frac{1}{e^{\beta(optc(s,\vec{o},g) - optc(s,g))}} \cdot Prob \text{ for } g \in G, \tag{2}$$

where α is the normalisation constant and β is a rate parameter, which changes the shape of a distribution without changing the rankings: the lower the value of β, the flatter the distribution.

We note that, in [26], the β parameter is made self-modulating and is used to represent an observer's confidence in its prediction. We also use this parameter to model confidence but, rather than basing that on whether a plan has become unnecessarily suboptimal/irrational (as in [26]), we will base confidence on how closely the most recent observation conforms to *expectation* (see 4.1).

Information that is irrelevant tends to be forgotten [12]. We will adapt the rationality measure (RM) also from [26] (at (3) below) to model relevance. The RM is designed to evaluate rational behaviour in situations—such as GR or deception—when the real goal is unknown. It quantifies the degree to which an observation sequence corresponds to an optimal plan for *any* of multiple possible goals. When observed behaviour is fully optimal, the RM equals 1.

$$RM_{MS}(s, \vec{o}, G) = \max_{g \in G} \frac{optc(s, g)}{optc(s, \vec{o}, g)}. \tag{3}$$

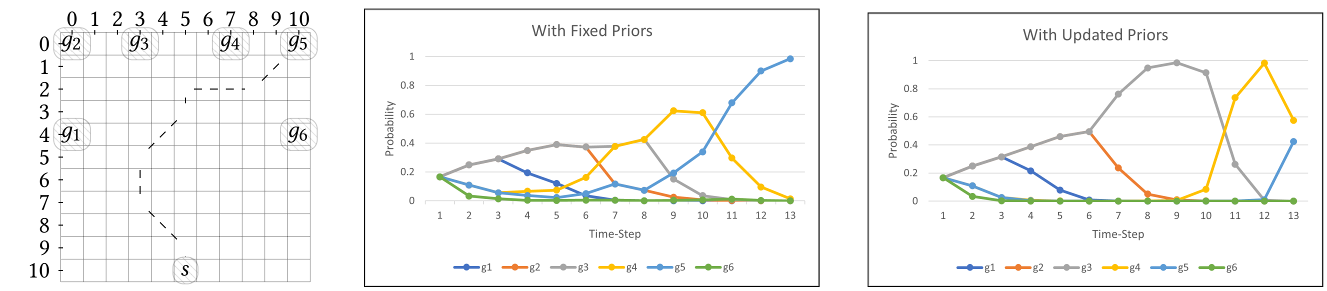
We will use the RM to estimate the *relevance* of an observation to help determine how likely it is to be remembered (see 4.2).

Goal recognition provides our perceptual model for strategic deception. Indeed, deception can be regarded as an inversion of GR. Whereas GR determines an agent's intent by observing its behaviour, deception involves generating behaviour such that an observer is *unable* to determine the agent's intent. Our definition of deception is adapted from [24] which states that an observation sequence is deceptive if—using any suitable probabilistic model of GR—the probability of the real goal, given those observations, is less than or equal to the probability of any other possible goal.

Definition 2. *Given an intended goal $g_r \in G$, observations $\vec{o} \in \Omega^*$ are **deceptive** iff $Pr(g_r \mid \vec{o}) \leq Pr(g \mid \vec{o})$ for $g \in G \setminus \{g_r\}$.*

We conclude this section by noting that, making no distinction between types of GR—one to determine a likely goal, the other to determine an observing agent's *belief* about the goal—the large body of work that builds on "plan recognition as planning" [29, 30] has established that there is a cost-radius (Radius of Maximum Probability, RMP) within which an agent's "real" goal is inevitably the most likely goal [24, 25] and, similarly, that there is a calculable limit to the number of steps in a plan (Worst-Case Distinctiveness, WCD) beyond which the real goal can be distinguished from other goals [14–16]. We will show that, considered from the more human-like perspective afforded by extended goal recognition, those apparent constraints can be overcome.

[1] For a detailed explanation, refer to [25].



(a) Noisy walk from s to g_5. The probability distribution across goals is calculated at every time-step, cell-to-cell.

(b) As originally calculated. When prior probabilities across goals are fixed, the posterior probability of each goal changes depending on the agent's most recently observed location.

(c) When priors are updated, posterior probabilities are mediated by previously-held beliefs—which linger even after the destination has been reached.

Figure 1: Noisy Walk Revisited (from [30]). Given a set of possible goals $G = \{g_1, g_2, ..g_6\}$, and observations at every time-step (including an assumed observation at g_5), probabilities were calculated using Equation (2) with $\beta = 1$. When the prior probability distribution across goals *Prob* is updated after every observation, rather than being taken as a fixed, unchanging parameter, the posterior probability distribution (shown in the graphs) changes dramatically.

3 PRIOR PROBABILITIES AND FALSE BELIEF

In this section, we examine the way that contemporary models of GR, such as that set out in Section 2 above, have handled prior probabilities and show that, counter-intuitively, when priors are handled *more* carefully they may generate *less* accurate results which are, nevertheless, *more* consistent with human-like reasoning.

The probability distribution at Equation (2) is derived from Bayes Rule. Thus, as expected, to obtain the probability of a goal given the observations, we must not only consider the likelihood of observations given the goal (obtained here by inverting cost difference) but also the *prior* probability of the goal. Observe however that, as set out, prior probabilities *Prob* in (2) is a given, unchanging parameter, supplied as part of the problem definition. Typically moreover, for convenience, priors are assumed equal [e.g., 23, 30], which means they cancel out on normalisation and can be ignored. *The term appears in the formula but is never called upon to do any work.*

Now, the above assumption (of one-off/unchanging priors) is reasonable if the observation sequence can be regarded as a single behaviour—compared with a similarly atomic *optimal* behaviour—but if instead observations are assumed to occur at separate time-steps, the probabilities calculated at one time-step ought properly to become the prior probabilities considered at the next.

Ramirez and Geffner [30] provide the example of a noisy walk, reproduced at Figure 1a. Probabilities are calculated at each time-step but always on the assumption that priors are equal. As that paper states, "The challenge in this formulation is the definition of the likelihoods"(p.1123), that is *not* the priors, and this emphasis—or lack of emphasis—appears to have carried forward into the work of subsequent authors extending the models from [29] and [30], [e.g., 14, 23, 32, 34]. And this has occurred in spite of at least three clear indicators that the issue should be more carefully examined.

Prior Probability as Belief. Baker et al. [2] also build on Bayes' Theorem. Setting out to model human-like reasoning, they use a POMDP framework to represent a Belief-Desire-Intent paradigm within which desire is represented by the cost/reward for performing certain actions in certain states and belief is represented by prior probability. Comparing their model with less expressive examples, they demonstrate not only that human goal recognition may be inconsistent with the ground truth (important to our work) but also that it involves *updating prior probabilities after each observation*: that is, assessing prior probability once only at the initial state—as implied by formulations such as Equation (2)—*does not match* to their observations of human reasoning.

Counter-intuitive Results. Masters and Sardina [23] developed an efficient alternative to the cost difference formula given at (1) for use in path-planning and other fully observable domains. The alternative formulation is based not on a full observation sequence but on the *single most recent* observation. By directly plugging this formula into the original probability distribution formula from [30]—and assuming that prior probabilities are equal and therefore irrelevant—they obtain identical goal rankings to the original. The finding, as the authors state, is counter-intuitive since it appears that they are able to predict where an agent is going without knowing where it has been. Now, although the alternative cost difference formula is correct and predicts the most likely goal as accurately as [30], intuitively we recognise that, correct or not, the agent's immediate history *ought* to be relevant to our (human) assessment of an agent's most probable goal.

Human Experiment. Vered et al. [34] builds on [29] to develop an explicitly human-like GR model based on "goal mirroring", which determines another agent's purpose by asking, "If I were doing what they are doing, what would my goal be?" Goal mirroring uses *online* goal recognition. That is, observations are assumed to be processed and evaluated incrementally. Even so, referencing the scholarship of Ramirez and Geffner, their custom probability distribution formula makes no adjustment for prior probability. When evaluated against human participants, researchers identify an unexpected result: humans tend to rule-out goals too quickly, commit to them too soon and persist in believing in a goal even when it seems probabilities ought to flip against it. To explain this, the authors suggest participants may be using knowledge outside the problem definition. It is also possible, however, that some of this "outside knowledge" is simply the incremental application of prior probability: a potentially erroneous but predictable aspect of human reasoning.

Returning to the noisy walk example—which, note, is played out in a fully observable path-planning domain—the differences between Figures 1b and 1c are stark. When prior probabilities are taken into consideration—and with no other change to the underlying model—the probability of g_4 *exceeds* the probability of g_5, *even after* g_5 *has been achieved*, shattering assertions cited in Section 2 with respect to the RMP [25] and WCD [14].

Finally, we note that recent planning based models of GR which, instead of cost, depend on landmarks—that is, facts that *must* be true or actions that *must* occur in order for a goal to be achieved—do not necessarily evaluate observations sequentially. In such cases, prior probabilities cannot be considered at all [e.g., 28].

4 EXTENDED GOAL RECOGNITION

In this section, we present our main contribution: a GR framework that supports strategic deception by formalising three notions: confirmation bias, selective attention and memory decay.

Extended goal recognition is the problem of determining from observed behaviour, not necessarily the most likely goal (though it may be), but the goal that will be *believed* to be most likely by a predictably fallible observer. Therefore, we now extend the Section 2 framework to support three notions: (1) that observation sequences are time-sensitive; (2) that an agent may be faced with multiple competing observables, *only one of which* it is capable of fully attending at any one time; (3) previously remembered observations may be forgotten.

Rather than a wholesale reimagining of GR, the extensions are presented as a series of modifications to the Section 2 framework but might equally be applied—in combination or independently—to other models.

The assumed domain is now partially observable, not because information is withheld or because the agent's sensors are faulty, but because the model aims to capture a boundedly-rational agent unable to process and retain all the observations made available. Furthermore, in addition to assumptions from Definition 1 that the domain must support costed state-to-state transitions associated with observable phenomena, we now also assume that observations include sufficient information for ongoing costs to be calculated. For example, if costs within the domain are calculated in terms of distance, then each observation is assumed to include positional data; if costs depend on spending, observations include remaining funds. This proviso enables us to calculate optimal costs that would previously have been measured by reference to a fully observable initial state by reference instead to the *first remembered observation*.

Definition 3. *An* ***extended GR problem*** *is a tuple* $\mathcal{P}_x = \langle \mathcal{D}, \Omega, mag^*, \vec{Q}, G, s_0, Prob \rangle$ *where:*

- $\mathcal{D}$ *(as before) is a model of the GR domain which defines states, transitions between states and their cost;*
- Ω *is the set of all observable phenomena in the domain;*
- $mag^* : \Omega \mapsto \mathbb{R}$ *is the base magnitude of each observable;*
- $\vec{Q} = O_1, O_2, ..O_n$ *is a time-ordered sequence of sets, where each set* $O_i \subseteq \Omega$ *comprises all observable phenomena available to the agent's sensors at a particular time-step* $i \in \{1, 2, .., n\}$*;*
- *G (as before) is the set of candidate goals;*
- s_0 *is the initial observation, which includes all cost-relevant data; and*
- *Prob (as before) is the prior probability distribution over G.*

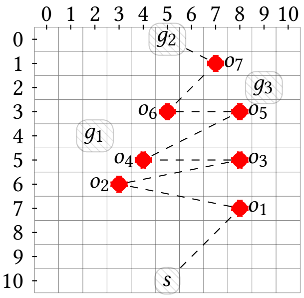
The above model differs from Definition 1 in the following ways, which we explain further in Sections 4.1 to 4.3 below.

(1) $\mathcal{P}_x$ is an *online* problem. That is, observations are delivered incrementally at distinct time-steps.
(2) Each observable element $o \in \Omega$ has a base magnitude $mag^*(o)$, which may be infinite but is otherwise subject to decay. Note, moreover, that the cost of "achieving" an observation is the cost of the event that gave rise to it (a property of the domain). As a result, at any given moment, an observation has an *effective* magnitude $mag(o)$, which may differ from its base magnitude. (As we see later, at 4.2.1, a deceptive agent can pay more to increase or decrease the intensity of an observation for the purpose of misdirection.)
(3) Instead of a sequence of individual observations $\vec{o} = o_1, .., o_n$, $\vec{Q} = O_1, .., O_n$ is a sequence of *sets*, where each set $O_i = \{o \mid$ occurred at time $i\}$. That is, each set comprises all potential observations newly available (or refreshed) at the current time-step, only one of which, denoted o_t, is ultimately encoded and remembered (see Section 4.2 and Figure 4). Practically, the observing agent assembles its own observation sequence, selecting one observation at each time-step from the set of all ***potential observations*** O_t and adding it to a sequence that comprises all those observations it has previously selected. Meanwhile, all previously held observations decay (see Section 4.3 and Figure 5). Thus, as one observation is added, one or more are likely to be forgotten/omitted so, at a given time t, only a distilled sequence of ***recalled observations***, which we denote $\vec{o}_t$, is available to the agent.
(4) There is no initial state as such. s_0 is the initial *observation*. Subsequently, s_t is taken to represent the first remembered observation at time-step t.

In summary, at time-step t: O_t is the set of all newly available observations (which could be a singleton); o_t is the observation selected/remembered from that set; $\vec{o}_t$ is the sequence of observations currently available (to which o_t has just been added); and s_t is the first observation in that sequence.

As before, the solution to an extended GR problem is a probability distribution across goals or—more properly since sets of observations are delivered incrementally—a *sequence* of probability distributions, just as presented w.r.t. the random walk (Figure 1), where the observer's beliefs about the likelihood of goals is tracked given the observations available at successive time-steps.

We now formalise three core notions that this model supports. Observe that, although the phenomena themselves are undisputed and supported by decades of research [3], there is less agreement about how to calculate their impact. Our strategy in each case, therefore, is to model the psychological principle as closely as possible while minimising changes made to an established approach.



(a) Observations $\vec{o} = o_1, o_2, .., o_7$ on a path from s to g_2. The probability of each goal is calculated at every observation and at g_2.

o_i	Generic GR: $Pr_{MS}(G \mid \vec{o})$ g_1	g_2	g_3	β
o_1	0.2687	0.3204	**0.4108**	0.1
o_2	**0.3844**	0.3257	0.2897	0.1
o_3	0.2463	0.3302	**0.4233**	0.1
o_4	**0.3515**	0.3431	0.3052	0.1
o_5	0.2111	0.3456	**0.4431**	0.1
o_6	0.2836	**0.3895**	0.3267	0.1
o_7	0.2083	**0.4026**	0.3890	0.1
g_2	0.2196	**0.4888**	0.2914	0.1

(b) Probabilities calculated for each goal $G = \{g_1, g2, g_3\}$ at each observation o_i and at g_2 using Equation (2). β is fixed, priors are always assumed equal (and ignored).

o_i	Extended GR: $Pr(G \mid \vec{Q})$ g_1	g_2	g_3	β
o_1	0.3332	0.3333	**0.3334**	0.0001
o_2	**0.3335**	0.3333	0.3331	0.0004
o_3	0.1133	0.2806	**0.6060**	0.3098
o_4	0.1134	0.2808	**0.6057**	0.0010
o_5	0.0001	0.0371	**0.9627**	1.0000
o_6	0.0001	0.0373	**0.9625**	0.0024
o_7	7.04E-06	0.0431	**0.9568**	0.4367
g_2	6.87E-06	0.0448	**0.9551**	0.0079

(c) Probabilities for each goal as at (b) but calculated using Equation (5). Priors are taken into account and β is updated to reflect the observer's confidence.

Figure 2: The Impact of Confirmation Bias. Under Generic GR, the goal regarded as most probable changes with every change of direction. Under Extended GR, the goal *first* believed most probable persists even after repeated zigzagging. Once confidence (β) has peaked at observation o_5 belief becomes very hard to shake, even when confidence later diminishes (at o_6 and g_2).

4.1 Confirmation Bias

"With respect to deception, one overwhelming conclusion stands out: It is far easier to lead a target astray by reinforcing the target's existing beliefs ... than to persuade a target to change his or her mind." [12, p.298]

As demonstrated in Section 3, prior probability—which stands in for an agent's previously held beliefs—is a critical factor in determining posterior probability. Although (in common with other models, for simplicity) we initially assume that prior probabilities across goals are equal, we also assume that, since observations are made online, they are iterative and cumulative. Thus, whereas [30], working offline, could arguably consider a complete sequence of observations as a single behaviour, online goal recognition implicitly considers each new observation with reference to those that have gone before: priors must be meaningfully considered.

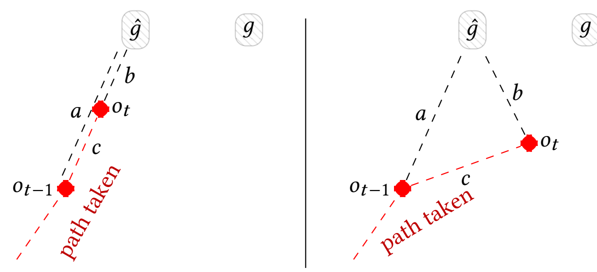
In formalising confirmation bias, however, prior probability is only part of the picture. Priors establish expectation but the full impact of confirmation bias as a vehicle for deception (including self-deception) arises when expectations are confirmed. When this occurs, the observer's confidence in the correctness of their prediction increases, their next prediction is made with even more certainty; and the effect can snowball.

In the context of a GR problem $\mathcal{P}_x$, we measure confidence in terms of progress made towards *the goal previously thought to be most probable* (i.e., the goal with the highest prior probability). Remembering that o_t is the most recently added observation at time-step t and allowing $P(\cdot)$ as a generic for probability, the definition measures the difference between optimal expected and actual progress, constrained between 0 and 1.

Definition 4. *Given a goal $\hat{g} \in G$ such that $P(\hat{g} \mid \vec{o}_{t-1}) \geq P(g \mid \vec{o}_{t-1})$ for all $g \in G \setminus \{\hat{g}\}$ (i.e., $\hat{g}$ was the most probable goal at time-step $t-1$), **confidence** is given by:*

$$conf(o_t, o_{t-1}, \hat{g}) = e^{optc(o_{t-1},\hat{g}) - optc(o_t,\hat{g}) - optc(o_{t-1},o_t)}. \quad (4)$$

The definition is illustrated at Figure 3 in terms of cost-distance. When the most recent observation o_t is on an optimal path towards



(a) $a - b - c = 0$
conf $= 1$

(b) $a - b - c < 0$
$0 <$ ***conf*** < 1

Figure 3: Distance equals cost. a = *optc*$(o_{t-1}, \hat{g})$, b = *optc*$(o_t, \hat{g})$, c = *optc*(o_{t-1}, o_t). $\hat{g}$ was the most probable goal at $t-1$. When the most recent observation is on an optimal path to $\hat{g}$, *confidence* $= e^0 = 1$. If not, $0 <$ *confidence* < 1.

the expected goal $\hat{g}$ (based on the previous observation o_{t-1}), confidence is maximised (Figure 3a); if the path deviates *un*expectedly, confidence is correspondingly reduced (Figure 3b).

Taking the two things together, the ***online solution*** to an extended GR problem $\mathcal{P}_x$ that accommodates the notion of confirmation bias is the distribution given by:

$$Pr(G \mid \vec{Q}) = \alpha \cdot \frac{1}{e^{\beta(optc(\vec{o}_t,g) - optc(s_t,g))}} \cdot Pr(g \mid \vec{o}_{t-1}) \text{ for } g \in G, \quad (5)$$

where α is a normalisation constant, $\beta = conf(o_t, o_{t-1}, \hat{g})$ and $\vec{Q}$, recall, is the sequence of *all* sets of observations, from which time-sensitive sequences $\vec{o}_t$ and $\vec{o}_{t-1}$ can be extrapolated (as later discussed, 4.2).

Observe that Equation (5) now explicitly incorporates prior probabilities by reference to the probability distribution calculated at the previous time-step, $Pr(g \mid \vec{o}_{t-1})$. Moreover, the β parameter has been made self-modulating to represent the observer's confidence.

Finally, notice that the above factors (expectation and confidence) work in tandem. An exaggeratedly confident prediction accentuates the probability of the most probable goal, potentially pushing it close to 1 and therefore pushing others close to 0. If, owing to the β value, an observer is sufficiently confident of a particular goal, the probability of that goal approaches 1. At the next time-step, that confident assessment becomes prior probability. Now other goals—regardless of implications arising from the most recent observation—may be overlooked.

Figure 2 illustrates the effect. An agent is observed seeming to zigzag between goals g_3 and g_1. Since g_3 is the first goal targeted (and therefore the first goal *believed* to be the real goal), probabilities continue to slightly favour that goal. Once confidence (β) reaches 1.0 at o_5, the balance has tipped so far that g_3 remains overwhelmingly the most likely goal, even when g_1 is approached quite closely (e.g., at o_6) and *after* g_2 has been achieved.

This matches our understanding from behavioural science and from [12]. Confirmation bias acts as a sort of inertia: people are inclined to continue to believe what they previously believed [13].

We have described confirmation bias in the context of an extended GR problem $\mathcal{P}_x$, within which time-dependent observations are explicitly defined. Time-dependent observation sequences are not uncommon, however. The concept can, therefore, be regarded as orthogonal to more standard approaches.

4.2 Selective Attention

Recall that extended GR is not concerned with identifying an agent's most likely goal but with demonstrating how easily an observer can be made to believe in a false goal. This brings us to the second core feature of our model: not everything available to an agent's fully-functioning sensors is necessarily stored and available for recall. When multiple observable phenomena occur simultaneously, a memory-constrained agent must decide which of them is most worthy of attention.

Drawing on recognised psychological principles of perceptual recognition involving bottom-up and top-down processing [3, 12], we model the selective attention process as the combined effect of an observation's magnitude or intensity (bottom-up) and its perceived relevance in the context of previously remembered observations (top-down).

4.2.1 Magnitude. Every observable element or event in a domain has a fixed *base* magnitude $mag^*(o)$ for $o \in \Omega$ (see Definition 3), which quantifies the degree to which it is inherently likely to attract attention. An explosion, for example, has a greater base magnitude than a firework, and a firework greater base magnitude than a cough. Moreover, although base magnitude is fixed, an observation's *effective* magnitude—as perceived and later recalled by an observer—is dynamic. This means that (a) the effective magnitude of an observation may diminish over time (see 4.3); and (b) its initial magnitude may itself be amplified or diminished by an agent willing to pay more for the event that gave rise to it. That is, given an observation o elicited by event e, with base magnitude $mag^*(o)$, to increase its initial magnitude by a factor of f, such that $mag(o) = mag^*(o) * f$, $cost(o) = c(e) * f$. Similarly, to diminish its initial magnitude such that $mag(o) = mag^*(o)/f$, $cost(o) = c(e) * f$.

To compare the relative magnitude of multiple simultaneous observations, we divide by their sum.[2]

[2] We have made the equation sigmoidal to future-proof against the possibility that all elements at the timestep might be unobservable.

Definition 5. *Given a set of potential observations O_t, the **comparative magnitude** of each $o \in O_t$ is given by:*

$$CM(o, O_t) = 0 \leq \frac{mag(o)}{1 + \sum_{o' \in O_t} mag(o')} < 1. \tag{6}$$

4.2.2 Relevance. "Information that does not fit neatly into an existing hypothesis tends to be ignored or misperceived." [12, p.300]

When multiple observable phenomena occur simultaneously, it is not only magnitude but also *relevance* that determines which of them will be remembered. That is, we notice and store for recall the observable phenomenon that *seems to make the most sense.* To formalise this, we build on the notion of a rationality measure (RM) from [26] as discussed in Section 2. Its documented purpose is to evaluate an agent's future expected degree of rationality, given their past behaviour. Here, we use it to evaluate the apparent rationality of observation sequences that would result from adding each of multiple *potential* observations (each $o \in O_t$) to the *recalled* observation sequence ($\vec{o}_{t-1}$) assembled so far. That is, given what we know, which potential observation provides the most rational continuation towards any one of the known possible goals.

Definition 6. *Given a set of possible goals G, potential observations O_t, and a sequence of previously attended observations $\vec{o}_{t-1}$, the **relevance** of an observation $o \in O_t$ is given by:*

$$rel(o, \vec{o}_{t-1}, G) = \max_{g \in G} \frac{optc(s_{t-1}, g)}{optc(s_{t-1}, \vec{o}_{t-1} \cdot o, g)}. \tag{7}$$

Observe that $\vec{o}_{t-1} \cdot o$ in the denominator of Equation (7) represents the observations available so far (i.e., at time-step $t-1$) to which each newly available observation $o \in O_t$ is appended.[3] Moreover, recall that s_{t-1} is the first remembered observation at time-step t. Thus Equation 7 is the RM of Equation 3, modified only to accommodate the variable observation sequence that the extended model allows.

Putting it together, we rely on both magnitude and relevance (multiplied to constrain within the range $[0, 1]$) to determine which of multiple potential observations is most likely to be attended, encoded, and available for future recall.

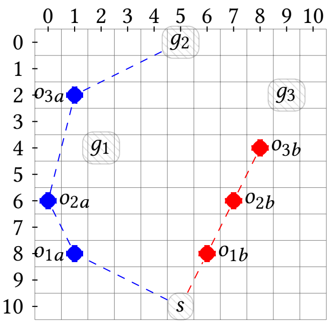
Definition 7. *Given a set of possible goals G, potential observations O_t, and a sequence of previously attended observations $\vec{o}_{t-1}$, the **attended observation** at time t, $o_t \in O_t$ is given by:*[4]

$$o_t(O_t, \vec{o}_{t-1}, G) = \underset{o \in O_t}{\arg\max}\, CM(o, O_t) \cdot rel(o, \vec{o}_{t-1}, G). \tag{8}$$

Example 1. Consider the example illustrated in Figure 4, which demonstrates the power of selective attention as a deceptive strategy. Starting at s, an agent wishes to reach g_2 without being observed. It constructs two plans to exploit the selective attention of the observer: one that it will follow ($\vec{o}_a = o_{1a}, o_{2a}, o_{3a}$, blue) and one that it wants its observer to believe ($\vec{o}_b = o_{1b}, o_{2b}, o_{3b}$, red). At each time-step, the agent contrives to generate two potential observations: one from each plan. Thus, in the context of an extended GR problem $\mathcal{P}_x$, the sequence of sets $\vec{Q} = O_1, O_2, ..$ available to the observer at each time-step is populated with two conflicting observations: $O_1 = \{o_{1a}, o_{1b}\}, O_2 = \{o_{2a}, o_{2b}\}, O_3 = \{o_{3a}, o_{3b}\}$.

[3] Relevance is measured w.r.t. known goals. If an agent's real goal g_r is unknown (i.e., $g_r \notin G$), the *relevance* of an observation that relates to it is minimised. This helps to explain why magicians rarely repeat a trick or disclose its ending [22, pp.135-8].

[4] If there are multiple such observations, selection may be randomised.



(a) By planning a suboptimal path ($\vec{o}_a$, blue) to its real goal g_2 and a false "optimal" path ($\vec{o}_b$, red) then contriving pairs of observations to be available simultaneously, an agent can manipulate an observer's selective attention.

$\vec{Q}$	max	Relevance g_1	g_2	g_3
o_{1a}	07984	0.7984	0.6903	6306
$\mathbf{o_{1b}}$	**1**	0.8973	0.9234	**1**
o_{2a}	0.6	0.6	0.4864	3728
$\mathbf{o_{2b}}$	**1**	0.6796	0.8578	**1**
o_{3a}	0.5775	0.5469	0.8009	0.4714
$\mathbf{o_{3b}}$	**1**	0.5469	0.8009	**1**

(b) A set of multiple observations is delivered at each time-step. $O_1 = \{o_{1a}, o_{1b}\}, O_2 = \{o_{2a}, o_{2b}\}$, etc. Evaluated using Equation (7), only the most relevant at each time-step (i.e., "max" above) is added to the sequence of encoded observations and used to predict the most likely goal.

o_i	Generic GR: $Pr_{MS}(G \mid \vec{o}_b)$ g_1	g_2	g_3	β
o_{1a}	**0.9158**	0.0642	0.0199	1.0
o_{2a}	**0.9193**	0.0765	0.0040	1.0
o_{3a}	**0.6519**	0.3380	0.0100	1.0
g_2	0.0003	**0.9940**	0.0056	1.0
$\vec{Q}$	**Extended GR: $Pr(G \mid \vec{Q})$** g_1	g_2	g_3	β
O_1	0.3240	0.3240	**0.3519**	0.1
O_2	0.0251	0.1456	**0.8292**	1.0
O_3	0.7.4E-05	0.0144	**0.9855**	1.0
g_2	0.0003	0.9940	**0.9854**	0.0004

(c) Generic GR cannot process simultaneous observations. If $\vec{o}_a$ is observable, probabilities calculated accordingly. Under extended GR, assuming equal intensity, probabilities are based on the most *relevant* observations.

Figure 4: Exploiting Selective Attention. An agent wishing to reach g_2 without being observed can use extended GR to plan two paths such that the false path $\vec{o}_b$ (red) is more relevant and/or has greater intensity than the true path $\vec{o}_a$ (blue).

If both observations in each set were of equal magnitude and evaluated as having equal relevance, then selection between them would be made randomly with a 50% chance of the agent being observed. From Equation (8), however, note that there are two ways that an observer's selective attention can be exploited by an agent wishing to avoid detection: by manipulating the magnitude of the observations so that those to be seen outweigh those to be hidden; and/or by fabricating the false sequence—in this case, $\vec{o}_b$—in such a way that its *relevance* exceeds that of the plan for the real goal ($\vec{o}_a$).

While the first approach (manipulating intensity) is the most obvious—effectively a distraction—the second (manipulating relevance) provides a long-term advantage. This is because, if an observable event seems to lack relevance, an *alternative* event will be encoded and remembered. That alternative becomes part of the (potentially falsified) recalled observation sequence against which the relevance of *all future observations* will be evaluated.

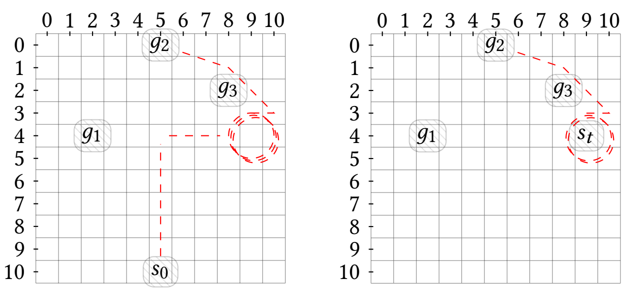
Referring again to Figure 4, a devious agent might adopt both strategies simultaneously, increasing the intensity of observations in $\vec{o}_b$ (e.g., by enlisting the assistance of a noisy confederate) while taking a path such as that shown in blue. As Tables 4b and 4c show, however, under extended GR, even if all observations are assumed to be of equal intensity, the more "relevant" observations on the red path $\vec{o}_b$ leave observations on the blue path $\vec{o}_a$ unattended.

Observe that this figure mirrors the common structure of a magic trick: the given story (to g_3) is shown to the audience while the hidden story is secretly contrived, arriving "as if by magic" at g_2.

Although our focus in this section has been on selective attention, it is worth noting the simultaneous impact of confirmation bias on the 'selectively' attended observations: priors achieved by faking the optimal (red) path $\vec{o}_b$ become so high that they confirm g_3 even after g_2 has been achieved.

4.3 Memory Decay

A memory-constrained agent (such as a human) cannot remember everything. Observed phenomena are not retained indefinitely and



(a) Traversed path. **(b) Remembered path at time t.**

Figure 5: The Impact of Memory Decay. An agent that remembers the full traversed path (a) retains the impression (confirmation bias) that the plan (in red) is targeting goal g_2. An agent impacted by memory decay (b) tries to solve the wrong problem, building expectation that the real goal is g_3.

our model adopts a conventional implementation of decay, familiar from work such as [33] involving pheromones.

Given an observation sequence $\vec{o}_t = o_1, o_2, ..o_n$, at every *subsequent* time-step, $t+1, t+2, etc.$, the effective magnitude of each element $mag(o_i)$ is multiplied by a decay factor $\delta < 1 \in \mathbb{R}^+$. If magnitude drops below some threshold of negligibility ϵ, the observation is removed from the sequence at the next time-step. Formally, $\vec{o}_t = o_1, o_2, ..o_n \mid mag(o_i) > \epsilon$ for each $i \in \{1..n-1\}$ and $o_n = o_t$ (the observation added at time t). Thus, $\vec{o}_t$ may represent a different observation sequence at every time-step. Typically, the sequence is first-in-first-out. That is, oldest observations are forgotten first. If an observation is particularly intense, however (i.e., has excessive initial magnitude), it may persist long after more standard observations have been forgotten so that an observation sequence at one time-step may even have different cardinality from that at another.

Although the general principle is well understood, Figure 5 illustrates the particular impact of memory decay in the context of deception. Recall from Definition 1 that the generic GR framework references a fully observable initial state but that in applying the extended model (list item 4, p.4), we substitute s_t, the first *remembered* state. Now, the core principle of cost-based GR is that of cost difference (Equation (1)): the difference in cost between an optimal plan and the plan that satisfies the observations—that is, what a fully rational agent *ought* to have done (optimal) versus what it actually appears to have done (observed). But to estimate *either* of those parameters, we need to know where the agent started from.

When the initial state changes, so does the problem to be solved. If the observing agent has forgotten the *real* initial state or is made to believe a *false* initial state, it is unknowingly trying to solve the wrong problem. This effect is at the core of many magic and confidence tricks (i.e., professional deceptions) where it is known as time or space displacement [27] and is a factor in priming, used in Ettinger and Jehiel's game theoretic account of deception [11], whereby an observing agent is "softened up" by being fed misleading information before the deception seems to have begun.

5 RELATED WORK

Our work is situated at the intersection of goal recognition, human-aware AI and deceptive planning.

We build on the tradition of cost-based goal recognition, pioneered by Ramirez and Geffner [29, 30] and Baker et al. [1, 2] and variously extended by Sohrabi et al. [32], Vered et al. [34] and Escudero-Martin et al. [10], amongst others. Our perspective on GR is its use as a model, not necessarily to determine an agent's goal but to determine a third party's probable *belief* about an observed agent's goal. Although some of the above authors implicitly make the distinction [e.g., 2, 34], the only "types" usually considered in the context of GR depend on whether or not the observed agent is *aware* of being observed [5]. The issue of third party belief has long been discussed in terms of "predictability" by the robotics community, notably by Dragan et al. [9] who demonstrates that optimal behaviour is not necessarily the most legible/predictable to a human observer. Rather, the key to maximising legibility is the elimination of ambiguity. Dragan et al. have also inverted that principle to develop deceptive motion which is *un*predictable [8]—much as Masters and Sardina [24] invert goal recognition to model deceptive path-planning.

With the growing interest in human-aware computing, ideas about legibility have been reinvigorated under the label "interpretable AI". In this context, the difference between the ground truth and human perception of the truth, which we explore, is critical. The distinction has been examined by Kulkarni et al. [19], whose model supports both cooperative and adversarial interaction, and by Chakraborti et al. [7], who have conducted a recent survey of the field. In other work, Chakraborti and Kambhampati [6] show how an understanding of the gap between reality and belief can be exploited for deceptive purposes in the context of explainable AI by an agent withholding information or providing an explanation that is *acceptable* rather than necessarily true.

Although much of the above research ostensibly deals with deception, there is a tendency—clear in the Chakraborti et al. survey and apparent also in Keren et al.'s work on goal recognition design [14–16]—to conflate deception with privacy. Privacy, moreover, is frequently associated with ambiguity [e.g., 18, 24]. This may be a response to Dragan et al.'s finding (mentioned above) that legibility demands the elimination of ambiguity: by maximising ambiguity, information can be made secure. Our work departs from this tradition by deliberately introducing misdirection as construed by professional deceivers (e.g., military strategists [4, 12] and professional magicians [21, 27]): deception that points *away* from the truth that we are trying to hide by pointing *towards* the falsehood we want to be seen.

Extended GR draws not only on professional deception but also on behavioural economics [13] and standard psychological texts [3]. In modelling some of the predictable and well-documented idiosyncrasies of the typical human-in-the-loop—aspects which are not usually considered—we see our work as relevant and, we hope, orthogonal to research in all three of the above domains.

6 CONCLUSION

Extended goal recognition is not attempting to determine the ground truth; that is not its purpose. Rather, it acts as a half-way-house towards human-aware planning: a rapidly expanding area of research, largely dominated by cooperative planning and explainable AI. It is in the context of deception, however, that the need for an improved model becomes obvious, owing to the many ways that people can be deceived which pre-existing models are unable to represent.

Deception in the context of AI planning is inevitably simplistic by comparison with that which occurs in the real world. From our lived experience, people may come to believe a falsehood for many reasons. For example, a contradictory event occurred in plain sight, but they did not see it. If they did see it, they failed to recognise its significance so did not remember it. If they did remember it, they later forgot it. Or perhaps, even as they saw it, they were already so convinced of an "alternative fact" that they did not believe it. Extended goal recognition supports all of these possibilities.

We have highlighted a potential limitation with respect to the implementation of prior probability in recent models of cost-based goal recognition when used to facilitate human-aware planning, and note that the problem may be exacerbated rather than fixed in landmark-based accounts. We suggest that, practically and conceptually, this contribution is important since it flags an issue that seems to have gone unnoticed in a body of work well-cited and frequently extended. We have built on well-respected, existing models to present a novel GR framework that formalises notions of confirmation bias, selective attention and memory decay. To achieve this our model incorporates a parameter that self-adjusts when expectations (evidenced by prior probabilities) are confirmed. It supports time-sequential *sets* of observations within which each observation has an initial magnitude that decays over time; and we have demonstrated how a "selectively attended" observation may be chosen from a set based on its magnitude and relevance. We have demonstrated that the model can be used to support—and, indeed, to explain—how easily false beliefs may arise, how difficult they can be to overturn and how readily they can be exploited in the context of strategic deception.

ACKNOWLEDGMENTS

We thank our reviewers and gratefully acknowledge the support of the Australian Research Council (DP180101215).

REFERENCES

[1] Chris L Baker, Rebecca R Saxe, and Joshua B Tenenbaum. 2009. Action understanding as inverse planning. *Cognition* 113, 3 (2009), 329–349.

[2] Chris L Baker, Rebecca R Saxe, and Joshua B Tenenbaum. 2011. Bayesian theory of mind: Modeling joint belief-desire attribution, In Proceedings of the Cognitive Science Society. *Proceedings of the Cognitive Science Society*, 2469–2474.

[3] DA Bernstein, LA Penner, A Clarke-Stewart, and EJ Roy. 2006. *Psychology*. Houghton Mifflin Company, Boston.

[4] J. Barton Bowyer. 1982. *Cheating: Deception in War & Magic, Games & Sports.* St Martin's Press. J. Bowyer Bell and Barton Whaley, under a pseudonym.

[5] Sandra Carberry. 2001. Techniques for plan recognition. *User Modeling and User-Adapted Interaction* 11, 1-2 (2001), 31–48.

[6] Tathagata Chakraborti and Subbarao Kambhampati. 2019. (When) can AI bots lie?. In *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*. 53–59.

[7] Tathagata Chakraborti, Anagha Kulkarni, Sarath Sreedharan, David E Smith, and Subbarao Kambhampati. 2019. Explicability? legibility? predictability? transparency? privacy? security? the emerging landscape of interpretable agent behavior. In *Proceedings of the International Conference on Automated Planning and Scheduling*, Vol. 29. 86–96.

[8] Anca D Dragan, Rachel M Holladay, and Siddhartha S Srinivasa. 2014. An Analysis of Deceptive Robot Motion.. In *Robotics: science and systems*. 10.

[9] Anca D Dragan, Kenton CT Lee, and Siddhartha S Srinivasa. 2013. Legibility and predictability of robot motion. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction*. 301–308.

[10] Yolanda Escudero-Martin, Maria D Rodriguez-Moreno, and David E Smith. 2015. A fast goal recognition technique based on interaction estimates, In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI). *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 761–768.

[11] David Ettinger and Philippe Jehiel. 2010. A theory of deception. *American Economic Journal: Microeconomics* 2, 1 (2010), 1–20.

[12] Richards J Heuer Jr. 1981. Strategic deception and counterdeception: A cognitive process approach. *International Studies Quarterly* 25, 2 (1981), 294–327.

[13] Daniel Kahneman. 2011. *Thinking, fast and slow.* Farrar, Straus and Giroux, U.S.A.

[14] Sarah Keren, Avigdor Gal, and Erez Karpas. 2014. Goal Recognition Design, In Proceedings of the International Conference on Automated Planning and Scheduling (ICAPS). *Proceedings of the International Conference on Automated Planning and Scheduling (ICAPS)*, 154–162.

[15] Sarah Keren, Avigdor Gal, and Erez Karpas. 2015. Goal Recognition Design for Non-Optimal Agents, In Proceedings of the National Conference on Artificial Intelligence (AAAI). *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 3298–3304.

[16] Sarah Keren, Avigdor Gal, and Erez Karpas. 2016. Privacy Preserving Plans in Partially Observable Environments, In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI). *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 3170–3176.

[17] Gustav Kuhn. 2019. *Experiencing the impossible: The science of magic.* MIT Press.

[18] Anagha Kulkarni, Matthew Klenk, Shantanu Rane, and Hamed Soroush. 2018. Resource Bounded Secure Goal Obfuscation. In *AAAI Fall Symposium on Integrating Planning, Diagnosis and Causal Reasoning.*

[19] Anagha Kulkarni, Siddharth Srivastava, and Subbarao Kambhampati. 2018. A unified framework for planning in adversarial and cooperative environments. In *ICAPS Workshop on Planning and Robotics.*

[20] Anagha Kulkarni, Siddharth Srivastava, and Subbarao Kambhampati. 2020. Signaling Friends and Head-Faking Enemies Simultaneously: Balancing Goal Obfuscation and Goal Legibility. *Published independently online* (2020).

[21] Peter Lamont and Richard Wiseman. 2005. *Magic in theory: An introduction to the theoretical and psychological elements of conjuring.* Univ of Hertfordshire Press.

[22] Stephen Macknik, Susana Martinez-Conde, and Sandra Blakeslee. 2010. *Sleights of mind: What the neuroscience of magic reveals about our everyday deceptions.* Henry Holt and Company.

[23] Peta Masters and Sebastian Sardina. 2017. Cost-based goal recognition for path-planning. In *Proceedings of Autonomous Agents and Multi-Agent Systems (AAMAS).* International Foundation for Autonomous Agents and Multiagent Systems, 750–758.

[24] Peta Masters and Sebastian Sardina. 2017. Deceptive Path-Planning. In *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI).* 4368–4375.

[25] Peta Masters and Sebastian Sardina. 2019. Cost-Based Goal Recognition in Navigational Domains. *Journal of Artificial Intelligence Research (JAIR)* 64 (2019), 197–242.

[26] Peta Masters and Sebastian Sardina. 2019. Goal Recognition for Rational and Irrational Agents. In *Proceedings of Autonomous Agents and Multi-Agent Systems (AAMAS).* 440–448.

[27] Darwin Ortiz. 2006. *Designing Miracles: Creating the Illusion of Impossibility.* Ortiz Publications.

[28] Ramon Fraga Pereira, Nir Oren, and Felipe Meneguzzi. 2020. Landmark-based approaches for goal recognition as planning. *Artificial Intelligence* 279 (2020), 103217.

[29] Miquel Ramirez and Hector Geffner. 2009. Plan recognition as planning, In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI). *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 1778–1783.

[30] Miquel Ramirez and Hector Geffner. 2010. Probabilistic plan recognition using off-the-shelf classical planners, In Proceedings of the National Conference on Artificial Intelligence (AAAI). *Proceedings of the National Conference on Artificial Intelligence (AAAI)*, 1121–1126.

[31] S. Russell and P. Norvig. 2013. *Artificial Intelligence: A Modern Approach* (3 ed.). Prentice Hall, New Jersey, USA.

[32] Shirin Sohrabi, Anton V. Riabov, and Octavian Udrea. 2016. Plan Recognition as Planning Revisited., In Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI). *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, 3258–3264.

[33] H Van Dyke Parunak, Sven Brueckner, Robert Matthews, John Sauter, and Steve Brophy. 2007. Real-time agent characterization and prediction. In *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems.* 1–8.

[34] Mor Vered, Gal A. Kaminka, and Sivan Biham. 2016. Online goal recognition through mirroring: Humans and agents. In *Conference on Advances in Cognitive Systems.*