# A New Metric and Method for Goal Identification Control

**Kai Xu**[1], **Quanjun Yin**[1], **Qi Zhang**[1]
1. College of System Engineering
National University of Defense Technology, Changsha, 410073, China.
{xukai09@, yinquanjun@, zhangqiy123@}nudt.edu.cn

## Abstract

Recent research has found situations where the identification of agent goals could be purposefully controlled, either by changing the underlying environment to make it easier, or exploiting it during agent planning so as to delay the opponent's goal recognition. The paper tries to answer the following questions: what kinds of actions contain less information and more uncertainty about the agent's real goal, and how to describe this uncertainty; what is the best way to control the process of goal identification. Our contribution is the introduction of a new measure we call *relative goal uncertainty (rgu)* with which we assess the goal-related information that each action contains. The *rgu* is a relative value associated with each action and represents the goal uncertainty quantified by information entropy after the action is taken compared to other executable ones in each state. After that, we show how goal vagueness could be controlled either for one side or for both confronting sides, and formulate this goal identification control problem as a Mixed-Integer Programming problem. Empirical evaluation shows the effectiveness of the proposed solution in controlling goal identification process.

## 1 Introduction

Goal recognition, the ability to recognize the plans and goals of other agents, enables humans, AI agents or command and control systems to reason about what the others are doing, why they are doing it, and what they will do next [Sukthankar *et al.*, 2014]. Until now, goal recognition system works well in many applications like human-robot interaction [Hofmann and Williams, 2007], intelligent tutoring [Min *et al.*, 2014], system intrusion detection [Geib and Goldman, 2001] and security applications [Jarvis *et al.*, 2005].

Though this technique has been successfully applied to many application domains, new problem arises when goal recognition encounters with goal uncertainty discovered in certain environmental settings. Figure 1 offers a simple example that is first mentioned in [Keren *et al.*, 2014] and will help to clarify the concepts of our work.

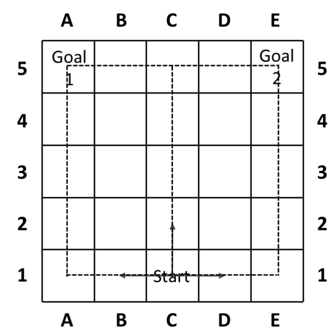



Figure 1: An example where goal identification process could be controlled

The model consists of a simple room (or airport) with a single entry point, marked as 'Start' and two possible exit points (boarding gates), marked as 'Goal 1' (domestic flights) and 'Goal 2' (international flights). An agent can move vertically or horizontally from 'Start' to one of the goals. Notice that in this model, the goal of the agent becomes clear once turning left or right, while moving vertically would impose goal uncertainty on goal recognizers and postpone the goal identification. Therefore, in the worst case an optimal agent can move up 4 steps before it is obliged to turn towards its goal.

The goal uncertainty showed in the above example could be viewed as an inherent property of one particular goal recognition task, and pose a new challenge to effective goal reasoning. The problem finds itself relevant in many of the similar applications of goal recognition tasks where goal uncertainty exists. Also, it should be noted that, apart from the friendly situation in Figure 1 and works in [Keren *et al.*, 2014], the goal uncertainty could further be used as a potential backdoor for the adversary to delay goal identification.

Therefore, this paper focuses on effective control of the goal identification process. As a first stage of our exploration, we assume agents are optimal and that the actions of the agent are fully observable and deterministic. In order to achieve our objective, we introduce a new concept called *relative goal uncertainty (rgu)*, with which we assess the goal-related information that each action contains. The *rgu* is a relative value associated with each action that agent would take and represents the goal uncertainty quantified by information entropy after the action is taken compared to other executable ones in each state.

The *rgu* value associated with each action helps in assessing the uncertainty that exists in goal recognition task, and our goal is to provide methods for agents to control it. Towards this end, we define two optimization models based on Mixed-Integer Programming, one of which reduces the goal uncertainty through limiting the set of available actions an agent can perform, the other one delays goal identification through balancing the reward and goal uncertainty during mission planning.

Goal identification control, while relevant to goal recognition [Sukthankar *et al.*, 2014] or goal reasoning, is a different task. While goal recognition aims at discovering the goals of an agent according to observations, goal identification control rather offers an offline and online combined solution for assessing and controlling the goal uncertainty in order to assure the goal of any optimal agent in the system is recognized. Also, different from goal recognition design problem [Keren *et al.*, 2014; Son *et al.*, 2016; Wayllace *et al.*, 2017] which purely focuses on facilitating goal recognition offline, goal identification control provides a more compact solution for agents to control the goal uncertainty and allows this additional information to be incorporated into ongoing missions.

Finally, considering the situation where both confronting sides are trying to control the goal identification process (*i.e.*, both sides are behaving in a strategic game-theoretic manner), the paper then proposes the Shortest-Path Largest-Uncertainty Network Interdiction model (SPLUNI), which could be viewed as a bilevel mixed-integer program, and is also a special case of a static Stackelberg game. The SPLUNI model is transformed into its dual form using KKT conditions and computed using mixed-integer programming method.

The paper is organized as follows. We start by providing the necessary background on probabilistic goal recognition. We continue by introducing the formal model representing the goal identification control problem, and the *rgu* value. The following sections present the methods we developed for calculating *rgu* and controlling the goal uncertainty of a given goal identification control problem. We conclude with an empirical evaluation, a discussion of related work, and a conclusion.

## 2 Probabilistic Goal Recognition

[Xu *et al.*, 2017] shows a simple yet effective probabilistic goal recognizer based on Markov Decision Process (MDP). The probabilistic goal recognition model is defined by a tuple $D = \langle S, s_0, A, f, G, e, O\rangle$, where $S$ is a finite and discrete state space; $s_0$ is the start state of the agent; $A$ is the set of actions; $f: S \times A \rightarrow S$ is a deterministic state transition function; $G$ is a set of goal states; $e$ is the goal termination variable and $O$ is a non-empty observation set. Essentially, the model is a dynamic Bayesian network, in which all causalities could be depicted. We introduce a full DBN structure depicting two time slices is presented in Figure 2. The behaviors of system evolution are described using functions or parameters.

- state transition function $T: S \times A \times S \rightarrow [0,1]$ is $P_{s_\tau} = p(s_\tau | s_{\tau-1}, a_\tau)$,
- observation function $S \times O \rightarrow [0,1]$ is $P_{o_\tau} = p(o_\tau | s_\tau)$,
- agent action policy $P_{a_\tau} = p(a_\tau | s_{\tau-1}, g_\tau)$,
- goal transition probability $P_{g_\tau} = p(g_\tau | e_{\tau-1}, g_{\tau-1})$,
- goal termination probability $P_{e_\tau} = p(e_\tau | g_\tau, s_\tau)$.

Recognizing the evader's goal is an inference problem trying to find the real goal behind agent actions based on observations online. In essence, the task is to compute the posterior distribution $P(g_\tau | o_\tau)$ of goal $g_\tau$ given observation $o_\tau$. This could be achieved either by accurate inference or by approximate methods.

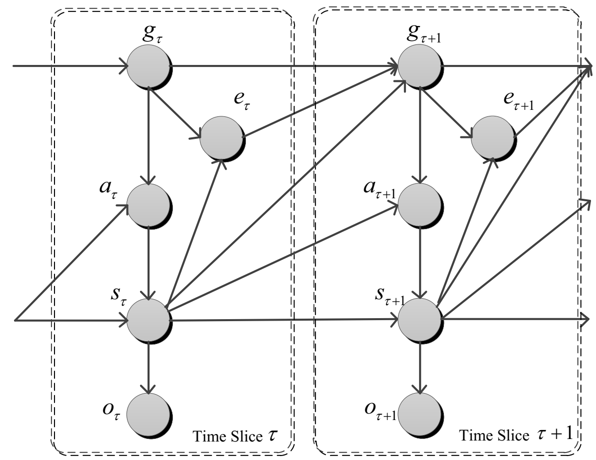



Figure 2: The DBN structure of the model

## 3 Relative Goal Uncertainty

Given the set of possible goal states $G$ and the current observation $o_\tau$, the probability distribution over possible goals $P(G|o_\tau)$ actually reflects the goal uncertainty the agent will encounter after selecting actions $A(s_\tau)$ and reaches the next state $o_\tau$ (assuming fully observable and deterministic) according to $f$. This enlightens us to use the goal inference information one step ahead to measure the goal uncertainty associated with available actions at each state, specifically the adjacent edges directing out of each node in the road network example.

Defined over each possible action $a_i \in A(s)$ at state $s$, the metric *rgu* is a relative value describing the overall uncertainty performance of state $s$ before and after the action $a$'s interdiction. Intuitively, this could be done by summing up the information entropy of the pre-computed, next-step goal probability which are reasoned about in the goal recognition problem $D = \langle S, s_i, A, f, G, e, O = \{s_i, s_j\}\rangle$, where $s_i \in S$ and $s_j \in FS(s_i)$, the set of states that could arrive at according to $f(s_i)$.

**Definition 1. (Entropy-based *rgu*)** *The* Entropy-based *rgu* *($rgu_E$) over a goal identification control problem is defined as:*

$$rgu_E(s, a_i) = \frac{\sum_{a_j \in A'(s)} H(G|a_j, s)}{|A'(s)|} \tag{1}$$

*where $a_i \in A(s), a_j \in A'(s)$, and $A'(s) = A(s) \backslash a_i$. The entropy $H(G|a_j, s)$ computes the goal uncertainty contained in the reasoning result, and $H(G|a_j, s) = -\sum p(G|a_j, s) \cdot \log p(G|a_j, s)$. Technically, we also need to remove zero entries of $p(G|a_j, s)$, followed by normalization guaranteeing that $sum(p) = 1$.*

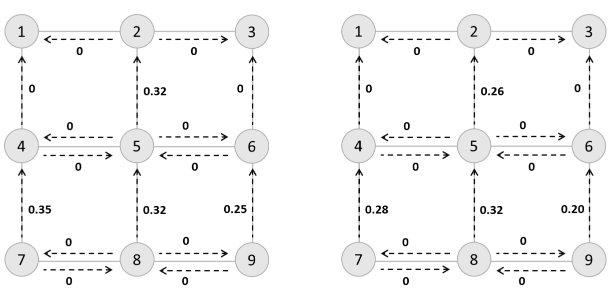
The $rgu_E$ quantified by entropy $H$ evaluates the goal uncertainty in a natural and compact way. More uncertainty introduced after the agent takes its action gives less information to goal recognizer and postpones the goal identification process. As the example in Figure 1 and assuming the agent's are fully optimal, the $rgu_E$ for three actions leading to states $(2, C)$, $(1, B)$ and $(1, D)$ are $0.32$, $0.00$ and $0.00$ respectively. Also, the definition of $rgu_E$ could easily be applied to tasks where $|G| > 2$.

It should be noted that, the definition of $rgu_E$ makes sure that this metric can be naturally applied to the situation where the agent changes its goal during the midway, as shown in Theorem 1.

**Theorem 1.** *The $rgu_E$ definition is applicable to goal changing situations.*

*Proof.* As $rgu_E$ is defined using posterior $p(G|a_i, s)$ computed in goal recognition task $D$, thus it only depends on the problem structure $D = \langle S, s_0 = s_i, A, f, G, e, O = \{s_i, s_j\}\rangle$, $\forall s_i \in S$. Thus $rgu_E$ value has nothing to do with agent's particular starting point. □

To illustrate how $rgu_E$ would be used in goal identification control problem and for clarity, we compute the values upon a $3 * 3$ grid network (Figure 3 (a)) with nodes representing grids and dashed edges connecting nodes.




(a) (b)

Figure 3: The $rgu_E$ and Discounted *rgu* in a 3*3 grid network

Selecting absorbing nodes No.1 and No.3 as possible goals, we compute $rgu_E$ for each action depicted as dashed line, as shown in Figure 3 (a). Higher $rgu_E$ means greater uncertainty associated with that action. Here we discuss two path planning tasks $T$ and $T'$, where the agent has different starting points ($s_0 = 9$ or $s_0 = 8$). The tasks are defined as $T = \langle S, A, f, C', G = \{1, 3\}, s_0\rangle$, with $C' = C + rgu_E$ to model the road barrier, patrolling police or other methods the goal recognizer could use in order to control the goal identification process.

For the first task, the agent would choose $h = \langle 9, 8, 7, 4, 1\rangle$ for goal No.1 and $h = \langle 9, 6, 3\rangle$ for goal No.2. While for the second task, using the definition of $rgu_E$, the agent would choose one of the two optimal routes ($h = \langle 8, 7, 4, 1\rangle$ or $h = \langle 8, 5, 4, 1\rangle$), other than $h = \langle 8, 5, 2, 1\rangle$ for goal No.1, and $h = \langle 8, 9, 6, 3\rangle$ or $h = \langle 8, 5, 6, 3\rangle$ other than $h = \langle 8, 5, 2, 3\rangle$ for goal No.3. For the latter case, we find that $rgu_E$ cannot help the agent capture the fact in goal identification control that early interdiction would be more important than a later one.

In order to solve this problem, we define a Discounted *rgu* so as to reduce the *rgu* value along with the increase of the number of steps that agent has taken from the start. For agent changing its goal during the midway, we could recompute this value from the original $rgu_E$. Case 3.1 shown in Section 4.1, which is talked about in the next section, gives a good indication of this situation occurrence.

**Definition 2.** **(Discounted *rgu*)** *The discounted rgu ($rgu_{dis}$) of a goal identification control problem is defined as:*

$$rgu_{dis}(s, a_i) = rgu(s, a_i) * \beta^d \tag{2}$$

*where $\beta$ is the discount factor and $d$ is the number of steps that agent has taken from the start.*

With the discount rate equals to $0.8$, Figure 3 (b) shows the discounted *rgu* values which have been adjusted according to its importance to the goal identification control task.

## 4 Goal Identification Control

In this section, we show how to control the *rgu* using the optimization technique. Instead of requiring the cost of cost-minimal plans to achieve each goal $g \in G$ be the same before and after removing the subset of actions, as researched about in Goal Recognition Design problem [Keren *et al.*, 2014], we are more interested in deliberate changing goal uncertainty under adversarial settings where the resource constraint compared to optimality conservation is more applicable. Also, the original way of removing actions permanently is changed "softer" by adding additional costs to actions.

### 4.1 Reduction and Improvement of *rgu*

The goal identification control models based on optimization techniques are introduced. We first present models which are individually used for reducing and improving goal uncertainty for the different sides of the goal recognition task. Then we talk about the applicability of the offline-computed *rgu* in the online goal identification control.

The models could be transformed into the problem of maximizing (minimizing) the expectation of $s - t$ path length, with the *rgu* being proportionally added to the original length. The mathematical-programming formulation of our new goal identification control model is as follows:

**Problem:** Maximize (Minimize) the expectation $s - t$ path length in a directed network by interdicting arcs,

**Indices:** $i \in N$, nodes in $G$ ($s$ is the current source node, $t$ is the terminus ),
$k = (i, j) \in A$, arcs in $G$,
$k \in FS(i)(k \in RS(i))$, arcs directed out of (into) node $i$,

**Data:** $c_k = 1$, normal length of arc $k$ (vector form $\mathbf{c}$),
$d_k = 1$, added integer delay if arc $k$ is interdicted (vector form $\mathbf{d}$),
$rgu_k$, relative goal uncertainty of arc $k$ (vector form $\mathbf{rgu}$)
$r_k = 1$, resource required to interdict arc $k$ (vector form $\mathbf{r}$),
$R$, total amount of interdiction resource,

**Variables:** $x_k = 1$ if arc $k$ is interdicted by the interdictor; else $x_k = 0$,
$y_k = 1$ if arc $k$ is traversed by the evader; else $y_k = 0$

The formulation of *rgu* reduction problem (**[RGUR-P]**) for the observer is:

$$\textbf{[RGUR-P]} \quad \max_{\mathbf{x} \in X} \sum_{k \in A} (c_k + x_k d_k (1 + rgu_k)) y_k$$

$$\sum_{k \in FS(i)} y_k - \sum_{k \in RS(i)} y_k = \begin{cases} 1 & \text{for} \quad i = s \\ 0 & \forall i \in N \backslash \{s, t\} \\ -1 & \forall i = t \end{cases} \tag{3}$$

$$x_k, y_k \in \{0, 1\}, \quad \forall k \in A \tag{4}$$

where $X = \{\mathbf{x} \in \{0, 1\}^{|A|} | \mathbf{r}^T \mathbf{x} \leq R\}$, Eq. (3) is the flow-balance constraint.

While the formulation of *rgu* improvement problem (**[RGUI-P]**) for the observed agent is:

$$\textbf{[RGUI-P]} \quad \min_{\mathbf{y}} \sum_{k \in A} \frac{(c_k + x_k d_k) y_k}{1 + rgu_k}$$

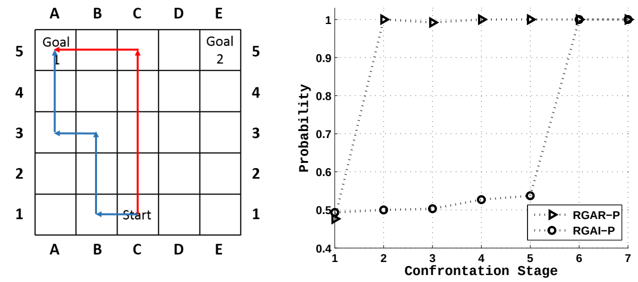
$$\sum_{k \in FS(i)} y_k - \sum_{k \in RS(i)} y_k = \begin{cases} 1 & \text{for} \quad i = s \\ 0 & \forall i \in N \backslash \{s, t\} \\ -1 & \forall i = t \end{cases} \tag{5}$$

$$x_k, y_k \in \{0, 1\}, \quad \forall k \in A \tag{6}$$

where $X = \{\mathbf{x} \in \{0, 1\}^{|A|} | \mathbf{r}^T \mathbf{x} \leq R\}$. Thus with additional goal uncertainty information, the problem could be transformed into network interdiction models, which are of a typical mixed-integer program (MIP) and solved using linear programming algorithms.

Figure 4 shows two moving traces of the observed agent under [RGUR-P] and [RGUI-P] problems. The observer could use [RGUR-P] to reduce the enemy's goal uncertainty and accelerate situation awareness, while the observed agent, after analyzing the observer's goal recognition task, could actually use [RGUI-P] to cover its real intention to the maximum extent. General results would be given in Section 5.

Note that, until now, the metric *rgu* used for assessing goal uncertainty is defined over actions available at each state in an offline manner, *i.e.*, $p(G|a_i, s)$ are computed according to $P(G|O) = \alpha P(O|G) P(G)$, with prior probability $P(G)$



(a) (b)

Figure 4: The moving traces (blue line for agents moving after the observer's interdiction; red for the observed agent selecting the most ambiguous path) and their corresponding recognition results $p(G|O)$ for the example shown in Figure 1 (a).

following Uniform distribution, instead of over agent's full course of actions where the real-time $P'(G)$ contains history information, this may incur uncertainty inconsistency among individual states and the whole task. Fortunately, our definition of *rgu* has no effect on problem solving of the goal identification control,

Case 1: If the real-time prior $P'(G)$ cannot help observer distinguish the right goal, meaning that $H(P'(G))$ is relatively high. Then it is naturally established for agents choosing actions with either high or low uncertainty.

Case 2: If $P'(G)$ distinguishes the goals clearly where $H(P'(G))$ is relatively low, and agent chooses the action that has high uncertainty, *i.e.*, *rgu* $\neq 0$, then highly uncertain action would input little information to the original beliefs about agent goals. The agent's belief of the right goal would still maintains. However, in this case, resource needed for action interdiction would be wasted as little information would be given and agent's intention usually maintains for a period.

Case 3: Also if $P'(G)$ distinguishes the goals clearly, *e.g.* the prior $P'(g)$ for goal $g$ with the largest value, while agent chooses the action that has low uncertainty *rgu*, then two cases should be considered. Firstly, agent's belief of the right goal has changed according to the posterior $P(G|o)$ (Case 3.1). This happens when the agent is changing its goal from $g$ to $g'$ during the midway, and thus chooses the action that will lead to $g'$. Secondly, the computed posterior $P(G|o)$ still capture the right goal $g$ (Case 3.2). For both cases, the *rgu* works properly.

## 4.2 The SPLUNI Model and its Dual Form

After individually talk about *rgu* reduction and improvement, in this section we propose a Shortest-Path Largest-Uncertainty Network Interdiction model (SPLUNI), where uncertainty becomes another optimization objective along with path length. The SPLUNI model could be viewed as a bilevel mixed-integer program (BLMIP), and is a special case of static Stackelberg game.

The problem description of SPLUNI is similar to [RGUR-P] and [RGUI-P], whereas it is described as maximizing the expectation of the shortest $s - t$ path length while minimizing the largest uncertainty and $0 \leq c_k, d_k, r_k < \infty$. Its mathematical-programming formulation is given as:

$$\textbf{[SPLUNI-P]} \quad \max_{\mathbf{x}\in X} \min_{\mathbf{y}} \sum_{k\in A} \frac{(c_k + x_k d_k(1+\alpha * rgu_k))y_k}{1+\beta * rgu_k}$$

with the same constraints as Eq. (3), Eq. (4) and $\mathbf{r}^T\mathbf{x} \leq R$. $\alpha, \beta$ are parameters controlling the importance of goal uncertainty between two objectives.

As a BLMIP problem, which cannot be solved directly using the MIP approaches, we propose a dual reformulation of SPLUNI. Fix the outer variable $x$ and take the dual of the inner minimization using KKT conditions, the final MIP formulation is given as:

$$\textbf{[SPLUNI-D]} \quad \max_{\mathbf{x}\in X, \vec{\pi}} \mathbf{b}^T\vec{\pi}$$

$$s.t. \quad \mathbf{K}^T\vec{\pi} = \mathbf{c}' + \mathbf{D}'\mathbf{x} \tag{7}$$

$$\pi_s = 0 \tag{8}$$

where $\mathbf{K}$ is the network matrix controlling $\mathbf{y}$ as in Eq. (3), $\mathbf{r}^T\mathbf{x} \leq R$, $\mathbf{b} = (1, 0, \cdots, 0, -1)^T$, $c'_k = c_k/(1+\beta * rgu_k)$ and $d'_{kk} = d_{kk} * (1+\alpha * rgu_k)/(1+\beta * rgu_k)$.

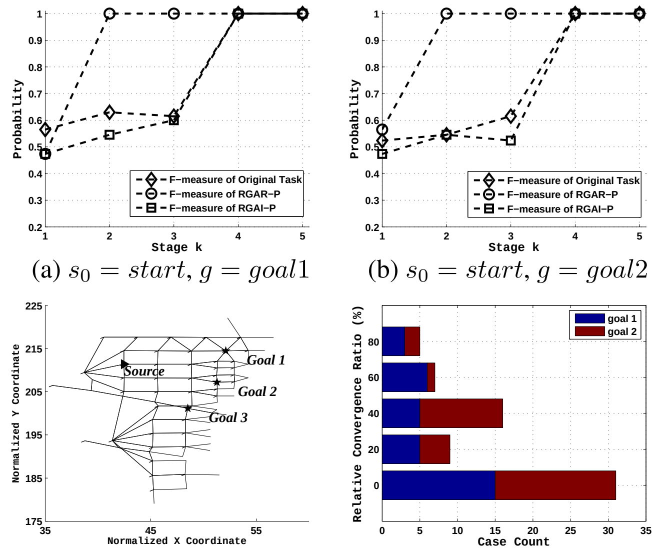
# 5 Experiments

For simplicity, we select a reduced Chicago Sketch Road Network [Xu *et al.*, 2017] expanded from the vertex No.368 to its neighbours and neighbours' neighbours for 5 times, consisting 51 vertexes and 113 edges. The computations of the RGUR and RGUI are formulated into a BLMIP, and SPLUNI into a BLMIP and solved using the MIP solvers of CPLEX 11.5 and YALMIP toolbox of MATLAB [Lofberg, 2005]. Nontrivial details are omitted.

## 5.1 Tests on Uncertainty Reduction

Upon the reduced Chicago sketch road network, a goal recognition task is defined where $s_0 = start$, $G = \{goal1, goal2\}$ as shown in Figure 5 (c). We first show the influence of uncertainty on the goal recognition system and how *rgu* reduction model could be used in advancing situation awareness. A test dataset consisting of 100 labeled traces for each goal is collected using agent decision model.

Evaluated using *F*-measure, which is a metric that has been frequently used to measure the overall accuracy of recognizer [Sukthankar *et al.*, 2014], goal uncertainty associated with certain domains indeed seriously impedes effective recognition (Figure 5 (a) and (b)). Under [RGUR-P] model, the values of *F*-measure after the network redesignation increases to almost 1 the moment agent selects its first action, compared to values under tasks with no changes and the most ambiguous situation ([RGUI-P]). Also, we statistically evaluate the maximum early prediction that our model enables according to *Convergence Point* metric [Sukthankar *et al.*, 2014], defined as $0 < \tau = CPoint \leq T,\ s.t.\ p_\tau(g_{true}|o_\tau) \geq \beta$. For tests in Section 5, $\beta = 0.8$. We test recognition tasks $D = \{s_0, G = \{goal1, goal2\}\}$ for $\forall s_0 \in S\backslash\{G, U\}$, where $U$ is the set of nodes having no available paths to $G$, and



(a) $s_0 = start, g = goal1$ (b) $s_0 = start, g = goal2$

(c) 5-layer Chicago Roadmap (d) Convergence Ratio

Figure 5: The statistical evaluation of uncertainty influence on goal recognition upon a reduced sketch road network.

compare early convergence between [RGUR-P] and [RGUI-P] using

$$RelConvergeRatio = \frac{CPoint_I - CPoint_R}{|h = \langle s_0, \ldots, g_i\rangle| - 1} \tag{9}$$

Clearly, both for tasks where $g_{true}$ equals to $goal1$ and $goal2$, there are considerable reductions of goal uncertainty. Also it should be noted that, although the uncertainty widely exists in many recognition tasks, it does not exist in all cases. Further for generality, we test 3 more cases where $G_1 = \{goal2, goal3\}$, $G_2 = \{goal1, goal3\}$, $G_3 = \{goal1, goal2, goal3\}$, as shown in Table 1.

Table 1: The number of cases (discard states unreachable to $g_{true}$) fall in intervals of Relative convergence ratio ($rcr(\%)$). "/" separates results with elements in $G$ being the true goal.

| $G\backslash rcr(\%)$ | 0 | $\leq 20$ | $\leq 40$ | $\leq 60$ | $\leq 80$ |
|---|---|---|---|---|---|
| $\{goal2, goal3\}$ | 23/24 | 1/3 | 3/3 | 2/0 | 1/0 |
| $\{goal1, goal3\}$ | 25/31 | 2/4 | 6/0 | 0/0 | 0/0 |
| $\{g1, g2, g3\}$ | 15/21/19 | 3/6/1 | 4/1/4 | 3/1/1 | 0/0/0 |

## 5.2 Performance of SPLUNI

Finally we test the performance of SPLUNI (set $\alpha = \beta = 1$), which is the game between the observer (interdictor) and the observed agent (evader), in maximizing the expectation of the shortest $s - t$ path length while minimizing the largest uncertainty, as shown in Table 2. We first define the path interdiction efficiency as $E = \Delta L/(\mathbf{d}^T * \mathbf{x})$ where $\Delta L$ is the increased path length for evader after interdiction and $\mathbf{d}^T * \mathbf{x}$ is the total length increase in the network. The tests are conducted over those ambiguous cases in three goal settings.

$E(E)$ and $E(rcr)$ help us understand the impact of SPLUNI model on both network interdiction and goal uncertainty reduction.

Table 2: The expectations of path interdiction efficiency and relative convergence ratio ($rcr(\%)$) for different goal setting.

| | $\{g1, g2\}$ | $\{g1, g3\}$ | $\{g2, g3\}$ |
|---|---|---|---|
| $E(E)$ | 56.0/73.6 | 63.4/67.0 | 77.4/90.7 |
| $E(rcr)$ | 19.7/13.3 | 7.4/9.1 | 2.3/1.9 |

Clearly, SPLUNI model works well for observer by adding the evader's path length and at the same time reducing the goal uncertainty during the process. This is of value for security domains where high-value targets need to be timely recognized meanwhile the intruder's actions be delayed.

# 6 Background and Related Work

## 6.1 Model-based Probabilistic Goal Recognition

The goal recognition problem has been formulated and addressed in many ways, as a matching problem over a suitable AND/OR graph [Avrahami-Zilberbrand and Kaminka, 2005], a parsing problem over grammar [Pynadath and Wellman, 1998], a probabilistic inference task over a dynamic Bayesian network [Bui *et al.*, 2002] and an inverse planning problem over planning models [Baker *et al.*, 2009; Ramırez and Geffner, 2011]. Among those approaches, two formulations solve the goal or plan recognition problem from different perspectives. One focuses on constructing a suitable library of plans or policies, while another one replaces that by an agent action model and a set of possible goals. The advantage of the latter formulation has been discussed about in [Xu *et al.*, 2017].

Hidden Markov models (HMMs) are widely used in probabilistic goal recognition. [Bui *et al.*, 2002] proposed an abstract hidden Markov model (AHMM) to recognize an agent's behavior in dynamic, noisy, uncertain domains, and across multiple levels of abstraction. Comparing to the HMM, MDP can describe agent actions and interactions between agents and the environment. [Baker *et al.*, 2009] consider the goal recognition problem over a MDP setting where actions are assumed to be stochastic and states fully observable. [Ramırez and Geffner, 2011] extend their work to POMDP settings where states are partially observable. In most model-based probabilistic goal recognition research [Baker *et al.*, 2009; Ramırez and Geffner, 2011], posterior goal distribution $P(G|O)$ is usually obtained from the Bayes rule

$$P(G|O) = \alpha P(O|G)P(G), \tag{10}$$

where $\alpha$ is a normalizing constant.

Apart from the above models, recent machine learning methods including reinforcement learning [Yue *et al.*, 2016], deep learning [Bisson *et al.*, 2015] and inverse reinforcement learning [Zeng *et al.*, 2018] have already been successfully applied in learning the agents' decision models for goal recognition tasks. These efforts once again extend the usability of model-based probabilistic goal recognition methods by constructing agents' behavior models from real data.

## 6.2 Goal Recognition Design

Goal recognition design was first introduced in [Keren *et al.*, 2014] to reduce goal uncertainty and advance the correct recognition through redesigning the underlying environment. Since then, lots of research has been carried out. [Keren *et al.*, 2014; 2016] extends the previous work by accounting for agents that behave non-optimally and non-observably. [Wayllace *et al.*, 2016] further allows the outcomes of agents' actions to be non-deterministic, and proposes a *Stochastic GRD* problem. Apart from the relaxation of assumptions, researchers try to solve the GRD from different aspects, or use newly established metrics. [Son *et al.*, 2016] addresses the same problem based on Answer Set Programming, resulting in higher scalability and efficiency. [Mirsky *et al.*, 2017] extends GRD to the Plan Recognition Design (PRD), which is the task of designing a domain using plan libraries in order to facilitate fast identification of an agent's plan. Noticing the inconsistency between the original *wcd* (*worst-case distinctiveness*) and Stochastic GRD model, a new metric, namely *all-goals wcd* ($wcd_{ag}$), is proposed by [Wayllace *et al.*, 2017].

# 7 Conclusion

We present a new perspective to goal identification control using off-the-shelf probabilistic goal recognizer and introduce the *relative goal uncertainty* (*rgu*) value for actions in a goal recognition task. We present ways of controlling goal uncertainty, followed by a presentation of method in an adversarial setting using a Shortest-Path Largest-Uncertainty Network Interdiction model. Empirical evaluation shows the effectiveness of the proposed solution in controlling goal uncertainty in recognition tasks.

Currently, as there exists inconsistency of goal uncertainty between the offline $P(G)$ which follows Uniform distribution and the real-time $P'(G)$ assessed online according to the agent's observations, it is still an open question as to how to incorporate the goal uncertainty information into the goal identification control process online.

# References


[Avrahami-Zilberbrand and Kaminka, 2005] Dorit Avrahami-Zilberbrand and Gal A Kaminka. Fast and complete symbolic plan recognition. In *IJCAI*, pages 653–658, 2005.

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

[Bisson *et al.*, 2015] Francis Bisson, Hugo Larochelle, and Froduald Kabanza. Using a recursive neural network to learn an agent's decision model for plan recognition. In *IJCAI*, pages 918–924, 2015.

[Bui *et al.*, 2002] Hung Hai Bui, Svetha Venkatesh, and Geoff West. Policy recognition in the abstract hidden markov model. *Journal of Artificial Intelligence Research*, 17:451–499, 2002.

[Geib and Goldman, 2001] Christopher W Geib and Robert P Goldman. Plan recognition in intrusion detection systems. In *DARPA Information Survivability Conference & Exposition II, 2001. DISCEX'01. Proceedings*, volume 1, pages 46–55. IEEE, 2001.

[Hofmann and Williams, 2007] Andreas G Hofmann and Brian C Williams. Intent recognition for human-robot interaction. In *Interaction Challenges for Intelligent Assistants*, pages 60–61, 2007.

[Jarvis *et al.*, 2005] Peter A Jarvis, Teresa F Lunt, and Karen L Myers. Identifying terrorist activity with ai plan recognition technology. *AI Magazine*, 26(3):73, 2005.

[Keren *et al.*, 2014] Sarah Keren, Avigdor Gal, and Erez Karpas. Goal recognition design. In *ICAPS*, 2014.

[Keren *et al.*, 2016] Sarah Keren, Avigdor Gal, and Erez Karpas. Goal recognition design with non-observable actions. In *AAAI*, pages 3152–3158, 2016.

[Lofberg, 2005] Johan Lofberg. Yalmip: A toolbox for modeling and optimization in matlab. In *Computer Aided Control Systems Design, 2004 IEEE International Symposium on*, pages 284–289. IEEE, 2005.

[Min *et al.*, 2014] Wookhee Min, Eunyoung Ha, Jonathan P Rowe, Bradford W Mott, and James C Lester. Deep learning-based goal recognition in open-ended digital games. *AIIDE*, 14:3–7, 2014.

[Mirsky *et al.*, 2017] Reuth Mirsky, Roni Stern, Meir Kalech Ya'akov (Kobi) Gal, and Meir Kalech. Plan recognition design. In *AAAI*, pages 4971–4972, 2017.

[Pynadath and Wellman, 1998] David V Pynadath and Michael P Wellman. Generalized queries on probabilistic context-free grammars. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1):65–77, 1998.

[Ramırez and Geffner, 2011] Miquel Ramırez and Hector Geffner. Goal recognition over pomdps: Inferring the intention of a pomdp agent. In *IJCAI*, pages 2009–2014. IJCAI/AAAI, 2011.

[Son *et al.*, 2016] Tran Cao Son, Orkunt Sabuncu, Christian Schulz-Hanke, Torsten Schaub, and William Yeoh. Solving goal recognition design using asp. In *AAAI*, pages 3181–3187, 2016.

[Sukthankar *et al.*, 2014] Gita Sukthankar, Christopher Geib, Hung Hai Bui, David Pynadath, and Robert P Goldman. *Plan, activity, and intent recognition: Theory and practice*. Newnes, 2014.

[Wayllace *et al.*, 2016] Christabel Wayllace, Ping Hou, William Yeoh, and Tran Cao Son. Goal recognition design with stochastic agent action outcomes. In *IJCAI*, 2016.

[Wayllace *et al.*, 2017] Christabel Wayllace, Ping Hou, and William Yeoh. New metrics and algorithms for stochastic goal recognition design problems. In *Twenty-Sixth International Joint Conference on Artificial Intelligence*, pages 4455–4462, 2017.

[Xu *et al.*, 2017] Kai Xu, Kaiming Xiao, Quanjun Yin, Yabing Zha, and Cheng Zhu. Bridging the gap between observation and decision making: Goal recognition and flexible resource allocation in dynamic network interdiction. In *IJCAI*, volume 4477, 2017.

[Yue *et al.*, 2016] Shiguang Yue, Kristina Yordanova, Frank Krüger, Thomas Kirste, and Yabing Zha. A decentralized partially observable decision model for recognizing the multiagent goal in simulation systems. *Discrete Dynamics in Nature and Society*, 2016, 2016.

[Zeng *et al.*, 2018] Yunxiu Zeng, Kai Xu, Quanjun Yin, Long Qin, Yabing Zha, and William Yeoh. Inverse reinforcement learning based human behavior modeling for goal recognition in dynamic local network interdiction. In *AAAI Workshop on Plan, Activity, and Intent Recognition*, 2018.