

Plan Selection for Probabilistic BDI Agents

Jianbing Ma¹, Weiru Liu², Jun Hong², Lluís Godo³, Carles Sierra³

¹Faculty of Science and Technology, Bournemouth University, Bournemouth, BH12 5BB, UK

²EEECES, Queen's University of Belfast, Belfast, BT1 7NN, UK

³IIIA - CSIC, Campus de la UAB s/n, 08193 Bellaterra, Spain

¹jma@bournemouth.ac.uk, ²{w.liu,j.hong}@qub.ac.uk,

³{godo, sierra}@iiia.csic.es

Abstract—When an agent wants to fulfill its desires about the world, the agent usually has multiple plans to choose from and these plans have different pre-conditions and additional effects in addition to achieving its goals. Therefore, for further reasoning and interaction with the world, a plan selection strategy (usually based on plan cost estimation) is mandatory for an autonomous agent. This demand becomes even more critical when uncertainty on the observation of the world is taken into account, since in this case, we consider not only the costs of different plans, but also their chances of success estimated according to the agent's beliefs. In addition, when multiple goals are considered together, different plans achieving the goals can be conflicting on their preconditions (contexts) or the required resources. Hence a plan selection strategy should be able to choose a subset of plans that fulfills the maximum number of goals while maintaining context consistency and resource-tolerance among the chosen plans. To address the above two issues, in this paper we first propose several principles that a plan selection strategy should satisfy, and then we present selection strategies that stem from the principles, depending on whether a plan cost is taken into account. In addition, we also show that our selection strategy can partially recover intention revision.

Keywords-Plan Selection; BDI Agents; Uncertainty Reasoning; Intention Revision

I. INTRODUCTION

Since Bratman proposed his planning theory of intentions in [1] and the subsequent seminal work of Cohen and Levesque in [3] formalizing Bratman's framework, modelling agent systems using theories of intention has obtained more and more attention. Among these theories, the BDI (Belief, Desire, Intention) approach, introduced by Rao and Georgeff in [15], [16], might be the most well-known and commonly accepted framework for modelling and implementing multi-agent scenarios. In BDI systems, the idea that rational agents should not pursue impossible goals since their beliefs forbid any intention to achieve these goals is readily reflected in the logic frameworks of Belief, Desire, and Intention systems [15]. Moreover, recently, in [18], it is proposed that an agent should not have a set of conflicting intentions, which cannot be achieved simultaneously. When this situation happens, i.e., there exists a set of conflicting intentions, intention revision should be applied to resolve this matter. Actually, intention revision, or more extensively, intention reconsideration that revisits the commitments to planned activity held by an agent is an important notion in BDI systems [1].

As intentions are represented as sets of hierarchically related plans [16], in this sense plan selection can be seen as an

extended variant of intention revision. That is, given multiple goals and each goal may have a set of plans to achieve it, how to choose a *suitable* subset of plans among all the plans? In the literature, there are few papers considered plan selection in BDI systems. In [4], the main focus is to measure the cost of a plan; [6] studies cooperative plan selection through trust among multi-agents; Other works like [20], [19] apply machine learning to derive the probability for the condition of a plan and use it for plan selection. In the meantime, intention revision, which is closely related to plan selection w.r.t the revision/selection methods, is also investigated by recent works in BDI systems, e.g., [22], [21], [7], [5], [18]. While these works studies on plan selection or intention revision, they seldom consider uncertainty that can emerge in agent's beliefs (see [13], [9], [12]). However, in multi-agent systems, especially in BDI systems, uncertainty has been discussed in many research articles (e.g., [14], [17], [8], [2]) in relation to both incomplete information about an environment and an agent's uncertain beliefs about some observations. As a consequence, it is reasonable to expect these two trends be combined together, i.e., to study plan selection strategies in uncertainty-enabled BDI systems.

In this paper, we focus on the uncertainty of the agent's beliefs, which leads to the uncertainty for determining the satisfiability of the preconditions of plans. More precisely, suppose there are multiple goals G_1, \dots, G_n that an agent wants to achieve, and each goal is associated with several alternative plans that can achieve it. When uncertainty is not considered, an optimal plan selection strategy is to choose a set of plans that can achieve the maximum number of goals (or minimizing the costs when cost information is available for the plans). When uncertainty is taken into account, i.e., the precondition for a plan cannot be determined as simply YES or NO, but with a probability value, an optimal plan selection strategy should also consider the probabilities that the plans can be successfully executed, and clearly if we cannot meet both the maximal goals and the greatest probability, then we have to make an appropriate compromise between the two.

This observation, i.e., maximizing the chances that plans can be executed when uncertainty is present, is equally applicable to intention revision. In recent works about intention revision (e.g., [18]), intentions generated from pre-ranked plans are used for achieving goals. For example, an intention from ranked plans can be of the form: $!G : \phi_1 \leftarrow P_1 : \dots : \phi_n \leftarrow P_n$ which means that, for achieving goal G , if ϕ_1

holds, then P_1 (the plan ranked first) is executed; if P_1 is not successfully executed, and if ϕ_2 holds¹, then P_2 is chosen for execution; and so on. Here it should be pointed out that this form of plan definition actually integrates a set of plans as: $!G : \phi_i \wedge \bigwedge_{j=1}^{i-1} \neg\phi_j \leftarrow P_i, 1 \leq i \leq n$. In these plans, the contexts (preconditions, i.e., $\phi_i \wedge \bigwedge_{j=1}^{i-1} \neg\phi_j$) of them are exclusively and iteratively ordered. To overcome this limitation, in our framework, we extend the plans (and their corresponding conditions) to be separated as simply: $!G : \phi_i \leftarrow P_i, 1 \leq i \leq n$, so that the probabilities of conditions for the plans can be evaluated simultaneously. Clearly if all the ϕ_i s are exclusive, then this set of plans can be rewritten into the pre-ranked plan format $!G : \phi'_1 \leftarrow P_1 : \dots : \phi'_n \leftarrow P_n$ such that $\phi'_i = \phi_i \wedge \bigwedge_{j=1}^{i-1} \neg\phi_j \equiv \phi_i$ (since the exclusiveness of ϕ_i s), $1 \leq i \leq n$. That is, the classic pre-ranked plan formats are recovered when our framework is reduced to the situation that all the ϕ_i s are exclusive. Moreover, instead of focusing on crisp formulas (i.e., ϕ_i) as done previously, in our framework, each formula is attached with a probability according to an agent's mental states of the worlds. This probabilistic plan selection framework is also a generalization of ranking-based plan representation, since the latter can be recovered by ranking plans based on probabilities of their preconditions.

As an example, consider a logistics company delivering goods from city A to city B. The delivery plan is sensitive to weather condition. On a day with good weather, each trolley only needs one driver while in bad weather, two drivers are needed. As plans are designed before weather conditions are known (e.g., plans are drawn one day before delivery, etc.), the uncertainty on weather conditions should be taken into account. The company will choose different set of plans to execute based on different expectations of weather conditions. Here the variation on weather expectation can stem from the change of agent's beliefs on weather conditions when time elapses. Therefore, a framework that is able to represent and reason with uncertainty in BDI plan selection (i.e., selecting the best set of plans according to some strategy/principles) would be desirable.

Motivated by the above observations, in this paper, we investigate how to perform plan selection when agents' belief state is uncertain. We consider three basic principles for plan selection under uncertainty. The first principle is *Mind Consistency*. That is, the set of plans an agent wants to execute together should be consistent in the sense that the set of preconditions of the set of plans are feasible, and the joint preconditions should be as plausible as possible. The second principle is *Resource Tolerance*, which means that the set of plans to be executed together should not have any conflict in resources (human resources, funds, etc.). For instance, if an agent does not have enough funds to execute all the plans in the chosen set, then it violates resource tolerance. The third one is *Maximizing Goals*, which states that an agent should aim to achieve as many goals as possible.

¹Note that in general ϕ_1 and ϕ_2 are not necessarily inconsistent, given that they are preconditions of different plans. This will be further illustrated in our proposed *Mind Consistency* principle in later sections.

We first develop our framework to deal with plan selection when no cost/reward information is available. We show that in this case we can choose the best plan set with maximum degree of certainty and achieving the maximum number of goals. We also show that this step of the framework partially recovers intention revision strategies introduced in [18]. Furthermore, we study cases when cost/reward information is attached to plans, and demonstrate that our framework can still choose the best set of plans that maximize expected profits (rewards minus costs) while satisfying the maximum number of goals.

The remainder of this paper is organized as follows. In Section 2, we provide some preliminaries on notations and definitions for our framework. We then propose our principles on plan selection in Section 3. Next, we consider strategies for plan selection. For that firstly in Section 4 we study plan selection without taking into account the reward/cost information for goals and plans and secondly we add the rewards/costs issue into plan selection which will be addressed in Section 5. In Section 6, we compare our strategies with intention revision methods. In Section 7, we compare our framework with some relevant works and finally we conclude our paper in Section 8.

II. PRELIMINARIES

In this paper, we use a variant of the CAN (Conceptual Agent Notation) BDI programming language which is in line with the well-known Agent-Speak language developed by Rao and Georgeff [15], [16]. The CAN language provides three types of atoms: events or event-goals (denoted by e), beliefs (denoted by b), and actions (denoted by a). Here an action atom is presented with its precondition and consequence, i.e., we write $a : \phi \leftarrow \psi$, where the precondition ϕ is a conjunction of belief literals and the consequence ψ is a list of adding or removing of belief atoms in an agents' belief base. All actions are assumed to be deterministic. To represent uncertainty (in this paper we use probability theory to represent uncertainty), we also introduce probability values for uncertain beliefs.

We use L_B to denote a propositional language over basic belief atoms formed in the usual manner. Lowercase Greek letters, possibly with decorations, will be used to denote a sentence of L_B .

We use Λ to denote an action description library that contains descriptions about primitive actions and their effects.

We use \mathcal{P} to denote a plan library that contains a set of plans in the domain. Each plan is associated with an event goal e that the plan aims to achieve, a precondition ϕ which is a conjunction of belief literals and a plan body P to realize e when condition ϕ is believed to be true. P can be a primitive action (e.g., a), adding or deleting of a belief atom (e.g., $b, \neg b$), testing for a condition (e.g., $?\phi$), posting a goal (e.g., $!e$), or a sequence of plans (e.g., $P1; P2$). In addition, in this paper we also introduce costs/rewards (and hence profits) for plans which are numerical values.

In CAN, a BDI agent is specified by an initial belief base \mathcal{B} , a plan library \mathcal{P} , and an action description library Λ . \mathcal{B} is a set of atoms representing the agent's initial beliefs about the world (i.e., for which the agent believes to be true).

In our framework, we suppose there is a probability distribution² on the language L_B (in the usual possible-world semantics), and for each plan $e : \phi \leftarrow P$, according to the probability distribution, we can always obtain $p(\phi)$ as the probability value of ϕ .

The intention base Π contains the intentions that the agent has already committed to for handling previously posted events. As mentioned in [16], an intention is a set of hierarchically related plans, (e.g., the aforementioned G).

A. Semantics

In [15], [16], Rao and Georgeff introduced a BDI configuration for rational agents, which intuitively can be written as a 5-ary tuple $C_{RG} = \langle \mathcal{G}, \Lambda, \mathcal{A}, \mathcal{B}, \Pi \rangle$, where components \mathcal{G} , Λ , \mathcal{B} and Π are the goal library, action library, belief base and intention base, respectively, and \mathcal{A} is the sequence of actions executed so far. Here we also put the plan library \mathcal{P} into the tuple (and makes it a 6-ary tuple) as $C = \langle \mathcal{G}, \mathcal{P}, \Lambda, \mathcal{A}, \mathcal{B}, \Pi \rangle$.

Like in [18], in this paper an agent configuration is used to depict the internal state of the agent. To link the internal state to the external event, apparently we also need a notion to describe the environment which interacts with the internal state to determine which external actions should be considered by the agent [18]. Formally speaking, let C denote the set of all possible agent configurations and E be the set of all possible external events in the domain, i.e., adding or removing beliefs, and posting a goal, etc., the environment ϵ is defined as a function $\epsilon : C \rightarrow 2^E$ [18]. That is, within a given environment and a configuration, the set of actions an agent is about to take (which lead to sets of events in E) are determined. For example, given a configuration C and an environment ϵ , $\epsilon(C)$ might contain a belief update that the time cost for a journey from city A to city B has dropped 20% (e.g., due to the deployment of a high-speed train). Furthermore, each time the agent takes an action in some environment, its configuration is changed accordingly. That is, w.r.t. an environment ϵ , an agent configuration C can be transformed into another configuration C' . To represent this, the transition relation \rightarrow_ϵ for each environment ϵ is defined such that $C \rightarrow_\epsilon C'$ indicates configuration C' is resulted from taking one single action on configuration C , within environment ϵ .

As we consider rational agents, an intuitive restriction on environments is that any environment should be consistent with any configuration in the sense that for any agent configuration, an environment cannot lead the agent to consider two contradictory actions. Formally, for any belief atom b , configuration C and environment ϵ , at most one of $+b$ and $-b$ are in $\epsilon(C)$ (Recall that $\epsilon(C)$ is a set of external events).

Similarly to [18], as we will focus on dealing with \mathcal{P} , we denote $C_- = \langle \mathcal{G}, \Lambda, \mathcal{A}, \mathcal{B}, \Pi \rangle$ as a so-called *diminished configuration*. If there is no confusion w.r.t. context, we also call C_- a configuration.

For each goal G , let $P = \{P^1, \dots, P^m\}$ be the set of plans to achieve G .

III. PRINCIPLES OF PLAN SELECTION

In this section, we explore the principles that a plan selection strategy should satisfy when uncertainty is taken into account.

Plan selection aims to choose the best subset of plans from a given plan base that an agent needs to consider. This can happen when a new goal is added (and hence several plans for the goal are added as well), the agent's belief is changed, the agent discovers an alternative plan to achieve a goal, or the agent detects that there is some internal inconsistency among some plans that the agent wants to execute. Clearly, we cannot randomly choose several plans and wait until the execution of some plan has been blocked, since it wastes resources. Therefore, we should provide some strategies to enable the agent to choose the best subset of plans in these cases. In addition, in this paper, we focus on plan selection strategies without looking into plan execution details (e.g., actions or subgoals).

To achieve different goals, an agent generates different sets of plans. These plans have pre-conditions which may lead to a subset of plans that cannot be executed together since their pre-conditions are inconsistent. For instance, to achieve a goal "Safe-Operation", there are two plans for a train agent. P_1 : *if there is a train ahead of it within 1 mile (a safe distance limit), then it should stop until the preceding train has moved to a safe enough distance*; and P_2 : *if there is no train ahead or the preceding train is not within 1 mile, then it can start moving*. In addition, the train agent has another goal "MovingFast", which gives rise to a plan P_3 : *if there is no train ahead within 1 mile, it keeps moving*. Then plans P_1 and P_3 cannot be executed (or even triggered) together since their pre-conditions are in conflict, whilst P_2 and P_3 can be executed together. This is what we call the principle of *Mind Consistency* as follows:

Mind Consistency If an agent wants to execute multiple plans together, he/she believes that it is possible to trigger them together. That is, the pre-conditions of the plans are consistent.

Mind Consistency only focuses on agent's beliefs. An agent should be rational in defining and choosing the best set of plans. Moreover, when uncertainty is taken into account, this principle has a further intuition. That is, the agent favors those plans (while they are consistent) with greatest possibility. So if there are multiple sets of plans whose pre-conditions are consistent, the agent prefers the set that is the most plausible. This is reasonable, since the agent has to make sure the plans can be triggered with a probability as high as possible, otherwise if plans can only be rarely triggered, then obviously the goals cannot be assured.

In this paper, for a set of plans, we use the joint probability of its pre-conditions to measure the consistency degree of agents' mind of the set of plans. Note that ensuring the greatest joint probability guarantees that the set of plans are logically consistent, since for logically inconsistent plans, the joint probability is 0.

More precisely, when we consider probabilities attached on preconditions, then Mind Consistency should be extended to:

Mind Consistency If an agent wants to execute multiple plans together, he/she believes that it is possible to

²For obtaining a probability distribution from a knowledge base, see [11].

trigger them together, and will choose the set of plans with the greatest joint probability.

Further more, if each plan is associated with a utility, then maximizing expected utility will be considered to replace the consideration on joint probability and Mind Consistency is adapted to:

Mind Consistency The agent wants to obtain as many expected utilities as possible if expected utilities can be computed.

The use of the three Mind Consistency principles will be automatically clarified by the context (depending on whether probability or utility is used).

The second principle considers the resources that the execution of plans requires. Intuitively, if some set of plans are going to be executed, there should be enough resources to satisfy their needs. Here the resources include (but are not limited to) time, monetary, and physical human resources, etc. Similar to the Mind Consistency principle, we have the following *Resource Tolerance* principle.

Resource Tolerance If an agent wants to execute multiple plans together, he/she knows that there is enough resources to execute them together.

As an example, to achieve a goal “GoingHome” when an agent is off work at 5pm, it can either take a bus which takes 15 minutes or walk which takes 1 hour. If it also wants to achieve a goal “HavingGoodDinner no later than 6pm”, which requires a plan of cooking for 30 minutes, then obviously the plan walking home is not resource (e.g., time) tolerant with the plan of cooking, whilst taking a bus home and cooking are resource tolerant.

It is worth to point out that for resource tolerance, there is an implicit assumption of resource conflict. For instance, a person may not be able to do two specific things simultaneously such as driving and cooking, but can both drive and listen to music (two actions), or a fixed amount of money (e.g., \$100) cannot cover two kinds of costs (e.g., \$60 and \$70), etc. We assume resource dependency (or tolerance) is specific and known to individual applications, so we do not consider about resource inconsistency further in this paper.

This principle is similar to the Environment Tolerance principle in [18]. However, since in our paper we consider uncertain beliefs, preconditions of plans are then not in a certain state but are associated with probabilities. So tolerance on preconditions does not apply and hence environment tolerance therefore reduces to resources tolerance here. In addition, Environment Tolerance principle in [18] is used for intention revision, while Resource Tolerance here is used for plan selection.

There still could be more than one set of plans that satisfy both the aforementioned principles. A simple and intuitive example is that: in the no utility case, if a set \mathcal{P} of plans satisfies the aforementioned principles, then any subset of \mathcal{P} also satisfies them. Clearly in this case we prefer \mathcal{P} to any of its (strict) subset. To deal with this issue, we need the following principle.

Maximizing Goals The agent wants to achieve as many goals as possible.

With the Maximizing Goals principle, clearly we do not need to worry about the aforementioned problem since \mathcal{P} satisfies more goals than any of its subsets.

IV. PLAN SELECTION WITHOUT REWARDS/COSTS

In this section, we describe the process of plan selection according to the agent’s belief change. Here we do not consider rewards of goals or costs of plans in the selection process, and hence we do not calculate expected utilities.

Let us have a look at the following example.

Example 1: Suppose a logistics company delivers goods from city A to city B (and it takes one day to go from A to B), or inside city A . The delivery is sensitive to the weather condition. If it is a good day, then each trolley only needs one driver to drive and look after it while in bad weather, two drivers are needed in one trolley for cross-city delivery. Suppose the company has two drivers, Alex and Bob, and the goods occupy two trolleys. To differentiate, we name the goods in two trolleys X and Y , respectively. X should be delivered from city A to city B , and Y is delivered inside A after X is delivered due to some reliance relation.

Then we have the following goals and plans for the company:

$!deliverX : goodWeather \leftarrow Go(Alex). (P_1)$
 $!deliverX : \neg goodWeather \leftarrow Go(Alex) \& Go(Bob). (P_2)$
 $!deliverY : true \leftarrow Go(Bob). (P_3)$

That is, in good weather, Alex will be sent out to deliver X and Bob to deliver Y . They can go together. However, in bad weather, Alex and Bob will have to deliver X together, so in this case X and Y cannot be delivered together. Given that it takes one day to deliver a trolley of goods, one trolley has to be delivered the other day.

Clearly if the condition of good weather does not hold, then goals $deliverX$ and $deliverY$ have resource conflict, i.e., there are not enough human resources to deliver.

Here comes the uncertainty that the weather condition is not certain. It cannot be fully determined before the delivery starts. In this case we should model agent’s beliefs with uncertainty, and hence the preconditions of the goals are also uncertain.

Now if we know the probability on the weather, then which subset of plans should be chosen?

In situations like Example 1, according to Resource Tolerance, a good plan selection strategy should guarantee that the selected plans can be executed together, without resource conflict (e.g., human resource conflict between P_2 and P_3). Then according to the principle of Maximizing Goals, the strategy should be able to select the largest number of plans that achieve the maximal number of goals³. Finally, according to Mind Consistency, the strategy should pay attention to the set of plans (obtained up to now) with the greatest probability.

A natural question here is whether the strategy always provide a resulting plan set. The answer is positive, at least in the sense of an empty set of plans. To this end, similar to [18], we denote $Success(\mathcal{P}, C_-, \epsilon)$ to be an undefined primitive in our framework indicating that \mathcal{P} can be successfully executed

³We will explain it later.

in environment ϵ with configuration C_- . Actually we have mentioned that for resources, usually there is an implicit assumption for resource conflict, which is self-indicative in the application domain but cannot be described in general. So the primitive $Success(\mathcal{P}, C_-, \epsilon)$ can be viewed as a domain specific concept indicating resource consistency.

First note that an empty set of plans is always executed successfully w.r.t. any configuration in any environment.

Axiom 1: (Null Success) $\forall C, \epsilon. Success(\emptyset, C_-, \epsilon)$.

This axiom setting actually provides us a trivial answer to plan selection. That is, if we do not consider the principle of Maximal Goals, then we can always obtain an empty plan set which satisfies Mind Consistency and Resource Tolerance. This shows why the principles of Maximal Goals and Mind Consistency are not only intuitive, but also necessary. In addition, this axiom also ensures that our system will not fail.

When we mention that the strategy should be able to select the largest number of plans that achieve the maximal number of goals, we should note that one goal is only needed to be achieved once. That is, if we have multiple plans to achieve the same goal, it is sufficient to choose one specific plan to achieve the goal. Formally, for any plan P , let G_P be the goal it aims to achieve and \mathcal{P}_P be the set of plans for goal G_P . Now we define single choice plan sets that a goal can be achieved by a single plan.

Definition 1: (Single Choice Plan Sets) $SC(\hat{\mathcal{P}}) \stackrel{def}{=} \nexists P_i, P_j \in \hat{\mathcal{P}}. P_j \in \mathcal{P}_{P_i}$.

Note that here $P_j \in \mathcal{P}_{P_i}$ also implies $P_i \in \mathcal{P}_{P_j}$. More precisely, it means that plans P_i and P_j are designed to achieving the same goal, and hence \mathcal{P}_{P_i} and \mathcal{P}_{P_j} are the same set. Therefore, the non-existence of such P_i, P_j indicates that in a single choice plan set, for any goal G , there is at most one plan that can achieve it.

Evidently, for a plan set, we require that this set of plans can be executed together, without any resource conflict. In this sense, now we define feasible plan sets which stem from the resource tolerance principle that the sets of plans should be possible in some environments, or there should be enough resources to execute the set of plans together.

Definition 2: (Possible Feasible Sets) $Poss(\hat{\mathcal{P}}, \mathcal{P}, C_-) \stackrel{def}{=} \hat{\mathcal{P}} \subseteq \mathcal{P} \wedge SC(\hat{\mathcal{P}}) \wedge \exists \epsilon. Success(\hat{\mathcal{P}}, C_-, \epsilon)$.

This definition precludes any resource conflict. For instance, in Example 1, we have both $Poss(\{P_1\}, \mathcal{P}, C_-)$ and $Poss(\{P_1, P_3\}, \mathcal{P}, C_-)$ hold while $Poss(\{P_2, P_3\}, \mathcal{P}, C_-)$ does not hold since both P_2 and P_3 require $Go(Bob)$.

Right now, we have applied the Resource Tolerance principle, and obtained possible feasible plan sets afterwards. Now it is time to apply the Maximizing Goals principle on the obtained plan sets. Clearly with the Maximizing Goals principle, an agent would like to achieve as many plans together as possible, hence a possible feasible set should be extended to a maximal state.

Definition 3: (Maximum Feasible Sets) $Max(\hat{\mathcal{P}}, \mathcal{P}, C_-) \stackrel{def}{=} Poss(\hat{\mathcal{P}}, \mathcal{P}, C_-) \wedge \nexists \mathcal{P}'. [\mathcal{P}' \supset \hat{\mathcal{P}} \wedge Poss(\mathcal{P}', \mathcal{P}, C_-)]$.

For instance, in Example 1, we have both $Max(\{P_2\}, \mathcal{P}, C_-)$

and $Max(\{P_1, P_3\}, \mathcal{P}, C_-)$ hold while $Max(\{P_1\}, \mathcal{P}, C_-)$ does not hold. That is, both $\{P_2\}$ and $\{P_1, P_3\}$ are maximum feasible plan sets while $\{P_1\}$ is not (since $\{P_1\} \subset \{P_1, P_3\}$).

Note that together with the single choice definition, here the number of plans in the set just equals the number of achieved goals.

Let $MFS(\mathcal{P}, C_-)$ be the set of all maximum feasible sets.

Let Pre_P be the precondition of plan P .

Definition 4: (Probability Of a Single Choice Plan Set) $\forall \hat{\mathcal{P}}. SC(\hat{\mathcal{P}}), Prob(\hat{\mathcal{P}}) \stackrel{def}{=} Pr(\bigwedge_{P \in \hat{\mathcal{P}}} Pre_P)$.

Here we have used the Mind Consistency principle. That is, if the preconditions of plans in $\hat{\mathcal{P}}$ are inconsistent, then we have $Prob(\hat{\mathcal{P}}) = 0$ which means that the agent will ignore this set of plans.

Based on the above definitions, we can define our selection results.

Definition 5: (Selection Result) $SR(\hat{\mathcal{P}}, \mathcal{P}, C_-) \stackrel{def}{=} \hat{\mathcal{P}} \in MFS(\mathcal{P}, C_-) \wedge \nexists \mathcal{P}'. [\mathcal{P}' \in MFS(\mathcal{P}, C_-) \wedge Prob(\mathcal{P}') > Prob(\hat{\mathcal{P}})]$.

This definition simply reflects our intuition of the selection strategy that we are looking for a maximal feasible plan set with the greatest joint probability.

For instance, in Example 1, if we have $Pr(goodWeather) = 0.8$, then according to Def. 5, the selection result is $\{P_1, P_3\}$ while if $Pr(\neg goodWeather) = 0.8$, then the selection result changes to $\{P_2\}$.

The next proposition shows that our system will not fail. That is, it always provide a selection result.

Proposition 1: $\forall \mathcal{P}, C_-. \exists \hat{\mathcal{P}}. SR(\hat{\mathcal{P}}, \mathcal{P}, C_-)$.

While Proposition 1 ensures there will be a selection result, it does not say whether the result is an emptyset or not. To answer this question, the following property shows that as long as there is one plan possible, the selection result is not empty.

Proposition 2: $\forall \mathcal{P}, C_-. [\exists \mathcal{P}' \subseteq \mathcal{P}. Poss(\mathcal{P}', \mathcal{P}, C_-) \wedge Prob(\mathcal{P}') > 0] \equiv [\forall \hat{\mathcal{P}}. SR(\hat{\mathcal{P}}, \mathcal{P}, C_-) \Rightarrow \hat{\mathcal{P}} \neq \emptyset]$.

As in real-life applications, the agent usually maintains at least one possible plan, then our system will provide a non-empty plan set.

Given the selection result, we should show that it is the best plan set we can have. That is, we need to show that any possible plan set, compared to the selection result, either it cannot achieve as many goals as the selection plan set can, or it is not as plausible as the selection result.

Proposition 3: $\forall \mathcal{P}, C_-, \mathcal{P}' \subseteq \mathcal{P}. \hat{\mathcal{P}}. [Poss(\mathcal{P}', \mathcal{P}, C_-) \wedge SR(\hat{\mathcal{P}}, \mathcal{P}, C_-) \Rightarrow \{|\mathcal{P}'| \leq |\hat{\mathcal{P}}| \vee Prob(\mathcal{P}') \leq Prob(\hat{\mathcal{P}})\}]$.

With the above properties, obviously our plan selection strategy does provide satisfactory selection result.

V. PLAN SELECTION WITH REWARDS/COSTS

In this section, we describe the strategy of plan selection when we consider rewards of goals and costs for executing plans in the selection process. Here we suppose that rewards and costs are measurable numerical values (interested readers can refer to [4] for the calculation of the cost of a plan). The approach proposed here can be extended to deal with situations

that the $(reward, cost)$ pair can be compared in a pairwise sense.

The plan selection strategy is similar to the one introduced for plan selection without rewards/costs. However, since we have rewards/costs information and hence expected utility information of a plan (or a plan set), we use the expected utility information to replace simply the probability information.

Again let G_P be the goal that plan P is designed to achieve. To calculate the expected utility of a plan, we first define the profit of a plan as follows.

Definition 6: (Profits of a plan) $Prof(P) \stackrel{def}{=} Rew(G_P) - Cost(P)$.

Here we suppose all profits are positive, otherwise we do not need such a plan⁴.

Recall that Pre_P stands for the precondition of P . Now we can calculate the expected utility in the standard way.

Definition 7: (Expected Utility Of a plan) $EU(P) \stackrel{def}{=} Pr(Pre_P)Prof(P)$.

Actually what we want is the expected utility of a plan set instead of a plan. So we define the expected utility of a single choice plan (SCP) set as follows.

Definition 8: (Expected Utility Of a SCP Set) $\forall \hat{\mathcal{P}}. SC(\hat{\mathcal{P}}), EU(\hat{\mathcal{P}}) \stackrel{def}{=} \sum_{P \in \hat{\mathcal{P}}} EU(P)$.

Now we define the selection result as the maximal feasible sets that have the greatest expected utility.

Definition 9: (Selection Result) $SRP(\hat{\mathcal{P}}, \mathcal{P}, C_-) \stackrel{def}{=} \hat{\mathcal{P}} \in MFS(\mathcal{P}, C_-) \wedge \nexists \mathcal{P}'. [\mathcal{P}' \in MFS(\mathcal{P}, C_-) \wedge EU(\mathcal{P}') > EU(\hat{\mathcal{P}})]$.

Compared to Def. 5, this definition uses expected utility to replace the probability value, since intuitively the expected utility is more desirable than just the probability.

Similar to the no rewards/costs case, we have the following properties for the selection strategy. The first says that this strategy does not fail.

Proposition 4: $\forall \mathcal{P}, C_-. \exists \hat{\mathcal{P}}. SRP(\hat{\mathcal{P}}, \mathcal{P}, C_-)$.

Second, to avoid the empty selection result, the following property shows that as long as there is one plan possible, the selection result is not empty.

Proposition 5: $\forall \mathcal{P}, C_-. [\exists \mathcal{P}' \subseteq \mathcal{P}. Poss(\mathcal{P}', \mathcal{P}, C_-) \wedge Prob(\mathcal{P}') > 0] \equiv [\forall \hat{\mathcal{P}}. SRP(\hat{\mathcal{P}}, \mathcal{P}, C_-) \Rightarrow \hat{\mathcal{P}} \neq \emptyset]$.

And finally we show that any possible plan set, compared to the selection result, either it cannot achieve as many goals as the selection result can, or it does not bring as much expected utility as the selection result.

Proposition 6: $\forall \mathcal{P}, C_-, \mathcal{P}' \subseteq \mathcal{P}, \hat{\mathcal{P}}. [Poss(\mathcal{P}', \mathcal{P}, C_-) \wedge SRP(\hat{\mathcal{P}}, \mathcal{P}, C_-) \Rightarrow \{|\mathcal{P}'| \leq |\hat{\mathcal{P}}| \vee EU(\mathcal{P}') \leq EU(\hat{\mathcal{P}})\}]$.

Note that plan selection with rewards/costs is not a simple extension of selection without rewards/costs. That is, if we define all rewards as 1 and all costs as 0, then the two kinds of selection results do not match.

Example 2: In Example 1, let $Pr(\neg goodWeather) = 0.6$, then by the first selection result we get $\{P_2\}$, and by the second selection result we get $\{P_1, P_3\}$.

⁴However, if we cannot measure the reward (and set it to 0) and hence only use cost (then $Prof(P)$ is negative), our system still works, with $Prof(P)$ changed to $-Prof(P)$ subsequently.

The reason is that for overlapped preconditions, the joint probability is not the sum of the probabilities of such preconditions, whilst the joint profit is always the sum of the profits of all participated plans.

VI. RELATION TO RANKED INTENTION REVISION

Intention revision is an important task in BDI programming. It aims to choose the best subset of intentions from a given intention base that the agent needs to reconsider. This can happen when a new goal is added, the agent's belief is changed, or the agent finds that there is some internal inconsistency among some intentions that the agent wants to execute (which is similar to situations for plan selection).

In the recent paper about intention revision [18], ranked intentions represented as $!G : \phi_1 \leftarrow P_1 : \dots : \phi_n \leftarrow P_n$ are revised. The semantics for this construct is that if ϕ_1 is satisfied, then P_1 is executed to achieve G , otherwise if ϕ_2 is satisfied, then P_2 is chosen, and so on. Here for convenience we call each $\phi_i \leftarrow P_i$ a *ranked item*. [18] proposes two revision methods to make the intention base consistent: i) dropping some intentions; ii) modifying some intentions (i.e., removing some ranked items in the modified intentions)⁵; to make the intention base consistent.

Now we reduce the uncertainty of plans into preferences on intentions. The transformation is defined as:

Definition 10: For a ranked intention $!G : \phi_1 \leftarrow P_1 : \dots : \phi_k \leftarrow P_k$, suppose we can attach probabilities as $Pr(\phi_1) \geq Pr(\phi_2) \geq \dots \geq Pr(\phi_k)$ and $Pr(\phi_1) - Pr(\phi_k) < \rho$ (ρ is a small enough positive real). Then we can then obtain a set of plans to achieve G as $!G : \phi_i \leftarrow P_i, 1 \leq i \leq k$, and we set the profit of each such plan by 1. We call it this set of plans an *Uncertainty Replacement* of the ranked intention.

For an intention base Π , we denote by $UR(\Pi)$ be the union of the uncertainty replacements of intentions in Π .

Note that with this transformation, our definition of single choice plan set becomes a distinct intention set as defined in paper [18].

Our revision result $SRP(\hat{\mathcal{P}}, \mathcal{P}, C_-)$ reduces to modification revision (Rev^M) defined in [18].

Proposition 7: $\forall \Pi, C_-. SR(UR(\Pi), UR(\Pi), C_-) \Rightarrow Rev^M(\hat{\Pi}, \Pi, C_-)$.

Here C'_- is a diminished configuration.

Example 3: In Example 1, if we introduce a ranked intention as:

$!deliverX : \neg goodWeather \leftarrow Go(Alex) \& Go(Bob) : goodWeather \leftarrow Go(Alex)$.

And we transform it as P_1 and P_2 (Recall P_1, P_2, P_3 in Example 1), where we set $P(\neg goodWeather) = 0.55$ and $Prof(P_1) = Prof(P_2) = Prof(P_3) = 1$.

By Def. 9, the plan selection result is $\{P_1, P_3\}$, and we have the modification revision result is just $\{P_1, P_3\}$.

In contrast, if we view the ranked intention $!G : \phi_1 \leftarrow P_1 : \dots : \phi_k \leftarrow P_k$ as a plan with precondition ϕ_1 and probability $Pr(\phi_1)$, then the result of applying our approach leads to a subset of the result of the dropping-intention revision (Rev^D)

⁵The definitions of the two revision methods are complicated. So we omit them here.

in [18]. For any intention base Π , let $trans(\Pi)$ denote the set of all transformed plans.

Proposition 8: $\forall \Pi, C_-$. if $SR(trans(\hat{\Pi}), trans(\Pi), C_-)$ and $Rev^D(\hat{\Pi}, \Pi, C'_-)$, then we have $trans(\Pi) \in \hat{\Pi}$.

Example 4: In Example 1, if we introduce a ranked intention as (assume $Pr(gW) < 0.5$, gW is short for *goodWeather*):

$\text{!deliver}X : \neg gW \leftarrow Go(Alex) \& Go(Bob) : gW \leftarrow Go(Alex)$.
(P_4)

Then together with P_3 (Recall P_1, P_2, P_3 in Example 1), by Def. 5, the plan selection result is $\{P_3\}$, and the dropping revision result is $\{\{P_4\}, \{P_3\}\}$.

Now suppose $Pr(gW) > 0.5$, if we introduce another ranked intention as:

$\text{!deliver}X : gW \leftarrow Go(Alex) : \neg gW \leftarrow Go(Alex) \& Go(Bob)$.
(P_5)

Then together with P_3 , by Def. 5, the plan selection result is $\{P_3, P_5\}$, and the dropping revision result is $\{\{P_3, P_5\}\}$.

VII. RELEVANT WORKS

In [4], repair plan selection strategies are discussed where plans to be selected are for different goals (intentions), whilst in our approach, different plans can be used to achieve the same goal and hence should be made choice among them. Also, the main focus of [4] is to calculate the cost of a plan.

In [6], cooperative plan selection strategies based on trust information among multi-agents is studied. It stresses that individual agents must interact for the overall system to function effectively and it considers the information needed by an agent to be able to assess the degree of risk involved in a particular course of action. This information is then used for plan selection. That is, the plan selection strategy of [6] focuses on interaction information (i.e., trust) between multi-agents which is beyond our paper.

Papers [20], [19] address the problem that preconditions of plans are restricted to be boolean formulas that are to be specified at design/implementation time. So they propose methods to allow agents to learn the probability of success for plans based on previous execution experiences, using the decision tree model. While they also consider uncertainty in plan selection, they focus on the success rate of plans.

VIII. CONCLUSIONS

In this paper, we have discussed plan selection under uncertainty. We also have proposed three principles on plan selection with uncertainty. Furthermore, we have provided strategies to deal with plan selection cases where rewards/costs are considered for goals or not. We have also compared our strategies with some intention revision approaches with ranked intentions.

As we know, beliefs can be revised upon receiving new information which is not completely consistent with the prior beliefs. In this case, the set of plans that will be chosen should also be changed. If the beliefs are revised before any plan has been executed, then we just need to follow the revised beliefs and use the framework proposed in this paper. However, if the beliefs are revised when a subset of plans have

been partially executed, then more complicated plan selection methods should be proposed. This remains an open problem and is an interesting future direction of our work.

In the future, we also want to explore uncertainty cases for cascaded plan sets directly and apply our strategies to intention revision. In addition, another interesting future direction is to consider uncertainty on the degree that a plan can fulfill the goal. Actually in real life, we cannot ensure that a goal is always 100% achieved when we have executed a plan for it. Furthermore, considering uncertainty on desires, action results, etc., could also be promising.

A weakness about the current plan selection approaches is that it only concerns with a one-shot selection, rather than an iterative selection process. This is because the selection function does not change the environment which, in an iterative selection process, should be changed according to actions taken and external intervention. In this case, a solution might be to replace environments with epistemic states [10] which can be updated consistently. This is also an interesting problem for future work.

REFERENCES

- [1] M. E. Bratman. *Intentions, Plans, and Practical Reason*. Harvard University Press, 1987.
- [2] A. Casali, L. Godo, and C. Sierra. Graded bdi models for agent architectures. In *Proceedings of the 5th International Workshop on Computational Logic in Multi-Agent Systems (CLIMA V)*, pages 126–143. Springer, 2004.
- [3] P. R. Cohen and H. J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42:213–261, 1990.
- [4] K. H. Dam and M. Winikoff. Cost-based bdi plan selection for change propagation. In *Proc. of 7th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2008)*, Estoril, Portugal, 2008.
- [5] J. Grant, S. Kraus, D. Perlis, and M. Wooldridge. Postulates for revising bdi structures. *Synthese*, 175:127–150, 2010.
- [6] N. Griffiths and M. Luck. Cooperative plan selection through trust. In *Multi-Agent System Engineering: Proceedings of the Ninth European Workshop on Modelling Autonomous Agents in a Multi-Agent*, pages 162–174. Springer-Verlag, 1999.
- [7] T. Icard, E. Pacuit, and Y. Shoham. Joint revision of beliefs and intentions. In F. Lin, U. Sattler, and M. Truszczynski, editors, *Proceedings of Principles of Knowledge Representation and Reasoning (KR)*, page 572C574. AAAI Press, 2010.
- [8] J. Lang, L. van der Torre, and E. Weydert. Hidden uncertainty in the logical representation of desires. In *International Joint Conference on Artificial Intelligence, IJCAI 03*, Acapulco, Mexico, 2003.
- [9] J. Ma and W. Liu. A framework for managing uncertain inputs: an axiomization of rewarding. *International Journal of Approximate Reasoning*, 52(7):917–934, 2011.
- [10] J. Ma, W. Liu, and S. Benferhat. A belief revision framework for revising epistemic states with partial epistemic states. In *Proc. of AAAI'10*, pages 333–338, 2010.
- [11] J. Ma, W. Liu, and A. Hunter. Inducing probability distributions from knowledge bases with (in)dependence relations. In *Procs. of AAAI*, pages 339–344, 2010.
- [12] J. Ma, W. Liu, and P. Miller. A characteristic function approach to inconsistency measures for knowledge bases. In *Proceedings of international conference on Scalable Uncertainty Management (SUM)*, pages 473–485. Springer, 2012.
- [13] J. Ma, W. Liu, P. Miller, and W. Yan. Event composition with imperfect information for bus surveillance. In *Procs. of AVSS*, pages 382–387. IEEE Press, 2009.
- [14] S. Parsons, C. Sierra, and N. R. Jennings. Agents that reason and negotiate by arguing. *Journal of Logic and Computation*, 8(3):261–292, 1998.
- [15] A. S. Rao and M. P. Georgeff. Modeling rational agents with a bdi-architecture. In *Proceedings of Principles of Knowledge Representation and Reasoning (KR)*, pages 473–484. Morgan Kaufmann, 1991.

- [16] A. S. Rao and M. P. Georgeff. An abstract architecture for rational agents. In C. Rich B. Nebel and W. R. Swartout, editors, *Proceedings of Principles of Knowledge Representation and Reasoning (KR)*, pages 438–449. Morgan Kaufmann, 1992.
- [17] M. Schut, M. Wooldridge, and S. Parsons. Reasoning about intentions in uncertain domains. In *Proceedings of Symbolic and Quantitative Approaches to Reasoning with Uncertainty, 6th ECSQARU*, pages 84–95, Toulouse, France, 2001.
- [18] S. Shapiro, S. Sardina, J. Thangarajah, L. Cavedon, and L. Padgham. Revising conflicting intention sets in bdi agents. In *Procs. of AAMAS*, pages 1081–1088, 2012.
- [19] D. Singh, S. Sardina, and L. Padgham. Extending bdi plan selection to incorporate learning from experience. *Robot. Auton. Syst.*, 58(9):1067–1075, 2010.
- [20] D. Singh, S. Sardina, L. Padgham, and S. Airiau. Learning context conditions for bdi plan selection. In *Proc. of 9th Int. Conf. on Autonomous Agents and Multiagent Systems (AAMAS 2010)*, pages 325–332, Toronto, Canada, 2010.
- [21] W. van der Hoek, W. Jamroga, and M. Wooldridge. Towards a theory of intention revision. *Synthese*, 155(2):265–290, 2007.
- [22] W. Wobcke. Plans and the revision of intentions. In C. Zhang and D. Lukose, editors, *Distributed Artificial Intelligence: Architecture and Modelling, volume 1087 of Lecture Notes in Computer Science (LNCS)*, pages 100–114. Springer, 1995.