Decisions with multiple simultaneous goals and uncertain causal effects

Paulo Trigo and Helder Coelho

Abstract A key aspect of decision-making in a disaster response scenario is the capability to evaluate multiple and simultaneously perceived goals. Current competing approaches to build decision-making agents are either mental-state based as BDI, or founded on decision-theoretic models as MDP. The BDI chooses heuristically among several goals and the MDP searches for a policy to achieve a specific goal. In this paper we develop a preferences model to decide among multiple simultaneous goals. We propose a pattern, which follows a decision-theoretic approach, to evaluate the expected causal effects of the observable and non-observable aspects that inform each decision. We focus on yes-or-no (i.e., pursue or ignore a goal) decisions and illustrate the proposal using the RoboCupRescue simulation environment.

1 Introduction

The mitigation of a large-scale disaster, caused either by a natural or a technological phenomenon (e.g., an earthquake or a terrorist incident), gives rise to multiple simultaneous goals that demand the immediate response of a finite set of specialized agents. In order to act rationally the agent must evaluate multiple and simultaneous perceived damages, account for the chance of mitigating each damage and establish a preferences relation among goals. The belief-desire-intention (BDI) mental-state architecture [7] is widely used to build reasoning agents, equipped with a set of beliefs about the state of the world and a set of desires which, broadly speaking, identify those states that the agent has as goals. From its beliefs and desires, and via

Paulo Trigo
Instituto Superior de Engenharia de Lisboa – ISEL, DEETC and LabMAg, GuIAA; Portugal
e-mail: `ptrigo@deetc.isel.ipl.pt`

Helder Coelho
Faculdade de Ciências da Universidade de Lisboa – FCUL, DI and LabMAg; Portugal
e-mail: `hcoelho@di.fc.ul.pt`

deliberation, the agent formulates an intention that can be seen as the goal, or desire, the agent commits to bring about. Although one side of rational behavior is the capability to establish preferences among simultaneous goals, current BDI theory and systems do not provide a theoretical or architectural framework for deciding how goals interact and how an agent decides which goals to pursue. When faced with multiple simultaneous goals, the intention selection (decision) follows a heuristic approach, usually coded by a human designer [4]. Additionally, BDI models find it difficult to deal with uncertainty, hence hybrid models have been proposed combining BDI and Markov decision process (MDP) approaches [5, 6]; however, hybrid models usually assume that the goal has already been chosen and tackle the stochastic planning problem (in order to achieve the chosen goal).

In this paper we take the decision-theoretic notion of rationality to estimate the importance of goals and to establish a preferences relation among multiple goals. We propose a preferences model that allows agent developers to design the relationships between perceived (certain) and uncertain aspects of the world in an easy and intuitive manner. The design is founded on the influence diagram [2] (ID) framework that combines uncertain beliefs and the expected gain of decisions. The proposal's practical usefulness is experimentally explored in a fire fighting scenario in the RoboCupRescue [3] domain. The decision model incorporates general fire fighting principles in a way that considerably simplifies the specification of a preferences relation among goals. Despite such simplification, the attained results are consistent with the initial fire fighting principles.

The next section describes the preferences model, which is instantiated and evaluated in section 3; the section 4 presents our conclusions and future goals.

2 The preferences model

The premisse of the preferences model is that the relation among simultaneous goals follows from the expected utility of the available decisions. The expected utility of a decision combines two elements: i) the value of the state under observation, and ii) the likelihood of success of that decision. Given a set of available decisions, $\mathscr{D}$, a set of states, $\mathscr{S}$, an utility function, $u : \mathscr{S} \to \mathbb{R}$, and the probability, $P(s \mid d)$, to achieve $s \in \mathscr{S}$ after decision $d \in \mathscr{D}$, the expected utility, $eu : \mathscr{D} \to \mathbb{R}$, of decision-making is described by: $eu(D = d) = \sum_{s \in \mathscr{S}} P(s \mid D = d) u(s)$, where D is a variable that holds an available decision. Given any goal there are always two available decisions: i) pursue the goal, or ii) ignore the goal. Thus, $\mathscr{D} = \{\, \texttt{yes}, \ \texttt{no} \,\}$, is such that $D_g = \texttt{yes}$ and $D_g = \texttt{no}$ represent, respectively, the decision to pursue or to ignore goal $g \in \mathscr{G}$.

The utility of a goal, g, measures the importance, assigned by the agent, to the goal g. The "importance" is a criterion related to a valuation in terms of benefits and costs an agent has of a mental state situation [1]. The mental state is materialized by the agent beliefs regarding the perceived states and the desire to pursue, or ignore, each goal. Also, the goal achievement payoff is estimated by the difference between the expected utility on pursuing and ignoring that goal. Thus, the goal utility func-

tion, $u_{\mathscr{G}}$, for each $g \in \mathscr{G}$, is defined by,

$$u_{\mathscr{G}}(g) = eu(D_g = \text{yes}) - eu(D_g = \text{no}) \tag{1}$$

The utility function, $u_{\mathscr{G}}$, is used to establish the preferences about the set of goals $\mathscr{G}$. The preferences, $\forall\, g_1, g_2 \in \mathscr{G}$, are: i) $g_1 \succ g_2$, if the agent prefers g_1 to g_2, or ii) $g_1 \sim g_2$, if the agent is indifferent between g_1 and g_2. The rules, used to establish the preferences' total order among goals, are described by,

$$g_1 \succ g_2 \begin{cases} u_{\mathscr{G}}(g_1) > u_{\mathscr{G}}(g_2) \\ \vee \\ u_{\mathscr{G}}(g_1) = u_{\mathscr{G}}(g_2) \ \wedge \ eu(D_{g_1} = \text{yes}) > eu(D_{g_2} = \text{yes}) \end{cases} \tag{2}$$

$$g_1 \sim g_2 \ \left\{\ u_{\mathscr{G}}(g_1) = u_{\mathscr{G}}(g_2) \ \wedge \ eu(D_{g_1} = \text{yes}) = eu(D_{g_2} = \text{yes}) \right. \tag{3}$$

From expression 2 the agent prefers goals with higher payoff and when even, prefers goals that, when achieved, give higher expected advantage (i.e., higher $eu(D_g = \text{yes})$ value); the expression 3 indicates that in sight of equality the agent is indifferent between goals, thus taking, for instance, a random decision.

2.1 The causal effect pattern

The causal effects (consequences) of each decision are unknown, therefore our aim is to choose the decision alternative (goal) that minimizes the eventual disadvantageous consequences of such decision. The ID framework combines uncertain beliefs to compute the expected utility of decisions, thus rationality is a matter of choosing the alternative that leads to the highest expected utility, given the evidence of available information. The ID extends the, Bayesian network, chance nodes with two additional nodes: decisions and utilities, and two additional arcs: influences and informational. As in belief networks, chance nodes represent random variables, i.e., the agent's uncertain beliefs about the world. A decision node holds the available choices, i.e., the possible actions to take. An utility node represents the agent's preferences. The links between the nodes summarize their dependency relations.

We propose the following guidelines, to structure the multiple and simultaneous goals decision problem, using the ID framework:

i. the current state is characterized by a set of variables that are observable at the decision-making time instant,

ii. the decision outcome is characterized by a set of variables that are non-observable at the decision-making time instant,

iii. the observable variables inform the decision nodes and the decision nodes influence the non-observable variables,

iv. the observable variables condition the non-observable variables,

   v. all dependencies among observable variables, or among non-observable variables are valid (whilst not generating any dependency loop),

  vi. the set of observable variables influences a set of utility nodes,

 vii. the set of non-observable variables influences a set of utility nodes,

viii. the two sets of utility nodes (cf. items vi vii) are disjoint, and

  ix. a decision influences both sets of utility nodes (cf. items vi vii).

Figure 1 illustrates the above guidelines using the regular ID symbols; circle is a chance node, rectangle is a decision node and the lozenge is an utility node.
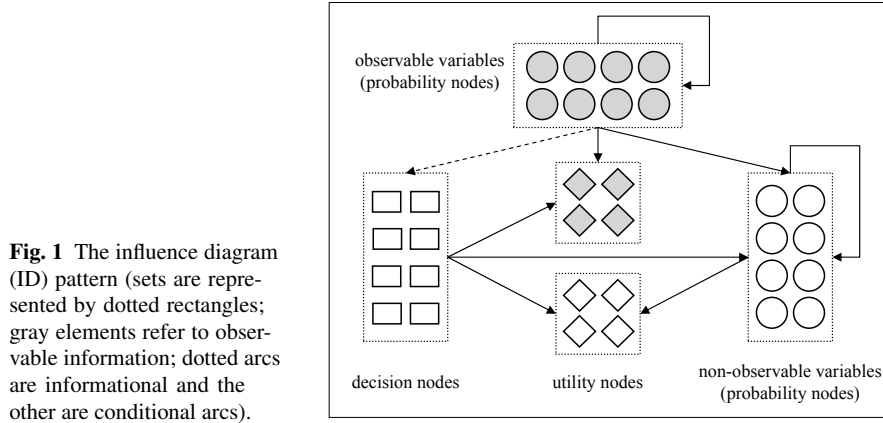

Fig. 1 The influence diagram (ID) pattern (sets are represented by dotted rectangles; gray elements refer to observable information; dotted arcs are informational and the other are conditional arcs).

The gray filling (cf. figure 1) has special meaning: i) the gray chance node indicates information availability, i.e., an observable variable (cf. item i above), and ii) the gray utility node indicates a dependency from a gray chance node, i.e., the utility of some observable variables (cf. item vi above). The sets of nodes with similar characteristics are aggregated by a dotted rectangle. The arcs connect sets of nodes (instead of individual nodes), therefore attaining an ID pattern, i.e., a template from which to build several different instances with the same overall structure.

2.2 The ID pattern usage

The ID pattern (cf. figure 1) is used to support the construction of the goal utility function, $u_{\mathscr{G}}$ (cf. equation 1). Therefore, we propose the following method to specify the decision nodes:

i. identify the largest subsets of goals, $\mathscr{G}_i \subseteq \mathscr{G}$ such that $\cup_i \mathscr{G}_i = \mathscr{G}$ and all the goals $g \in \mathscr{G}_i$ are characterized by the same set of observable variables,

ii. for each $\mathcal{G}_i$ (cf. item i) specify a decision node labeled "D_i" and add the corresponding information arcs (from observable variables to "D_i"),

iii. for each decision node, "D_i", set its domain to "yes" and "no" values to represent, respectively, the decision to pursue, or ignore, a goal $g \in \mathcal{G}_i$; the goal, g, occurs after the observation of the variables that inform "D_i".

For concreteness and to illustrate the design of the decision problem, the next section materializes the preferences model in a simulated scenario.

3 Experimental setup

We used the RoboCupRescue environment to devised a disaster scenario that evolves at the Nagata ward in Kobe, Japan. Two buildings, B_1 and B_2, not far from each other (about 90 meters) catch a fire. The B_1 is relatively small and is located near Kobe's harbor, in a low density neighborhood. The B_2 is of medium size and it is highly surrounded by other buildings. As time passes, the fires' intensity increase so a close neighbor is also liable to catch a fire.

Figure 2 shows the disaster scenario; each opaque rectangle is a building and a small circle is positioned over B_1 and B_2. The two larger filmy squares define the neighborhood border of B_1 and B_2 within a d distance (in meters). The ground neighborhood area of a building is given by $ngb(d) = (2 \times d)^2$, for a distance d, and the set of buildings contained within such area is denoted as $\mathcal{N}_{B_i,d}$; we set $d = 250$ m, thus a ground neighborhood area of 250.000 m^2.
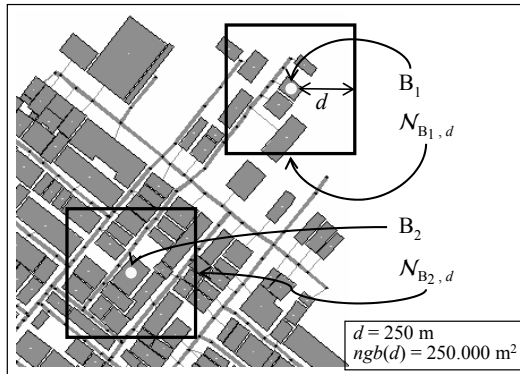

Fig. 2 Fire scenario in the buildings labeled B_1 and B_2 (the set of buildings contained within each building's neighborhood, $ngb(d)$, is represented by $\mathcal{N}_{B_i,d}$).

To simplify we assume that: i) buildings use identical construction materials, ii) buildings are residential (neither offices nor industries inside the buildings), and iii) there are no civilians, caught by fires, inside the buildings. We also assume that agents get informed about the fires the moment it starts; we are not concerned on how (through which communication channel) the agent gets such information. We

now carry on with the design of the multiple simultaneous goal decision problem, in the context of this illustrative domain.

3.1 The ID pattern instance

In order to apply the ID pattern (cf. figure 1) to the illustrative scenario (cf. figure 2) we considered, for each building, the following observable variables:

- The building's fire intensity, *fireIntensity* (we adopted the RoboCupRescue names), perceived by the agent with three values: i) 1, an early fire, ii) 2, an increasing intensity fire, and iii) 3, a high intensity fire.
- The building's total area, *allFloorsArea*, given by the building's ground area times the number of floors, with three values: i) `low`, ii) `medium`, and iii) `high`. Each value is $\frac{1}{3}$ of the normalized total area, i.e., the building's total area divided by the maximum total area of the buildings in the scenario ; e.g. for B_1 we have $\frac{7.668}{57.866} = 0,13$ (`low`) and for B_2 we have $\frac{19.454}{57.866} = 0,34$ (`medium`).
- The building's neighborhood density, *neighbourhoodDensity*, computed as the ratio between the summation of the ground area, *floorArea*(b), of each building within distance *d* of B_i neighborhood (i.e., each $b \in \mathcal{N}_{B_i,d}$), and the total area of that neighborhood (i.e., *ngb*(*d*)); the ratio is thus given by, $\frac{\sum_{b \in \mathcal{N}_{B_i,d}} floorArea(b)}{ngb(d)}$), and the *neighbourhoodDensity* has the following three values: i) `low`, ii) `medium`, and iii) `high`. Each value is $\frac{1}{3}$ of that ratio; e.g. for B_1 we have $\frac{39.900}{250.000} = 0,16$ (`low`) and for B_2 we have $\frac{138.500}{250.000} = 0,55$ (`medium`).

The non-observable variable, *destruction*, describes the destruction inflicted by the fire with three values, `low`, `medium`, and `high`, each representing, respectively, the intervals $]0;0,2]$, $]0,2;0,7]$ and $]0,7;1]$ of the destruction percentage.

The goals are *extinguished*(*B*) $\in \mathcal{G}' \subseteq \mathcal{G}$, where *B* is a building in fire. For readability, the subset $\mathcal{G}'$ will be named as *extinguish*. Hence, we specify a decision variable, *extinguish* (cf. section 2.2), that evaluates each goal, *extinguished*(*B*), whereas all the aspects that influence the decision (extinguish or ignore the fire in *B*), are represented through the observable variables: *fireIntensity*, *allFloorsArea* and *neighbourhoodDensity*.

To specify the utility nodes we follow *three general fire attack strategies* that, although intuitive, were acquired after the RoboCupRescue experimentation:

- the earlier a fire is attacked, the easier it is to extinguish the fire,
- the smaller the building, the less time it takes to extinguish the fire, and
- the higher the neighborhood density, the higher the need to extinguish the fire.

The above strategies are used to specify the utility nodes: *U1* and *U2*. The utility node *U1* is influenced by the observable variables and represents the agent's evaluation of the fire intensity impact on the neighbor building. For example, a fire may

cause higher damages in a high density than in a low density neighborhood (given an identical fire intensity); thus, the higher utility values are ascribed to high intensity fires that occur in high density neighborhoods. The utility node *U2* is influenced by the non-observable variable and represents the agent's evaluation of the building's expected final destruction. For example, an early fire is expected to cause a lower destruction than a high intensity fire (given equivalent total areas and neighborhood density); thus, higher utility values are ascribed to early low intensity fires.

Figure 3 presents the ID that assembles all the above analysis: observable and non-observable variables, decision and utility nodes.
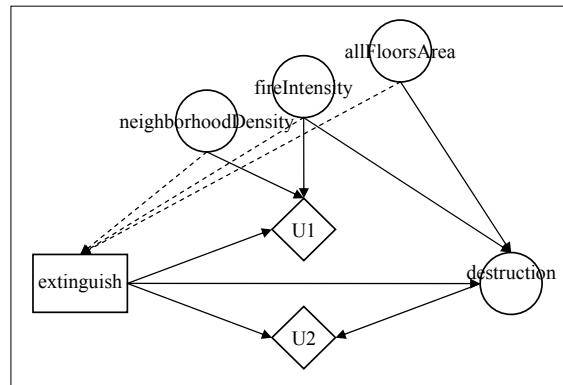

Fig. 3 Influence diagram for the *extinguish* set of goals (the construction follows the ID pattern, depicted in figure 1, thus adopting the terminology thereby defined).

The ID (cf. figure 3) is an instance of the proposed causal effect pattern (cf. figure 1) and digests the analysis of the illustrative scenario (cf. figure 2). The figure 3 intelligibility stresses that the ID is very handy in revealing the structure (the influence among the decision constituents) of the decision problem.

3.2 The preferences relation

After the ID structure we built the conditional probability and utility tables (CPT and UT) attached, respectively, to each chance and utility node. The CPT represents the probabilistic knowledge about the causal relations among the state variables. The UT specifies a decision-making strategy.

Our strategy follows the *three general fire attack strategies*. Figure 4 shows the *extinguish* expected utility, and each situation is represented by a vector,

$$v \equiv \langle neighbourhoodDensity, allFloorsArea, fireIntensity \rangle,$$

of perceived values (of the observable variables), for each building with a fire.

Each *v* variable is graphically discriminated as: i) the *neighbourhoodDensity* is a circle that becomes larger as the neighborhood density increases, ii) circles in

vertical lines follow the *allFloorsArea* value, upper circles having lower areas, and iii) the clustering of *fireIntensity* is marked in the graphic. For example, the B_1 and B_2 vectors are, respectively, $\langle \texttt{low},\texttt{low},1 \rangle$ and $\langle \texttt{medium},\texttt{medium},1 \rangle$.

Fig. 4 Decision *extinguish*, given the observation of *v* (*neighbourhoodDensity*, *allFloorsArea* e *fireIntensity*); the buildings B_1 and B_2 are labeled in the graphic.
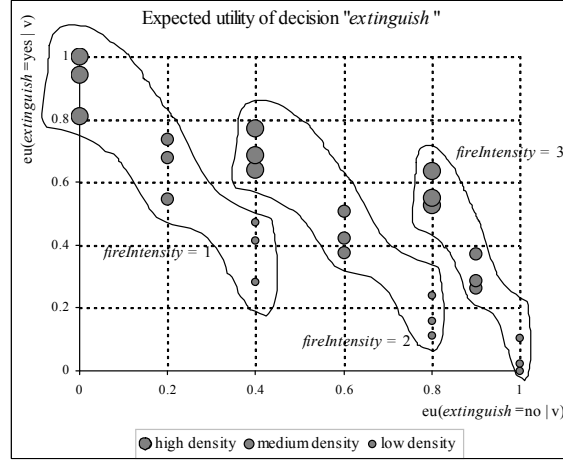
The expected *extinguish* utility, given *v* observation, is depicted in the abscissa and ordinate axis (cf. figure 4), respectively, by $eu(\,extinguish_g = \texttt{no} \mid v\,)$ and $eu(\,extinguish_g = \texttt{yes} \mid v\,)$, where subscript, *g*, is the goal of extinguishing the fire in the building where *v* was perceived, i.e., *g* represents the desired situation for the building (for readability symbol *g* is not plotted in figure 4).

Figure 4 shows that as the fire intensity increases the utility to extinguish decreases and the utility to ignore the fire increases although being interleaved with the various neighborhood densities (thus accounting the fire spreading effect). The highest utility (to extinguish) is assigned to lower area buildings (given equal values for other variables) as they are simpler and faster to control.

Figure 5 plots the agent preferences (given by expressions 2 and 3); a small diamond represents each previously shown *v* instance (cf. figure 4); a line goes through all diamonds in a course that links two adjacent priority situations (the darker segment highlights the B_2 to B_1 path). The highest preference is $\langle \texttt{high},\texttt{low},1 \rangle$, i.e., an early fire in a small building in a high density neighborhood; the lowest preference is $\langle \texttt{low},\texttt{high},3 \rangle$, i.e., a high intensity fire in a big building in a low density neighborhood. Table 1 details B_2 to B_1 preferences and shows that the three early fires are interleaved with higher intensity fires located in increasing density neighborhoods or decreasing area buildings. It is also interesting to note that the 2nd and 5th buildings only differ in their dimension (*allFloorsArea*) and the two buildings between them (3rd and 4th) have increasing neighborhood density and fire intensity and decreasing total area.

The interleaving of situations (shown in figure 5 and detailed in table 1), represents the trade-off, obtained from applying the expression 2, among our *three ge-*
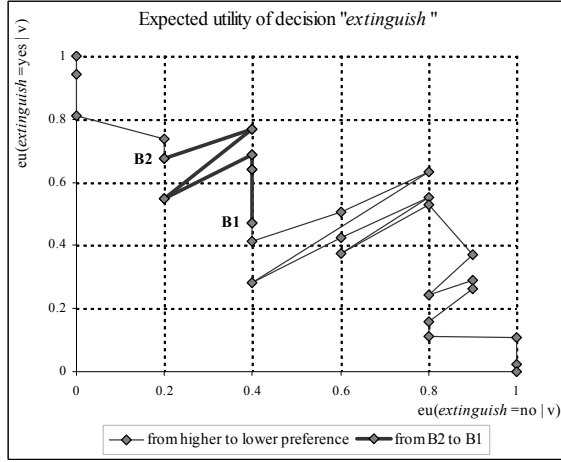
Fig. 5 A preference relation for the decision *extinguish*.

Table 1 Preferences detail (from B_2 to B_1); the first column identifies the buildings and the last column shows the utility value $u_{\mathcal{G}} \in [0,1]$; the first line represents B_2.

	Order	neighbourhoodDensity	allFloorsArea	fireIntensity	$\frac{u_{\mathcal{G}} - \min u_{\mathcal{G}}}{\max u_{\mathcal{G}} - \min u_{\mathcal{G}}}$
B_2	1st	medium	medium	1	0,73
...	2nd	high	low	2	0,69
	3rd	medium	high	1	0,66
	4th	high	medium	2	0,65
...	5th	high	high	2	0,62
B_1	6th	low	low	1	0,53

neral fire attack strategies. The rationality of those strategic guidelines may be disputed by a domain specialist for their verisimilitude with the real-world fire brigade strategies. Such dispute is a relevant contribution to adjust and mature the ID design but it is not the central discussion of this paper.

3.3 The decision design complexity

To apply the *three general fire attack strategies* (or any strategy set) a human designer would traduce its rationality into a total order relation among the state space. However, building a total order quickly becomes too complex. For example, our illustrative scenario has 4 variables, each with 3 values, thus a total of $3^4 = 81$ situations. To establish a total order, the human must compare each situation with all the others; in the worst case $\sum_{i=1}^{81-1} i = \frac{80 \times 81}{2} = 3240$ comparisons; in the best case (if able to apply a divide-and-conquer method), $81 \times \log_2 81 = 514$ comparisons. It is not likely that a human designer fulfils all those comparisons to establish a total order among all possible situations.

Our proposed ID design is much simpler. Assign, to each decision, the utility of observable and non-observable variables: $3^2 \times 2 + 3 \times 2 = 24$ assignments. This is an important complexity reduction: about 95% (from 514 to 24) in the above best case and about 99% (from 3240 to 24) in the above worst case. Despite that reduction the results (cf. figure 5 and table 1) exhibit a plausible translation of the general strategies used to guide the decision model design.

4 Conclusions and future work

This paper addresses a shortcoming, of current work, in the design of agents that act in complex domains: the evaluation of multiple simultaneous goals with observable and non-observable world state aspects. We propose a pattern, based on the influence diagram framework, to specify both the uncertainty of causal effects and the expected gain with regard to the decision of whether to pursue or ignore each goal. Practical experiences indicate that the ID pattern considerably simplifies the specification of a decision model (in RoboCupRescue domain) and enabled to established a preferences order among goals that is consistent with the initial, domain expert, very general strategies. This work represents the ongoing steps in a line of research that aims to develop decision-making agents that inhabit complex environments (e.g., the RoboCupRescue). Future work will apply the preferences model to the problem of coordinating teamwork (re)formation [6] from a centralized perspective.

References

1. Corrêa, M., Coelho, H.: Collective mental states in extended mental states framework. In: Proceedings of the IV International Conference on Collective Intentionality. Certosa di Pontignano, Siena, Italy (2004)
2. Howard, R., Matheson, J.: Influence diagrams. In: Readings on the Principles and Applications of Decision Analysis, vol. 2, pp. 721–762. Strategic Decision Group, Menlo Park, CA (1984)
3. Kitano, H., Tadokoro, S.: RoboCup Rescue: A grand challenge for multi-agent systems. Artificial Intelligence Magazine **22**(1), 39–52 (2001)
4. Pokahr, A., Braubach, L., Lamersdorf, W.: A goal deliberation strategy for BDI agent systems. In: Proceedings of the Third German Conference on Multi-Agent System Technologies (MATES-2005), pp. 82–94. Springer (2005)
5. Simari, G., Parsons, S.: On the relationship between MDPs and the BDI architecture. In: Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS-06), pp. 1041–1048. ACM Press, Hakodate, Japan (2006)
6. Trigo, P., Coelho, H.: Decision making with hybrid models: the case of collective and individual motivations. In: Proceedings of the EPIA-07 International Conference (New Trends in Artificial Intelligence), pp. 669–680. Guimarães, Portugal (2007)
7. Wooldridge, M.: Reasoning About Rational Agents, chap. Implementing Rational Agents. The MIT Press (2000)