BDI and QDT: a comparison based on classical decision theory

Mehdi Dastani and **Joris Hulstijn** and **Leendert van der Torre**
Department of Artificial Intelligence, Division of Mathematics and Computer Science,
Faculty of Sciences, Vrije Universiteit Amsterdam
De Boelelaan 1081a 1081 HV Amsterdam The Netherlands
{mehdi,joris,torre}@cs.vu.nl

Abstract

In this paper we compare Beliefs-Desire-Intention systems (BDI systems) with Qualitative Decision Theory (QDT). Our analysis based on classical decision theory illustrates several issues where one area may profit from research in the other area. BDI has studied how intentions link subsequent decisions, whereas QDT has studied methods to calculate candidate goals from desires, and how to derive intentions from goals. We also discuss the role of goals and norms in both approaches.

Introduction

In recent years the interest in models of decision making for autonomous agents has increased tremendously. Different proposals have been made, rooted in different research traditions and with different objectives. We are interested in the relation between two approaches to decision making. The first is based on an abstract model of the mental attitudes of an agent: beliefs, desires and intentions (BDI) (Bratman 1987; Rao & Georgeff 1991b; Cohen & Levesque 1990). The second is a qualitative extension of decision theory (QDT) (Pearl 1993; Boutilier 1994). Are BDI and QDT alternative solutions to the same problem, or are they complementary? In this paper we want to compare the two approaches, and find out what they can contribute to each other.

Both approaches criticize classical decision theory, which will therefore be our starting point in the comparison. Naturally, the discussion has to be superficial and cannot do justice to the subtleties defined in each approach. We therefore urge the reader to read the original papers we discuss. To complicate the question, each approach has different versions, with different objectives. In particular, there is a computational BDI, which is applied in software engineering (Jennings 2000), and explicitly considers architectures and implementations. There is also a cognitive theory of BDI, which models social and cognitive concepts in decision making (Conte & Castelfranchi 1995). Finally, there is a logical formalization of BDI, which we will use when we compare BDI with QDT (Rao & Georgeff 1991b; 1992). QDT developed out of models for reasoning under uncertainty. It is focused on theoretical models of decision making with potential applications in planning. To restrict the scope of our comparison, we focus on BDI work by Rao and Georgeff, which explicitly addresses the link with decision theory (Rao & Georgeff 1991a). We use their terminology where possible. For reasons of presentation, we restrict ourselves to formal logics, and can not go into architectures or into the underlying social concepts.

Consider Rao and Georgeff's initial BDI model (Rao & Georgeff 1991b). There are three types of data about respectively beliefs (B), goals (G) and intentions (I). These three types of data are related to each other by desires (D) and commitments (C) relations. These relations can be considered as constraints since they reduce the contents of the attitudes; in decision-theoretic terms they perform a kind of conditioning. This model is illustrated in figure 1. The content of the boxes corresponds to formulas in some logical language, with modal operators B, G and I. The formulas are interpreted as propositions: sets of possible worlds. Using modal operators we can specify axioms that correspond to particular semantic properties. For example, $Bp \rightarrow Gp$ to express *realism* (all potential goals p are believed feasible). Different combinations of axioms define different classes of models, or different agent types.

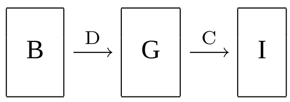

Figure 1: Relations between beliefs, goals, and intentions.

Work in qualitative decision theory is at least as diverse as the work on BDI, and we therefore again focus on a limited set of papers. There are two main issues. The first concerns the nature of the *qualitative decision rule*, for example to maximize expected utility. Given probabilities (corresponding to the beliefs of an agent) and valuations (corresponding to the desires of an agent), the rule selects which action to take. The second issue, is concerned with the *role of knowledge* in decision making. It investigates different ways to represent the data needed to apply a decision rule, and to deal with potential conflicts. Often representations developed for reasoning under uncertainty are used for this purpose. Looking at figure 1, we conclude that where BDI simply assumes selection functions (for instance to be supplied by a planning algorithm), QDT provides explicit processes.

We want to investigate how BDI and QDT are related. The layout of this paper is therefore as follows. First we discuss a version of decision theory. Then we discuss Rao and Georgeff's BDI approach and its relation to decision theory. In the third section we discuss qualitative decision theory, focusing on decision rules and on the role of knowledge in decision making. In the fourth section we deal with the role of goals in these approaches. In section 5 we consider the possibility of adding a fourth attitude, obligations.

Decision theory

There are many different ways to represent classical decision theory. Here we use a quasi-logical approach inspired by Jeffrey's classical logic of decision(Jeffrey 1965), which will be useful for our future comparisons. The typical construct here is the sum of products of probabilities and payoff values. This construct can be used in two different places, corresponding to two different sources of uncertainty. The first concerns uncertainty about the facts, the second uncertainty about what the agent desires.

- An action $a \in A$ is identified with its effects, i.e. with a probability function p_a from worlds $w \in W$ to the real interval $[0, 1]$ such that $\sum_{w \in W} p(w) = 1$ (reflecting that the probability of a set of worlds is the sum of the probabilities of the individual worlds). Preferences are represented by a real-valued function v, called payoff function, on the worlds, and the utility of an action is the sum of the products of probability and payoff of the worlds, $u(a) = \sum_{w \in W} p_a(w) \times v(w)$.
- Given the uncertainty about which world v a decision maker is in, represented by a probability distribution p on the worlds, the utility of an action as defined above is relative to this state, i.e. $u(a(v))$. Now the *decision rule* of the decision maker is to choose the action which maximizes his expected utility, i.e. $\max_{a \in A} \sum_{v \in W} p(v) \times u(a(v))$.

It is a well-known result from decision theory that a decision problem with both types of uncertainty can be rewritten to a decision problem with only uncertainty about the state the decision maker is in. The trick is to define a split each world the decision amker may be in, into a set of worlds, one for each possible effect of the actions.

Decision theory is an active research area within economics, and the number of extensions and subtleties is too large to address here. However, we do wish to note the following aspects:

Representability. Representation theorems, such as the most famous one of Savage (Savage 1954), typically prove that each decision maker obeying certain innocent looking postulates (about weighted choices) acts *as if* he applies the maximize-expected-utility decision rule with some probability distribution and utility function. Thus, he does not have to be aware of it, and his utility function does not have to represent selfishness. In fact, exactly the same is true for altruistic decision makers, they also act as if they maximum expected utility; they just have another utility function.

Preference elicitation. Given the fact that in general the decision maker is not aware of the probability distribution and his utility function, several techniques have been developed to make them explicit. This is typically an iterative process.

Decision rule. Several other decision rules have been investigated, including qualitative ones, such as maximin, minimax, minregret etc.

One shot. Classical decision theory has been extended in several ways to deal with multiple choices and plans, such as the well known Markov Decision Processes (MDPs).

Small Worlds. A "small world" in the sense of (1954) is derived from the real "grand world" by neglecting some distinctions between possible states of the world and thus by drawing a simplified picture. Of course, the problem of small worlds is that an analysis using one small world may fail to agree with an analysis using a more refined small world.

Independence. By introducing decision variables and independence assumptions the probability distribution can be represented concisely, as studied in Bayesian networks (Pearl 1993).

Multiple criteria. An important extension of decision theory deals with conflicting objectives, which can be represented by independence assumptions on the utility function, leading to ceteris paribus preferences (Keeney & Raiffa 1976).

Risk. Though classical decision theory is thought to incorporate some notion of risk (e.g. neutrality or risk averseness), for handling extreme values different constructs have been proposed.

One particular extension of interest here is decision-theoretic planning (Boutilier, Dean, & Hanks 1999), a response to the limitations of traditional goal-based planning. Goals serve a dual role in most planning systems, capturing aspects of both intentions and desires (Doyle 1980). Besides expressing the desirability of a state, adopting a goal represents some commitment to pursuing that state. For example, accepting a proposition as an achievement task commits the agent to finding some way to accomplish this objective, even if this requires adopting some subtasks that may not correspond to desirable propositions themselves (Dean & Wellman 1991). In realistic planning situations objectives can be satisfied to varying degrees, and frequently goals cannot be achieved. Context-sensitive goals are formalized with basic concepts from decision theory (Dean & Wellman 1991; Doyle & Wellman 1991; Boutilier 1994). In general, goal-based planning must be extended with a mechanism to choose between which goals must be adopted and which ones must be dropped.

On an abstract level, decision-making with flexible goals has split the decision-making process in two steps. First a decision is made which goals to adopt, and second a decision is made how to reach these goals. At first sight, it seems that we can apply classical decision theory to each of these two sub-decisions. However, there is a caveat. The two sub-decisions are not independent, but closely related! For example, to decide which goals to adopt we must know which goals are feasible, and we thus have to take the possible actions into account. Moreover, the intended actions constrain the candidate goals which can be adopted. Other complications arise due to uncertainty, changing environments, etc, and we conclude here that the role of decision theory in planning is complex, and that decision-theoretic planning is much more complex than classical decision theory.

BDI Logic

BDI logic is developed to formalize decision theoretic aspects of agent planning. The main advantage of BDI logic over existing decision theories is the role of prior intentions, which is essential for agent decision making and planning (Bratman 1987; Cohen & Levesque 1990; Rao & Georgeff 1991b). Prior intentions, plans or goals that agents have adopted or are committed to perform, constrain the options for future actions. Different principles are introduced to govern the balance between prior intentions and the formation of new intentions. For example, the principle that an agent should only abandon an intention when the motivating goal has become achieved or has become inachievable by some other cause. The key notions that are essential in agent planning and that are used to formalized BDI logic are beliefs, goals and intentions as well as time, actions, plans, probabilities, and payoffs.

BDI logic as defined by Rao and Georgeff is a multi-modal logic consisting of three modal operators: Beliefs (B), Goals (G), and Intentions (I). The belief operator B denotes *possible plans*, the goal operator G denotes *relevant plans*, and the intention operator I denotes *plans* the agent is *committed* to. These modalities are related to each other. A plan that an agent is committed to perform should be relevant ($Gp \rightarrow Ip$), and a relevant plan should be believed as possible ($Bp \rightarrow Gp$). Moreover, in the BDI logic two types of formulae are distinguished: state and path formulae. State formulae represent the effect of actions and path formulae represent plans, or combination of actions. In the semantics a so called branching time structure is imposed on states: a state at time t_n is followed by different states at time t_{n+1}, distinguished by different events. Path formulae then denote paths along states in the branching time structure.

Just like in decision theory, the effects of events involve uncertainties (determined by nature) and the effects of plans have a degree of desirability. Uncertainties about the effects of events are represented by formulae $PROB(\varphi) \geq a$ (probability of state formula φ is greater or equal to real number a). The desirability of plans is represented by formulae $PAYOFF(\psi) \geq a$ (desirability of the plan ψ is greater or equal to real value a). Their interpretation is defined in terms of beliefs and goals, respectively. In particular, given a model M, a world w at time t ($w_t \in M$), an assigned probability function $P(w_t)$, an assigned utility function $U(w_t)$, the set $\mathcal{B}^w$ of belief accessible worlds from w, and the set $\mathcal{G}^w$ of goal accessible worlds from w, then

$$M, w_t \models PROB(\varphi) \geq a \;\Leftrightarrow$$
$$P(\{w' \in \mathcal{B}^{w_t} \mid M, w'_t \models \varphi\}) \geq a.$$

$$M, w_t \models PAYOFF(\psi) \geq a \;\Leftrightarrow$$
$$\forall w' \in \mathcal{G}^{w_t}\ \forall x_i \quad M, x_i \models \psi \rightarrow U(x_i) \geq a$$
where x_i is a full path starting from w_t.

Based on these modalities and formulae, an agent may decide which plan it should commit to on the basis of a decision rule such as maximum expected utility.

Prior intentions are said to be essential. How are prior intentions accounted for in this approach? The role of prior intentions is established by the time structure imposed on sequences of worlds, combined with constraints that represent the *type of agent*. An agent type determines the relationship between intention-accessible worlds and belief and goal-accessible worlds, at different time points. The main types of agents are *blindly-committed* agents (intended plans may not change if possible plans or relevant plans change), *single minded* agents (intended plans may change only if possible plans change), and *open minded* agents (intended plans change if either possible plans or relevant plans change).

Given such relations between accessible worlds, intention is then defined as follows: Ip is true in a world w at time t if and only if p is true in all intention accessible worlds at time t. Now, assuming that intention accessible worlds are related to belief and goal accessible worlds according to the agent type, this definition states that a plan is true if and only if it is a continuation of an existing plan which has been started at time point t_0 and which is determined by the beliefs and goals according to the agent type.

BDI and decision theory

How are the issues discussed in BDI logic related to the issues discussed in decision theory? Clearly most issues of BDI logic are not discussed in classical decision theory as presented in this paper, because they are outside its scope. For example, prior intentions are said to be essential for the BDI logic, but they are irrelevant for one-shot decision theory. Instead, they are discussed in decision processes. Likewise, plans are not discussed in decision theory, but in decision-theoretic planning. In this section we therefore restrict ourselves to one aspect of BDI logics: the relation between the modal operators and $PROB$ and $PAYOFF$. We therefore consider the translation of decision trees to BDI logic presented in (Rao & Georgeff 1991a).

Rao and Georgeff emphasize the importance of decisions for planning agents and argue that the BDI logic can be used to model the decision making behaviour of planning agents by showing how some aspects of decision theory can be modeled by the BDI logic. In particular, they argue that a given decision tree, based on which an agent should decides a plan, implies a set of plans that are considered to be relevant and show how decision trees can be transformed into agent's goals. Before discussing this transformation we note that in decision theory there are two types of uncertainties distinguished: the uncertainty about the effect of actions and the uncertainty about the actual world. In Rao and Georgeff only the first type of uncertainty is taken into account. The fact that the decision tree is assumed to be given implies that the uncertainty about the actual world is ignored.

A decision tree consists of two types of nodes: one type of nodes expresses agent's choices and the other type expresses the uncertainties about the effect of actions. These two types of nodes are indicated by square and circle in the decision trees as illustrated in figure 2. In order to generate relevant plans (goals), the uncertainties about the effect of actions are removed from the given decision tree (circle in figure 2) resulting in a number of new decision trees. The uncertainties about the effect of actions are now assigned to the newly generated decision trees.

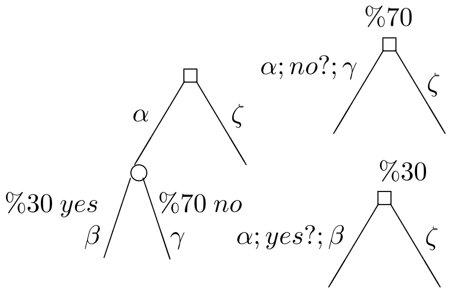

Figure 2: An example of a decision tree and how the uncertainty about agent's environment can be removed.

For example, consider again the decision tree illustrated in figure 2. A possible plan is to perform α followed by either β if the effect of α is yes or γ if the effect of α is no and suppose that the probability of yes as the effect of α is $\%30$ and that the probability of no as the effect of α is $\%70$. The transformation will generates two new decision trees: one in which event yes takes place after choosing α and one in which event no takes place after choosing α. The uncertainty $\%30$ and $\%70$ are then assigned to the resulted trees, respectively. Note that these new decision trees provides plans $P_1 = \alpha; yes?; \beta$ and $P_2 = \alpha; no?\gamma$ with probabilities $\%30$ and $\%70$, respectively. In these plans the effects of events are known.

The sets of plans provided by the newly generated decision trees are then considered as constituting goal accessible worlds. Note that a certain plan can occur in more than one goal world and that the payoffs associated with those plans remains the same. The probability of a plan that occur in more than one goal world is then the sum of the probabilities of different goal worlds in which the plan occurs. The agent can then decide on a plan by means of a decision rule such as maximum expected utility.

It is important to note that the proposed transformation is not between *decision theory* and *BDI logic*, but only between a *decision tree* and agent's *goal accessible worlds*. As noted, in Rao and Georgeff the uncertainty about agent's *actual world* is ignored by focusing only on one single decision tree that provides all *relevant* goal accessible worlds. However, an account of the uncertainty about the actual world implies a set of decision trees that provide a set of sets of

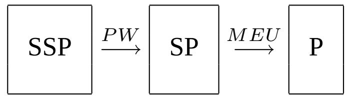

Figure 3: Relations between Beliefs (SSP), Goals (SP) and Intentions (P).

plans which in turn corresponds with a set of sets of probabilities distributions as it is known in decision theory.

In fact, a set of sets of probability (SSP) corresponds to agent's *beliefs*. Given such a set, an agent takes only one set of probabilities as being relevant. This can be done by taking the second source of uncertainty into account, which is formalized in decision theory by a probability distribution on the worlds (PW). The goals of an agent is then a set of probability distributions (SP) which correspond to a decision tree. Finally, given a set of probability distributions, an agent considers only one probability distribution (P) as being intended. This probability distribution is determined by applying one decision rule such as maximum expected utility (MEU). The application of a decision rule results in a plan that corresponds to agent's *intentions*. This view is illustrated in Figure 3.

Computational BDI

As we have seen, the BDI theory makes a number of claims related to the resource-boundedness of agents. First it claims to deal with incomplete information; this can be captured in DT by probability distributions. Second, it claims to deal with dynamic environments, by means of the stabilizing effect of intentions. Based on the BDI theory various proposals have been launched to implement a rational and resource-bounded agent that make decisions and plans (Rao & Georgeff 1991b; 1992). The proposed BDI agent consists of three data structures that represent beliefs, goals, and intentions of the agents together with an input queue of events. Based on these components, an interpreter is introduced by means of which the BDI agent can repeatedly observe the events from the environment, generates possible plans, deliberate on these plans, execute some decided plans, and evaluate its intended plans.

The decision process consists of two steps. At each step a decision is made on the basis of a selection mechanism and following some principles that govern the relations between beliefs, goals, and intentions. As noted, these principles are usually presented as constraints defined on relationship between different components.

At the first step, a BDI agent generates possible plans based on its observations, beliefs and using means-end analysis while it takes the previous chosen plans into account according to some principles such as ”blindly committed agents”, ”single-minded agents”, and ”open-minded agents”. Note that this step is an implementation of deciding relevant plans where uncertainty about actual world is taken into consideration. This is done by observing the environment and using beliefs and prior intentions. At the second step, a BDI agent applies some decision rules such as maximum expected utility to the set of possible plans that are generated at step one and decide a plan.

Qualitative decision theory

Qualitative decision theory developed in the area of reasoning about uncertainty, a sub-domain of artificial intelligence, which mainly attracts researchers with a background in reasoning about defaults and beliefs. Two typical and often cited examples are (Pearl 1993) and (Boutilier 1994). The research is motivated as follows:

- to do decision theory, or something close to it, when not all numbers are available.
- to provide a framework for flexible or context-sensitive representation of goals in planning.

Often the formalisms of reasoning about uncertainty are re-applied in the area of decision making, and therefore use concepts developed before. Thus, typically uncertainty is not represented by a probability function, but by a plausibility function, a possibilistic function, Spohn-type rankings, etc. Another consequence of this historic development is that the area is much more mathematically oriented than the planning community or the BDI community. The two main proponents of QDT in the last years have been Jon Doyle and Richmond Thomason, who wrote several papers (Thomason & Horty 1996), including some survey papers (Doyle & Thomason 1999), and who organized the AAAI 1997 spring symposium on qualitative decision theory. The reader is advised to consult their highly interesting work.

Now we consider two research areas which are related to the first and second step of decision making. The first step is concerned with the construction of the decision space, through the collection of the initial data. The data, or beliefs, help to constrain all possible courses of action for the agent, and implicitly select a limited number of candidate goals. On the basis of this data, we can apply a *decision rule*, which will select one or more particular actions. The research communities related to these steps have a different understanding of 'qualitative'. For the first step, typical issues are non-monotonicity of the representation formalism, and conflict resolution. Here one often uses symbolic (non-quantitative) representations, such as default logic. The numbers however do play a role in the semantics of these representation formalisms. In the second step, a qualitative version of the decision rule is applied. The decision rule can do without the exact numbers; it only requires differences expressed as orders of magnitude. Since this latter step also affects the data collection, we start by explaining it in more detail.

Decision rules

As we mentioned in the section on decision theory, examples of qualitative decision rules can already be found in decision theory itself, where besides the quantitative MEU decision rule also qualitative decision rules are studied. Examples are 'minimize the worst outcome' (pessimistic Wald criterion) or 'minimize your regret'. Some of these rules do not need exact numbers, but only qualitative orders. For example, to prevent the worst outcome we only have to know these worst outcomes; not how much worse they are compared to other outcomes. The same rules have been studied in qualitative decision theory. For example (Boutilier 1994) and (Dubois & Prade 1995) study versions of the Wald criterion.

However, there is a problem with a purely qualitative approach. It is unclear how, besides the most likely situations, also less likely situations can be taken into account. In particular we are interested in situations which are unlikely, but which have a high impact. i.e. an extreme high or low utility. For example, the probability that your house will burn down is not very high, but it is uncomfortable. Some people therefore decide to take an insurance. In a purely qualitative setting it is unknown how much expected impact such an unlikely situation has. There is no way to compare a likely but mildly important effect to an unlikely but important effect. Going from quantitative to qualitative we may have gained computational efficiency, but we lost one of the main useful properties of decision theory.

A solution proposed in qualitative decision theory is based on two ideas. First, the initial probabilities and utilities are neither represented by quantitative probability distributions and utility functions, nor by pure qualitative orders, but by something in between. This can be called *semi-qualitative*. Typically, one uses the representation formalisms developed for reasoning about uncertainty, such as possibilistic functions and Spohn-type rankings. Second, an assumption can be introduced to make the two semi-qualitative functions comparable. This has sometimes been called the *commensurability assumption*, see e.g. (Dubois & Prade 1995). Because they are crucial for QDT, we further illustrate these two ideas by Pearl's (1993) proposal.

First, we introduce semi-qualitative rankings. A belief ranking function $\kappa(w)$ is an assignment of non-negative integers to worlds $w \in W$ such that $\kappa(w) = 0$ for at least one world. Intuitively, $\kappa(w)$ represents the degree of surprise associated with finding a world w realized, and worlds assigned $\kappa(w) = 0$ are considered serious possibilities. Likewise, $\mu(w)$ is an integer-valued utility ranking of worlds. Second, we make the rankings commensurable by defining both probabilities and utilities as a function of the same ϵ, which is treated as an infinitisimal quantity (smaller than any real number). $\kappa(w)$ can be considered an order-of-magnitude approximation of a probability function $P(w)$ by writing $P(w)$ as a polynomial of some small quantity ϵ and taking the most significant term of that polynomial. Similarly, positive $\mu(w)$ can be considered an approximation of a utility function $U(w)$.

$$P(w) \sim C\epsilon^{\kappa(w)}, U(w) = O(1/\epsilon^{\mu(w)})$$

Taking $P'(w)$ as the probability function that would prevail after obtaining φ, the expected utility criterion $U(\varphi) = \Sigma_{w \in W} P'(w)U(w)$ shows that we can have for example likely and moderately interesting worlds ($\kappa(w) = 0, \mu(w) = 0$) or unlikely but very important worlds ($\kappa(w) = 1, \mu(w) = 1$), which have become comparable since in the second case we have ϵ/ϵ. There is one more subtlety here, which has to do with the fact that whereas κ rankings are positive, the μ rankings can be either positive or negative. This represents that outcomes can be either very desirable or very undesirable . For negative $\mu(w)$ we define

$U(w) = -O(1/\epsilon^{-\mu(w)})$. Besides the representation of the abstractions of the probability distribution and utility function there are many further issues discussed by Pearl, one of which we discuss now.

Actions and observations

Pearl's (1993) approach is based on two additional assumptions, inspired by deontic logic (the logic of obligations and permissions, see also the section on norms below). First, while theories of actions are normally formulated as theories of temporal changes, his logic suppresses explicit references to time. Second, whereas actions in decision theory are pre-designated to a few distinguished atomic variables, he assumes that actions 'Do(φ)' are presumed applicable to any proposition φ. In this setting, he explains that the expected utility of a proposition φ clearly depends on how we came to know φ. For example, if we find out that the ground is wet, it matters whether we watered the ground (action) or we happened to find the ground wet (observation). In the first case, finding φ true may provide information about the natural process that led to the observation φ, and we should change the current probability from $P(w)$ to $P(w|\varphi)$. In the second case, our actions may perturb the natural flow of events, and $P(w)$ will change without shedding light on the typical causes of φ. This issue is precisely the distinction between Lewis' conditioning and imaging, between belief revision and belief update, and between indicative and subjunctive conditionals. One of the tools Pearl uses for the formalization of this distinction are causal networks (a kind of Bayesian networks with actions).

Boutilier (1994) introduces a very simple but elegant distinction between consequences of actions and consequences of observations, by distinguishing between controllable and uncontrollable propositional atoms. This seems to make his logic less expressive but also simpler than Pearl's (but beware, there are still a lot of complications around). One of the issues he discusses is the distinction between goals under complete knowledge and goals under partial knowledge.

From desires to goals

In the first step we have a symbolic modal language to reason implicitly about the probability distributions and probability functions (or their abstractions), and this is where concepts like goals and desires come into the picture. Boutilier uses a modality $\mathcal{I}$ for Ideal. Its semantics is derived from Hansson's deontic logic with modal operator O for obligation (Hansson 1969). The semantics makes use of some minimality criterium.[1] If $M = \langle W, \leq, V \rangle$ with W a set of worlds, $\leq$ a binary reflexive, transitive, total and bounded relation on W and V a valuation function of the propositional atoms at the worlds, then $M \models \mathcal{I}(p|q)$ iff $\exists w\ M, w \models p \wedge q$ and $\forall w' \leq w\ M, w' \models (q \rightarrow p)$. In other words: $\mathcal{I}(p|q)$ is true if all minimal q-worlds are p-worlds. Note that the notion of minimal (or best, or preferred) can be represented by a partial order, by a ranking or by a probability distribution. A similar type of semantics is used by Lang (1996) for a modality D to model desire. An alternative approach represents conditional modalities by so called 'ceteris paribus' preferences, using additional formal machinery to formalize the notion of 'similar circumstances', see e.g. (Doyle & Wellman 1991; Doyle, Shoham, & Wellman 1991; Tan & Pearl 1994a; 1994b).

The consequence of the use of a logic is that symbolic formulas representing desires and goals put *constraints* on the models representing the data, i.e. the semantics. The constraints can either be underspecified, there are many models satisfying the constraints, or it can be overspecified, there are no models satisfying the constraints. Consequently, typical issues are how we can strengthen the constraints, i.e. in logical terminology how we can use non-monotonic closure rules, and how we can weaken the constraints, i.e. how we can resolve conflicts.

Non-monotonic closure rules The constraints imposed by $\mathcal{I}$ formulas of Boutilier (ideal) are relatively weak. Since the semantics of the $\mathcal{I}$ operator is analogous to the semantics of many default logics, Boutilier (1994) proposes to use non-monotonic closure rules for the $\mathcal{I}$ operator too. In particular he uses the well-known system Z (Pearl 1990). Its workings can be summarized as 'gravitation towards the ideal', in this case. An advantage of this system is that it always gives exactly one preferred model, and that the same logic can be used for both desires and defaults. A variant of this idea was developed by Lang (1996), who directly associates penalties with the desires (based on penalty logic (Pinkas 1995)) and who does not use rankings of utility functions but utility functions themselves. More complex constructions have been discussed in (Tan & Pearl 1994b; 1994a; van der Torre & Weydert 2000; Lang, van der Torre, & Weydert 2001b).

Conflict resolution Although the constraints imposed by the $\mathcal{I}$ operator are rather weak, they are still too strong to represent certain types of conflicts. Consider conflicts among desires. Typically desires are allowed to be inconsistent, but once they are adopted and have become intentions, they should be consistent. Moreover, in the logic proposed in (Tan & Pearl 1994b) a specificity set like $D(p), D(\neg p|q)$ is inconsistent (this problem is solved in (Tan & Pearl 1994a), but the approach is criticized as ad hoc in (Bacchus & Grove 1996)). A large set of potential conflicts between desires, including a classification and ways to resolve it, is given in (Lang, van der Torre, & Weydert 2001a). A radically different approach to solving conflicts is to apply Reiter's default logic to create extensions. This is recently proposed by Thomason (2000), and thereafter also used in the BOID architecture which will be explained below (Broersen *et al.* 2001).

Finally, an important branch of decision theory has to do with reasoning about multiple objectives, which may conflict, using multiple attribute utility theory: (Keeney & Raiffa 1976). This is also the basis of the 'ceteris paribus' preferences mentioned above. This can be used to formalize conflicting desires. Note that all the modal approaches

[1] In default logic, an exception is a digression from a default rule. Similarly, in deontic logic an offense is a digression from the ideal.

above would make conflicting desires inconsistent.

Goals and planning

Now that we have introduced decision theory, planning, BDI and QDT we want to say something about their perspective on goals. There is a distinction between goal-based planning and QDT. The distinction is that QDT defines the best action, whereas goal-based planning is a heuristic to find the best action (which may often give sub-optimal solutions). QDT is thus like decision theory and Rao and Georgeff's BDI approach, but unlike other BDI-like approaches like cognitive BDI which derive their intuitions from goal-based planning. We will also investigate how to construct goals. For this we use an extension of Thomason's BD logic.

What is a goal?

In the planning community a goal is simply the end-state of an adopted plan. Therefore the notion of goals incorporates both desirability and intention. As far as we know neither the relation between BDI and these goals nor the relation between QDT and these goals has been investigated. It has been observed in BDI that their notion of goals differs from the notion of goals in planning. This is now obvious: the first only plays the role of desirability, not of intention, because intention has been separated out.

QDT and goal-based planning

How can we derive intentions from desires and abilities? With only a set of desires, even with conditional desires, there is not much we can do. It seems therefore obvious that we have to extend the notion of desire. Two options have been studied: first an ordering on goals reflecting their priority, and secondly a local payoff function with each desire, representing its associated penalties (if violated) and rewards (if achieved). These concepts should not be taken too literary. In this section we use the latter approach. A more general account of penalty logic can be found in (Pinkas 1995).

Assume a set of desires with rewards (if the desire is achieved) and penalties (when a desire is not achieved), together with a set of actions the agent can perform and their effects. There are the following two options, depending on the number of intermediate steps.

1. (one-step) For each action, compute the sum of all rewards and penalties. The intended action is the one with the highest pay-off.
2. (two-step) First, conflicts among desires are resolved using rewards and penalties; the resulting non-conflicting desires are called goals. Second, the intended actions are the preferred actions of those that achieve the goals.

The first approach will give the best action, but the second is more efficient if the number of actions is much higher than the number of desires involved. This may be the reason that most existing planning systems work according to the second approach. If the rewards and penalties are interpreted as utilities and the first approach is extended in the obvious way with probabilities, then this first approach may be called a (qualitative) decision-theoretic approach (Boutilier 1994; Lang 1996; Pearl 1993). The utilities can be added and the desires are thus so-called additively independent (Keeney & Raiffa 1976).

Role of desires

Before we give our formal details we look a bit closer into the second approach. Here are some examples to illustrate the two approaches.

1. First the agent has to combine desires and resolve conflicts between them. For example, assume that the agent desires to be on the beach, if he is on the beach then he desires to eat an ice-cream, he desires to be in the cinema, if he is in the cinema then he desires to eat popcorn, and he cannot be at the beach as well as in the cinema. Now he has to choose one of the two combined desires, or optional goals, being at the beach with ice-cream or being in the cinema with popcorn.
2. Second, the agent has to find out which actions or plans he can execute to reach the goal, and he has to take all side-effects of the actions into account. For example, assume that he desires to be on the beach, if he will quit his job and drive to the beach, he will be on the beach, if he does not have a job he will be poor, if he is poor then he desires to work. The only desire and thus the goal is to be on the beach, the only way to reach this goal is to quit his job, but the side effect of this action is that he will be poor and in that case he does not want to be on the beach but he wants to work.

Now crucially, desires come into the picture two times! First they are used to determine the goals, and second they are used to evaluate the side-effects of the actions to reach these goals. In extreme cases, like the example above, what seemed like a goal may not be desirable, because the only actions to reach these goals have negative effects with much more impact than these goals. This shows that the balancing philosophy of BDI could be very helpful.

Comparison

We use an extension of Thomason's (2000) BD logic. We have $p \overset{\text{B}}{\to} q$ for 'if p then I belief q', $p \overset{\text{D}}{\to} q$ for 'if a then I desire x' and $a \overset{\text{B}}{\to} p$ for 'if I do a then p will result. Note that the a in $a \overset{\text{B}}{\to} p$ is an action whereas all other propositions are states. We do not get into details how this distinction is formalized. Simply assume that the action is a controllable proposition whereas the others do not have to be so (Boutilier 1994; Lang 1996). Otherwise think about the action is a complex proposition like $Do(s)$ whereas the others are $Holds(s)$, or the action is a modal formula See-To-It-That-p or $STIT : p$ whereas the others are not.

To facilitate the examples we work with explicit rewards for achieving a desire and penalties for not achieving a desire. That is, we write $a \overset{\text{D}}{\to} x^{r:p}$ for a desire $a \overset{\text{D}}{\to} x$ that results in a reward of $r > 0$ if it is achieved, i.e. if we choose an action that results in $a \wedge x$, and that results in a penalty $p < 0$ if it is not achieved, i.e. if we choose an action that results in $a \wedge \neg x$. For a total plan we add the rewards and penalties of all desires derived by it.

With these preliminaries we are ready to discuss some examples to illustrate the problem of planning with desires. Life is full of conflicting desires. You want to be rich and to do scientific research. You like to eat candy without becoming fat and wasting your teeth. You want to buy so many things but your budget is too limited. Resolving these kinds of conflicts is an important part of planning.

Example 1 *Let* $S = \{\top \overset{D}{\rightarrow} p^{3:-1}, \top \overset{D}{\rightarrow} \neg p^{2:-1}, q \overset{D}{\rightarrow} \neg p^{0:-10}, a \overset{B}{\rightarrow} p \wedge q, b \overset{B}{\rightarrow} \neg p \wedge \neg q\}$.

First we only consider the desires to calculate our goals. The only two desires which are applicable are the first two, the optional goals are p an $\neg p$. If we just consider the desire then we prefer p over $\neg p$, since the first has total of 3-1=2, whereas the latter has 2-1=1. There is only one action which results in p, namely action a which induces p and q.

However, now consider the actions and their results. There are two actions, a and b. The totals of these actions are 3-1-10=-8 and 2-1=1, such that the second action is much better. This action does not derive p however, but $\neg p$. Thus it does not derive the goal but its negation.

The example illustrates that the two stage process of first selecting goals and thereafter actions to reach goals may lead to inferior choices. In principle, the best way is to calculate the expected utility of each action and select the action with the highest utility. However, that is often too complex so we choose ythis goal based heuristic. This is the main problem of desire-based planning, but there are more complications. Consider the following example.

Example 2 *Let* $S = \{\top \overset{D}{\rightarrow} p, p \overset{D}{\rightarrow} q, q \overset{D}{\rightarrow} r, a \overset{B}{\rightarrow} p \wedge \neg r\}$.

In deriving goals we can derive $\{p, q, r\}$ by iteratively applying the first three desires. However, we can only see to it that p. The other derivations are useless.

The examples suggests that there is a trade-off between calculating an agent's goals from its desires, and calculating intended actions to achieve these goals. In some applications it is better to first calculate the complete goals and then find actions to achieve them, in other applications it is better to calculate partial goals first, look for ways to achieve them, and then continue to calculate the goals. An advantage of the second approach is that goals tend to be stable. We finally note that the approaches correspond to different types of agent behavior. The first is more reactive because it is action oriented whereas the second is more deliberative.

Norms and the BOID

Recently computational BDI has been extended with norms (or obligations), see e.g. (Dignum 1999), though it is still debated whether artificial agents need norms, and if they are needed then for what purposes. It is also debated whether they should be represented explicitly. Arguments for their use have been given in the cognitive approach to BDI, in evolutionary game theory and in the philosophical areas of practical reasoning and deontic logic.

Norms as goal generators. The cognitive science approach to BDI (Conte & Castelfranchi 1995; Castelfranchi 1998) argues that norms are needed to model social agents. Norms are important concepts for social agents, because they are a mechanism by which society can influence the behavior of individual agents. This happens through the creation of normative goals, a process which consists of four steps. First the agent has to believe that there is a norm, then it has to believe that this norm is applicable, then it has to decide that it accepts this norm – the norm now leads to a normative goal – and finally it has to decide whether it will fulfill this normative goal.

Reciprocal norms. The argument of evolutionary game theory (Axelrod 1984) is that reciprocal norms are needed to establish cooperation in repeated prisoner's dilemmas.

Norms influencing decisions. In practical reasoning, in legal philosophy and in deontic logic (in philosophy as well as in computer science) it has been studied how norms influence behavior.

The relation between QDT and norms is completely different from the relation between BDI and norms. Several papers mention formal relations between QDT and deontic logic. For example, (Boutilier 1994) observes that his ideality operator $\mathcal{I}$ is equivalent to Hansson's deontic logic, and (Pearl 1993) gave his paper the title 'from conditional oughts to qualitative decision theory'. The idea here seems to be that qualitative decision theory can be extended with normative reasoning, but that it *is* a kind of normative reasoning. Summarizing, and rather speculative, the distinction between the role of norms in BDI and QDT is that BDI considers norms as potential extensions, whereas QDT uses norms in its foundation. This can be explained by the normative character of decision theory, an aspect not stressed in BDI.

BOID architecture

Consider a BDI-like agent architecture. For environments that are essentially social environments, the discussion above suggests we could add a fourth component to represent obligations or norms. In that case we would get the following four attitudes:

belief information about facts, and about effects of actions

obligation externally motivated potential goals

intention committed goals

desire internally motivated potential goals

In case we represent these attitudes by separate reasoning components, we need principles to regulate their interaction. The so called BOID architecture (Broersen *et al.* 2001) does just that. It consists of a simple logic, an architecture and a number of benchmark examples to try and test various priorities among attitudes, resulting in different types of agents. For example, a social agent always ranks obligations over desires in case of a conflict. A selfish agent ranks desires over obligations.

How we can we specify such agent types by the BOID logic? To explain the architecture and the logic, we mention its main source of inspiration, Thomason's BDP. Thomason (Thomason 2000) proposes an interesting and novel approach, which is based on two ideas:

Planning. A logic of beliefs and desires is extended with more traditional goal-based planning, where the purpose of the beliefs and desires is to determine the goals used in planning. Planning replaces the decision rule of QDT.

Conflict-resolution. An important idea here is that the selection of goals is basically a process of conflict resolution. Thomason therefore formalizes beliefs and desires as Reiter default rules, such that resolving a conflict becomes choosing an extension.

The BOID as presented in (Broersen *et al.* 2001) extends Thomason's idea in the following two directions, though in the present state planning has not been incorporated (this is mentioned as one issue for further development).

Norms. In the BOID architecture the obligations are just like desires a relation between beliefs and candidate goals.

Intentions. If the intentions component commits to a desire (or an obligation), then it is internalized in the intentions component. Hereafter it overrides other desires during the extension selection, i.e. the conflict resolution phase.

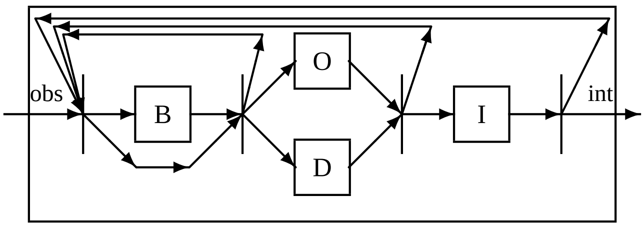

Figure 4: BOID Architecture

Each component is a high level abstraction of a process, of which only its input and output behavior are described. This is inspired by component based approaches in agent based software engineering, where such a component may for example be a legacy computer system with a wrapper built around it. The B component relates observations with their consequences, and actions with their effects. The O component relates observations and normative goals or obliged actions. The D component relates observations and goals or desired actions. The I component (which in the BDI terminology should have been called the C component – but BOCD does not sound that good) relates the goals or desired/obliged actions to intended actions.

We finally note several interesting properties of the BOID.

Relation by feedback. Different subprocesses and subdecisions are intimately related. As a first approximation, this is represented by the feedback loop in the architecture. Idea of architecture is that different types of agents can be represented by a single architecture, but with different control loops.

	BDI	QDT
area	software engineering	artificial intelligence
focus	application-oriented	theory-oriented
criticism	resource bounded	numbers unknown
intentions	yes	no
decision rules	no	yes
knowledge	yes	yes
desires	yes	yes
norms	possible extension	incorporated

Figure 5: Comparison

Types of agents. The same architecture can formalize different agent-types by formalizing different control of its components, which results in different ways in which the conflicts are resolved. This incorporates the distinction made in BDI between blind and open-minded agents, as well as the distinction between selfish and social agents (desires override obligations or vice versa).

Logical foundations. The formal process-oriented perspective is based on Reiter's default logic and its generalization in so-called input-output logic (Makinson & van der Torre 2000; 2001). This incorporates previous substudies in the creation and change of goals (van der Torre 1998a; 1999; 1998b) and in the logic of obligations (van der Torre & Tan 1999a).

Link. The logic and its architecture are relatively closely related.

Summary

In this paper we considered how the BDI and QDT approaches are related. We started with the observation that both are developed as criticisms on classical decision theory. In particular BDI introduced intentions, to make the decision-making process more stable. This may also be of value in QDT.

Both BDI and QDT have split the decision-making process into several parts. In BDI first desires are selected as goals, and second actions are selected and goals are committed to. In QDT first the input for the decision rule is constructed, and the decision rule is aplied. Here we noticed several ideas developed in QDT which can be used to enrich BDI: the use of different decision rules, the role of knowledge in decision making and the granularity of actions, and the construction process of goals from explicit (conditional) desires. The preliminary comparison of BDI and QDT can thus be summarized as follows.

We believe that a fruitful mutual influence and maybe even synthesis can be achieved. One way is to start with QDT and try to incorporate ingredients of BDI. Our BOID is one preliminary attempt, taking the process-oriented approach from Thomason's BDP and extending it with intentions. However, most elements of QDT discussed in this paper have not yet been incorporated in this rather abstract approach. Another approach is to extend BDI with the elements of QDT discussed here.

Finally, the comparison here has been preliminary, and we urge the interested reader to consult the original papers.

One direction for further study on the relation between BDI and QDT is the role of goals in these approaches. Another direction for further study is the role of norms. In BDI norms are usually understood as obligations from the society, inspired by work on social agents, social norms and social commitments (Castelfranchi 1998), whereas in decision theory and game theory norms are understood as reciprocal norms in evolutionary game theory (Axelrod 1984; Shoham & Tennenholtz 1997) that lead to cooperation in iterated prisoner's dilemmas and in general lead to an decrease in uncertainty and an increase in stability of a society.

Acknowledgments

Thanks to Jan Broersen and Zisheng Huang for many discussions on related subjects, and to the autonomous referees of this workshop for pointing us to relevant literature

References

Axelrod, R. 1984. *The evolution of cooperation*. Basic Books, New York.

Bacchus, F., and Grove, A. 1996. Utility independence in a qualitative decision theory. In *Proceedings of Fifth International Conference on Knowledge Representation and Reasoning (KR'96)*, 542–552. Morgan Kaufmann.

Boutilier, C.; Dean, T.; and Hanks, S. 1999. Decision-theoretic planning: Structural assumptions and computational leverage. *JAIR* 11:1–94.

Boutilier, C. 1994. Towards a logic for qualitative decision theory. In *Proceedings of the Fourth International Conference on Knowledge Representation and Reasoning (KR'94)*, 75–86. Morgan Kaufmann.

Bratman, M. E. 1987. *Intention, plans, and practical reason*. Harvard University Press, Cambridge Mass.

Broersen, J.; Dastani, M.; Huang, Z.; Hulstijn, J.; and van der Torre, L. 2001. The boid architecture. In *Proc. AGENTS'01*.

Castelfranchi, C. 1998. Modeling social agents. *Artificial Intelligence*.

Cohen, P., and Levesque, H. 1990. Intention is choice with commitment. *Artificial Intelligence* 42 (2-3):213–261.

Conte, R., and Castelfranchi, C. 1995. Understanding the effects of norms in social groups through simulation. In Gilbert, G., and Conte, R., eds., *Artificial societies: the computer simulation of social life*. London, UCL Press.

Dean, T., and Wellman, M. 1991. *Planning and Control*. San Mateo: Morgan Kaufmann.

Dignum, F. 1999. Autonomous agents with norms. *Artificial Intelligence and Law* 69–79.

Doyle, and Thomason. 1999. Background to qualitative decision theory. *AI magazine* 20:2:55–68.

Doyle, J., and Wellman, M. 1991. Preferential semantics for goals. In *Proceedings of the Tenth National Conference on Artificial Intelligence (AAAI'91)*, 698–703.

Doyle, J.; Shoham, Y.; and Wellman, M. 1991. The logic of relative desires. In *Sixth International Symposium on Methodologies for Intelligent Systems*.

Doyle, J. 1980. A model for deliberation, action and introspection. Technical Report AI-TR-581, MIT AI Laboratory.

Dubois, D., and Prade, H. 1995. Possibility theory as a basis for qualitative decision theory. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence (IJCAI'95)*, 1924–1930. Morgan Kaufmann.

Hansson, B. 1969. An analysis of some deontic logics. *Nous*.

Jeffrey, R. C. 1965. *The Logic of Decision*. McGraw-Hill, New York.

Jennings, N. R. 2000. On agent-based software engineering. *Artificial Intelligence* 117(2):277–296.

Keeney, R., and Raiffa, H. 1976. *Decisions with Multiple Objectives: Preferences and Value Trade-offs*. New York: John Wiley and Sons.

Lang, J.; van der Torre, L.; and Weydert, E. 2001a. Two types of conflicts between desires (and how to resolve them). In *This proceedings*.

Lang, J.; van der Torre, L.; and Weydert, E. 2001b. Utilitarian desires. *Autonomous Agents and Multi Agent Systems*. To appear.

Lang, J. 1996. Conditional desires and utilities - an alternative approach to qualitative decision theory. In *Proceedings of the Twelth European Conference on Artificial Intelligence (ECAI'96)*, 318–322. John Wiley and Sons.

Makinson, D., and van der Torre, L. 2000. Input-output logics. *Journal of Philosophical Logic* 29:383–403.

Makinson, D., and van der Torre, L. 2001. constraints for input-output logics. *Journal of Philosophical Logic*.

Pearl, J. 1990. systemz. In *Proceedings of the TARK'90*, 121–135. Morgan Kaufmann.

Pearl, J. 1993. From conditional ought to qualitative decision theory. In *Proceedings of the Ninth Conference on Uncertainty in Artificial Intelligence (UAI'93)*, 12–20. John Wiley and Sons.

Pinkas, G. 1995. Reasoning, nonmonotonicity and learning in connectionist network that capture propositional knowledge. *Artificial Intelligence* 77:203–247.

Rao, A., and Georgeff, M. 1991. . In *Proceedings of UAI'91*. Morgan Kaufmann.

Rao, A., and Georgeff, M. 1991. Modeling rational agents within a bdi architecture. In *Proceedings of Second International Conference on Knowledge Representation and Reasoning (KR'91)*, 473–484. Morgan Kaufmann.

Rao, A., and Georgeff, M. 1992. An abstract architecture for rational agents. In *Proceedings of Third International Conference on Knowledge Representation and Reasoning (KR'92)*, 439–449. Morgan Kaufmann.

Savage, L. J. 1954. *The foundations of statistics*. Wiley, New York.

Shoham, Y., and Tennenholtz, M. 1997. On the emergence of social conventions: modeling, analysis, and simulations. *Artificial Intelligence* 94(1-2):139–166.

Tan, S.-W., and Pearl, J. 1994a. Qualitative decision theory. In *Proceedings of the Thirteenth National Conference on Artificial Intelligence (AAAI'94)*.

Tan, S.-W., and Pearl, J. 1994b. Specification and evaluation of preferences under uncertainty. In *Proceedings of the Fourth International Conference on Knowledge Representation and Reasoning (KR'94)*, 530–539. Morgan Kaufmann.

Thomason, R., and Horty, R. 1996. Nondeterministic action and dominance: foundations for planning and qualitative decision. In *Proceedings of Theoretical Aspects of Reasoning about Knowledge (TARK'96)*, 229–250. Morgan Kaufmann.

Thomason, R. 2000. Desires and defaults: a framework for planning with inferred goals. In *Proceedings of Seventh International Conference on Knowledge Representation and Reasoning (KR'2000)*, 702–713. Morgan Kaufmann.

van der Torre, L., and Tan, Y. 1999a. Contrary-to-duty reasoning with preference-based dyadic obligations. *Annals of Mathematics and Artificial Intelligence* 27:49–78.

van der Torre, L., and Tan, Y. 1999b. Diagnosis and decision-making in normative reasoning. *Artificial Intelligence and Law* 51–67.

van der Torre, L., and Weydert, E. 2000. Parameters for utilitarian desires in a qualitative decision theory. *Applied Intelligence*.

van der Torre, L. 1998a. Labeled logics of goals. In *Proceedings of the Thirteenth European Conference on Artificial Intelligence (ECAI'98)*, 368–369. John Wiley and Sons.

van der Torre, L. 1998b. Phased labeled logics of conditional goals. In *Logics in Artificial Intelligence. Proceedings of the 6th European Workshop on Logics in AI (JELIA'98)*, 92–106. Springer, LNAI 1489.

van der Torre, L. 1999. Defeasible goals. In *Symbolic and Quantitative Approaches to Reasoning and Uncertainty. Proceedings of the ECSQARU'99*, 374–385. Springer, LNAI 1638.