Defeasible Preferences and Goal Derivation*

Craig Boutilier
Department of Computer Science
University of British Columbia
Vancouver, British Columbia
CANADA, V6T 1Z4
email: cebly@cs.ubc.ca

October 5, 1994

Abstract

We present a logic for representing and reasoning with qualitative statements of preference and normality and describe how these may interact in decision making under uncertainty. Our aim is to develop a logical calculus in which goals or objectives can be derived in defeasible settings. This system employs the basic elements of classical decision theory, namely probabilities, utilities and actions, but exploits qualitative information about these elements directly. Preferences and judgements of normality are captured in a modal/conditional logic called QDT, for which we present a semantics and sound, complete proof theory. A simple model of action is incorporated into QDT for the purpose of deciding appropriate courses of action. Without quantitative information, decision criteria other than maximum expected utility are pursued. We describe how techniques for conditional default reasoning can be used to complete information about both preferences and normality judgements, and we show how maximin and maximax strategies can be expressed in our logic. We also describe a qualitative analog of the notion of *value of information*.

Keywords: Goals, preferences, beliefs, defaults, decision theory, conditional logic, strategies, observations

*Some parts of this report appeared in preliminary form in "Toward a Logic for Qualitative Decision Theory," *Proc. of Fourth International Conf. on Principles of Knowledge Representation and Reasoning (KR-94)*, Bonn, pp.75–86 (1994).

1 Introduction

We typically expect a rational agent to behave in a manner that best furthers its own interests. However, an artificial agent might be expected to act in the best interests of a user (or designer) who has somehow communicated its wishes to the agent. In the usual approaches to planning in AI, a planning agent is provided with a description of some state of affairs, a *goal state*, and charged with the task of discovering (or performing) some sequence of actions to achieve that goal. This notion of goal can be found in the earliest work on planning (2) and persists in more recent work on intention and commitment (2). In most realistic settings, however, an agent will frequently encounter goals that it cannot achieve. As pointed out by Doyle and Wellman (2) an agent possessing only simple goal descriptions has no guidance for choosing an alternative goal state toward which it should strive.

Straightforward goal-driven behavior tends to be inflexible: an agent told to ensure that part A and part B are at location L by $5PM$ will be unable to do anything if it cannot locate B or if something prevents it from reaching L by 5PM. One might suppose that the agent should at least deliver A to L as close to 5PM as possible. While such partial fulfillment of deadline goals (2) undoubtedly arises frequently is practice, more general mechanisms will often be required. If A and B can't both be delivered, perhaps alternate parts C and D should be; or if the 5PM deadline can't be met, the agent should wait until next week. To this end, a recent trend in planning has been the incorporation of decision-theoretic methods for constructing optimal plans (2). Decision theory provides most of the basic concepts we need for rational decision making, in particular, the ability to specify arbitrary preferences over circumstances or outcomes. This allows desired outcomes or goals (and hence appropriate behaviors) to vary with context. For instance, if the most desirable outcome, say having both parts delivered by 5PM, is unachievable (or achievable with low probability) the appropriate course of action may be to strive for the less desirable "goal" of having different parts delivered.

Decision-theoretic methods ensure that proper goals and behaviors are derivable, accounting for the context-dependent nature of preferences, uncertainty of knowledge and factors that the decision maker is able or unable to control. This ranges from the classical "one step" decision making framework to multi-step optimal policy construction (2). The aim in all cases is to choose an action or course of behavior that has maximum expected utility (MEU). In other words, the goals of an agent can be derived, in a context-dependent fashion, from more basic (and robust) information.

Most decision-theoretic analysis is set within the framework of *maximum expected utility* (MEU). One impediment to the general use of such decision-theoretic tools is the requirement to have both numerical probabilities and utilities associated with the possible outcomes of actions. It is quite conceivable that such information is not readily available to the agent. We can often expect users to present information in a *qualitative* manner, including qualitative *preferences* over outcomes (one outcome or proposition is preferred to another) and qualitative *probabilities* (describing the relative likelihood of propositions or outcomes). The ability to reason *directly* with such qualitative constraints is therefore crucial. An appropriate knowledge representation scheme will allow the expression of constraints of this form and allow one to logically derive goals and reasonable courses of action, to the extent the given information allows.[1]

In this paper, we describe a logic and natural possible worlds semantics for representing and reasoning with qualitative probabilities and preferences, and suggest several reasoning strategies for qualitative decision making using this logic. We can represent *conditional preferences*, allowing (derived) goals to depend on context. Furthermore, these conditional preferences are *defeasible*: I might have a general preference for the proposition A (e.g., that parts be delivered to customers on time) but have a more specific "defeating" preference for $\neg A$ given B (e.g., a customer's account is past due). Semantically, preferences will be captured by an ordering over possible worlds, corresponding to an ordinal value function. The logic that captures such *default* preferences will exactly match existing conditional logics for default reasoning and belief revision (2; 2; 2). The component of the logic for capturing qualitative probabilities will be isomorphic, with a (separate) *normality* ordering on worlds representing their relative likelihood.

In order to strengthen possible conclusions, we will also present reasoning strategies for completing information about preferences and likelihoods, in essence, making assumptions about unstated constraints. To specify complete information about preferences can be an especially onerous task. This burden has been acknowledged in default reasoning, logic programming and reasoning about action. For instance, the *frame problem* (2) involves having a user specify the direct changes associated with an action without being forced to explicitly list things that do not change. In our setting, we want preferences for specific propositions to be specified naturally. However, since users will be indifferent to many propositions and outcomes, we want to ease the burden by assuming that the stated preferences are the *only* preferences held by the user. Our completion strategies do just this.

[1]While the foundations of decision theory are, in fact, based on such qualitative preferences (2; 2), the move to numerical utilities (and probabilities) requires that a preferences and likelihoods be calibrated by means of questions concerning acceptable exchanges between outcomes and lotteries. For an agent behaving according to the preferences of some user, this requires that either a) the user's preferences be so completely specified that such calculations can be made; or b) the user (or the source of preference information) be available to be queried about preference information as the need arises. Furthermore, a complete calibration of just the preference ranking, in the most fortunate circumstances, requires a number of queries at least as large as the number of possible worlds (exponential in the number of propositional atoms). Such a mechanism is also often criticized because the queries require answers to which a user does not have ready access or might be uncertain (2). While we do not question the need for such approaches — common in decision analysis and medical decision making domains — there will be many circumstances in which this information is inconvenient or impossible to obtain, or in which adequate decisions can be made without precise calibration of utilities and probabilities.

In addition, we describe several ways of making decisions with such completed information. These derivation strategies are motivated by the fact that the scales of normality and preference on which worlds are ranked are incomparable. This reflects the fact that user specified constraints provide qualitative information about the structure of the two rankings, not their relative magnitudes. We will discuss conditions under which decisions are reasonable in this framework.

We note that this paper deals primarily with single-step or "one-shot" decision making, rather than multistage, sequential problems. Goals are assumed to be reachable with a single action (which may of course be a compound "macro" action consisting of a sequence of primitive actions or the joint occurrence of several actions). While this is not a realistic assumption for many planning and decision problems, it does provide the foundations for extensions to qualitative multistage decision making, as we describe in the conclusion. We will also explain how this framework can be used to derive goals for use by a planning system, which plots a course of concrete actions.

In Section 2, we present the basic logic of preferences and its semantics, and show how existing techniques for conditional default reasoning can be used to make various assumptions about incomplete preference orderings. We briefly describe its relationship to *deontic logic*, developed for reasoning about permission and obligation. In Section 3, we add normality orderings to our semantics and describe a logic for dealing with both orderings. We describe the derivation of *ideal goal states*, roughly, the best situations an agent can hope for given certain fixed circumstances. This generalizes the usual notion of a goal in AI, for such goals are context-dependent and defeasible, and can be derived from more basic information rather than simply being asserted directly by a user. Such goals do not take into account the ability of an agent to change the fixed circumstances from which they are derived, nor the potential inability of an agent to achieve a goal. In Section 4, we explore a more realistic notion of goal that accounts for a simple form of ability. In planning, as in the decision theory, the ultimate aim is to derive appropriate actions to be performed that will achieve derived goal states. The ability of an agent to affect the world will have a tremendous impact on the *actual goal states* it attempts to achieve. One feature that becomes clear in our model is that, given incomplete knowledge, various behavioral *strategies* can emerge. In Section 5 we describe several strategies and show how these can be expressed in our logic. We also discuss *observations* and their role in improving decisions. Finally, in Section 6, we point out some related work, and on-going investigations into how the trade-offs between utility and probability can be captured in a qualitative manner. We also discuss the possible extension to multistage decision problems.

2 Conditional Preferences

A *goal* is typically taken to be some proposition that we desire an agent to make true. Semantically, a goal can be viewed as a set of possible worlds, those states of affairs that satisfy the goal proposition (2). Intuitively, if we ignore considerations of ability, the set of goal worlds should be those considered most desirable by an agent (or its designer). To achieve all goals is to ensure that the actual world lies within this desirable set.

Unfortunately, goals are not always achievable. My robot's goal to bring me coffee may be thwarted by a broken coffee maker. Robust behavior requires that the robot be aware of desirable alternatives ("If you can't bring me coffee, bring me tea"). Furthermore, goals may be defeated for reasons other than an inability to perform the actions that ensure them. It is often natural to specify general goals, but list exceptional circumstances that make the goal less desirable than the alternatives. For instance, a general preference for delivering parts within 24 hours may be overridden when the account is past due (which may in turn be overridden if the customer is important enough). To capture these ideas, we propose a generalization of standard goal semantics. Rather than a categorical distinction between desirable and undesirable situations, we will rank worlds according to their *degree of preference*. The most preferred worlds correspond to goal states in the classical sense. However, when such states are unreachable, a ranking on alternatives becomes necessary. Such a ranking can be viewed as an ordinal value function.

It is important to have a language in which such preferences can be specified. We typically expect an artificial agent to act in accordance with the desires of a user. We cannot expect a system designer to anticipate the needs of all users, thus an agent must be capable of accepting instructions — and often the most natural way to specify a goal is to ensure the agent is aware of the direct preferences that give rise to that goal, as well as "background" preferences that constrain how that goal is to be achieved (or perhaps modify the goal). Recent work in *software agents* for example has emphasized the important role of user preferences (2; 2). However, much of this work centers on learning user preferences. While certainly a vital component of intelligent systems, equally important is the ability to *tell* an agent what preferences should motivate and constrain its behavior. One typically does not want agents to go through hundreds or thousands of learning trials before having the ability to do anything useful.

The basic concept of interest will be the notion of *conditional preference*. We write $I(B|A)$, read "ideally B given A," to indicate that the truth of B is preferred, given A. This holds exactly when B is true at each of the most preferred of those worlds satisfying A. From a practical point of view, $I(B|A)$ means that if the agent (only) knows A, and the truth of A is fixed (or beyond its control), then the agent ought to ensure that B holds. Otherwise, should $\neg B$ come to pass, the agent will end up in a less than desirable A-world. The statement can be *roughly* interpreted as "If A, do B." Of course, this is a rough gloss as we will see. Often A is within the agent's power, and it may be better to "change A" than "do B." In addition, ensuring B is usually not sufficient to guarantee a best A-situation, for other preferences may also be applicable. We address these issues once we present the foundations for preference.

We propose a bimodal logic CO for conditional preferences using only unary modal operators. The presentation is brief. Further details can be found in (2; 2).

2.1 The Logic CO

We assume a propositional bimodal language $\mathbf{L}_B$ over a set of atomic propositional variables $\mathbf{P}$, with the usual classical connectives and two modal operators $\Box$ and $\overleftarrow{\Box}$. Our possible worlds semantics for preference is based on the class of *CO-models*, reflecting a preference ordering on situations.

Definition 2.1 A *CO-model* is a triple $M = \langle W, \leq, \varphi \rangle$, where: W is a set of possible worlds; valuation function $\varphi : \mathbf{P} \mapsto \mathbf{2}^W$ (where $\varphi(A)$ is the set of worlds at which A is true); and $\leq$ is a transitive connected binary relation on W.[2]

The relation $\leq$ is a total preorder over W, thus W consists of a set of $\leq$-equivalence classes or *clusters* of equally preferred worlds, with these clusters being totally ordered by $\leq$. We take $\leq$ to represent an ordering of preference: $v \leq w$ just in case v is at least as preferred as w.[3] This ordering is taken to reflect the desirability of situations, however this is to be interpreted (e.g., personal utility, moral acceptability, etc.). For our purposes, this ordering is assumed to reflect the preferences or utilities of some user, but below we describe other interpretations. We will speak of preferred situations as being more ideal or more acceptable than others. Figure 1 illustrates a typical CO-model.

We write $M \models_w A$ to indicate that A holds at world w, and denote by $\|A\|$ the set of such *A-worlds* (M is usually understood). We also us this notation for sets of sentences S, $\|S\|$ denoting those worlds satisfying each element of S. The satisfaction relation is straightforward for purely propositional sentences. The truth conditions for the modal connectives are

1. $M \models_w \Box\alpha$ iff for each v such that $v \leq w$, $M \models_v \alpha$.
2. $M \models_w \overleftarrow{\Box}\alpha$ iff for each v such that $w < v$, $M \models_v \alpha$.

[2]Relation $\leq$ is connected iff $w \leq v$ or $v \leq w$ for all v, w.

[3]While $w < v$ usually means v is a preferred outcome, the usual convention in AI is to "prefer" minimal models, hence we take $w < v$ to mean w is preferred.

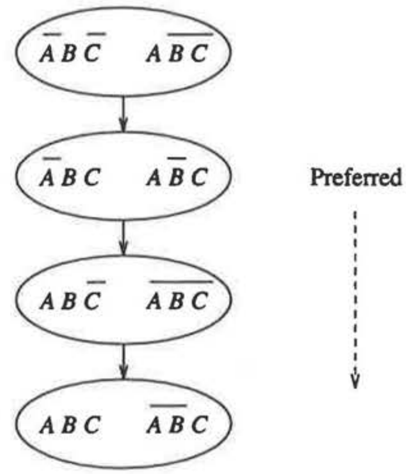


Figure 1: A CO-model

$\Box\alpha$ is true at a world w just in case α is true at all worlds at least as preferred as w, while $\overleftarrow{\Box}\alpha$ holds just when α holds at all less preferred worlds. The dual "possibility" connectives are defined as usual: $\Diamond\alpha \equiv_{\text{df}} \neg\Box\neg\alpha$ means α is true at some equally or more preferred world; and $\overleftarrow{\Diamond}\alpha \equiv_{\text{df}} \neg\overleftarrow{\Box}\neg\alpha$ means α is true at some less preferred world. $\overleftrightarrow{\Box}\alpha \equiv_{\text{df}} \Box\alpha \wedge \overleftarrow{\Box}\alpha$ and $\overleftrightarrow{\Diamond}\alpha \equiv_{\text{df}} \Diamond\alpha \vee \overleftarrow{\Diamond}\alpha$ mean α is true at all worlds and at some world, respectively. The logic CO is axiomatized in (2; 2) (see also Section 3).

2.2 Expressing Conditional Preferences

We now define a conditional connective $I(-|-)$ to express conditional preferences. $I(B|A)$ can be read as "In the most preferred situations where A holds, B holds as well," or "If A then ideally B." Intuitively, $I(B|A)$ should hold just when B holds at the most ideal A-worlds. These truth conditions can be expressed in $\mathbf{L}_B$ (see also (2; 2)):

$$I(B|A) \;\equiv_{\text{df}}\; \overleftrightarrow{\Box}\neg A \vee \overleftrightarrow{\Diamond}(A \wedge \Box(A \supset B)). \qquad (1)$$

Since nothing in our models forces the existence of minimal or most preferred A-worlds, the second clause of definition captures the condition that A holds at some world and that B holds at all A-worlds at least as preferred as that one. The first clause reflects the usual convention that the conditional holds vacuously when A is impossible. For ease of presentation, we will usually speak as if our models are well-founded, or that there exist minimal A-worlds for any proposition A. The set of minimal A-worlds, if it exists, is denoted $\min(A, \leq)$.

$I(B|A)$ can be thought of, as a first approximation, as expressing "If A then an agent ought to ensure that B," for unless B is true an agent cannot be in a best possible A-situation. We note that an absolute preference A, capturing the standard unconditional goal semantics, can be expressed as $I(A|\top)$, or equivalently, $\overleftrightarrow{\Diamond}\Box A$. We abbreviate this as $I(A)$ and read this as "ideally A". This can be read as expressing an unconditional desire for A to be true. Semantically, $I(A)$ ensures that A is true at all most preferred worlds $\min(\top, \leq)$. The model in Figure 1 satisfies $I(B|A)$ and $I(A \equiv B)$.

The dual of preference gives a notion of *toleration* or "don't care conditions." If $\neg I(\neg B|A)$ holds, then in the most preferred A-situations it is not required that $\neg B$; hence B is "tolerable" given A. We abbreviate this sentence $T(B|A)$. Loosely, we can think of this as asserting that an agent is *permitted* to do B if A. Unconditional toleration is denoted $T(A)$ and stands for $\neg I(\neg A)$, or equivalently, $\overleftrightarrow{\Box}\Diamond A$. In Figure 1, we have that $T(B|C)$, $T(\neg B|C)$ and $T(A)$.

The *relative* preference of two propositions can be also expressed directly in CO. We write $A \leq_P B$ to mean A is at least as preferred as B (intuitively, the best A-worlds are at least as good as the best B-worlds), and define it as:

$$A \leq_P B \equiv_{\mathrm{df}} \overleftrightarrow{\Box}(B \supset \Diamond A)$$

In our example, we have $C \leq_P \neg C$ since there is a world satisfying C that is preferable to any $\neg C$-world. Another useful notion is that of *strict preference*. If some proposition is more desirable than its negation no matter what other circumstances hold (e.g., deliveries to customer C *must* be on time), we can assert

$$\overleftrightarrow{\Box}(C \supset \Box C)$$

which ensures that every C-world is preferred to any $\neg C$-world. Of course, we cannot *a priori* abolish such strictly dispreferred situations, for they may occur due to events beyond an agent's control, and the relative preference of these strictly dispreferred $\neg C$ worlds is important. But in achieving stated goals condition $\neg C$ will be avoided if at all possible. In Figure 1, $A \equiv B$ is strictly preferred. Strict preferences can also be combined and prioritized (2).

The properties of the connective I are identical to those of the conditional connective $\Rightarrow$ defined in (2; 2) for default reasoning (see also Section 3). They are distinguished simply by their reading and the interpretation of the underlying ordering $\leq$. The crucial feature of the conditional connective is its defeasible and context-dependent nature. It is consistent to assert that one prefers coffee $I(C)$, but prefers tea if the coffee urn is empty and coffee will be delayed $I(\neg C|D)$. This conditional preference can in turn be defeated by a preference for coffee if it's late at night despite the wait, $I(C|D \wedge L)$.

As one should expect, absolute preferences, as well as preferences in any fixed context, must be consistent, for the following is a theorem of CO (for any possible A):

$$I(B|A) \supset \neg I(\neg B|A)$$

However, an agent's preferences needn't be complete, for $T(B|A) \wedge T(\neg B|A)$ is generally consistent. The property of *preferential detachment* holds in CO:

$$I(B|A) \wedge I(A) \supset I(B)$$

However, the principle of *factual detachment*

$$I(B|A) \wedge A \supset I(B)$$

is not valid. This has implications for the manner in which an agent should derive its actual preferences in a given situation, as we describe in the next section. Finally, we list a few other theorems associated with conditional preference:

RCM From $B \supset C$ infer $I(B|A) \supset I(C|A)$

LLE From $A \equiv B$ infer $I(C|A) \supset I(B|A)$

And $I(B|A) \wedge I(C|A) \supset I(B \wedge C|A)$

Or $I(C|A) \wedge I(C|B) \supset I(C|A \vee B)$

ID $I(A|A)$

RT $I(B|A) \supset (I(C|A \wedge B) \supset I(C|A))$

CM $I(B|A) \wedge I(C|A) \supset I(C|A \wedge B)$

RM $I(C|A) \wedge T(\neg C|A \wedge B) \supset I(\neg B|A)$

In the remainder of the paper, will take the models with which an agent evaluates its preferences to have two special properties, neither of which is crucial to the logic, but which enable some of our logical characterizations of goals to be expressed very compactly in our language. Intuitively, we expect the degree of preference associated with a world to depend only on the propositions that hold at that world. In other words, preference is a function of the *state of the world*. CO-models do not enforce this assumption — a model may contain two worlds w, v such that $\varphi(w) = \varphi(v)$, but one world is preferred to another. While such "duplicate" worlds have no bearing on the truth of conditional statements, they can influence compact characterizations of goals and they violate our intuitions. For the most part then we will consider only *functional models* satisfying the condition that

$$
\begin{array}{c}
\overline{D}\,R\,\overline{U} \\
\downarrow \\
D\,R\,\overline{U} \\
\downarrow \\
\overline{D}\,R\,U \\
\downarrow \\
\overline{D\,R\,U} \\
\\
\text{(a)}
\end{array}
\qquad\qquad
\begin{array}{c}
\overline{D}\,R\,\overline{U} \\
\downarrow \\
\overline{D}\,R\,U \\
\downarrow \\
D\,R\,\overline{U} \\
\downarrow \\
\overline{D\,R\,U} \\
\\
\text{(b)}
\end{array}
$$

Figure 2: Possible Interpretations of Preferences

for any two (distinct) worlds w, v, $\varphi(w) \neq \varphi(v)$. A second assumption we will make from time to time is that our language is *logically finite*, that is, the set of variables $\mathbf{P}$ upon which our language is based is finite. This ensures that there are a finite number of distinct states of affairs and, together with functionality, ensures our models are finite. This assumption is used primarily to allow deductively closed sets of beliefs and goals to be expressed in the language, but is not crucial. For logically finite, functional models, the truth conditions of a conditional preference can be reformulated as

$$M \models_w I(B|A) \text{ iff } \min(A, \leq) \subseteq \|B\|$$

2.3 Varieties of Preference

As noted above, the most important feature is that preferences are conditional and can vary with context. I can consistently assert $I(U|R)$ and $I(\overline{U}|\overline{R})$, that my agent should take an umbrella if it's raining, and leave it home if not. The potential goals U or $\neg U$ depend on context and need not be asserted categorically. Furthermore, these conditionals are *defeasible*: I can consistently assert that $I(\overline{U})$ without fear of contradicting $I(U|R)$. Notice that these two statements allow the conclusion $I(\overline{R})$ to be drawn — the agent can derive its (or *a user's*) preference for sunny weather.

This defeasibility also allows one to assert, together with the previous conditionals, $I(\overline{U}|R \wedge D)$, that an umbrella is not desired if I drive a car to work (D) instead of walking ($\neg D$) in the rain. Such a theory induces a partial structure like that illustrated in Figure 2(a). However, just as above, this entails $I(\neg D|R)$, that ideally I want to walk to work in the rain. Is this conclusion truly intended? On the surface, it seems reasonable to accept all three preference statements, but allow the assertion that I prefer to drive when it's raining. Yet $I(D|R)$ contradicts these other premises.

Suppose we actually intend $I(D|R)$. The (intuitive) source of the inconsistency is the statement $I(U|R)$. If I prefer to drive when it's raining, and prefer not to have an umbrella when I drive, it should be clear that the most preferred R worlds are those in which D and hence $\neg U$ hold. Therefore, I should not assert that at the most ideal R-worlds, U holds. Intuitively, the preference for U given R only holds when I do not drive; thus, $I(U|R \wedge \neg D)$ is a reasonable assertion, but $I(U|R)$ is not. Figure 2(b) shows a model capturing this intention. Figure 2(a), which validates $I(U|R)$, is appropriate when $I(\neg D|R)$ is intended, when I prefer walking to driving, even when it's raining.

We notice, however, that the assertion $I(U|R)$, I prefer an umbrella when it's raining, seems (potentially) appropriate even when Figure 2(b) is the intended model. Although, I might *like* to drive to work when it rains, the utterance "I want my umbrella if it's raining" does not strike one as inconsistent. This can be explained by considering the *variety* of preference statements one might make. In this case, my wish for an umbrella might reflect the fact that I am usually unable to drive to work. Even though I prefer to drive, I probably won't be able to, so my stated preference for U given R might reflect this fact. In this case, the *typical* R-world is one in which $\neg D$ holds, and hence one in which U should hold: my robot *should* bring an umbrella along. Very often stated preferences do not express *ideal* preferences. Rather, they may incorporate into the stated context (here, R) certain assumptions or default conclusions (such as $\neg D$), and thus express a preference conditioned on this extended context ($R \wedge \neg D$). The intended assertion $I(U|R \wedge \neg D)$ is perfectly consistent with Figure 2(b), but it may be abbreviated as $I(U|R)$ if the default conclusion $\neg D$ is understood. It is therefore crucial to realize that linguistically stated preferences can be come in different varieties:

- Conditional *ideal preferences* are expressed in the usual way. A statement $I(D|R)$ ensures that D is true in the best possible R-worlds. These have the semantics described above and are defeasible.
- *Strict preference* for a proposition C is expressed as above, $\overleftarrow{\Box}(C \supset \Box C)$. This expressed a desire for C "at all costs."
- *Expected preferences* express a desire for a proposition in a given context anticipating the expected consequences of that context. The preference for U given R above takes into account the likely truth of the (dispreferred) proposition $\neg D$ in the explicitly stated context R. The preference for U is based on this extended context. To analyze such statements, we require the additional machinery introduced in Sections 3 and 4.[4]

[4]Similarly, one can impose this alternate interpretation on direct statements of preference $A <_P B$, as Jeffrey (2) does.

2.4 Defeasible Reasoning with Preferences

The conditional logic of preferences we have proposed above is similar to the (purely semantic) proposal put forth by Hansson (2) for *deontic reasoning*, or reasoning about obligation and permission. In our logic, one may simply think of $I(B|A)$ as expressing a conditional obligation to see to it that B holds if A does.[5] Loewer and Belzer (2) have criticized this semantics "since it does not contain the resources to express actual obligations and no way of inferring actual obligations from conditional ones." In particular, they argue that any deontic logic should validate something like factual detachment, not just deontic detachment (the deontic analog of preferential detachment). The criticism applies equally well to our preference logic — one cannot logically derive actual preferences — because the principle of factual detachment does not hold. Factual detachment expresses the idea that if there is a conditional preference for B given A, and A is *actually* the case, then there is an actual preference for B. While the inference is a reasonable one, we do not expect, nor do we want it to hold logically because it threatens the natural defeasibility of our conditionals. For instance, if R and $I(U|R)$ entailed U or $I(U)$, so too would R, D, $I(U|R)$ and $I(\overline{U}|R \wedge D)$. Defeasible conditional preferences could not be expressed.

Various logics have been proposed to capture factual detachment in the deontic setting, and recently several complex default reasoning schemes have been applied to this problem (2; 2). We propose a simple solution based on the following observation: to determine preferences based on certain actual facts (propositions), we consider only the *most ideal* worlds satisfying those facts, rather than *all* worlds satisfying those facts. Let *KB* be a knowledge base containing statements of conditional preference and propositions. Given that such facts actually obtain, the ideal situations are those most preferred worlds satisfying *KB*. For instance, although $R \wedge I(U|R) \not\vdash U$, we are guaranteed that U holds at the most preferred R worlds in any model of these premises. This suggests a straightforward mechanism for determining actual preferences. We simply ask for those propositions α such that

$$\vdash_{CO} I(\alpha|KB)$$

This is precisely the preliminary scheme for conditional default reasoning suggested in (2). This mechanism unfortunately has a serious drawback: seemingly *irrelevant* factual information, or information about the consequences of actions, can paralyze the derivation of actual preferences.

Example 2.1 Let P denote that a certain part is painted, B that it's blemished, and S that it's destined

On our definition, $A <_P B$ means the *best* A-worlds are preferred, whereas Jeffrey defines such a statement to mean the expected utility of *all* A-worlds is greater than that for B.

[5]We elaborate on this connection below.

for shipment to a specific warehouse. Let D, E and F denote possible locations for a certain piece of equipment. If

$$KB = \{I(P|\overline{B}), \overline{B}\}$$

then the actual preference P is derivable using the scheme suggested above; that is, $\vdash I(P|KB)$. However, it is not derivable from $KB' = KB \cup \{S\}$. Because conditionals are defeasible, it is consistent with KB' to assert $I(\overline{P}|\overline{B} \wedge S)$. Although intuitively S is irrelevant to the preference for P, this irrelevance is not logically derivable.

Again consider KB with actual preference P. Suppose a painting action that achieves P requires the equipment in question to be moved, making either D, E or F true. Even though not stated, one can consistently assert $I(P \wedge \overline{D}|\overline{B})$, $I(P \wedge \overline{E}|\overline{B})$ or $I(P \wedge \overline{F}|\overline{B})$. Thus the agent cannot show that any of the moves D, E or F is tolerated — it cannot decide what to do. ∎

In this example, the fact that $I(P|\overline{B})$ is the only stated preference suggests that other factors are irrelevant to the relative preference of situations. Intuitively, these factors should be discounted. Unless stated otherwise, the part should be painted regardless of its destination (S); and the manner in which P is achieved (by moving either D, E or F) is not of concern.

One possible way to deal with this difficulty is to make certain assumptions about the preference ordering. In particular, it is possible to adopt the default reasoning scheme System Z (2) in this context. Given a set of conditional constraints, System Z enforces the assumption that worlds are assumed to be as preferred as possible consistent with these constraints. In other words, worlds are pushed down as far as possible in the preference ordering, "gravitating" toward absolute preference. In our example, the model induced by this assumption is shown in Figure 3. (For convenience we assume that $I(\overline{P}|B)$ holds and that propositions D, E and F are mutually exclusive.) In this model, *any* $\neg B$-world that satisfies P is deemed acceptable, regardless of the truth of the irrelevant factors. The technical details of System Z may be found in (2) and are summarized in Appendix B. In (2) we describe how the Z-model for any conditional theory can be axiomatized in CO. Intuitively, each world is assumed to be as acceptable as possible consistent with the constraints of the theory. The important features of this model are: a) the assumption induces a unique, "most compact" preference ordering; and b) the consequences associated with these assumptions can sometimes be efficiently computed (2; 2).

Is the assumption that worlds are preferred unless stated otherwise reasonable? Tan and Pearl (2) argue that worlds should gravitate toward "indifference" rather than preference. We cannot, of course, make sense of such a suggestion in our framework, since we do not have a bipolar scale

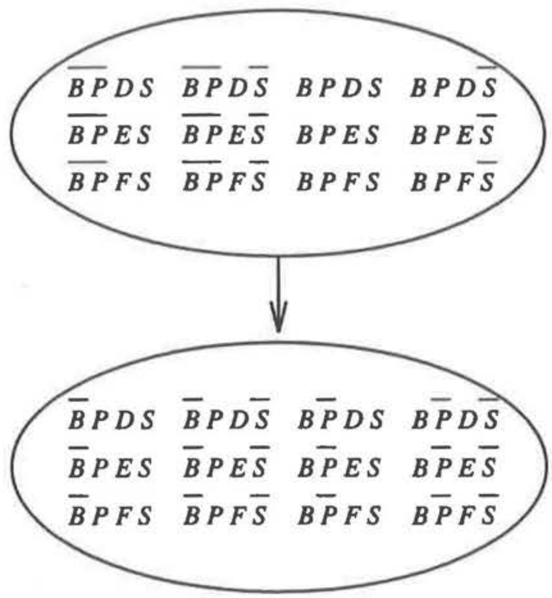


Figure 3: The Compact Preference Ordering

(where outcomes can be good, bad or neutral).[6] However, even if an "assumption of indifference" were technically feasible, we claim that the "assumption of preference" is the the right one in our setting.

Recall that we wish to use preferences to determine the set of goal states for a given context C. These are simply the most preferred C-worlds according to our ranking; call this set *Pref*(C). If the agent brings about *any* of these situations, it will have behaved correctly. A conditional preference $I(A|C)$ constrains the set *Pref*(C) to contain only A-worlds. Thus an agent will attempt to bring about some $A \wedge C$-world when C holds. But which $A \wedge C$-world is the right one? With no further information, System Z will set *Pref*$(C) = \|A \wedge C\|$; all $A \wedge C$-worlds will be assumed to be equally acceptable. This seems to be appropriate: with no further information, any course of action that makes A true should be judged to be as good as any other. Any other assumption, such as gravitation of worlds toward indifference, must make the set *Pref*(C) smaller than $\|A \wedge C\|$. For example, if we rule out worlds satisfying α from *Pref*(C), then *Pref*$(C) = \|A \wedge C \wedge \neg\alpha\|$. This requires that an agent striving for *Pref*(C) make $\neg\alpha$ true as well as A. This imposes unnecessary and unjustified restrictions

[6]Note that in classical decision theory, such distinctions do not exist. An outcome cannot be good or bad, nor can an agent be indifferent toward an outcome, in isolation; it can only be judged *relative* to other outcomes. An agent can adopt an attitude of indifference toward a *proposition*, as we explain below.

on the agent's goals, or on the manner in which it decides to achieve them. In our example, gravitation toward preference ensures that an agent is *not biased against* any of the alternatives D, E or F.

Notice that when *worlds* gravitate toward preference, our agent becomes indifferent toward most *propositions*. By maximizing the size of $Pref(C)$ (subject to the constraint that A be true), we minimize the number of propositions an agent will care about or attempt to make true in context C. In our example, if $A \wedge C \not\vdash \alpha$ and $A \wedge C \not\vdash \neg\alpha$, then $T(\alpha|C)$ and $T(\neg\alpha|C)$ will both be true in the Z-model. Again, in our example, one can see that System Z forces the agent to be indifferent regarding the propositions D, E and F, since no preference for any one was specified. Such indifference toward propositions in a given context seems to be the most appropriate assumption.

In (2; 2) we characterize System Z, in a default reasoning context, as embodying the principle of *conditional only knowing*. When certain beliefs are stated, either actual or conditional, System Z ensures that only propositions that can be shown to be believed (in a given context) are actually believed. We show this to be a generalization of the notion of *only knowing* often adopted in belief logics (2) that accounts for defeasible beliefs. In the preference setting, System Z captures the analogous assumption of "only preferring." Those preferences that can be derived in a given context C are assumed to be the *only* propositions the agent prefers or cares about in that context.

Certain problems with System Z have been shown to arise in default reasoning. These problems occur when reasoning about preferences as well. For example, if we have two independent (in this case, absolute) preferences $I(A)$ and $I(B)$, System Z will sanction both $T(A|\neg B)$ and $T(\neg A|\neg B)$; once the preference for B has been violated, one cannot ensure that A is still preferred. Various modifications to System Z have been proposed to deal with such problems, for instance, the "rule counting" systems of (2; 2). Such solutions can be applied in this setting as well, but the assumption of "only preferring" lies at the heart of these solutions as well.

We should point out that, while our presentation will assume a unique preference ordering, the definitions to follow do not require this assumption. We are typically given a set of conditional premises of the form $I(B|A)$, plus other modal sentences constraining the ordering. Unless these premises form a "complete" theory, there will be a space of permissible orderings. A defeasible reasoning scheme such as System Z can be used to complete this ordering, but we do not *require* the use of a single ordering — the definitions presented below can be re-interpreted to capture truth in all permissible orderings. One simply needs to derive the consequences of the premises in the logic QDT, the extension of CO presented in the following section.

2.5 Deontic Logic

Deontic logics have been proposed to model the concepts of obligation and permission (2; 2; 2; 2). While originally proposed using a unary modal connective, it has been widely recognized that this modal formulation has certain limitations. In particular, it is difficult to represent *conditional obligations* (2). For this reason, conditional deontic logic (CDL) has been introduced to represent the dependence of obligations on context (2). Obligation is then represented by a two-place conditional connective. The sentence $O(B|A)$ is interpreted as "It ought to be that B given A" or "If A then it is obligatory that B," and indicates a conditional obligation to do B in circumstances A. These logics can be interpreted semantically using an ordering on worlds that ranks them according to some notion of preference or ideality (2; 2). Such a ranking satisfies $O(B|A)$ just in case B is true at all most preferred of those worlds satisfying A. Thus, we can think of B as a *conditional preference* given A. Once A is true, the best an agent can do is B.

CO is presented semantically in much the same manner as the semantic system of Hansson (2). While Hansson does not present an axiomatization of his system DSDL3, we can show that CO provides a sound and complete proof theory for his semantics (2). We note also that our system is equivalent to Lewis's conditional deontic logic VTA (2). In particular, given a deontic interpretation, the connectives I and T capture notions of conditional obligation and permission respectively.

As described above, this conditional semantics for deontic statements has been criticized because it fails to validate the principle of factual detachment. However, as we have argued above for preferences, factual detachment is not required for the derivation of *actual* obligations. All that is required is the application of an appropriate defeasible reasoning scheme. Using CO we can express the conditional preferences involved in a number of classic deontic puzzles such as Chisholm's paradox and contrary-to-duty imperatives (2), and our defeasible reasoning scheme allows appropriate obligations to be deduced. For example, we can represent the preferences involved in the following account (2):

(a) It ought to be that Arabella buys a train ticket to vist her grandmother.

(b) It ought to be that if Arabella buys the ticket she calls to tell her she is coming.

(c) If Arabella does not buy the ticket, it ought to be that she does not call.

(d) Arabella does not buy the ticket.

We represent these sentences as $I(V)$, $I(C|V)$, $I(\neg C|\neg V)$ and $\neg V$. These give rise to no inconsistency in CO, and induce a natural ordering on worlds where only $V \wedge C$-worlds are most acceptable. Less preferred are $\neg V \wedge \neg C$-worlds, and still less preferred is any $\neg V \wedge C$-world. Notice that from this set we can derive $I(C)$, Arabella should (ideally) call; however, given the actual facts of the case (say

KB), we can actually infer that $I(\neg C|KB)$, Arabella should not call given the actual circumstances.

The preference logic described above should not be viewed simply as a means for representing the preferences (or personal utilities) of an agent. It can be used to represent any measure or ranking of desirability of any type, such as those reflecting moral or ethical norms, legal codes, or some combination. Thus, our proposal can be used more generally in these settings. Furthermore, the way in which we derive goals below, in particular, our use of default information and ability, also applies to the derivation of obligations in the deontic sense.

3 Default Knowledge

We should not require that goals be based only on "certain" beliefs in *KB*, but on reasonable default conclusions as well. Consider the following preference ordering with atoms R (it will rain), U (have umbrella) and C (it's cloudy). Assuming $\overline{C} \wedge R$ is impossible, we have:

$$\{\overline{C}\overline{R}\overline{U}, C\overline{R}\overline{U}\} < CRU < \{\overline{C}\overline{R}U, C\overline{R}U\} < CR\overline{U}$$

Suppose, furthermore, that it usually rains when it's cloudy. If $KB = \{C\}$, according to our notion of actual preference in the last section, the agent prefers $\overline{R}$ and $\overline{U}$ — in the best *KB*-world it doesn't rain despite the clouds. However, we cannot use factual preferences (given *KB*) directly to determine goals. Ideally, the agent would like to ensure that it doesn't rain and that it doesn't bring its umbrella. However, clearly the agent can do nothing to make sure $\overline{R}$ holds (we return to this in the next section). Given this, the "goal" $\overline{U}$ seems to be wrong. Once C is known, the agent should *expect* R and act accordingly.

As in decision theory, actions should be based not just on preferences (utilities), but also on the likelihood (probability) of outcomes. In order to capture this intuition in a qualitative setting, we propose a logic that has two orderings, one for preferences and one representing the degree of *normality* or *expectation* associated with a world.

3.1 The Logic QDT

The logic QDT, a step toward a *qualitative decision theory*, is characterized by the class of *QDT-models*.

Definition 3.1 A *QDT-model* is a quadruple $M = \langle W, \leq_P, \leq_N, \varphi \rangle$ where: W is a set of possible worlds; φ is a valuation function for W, $\mathbf{P}$; and $\leq_P$ and $\leq_N$ are each transitive, connected relations over W.

The ordering $\leq_P$ is a *preference ordering* on W of the type already discussed, while $\leq_N$ is a *normality ordering* on W. We take $w \leq_N v$ to mean w is at least as normal a situation as v (or is at least as expected). The submodels formed by restricting attention to either relation are clearly CO-models. The language of QDT contains four modal operators: $\Box_P$, $\overleftarrow{\Box}_P$ are given the usual truth conditions over $\leq_P$; and $\Box_N$, $\overleftarrow{\Box}_N$ are interpreted similarly using $\leq_N$. The conditional $I(B|A)$ is defined as previously, using the connectives $\Box_P$ and $\overleftarrow{\Box}_P$. A new *normative conditional* connective $\Rightarrow$ is defined in exactly the same fashion using $\Box_N$ and $\overleftarrow{\Box}_N$:

$$A \Rightarrow B \equiv_{\text{df}} \overleftrightarrow{\Box}_N \neg A \vee \overleftrightarrow{\Diamond}_N (A \wedge \Box_N (A \supset B)) \tag{2}$$

The sentence $A \Rightarrow B$ means B is true at the most normal A-worlds, and can be viewed as a default rule. This conditional is exactly that defined in (2; 2), and the associated logic is equivalent to a number of other systems (e.g., the qualitative probabilistic logic of (2; 2)). QDT can be axiomatized using the following axioms and inference rules for both the preference operators $\Box_P$, $\overleftarrow{\Box}_P$ and the normality operators $\Box_N$, $\overleftarrow{\Box}_N$:

K $\Box(A \supset B) \supset (\Box A \supset \Box B)$

K′ $\overleftarrow{\Box}(A \supset B) \supset (\overleftarrow{\Box} A \supset \overleftarrow{\Box} B)$

T $\Box A \supset A$

4 $\Box A \supset \Box\Box A$

S $A \supset \overleftarrow{\Box}\Diamond A$

H $\overleftrightarrow{\Diamond}(\Box A \wedge \overleftarrow{\Box} B) \supset \overleftrightarrow{\Box}(A \vee B)$

Nec From A infer $\overleftrightarrow{\Box} A$.

MP From $A \supset B$ and A infer B

We require the following axiom to capture their interaction:

PN $\overleftrightarrow{\Box}_N A \equiv \overleftrightarrow{\Box}_P A$

Theorem 3.1 *The logic QDT is sound and complete with respect to the class of QDT-models.*

A proof of completeness can be found in Appendix A.

3.2 Default Conclusions and Goal Derivation

Given a set of premises consisting statements of preference and normality as well as statements of actual fact, an agent is charged with the task of deciding what goals to achieve. For ease of presentation, we will assume that such a unique QDT-model has been determined reflecting an agent's (or user's) judgements of preference and expectation. This will seldom be true due to logical considerations alone, but methods such as System Z can be used to make the tacit assumptions about indifference and normality. What remains is to characterize exactly how the (propositional) statements of fact constrain the selection of goals.

Given a QDT-model M and a (finite) set of facts *KB*, we want goals to be based not just on *KB*, but also on the *expected* state of affairs induced by *KB*. We define the *default closure* of *KB* to be (where $\mathbf{L}_{CPL}$ is our propositional sublanguage)

$$Cl(KB) = \{\alpha \in \mathbf{L}_{CPL} : M \models KB \Rightarrow \alpha\}$$

As with preferences, the actual expectations of an agent are determined by considering only the most expected or normal worlds satisfying the given facts.[7] In other words, those propositions α that are normally true given *KB* form the agent's set of default conclusions. We assume (for simplicity of presentation) that $Cl(KB)$ is finitely specifiable and take it to be a single propositional sentence.[8]

Now we can assert that an agent's goals should be based on its expectations or default beliefs $Cl(KB)$, not *KB* alone. As a first approximation of a definition of goal, we define *ideal goals*:

Definition 3.2 Let M be a QDT-model M and $KB \subseteq \mathbf{L}_{CPL}$. A sentence $\alpha \in \mathbf{L}_{CPL}$ is an *ideal goal* (with respect to M, *KB*) iff

$$M \models I(\alpha | Cl(KB))$$

The *ideal goal set* (given M, *KB*) is the set of all such α.

Intuitively, the ideal goals are those sentences that must be true if the agent is to find itself in a best possible situation satisfying $Cl(KB)$. In our previous example, where $KB = \{C\}$, we have that $Cl(KB) \equiv C \wedge R$: if $Cloudy$ is known, $Rain$ is expected. Since $I(U|C \wedge R)$ is satisfied by the ordering above, the agent's ideal goals are exactly those sentences entailed by $C \wedge R \wedge U$; in particular, one of the agent's goals is to have an umbrella.

It should be clear that ideal goals are *conditional* and *defeasible*. For instance, if $KB = \{C\}$ is extended to include $\overline{R}$, the default conclusion R is no longer forthcoming and the agent has a different

[7]Once a unique model has been determined, statements of expectation and preference can be removed from *KB* without consequence. The fact that $KB \Rightarrow \alpha$ holds is unaltered — for a fixed model — if *KB* is restricted to its propositional component (similarly for $I(\alpha|KB)$).

[8]A sufficient condition for this property is that each "cluster" of equally normal worlds in $\leq_N$ corresponds to a finitely specifiable theory. This is the case in, e.g., System Z (2). The definitions below can easily be generalized semantically — though not syntactically — to deal with infinite sets of sentences.

ideal goal, namely $\overline{U}$. This defeasibility arises in two ways. First, addition of facts to *KB* (to form an extended knowledge base *KB'*) can cause certain default conclusions to become invalid, as when knowledge of $\overline{R}$ precludes conclusion of R. This causes the set of derived preferences to be based on a different context, in that $Cl(KB) \not\subseteq Cl(KB')$. This simply reflects the defeasibility of default inference. However, even if $Cl(KB) \subseteq Cl(KB')$ ideal goals can change due to the defeasibility of conditional preferences, as described in Section 2.

This formulation does not provide any indication as to what an agent should *do* in order to achieve these ideal goals. This will require the introduction of actions and ability (see the next section). For instance, notice that an ideal goal set is always deductively closed. We should not expect an agent to have to consider each member of this set individually or have an infinite set of "goals" in any practical sense. The notion of a *sufficient condition* for achieving all ideal goals can be defined in QDT and will prove useful later. The following definitions are evaluated with respect to some fixed QDT-model $M = \langle W, \leq_P, \leq_N, \varphi \rangle$.

Definition 3.3 Let X be some proposition. C is a *sufficient condition* given X iff $C \wedge X$ is satisfiable and $M \models \overleftarrow{\Box}_P(X \supset \overleftarrow{\Box}_P(X \supset \neg C))$.

Intuitively, a sufficient condition C guarantees that an agent is in some best possible X-world. Thus, if X is some fixed, unchangeable context, ensuring proposition C means the agent has done the best it could:

Proposition 3.2 *Let C be a sufficient condition given X and let $M \models_w C \wedge X$. Then $v <_P w$ only if $M \not\models_v X$.*

Proof By definition of sufficiency, $M \models_v X \supset \overleftarrow{\Box}_P(X \supset \neg C)$. Since $v <_P w$ and $M \models_w C \wedge X$, $M \models_v \overleftarrow{\Diamond}_P(C \wedge X)$; thus, $M \models_v \neg X$. ■

With respect to *Cl*(*KB*), ideal goals are necessary conditions for ensuring an agent lies in some best situation consistent with the set of (default) beliefs *Cl*(*KB*). A sufficient condition C for *Cl*(*KB*) guarantees the entire ideal goal set is satisfied.[9] In our example, U is a sufficient condition for $Cl(\{C\})$. This means the (infinite) set of ideal goals that must be made true, including such things as $C, R, U, C \supset U$, and $U \vee P_i$ for arbitrary propositions P_i, can be achieved simply by making U true.

Proposition 3.3 *C is sufficient for $Cl(KB)$ iff $M \models C \wedge Cl(KB) \supset \alpha$ for all ideal goals α.*

[9]Hector Levesque (personal communication) has suggested that sufficiency is the crucial "operator."

Proof That sufficient conditions guarantee ideal goals (only if) holds without qualification. The converse (if) requires the assumption of logical finiteness and functionality.

If C is a sufficient condition for $Cl(KB)$ then $M \models Cl(KB) \supset \overleftarrow{\Box}_P(Cl(KB) \supset \neg C)$. So if $M \models_w Cl(KB)$ then for all $w <_P v$, $M \models_v \neg(C \wedge Cl(KB))$ Thus $\|C \wedge Cl(KB)\| \subseteq \min(Cl(KB), \leq_P)$. Since $\min(Cl(KB), \leq_P) \subseteq \|\alpha\|$ for any ideal goal α, we have $M \models C \wedge Cl(KB) \supset \alpha$.

If $M \models C \wedge Cl(KB) \supset \alpha$ for each ideal goal α, then $\|C \wedge Cl(KB)\| \subseteq \min(Cl(KB), \leq_P)$ (assuming functionality and finiteness of the model). Thus $M \models Cl(KB) \supset \overleftarrow{\Box}_P(Cl(KB) \supset \neg C)$ and C is a sufficient condition. ■

We will explore a more detailed example in the next section. Notice also that the definition of ideal goals gives a certain "priority" to defaults over preferences. The belief set $Cl(KB)$ is constructed before the preference ranking is consulted. This stands in contrast with the expected utility framework of classical decision theory, where likelihood and utility are traded against one another. We explore the implications of this scheme in Section 4.

4 Ability in Goal Derivation

The definition of an ideal goal embodies the idea that an agent should attempt to achieve the best possible situation consistent with what it knows, as well as what it conjectures by default. However, as we have emphasized, this is suitable only when the agents beliefs *KB* and expectations are fixed and unchangeable. If the agent can change the truth of certain elements in *KB*, ideal goals may be too restrictive. For example, it may be that I have not arranged to be driven to work and it's raining. An ideal goal is to ensure that I have an umbrella. However, if it is within my power to arrange a drive — and I prefer that to walking in the rain — then this ideal goal is inappropriate. Thus, some notion of action and ability must come into play in goal derivation.

Actions must also play a role if we are to derive what an agent should *do*, rather than simply what it should *achieve*. Indeed, the term "goal" is often interpreted in this way, to denote the actions appropriate in a given context (as in "a goal to do A"). This becomes especially important when we notice that the set of propositions an agent should achieve will always be deductively closed. The ultimate aim of goal derivation is not the characterization of an infinite set of desirable propositions, but rather the discovery of a rather small set (or sequence) of actions that will achieve this end. Finally, actions must play a role in factoring out unachievable desires. For instance, an agent might prefer that it not rain; but this is something over which it has no control. Though it is an ideal outcome, to call this a goal is unreasonable.

These considerations actually point to something of a paradox in the roles of planning and goal derivation, and their interaction. If an agent is to plan (in the classical goal-directed sense), it must know what the goal is, what desirable states exist; and as we see the propositions over which an agent is able to exert control influence this choice of goal (for instance, if I *can* arrange a drive to work, my goal should be to actually do so). However, knowing whether it has control over a proposition might require that an agent actually attempt to derive a plan to make that proposition true; and if successful, the resulting plan may in turn affect other believed propositions. Ultimately, it appears that goal derivation and planning depend in a circular way on one another.

It seems unlikely that, prior to planning for a specific goal, an agent will have attempted to construct a plan for every proposition in every conceivable context, simply to determine its abilities. If that were the case, these plans could simply be stored and retrieved in a reactive architecture. We take ability to be a slightly higher level notion, one that exists prior to planning.

4.1 Controllable Propositions

To capture distinctions of this sort, we introduce a simple model of action and ability and demonstrate its influence on conditional goals. We ignore the complexities required to deal with the specification of effects, preconditions and such, in order to focus attention on the structure and interaction of ability and goals. We partition our atomic propositions into two classes: $\mathbf{P} = \mathcal{C} \cup \overline{\mathcal{C}}$. Those atoms $A \in \mathcal{C}$ are *controllable*, atoms over which the agent has direct influence. For the purposes of goal derivation, we take the only actions available to the agent to be of the form $do(A)$ and $do(\overline{A})$, which make A true or false, for every $A \in \mathcal{C}$. To keep the treatment simple, we assume actions have no effects other than to change the truth of A. In particular, all other atoms in $\mathbf{P}$ retain their truth values. The atom U (have umbrella) is an example of a controllable atom. Atoms in $\overline{\mathcal{C}}$ are *uncontrollable*, for example, R (it will rain).

An abstract action $do(A)$ should be viewed as expressing the fact that an agent has some idea that it can construct a plan to achieve a (high-level) proposition A without interfering with other such propositions. This plan may involve changing more detailed propositions, but should not affect the abstract facts involved in goal derivation. This plan construction may not be explicit — it may involve retrieval of a plan for A from an existing plan library. Indeed, such a plan may *structurally partial* (2), and the plan for achieving A may involve imply filling in certain details, or making low-level decisions regarding the actual means to achieve the end A.[10] A commitment to $do(A)$ entails a

[10]Leaving open several possibilities for achieving A may have a certain advantages, for example, allowing the *overloading of intentions* (2).

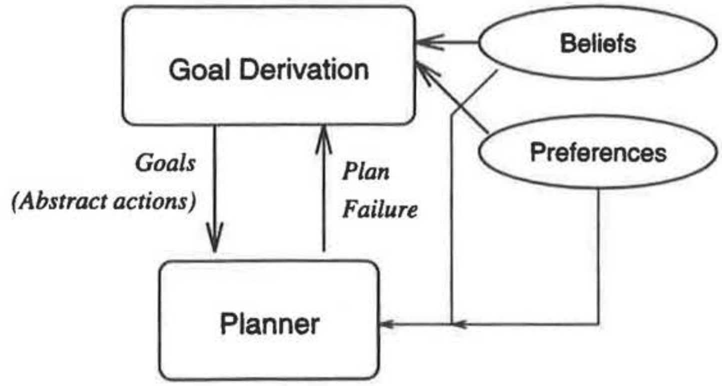


Figure 4: The Interaction of Planning and Goal Derivation

commitment to construct or instantiate a feasible plan to bring A about; and treating A as controllable entails a belief that such a commitment can (typically) be fulfilled in conjunction with other such commitments. For example, I may exploit the fact that I can plan to bring an umbrella or plan to arrange a drive independently and without (unresolvable) interaction.

This is not to say that these judgements of ability, and especially their independence, are always accurate; the nature of abstraction is such that the lack of detail must occasionally lead to difficulty. As we describe below, an agent may decide that it's goal includes making both A and B true, something which its abstraction admits as feasible. When the time comes to construct (or instantiate) a plan to achieve both, in a particular context, the agent may realize that their joint truth is not feasible. The original goal must be abandoned and a new goal — accounting for the impossibility of $A \wedge B$ — must then be derived. In this way, goal derivation uses an abstract (or approximate) notion of ability, and is in some sense an operation that exists prior to planning. However, a planner can influence the actual choice of goal by indicating the infeasibility of a goal derived by means of this abstraction. Figure 4 illustrates the interaction of goal derivation and planning. To the extent that the abstract judgements of ability are "accurate" however, goal rederivation and replanning will not be frequently required.[11]

We do not elaborate on this brief (and no doubt vague) sketch of the role of goal derivation and planning. A complete account will necessarily depend on more realistic accounts of ability in which the interactions of abilities is modeled, on the particular model of action chosen, and on the exact nature of the planner. However, the naive model of ability proposed here is sufficient for our purposes,

[11]This is not to say that replanning for other reasons is not common, for instance, in response to incorrect knowledge of the world or changes in the world.

namely, to illustrate how ability, beliefs and preferences interact in goal derivation. More elaborate models of ability will not change the fundamental nature of these interactions.

Given that atoms are categorized as controllable or uncontrollable, we can now distinguish three types of propositions.

Definition 4.1 For any set of atomic variables $\mathcal{P}$, let $V(\mathcal{P})$ be the set of valuations over this set (i.e., $v : \mathcal{P} \mapsto [0, 1]$). If $v \in V(\mathcal{P})$ and $w \in V(\mathcal{Q})$ for disjoint sets of variables $\mathcal{P}$, $\mathcal{Q}$, then $(v; w) \in V(\mathcal{P} \cup \mathcal{Q})$ denotes the extended valuation: $(v; w)(P) = v(P)$ if $P \in \mathcal{P}$; $(v; w)(P) = w(P)$ if $P \in \mathcal{Q}$.

Definition 4.2 A proposition α is *controllable* iff, for every $u \in V(\overline{\mathcal{C}})$, there is some $v \in V(\mathcal{C})$ and $w \in V(\mathcal{C})$ such that $v; u \models \alpha$ and $w; u \models \neg\alpha$.

A proposition α is *influenceable* iff, for some $u \in V(\overline{\mathcal{C}})$, there is some $v \in V(\mathcal{C})$ and $w \in V(\mathcal{C})$ such that $v; u \models \alpha$ and $w; u \models \neg\alpha$.

A proposition α is *uninfluenceable* iff it is not influenceable.

Intuitively, since atoms in $\mathcal{C}$ are within complete control of the agent, it can ensure the truth or the falsity of any controllable proposition α, according to its desirability, simply by bringing about an appropriate truth assignment. For instance, if $A, B \in \mathcal{C}$ then propositions $A \vee B$ and $A \wedge B$ are controllable — $do(A)$ ensures $A \vee B$ holds, $do(\overline{A})$ and $do(\overline{B})$ ensures it is false, and so on. If α is influenceable, we call an appropriate assignment u to $\overline{\mathcal{C}}$ a *context* for α; intuitively, should such a context hold, α can be controlled by the agent. For example, if $A \in \mathcal{C}$ and $X \in \overline{\mathcal{C}}$ then proposition $A \vee X$ is influenceable but not controllable: in context X the agent cannot do anything about the truth of $A \vee X$, but in context $\overline{X}$ the agent can make $A \vee X$ true or false through $do(A)$ or $do(\overline{A})$. Note that all controllable propositions are influenceable. Uninfluenceable propositions are those over which the agent has no control in any circumstance. If $X \in \overline{\mathcal{C}}$ then X is uninfluenceable.

The category of controllability into which a proposition falls is easily determined by writing it in minimal DNF. Let $PI(\alpha)$ denote the set of prime implicants of α. It is readily verified that:

Proposition 4.1 *a) α is controllable iff each clause in $PI(\alpha)$ contains some literal from $\mathcal{C}$ and some clause contains only literals from $\mathcal{C}$. b) α is influenceable iff some literal from $\mathcal{C}$ appears in $PI(\alpha)$. c) α is uninfluenceable iff no literal from $\mathcal{C}$ appears in $PI(\alpha)$.*

4.2 Complete Knowledge Goals

Given the distinction between controllable and uncontrollable propositions, we want to define goals so that an agent is required to do only those things within its control. A first attempt might simply be

$$
\begin{array}{cccc}
R\,C\,\overline{T\,H} & R\,\overline{C\,T\,H} & \overline{R\,C\,T}\,H & \overline{R\,C\,T\,H} \\
 & \downarrow & & \\
R\,C\,\overline{T}\,H & R\,\overline{C\,T}\,H & \overline{R\,C\,T}\,\overline{H} & \overline{R\,C\,T\,H} \\
 & \downarrow & & \\
R\,C\,T\,\overline{H} & R\,\overline{C}\,T\,\overline{H} & \overline{R}\,C\,T\,H & \overline{R}\,C\,\overline{T}\,H \\
 & \downarrow & & \\
R\,C\,T\,H & R\,\overline{C}\,T\,H & \overline{R}\,C\,T\,\overline{H} & \overline{R}\,C\,\overline{T\,H}
\end{array}
$$

Figure 5: User Preferences

to restrict the ideal goal set, as defined above, to controllable propositions. Thus an agent is simply charged with the task of bringing about those ideal goals within its control. The following example shows this to be inadequate.

Example 4.1 Consider the following scenario, using five atoms: O, it is overcast; R, it will rain; C, I have coffee; T, I have tea; and H, my office thermostat is set high. Of these variables, my robot can control C, T and H — it can bring me coffee or tea and can turn my thermostat up or down. The robot has the default information $O \Rightarrow R$ (it normally rains when it's overcast) and knows the facts $KB = \{O, \overline{H}, \overline{C}, \overline{T}\}$. Its default closure is $Cl(KB) = \{O, R, \overline{H}, \overline{C}, \overline{T}\}$. Finally, its preference ordering is designed to respect my preferences: when it's raining I prefer tea when I arrive and the thermostat set high, otherwise I prefer coffee and the thermostat set low. Thus, we have the preference ordering illustrated in Figure 5. (We assume O, R do not contribute directly to preference, and that priority has been given to C and T over H. We also allow the possibility that both C and T together satisfy a preference for either.) The robot has to decide what to do before I arrive at the office. ■

It should be clear that the robot should not determine its goals by considering the ideal situations satisfying $Cl(KB)$. In such situations, since $\overline{H}$ is known, $\overline{H}$ is "preferred" (indeed, it is a simple theorem of QDT that $I(\alpha|\alpha)$). Thus, the robot concludes that $\overline{H}$ *should be true*. This is clearly mistaken, for considering only the best situations in which one's knowledge of controllables is true prevents one from determining whether changing those controllables could lead to a better situation. Thus we do not require an agent to restrict attention to those situations where KB or $Cl(KB)$ is true. The fact that $\overline{H}$ is known should not unduly influence what are considered to be the best alternatives

— H can be made true if that is what's best.

Of course, the goals of an agent must still be constrained by *uninfluenceable* propositions it knows to be true. The agent should not reject all of its knowledge. For example, if the preference ordering above were modified to reflect my preference for clear weather ($\overline{O}$), my agent should not base its goals on this preference if it knows it is overcast (O). Making O false is beyond its control, and it goals should determined by restricting attention to $\overline{O}$-worlds. Thus we insist that the best situations satisfying known *uninfluenceable* propositions be considered.

Notice also that we should not ignore the truth of controllables when making default predictions. The prior truth value of a controllable might provide some indication of the truth of an uncontrollable; and we *must* take into account these uncontrollables when deciding which alternatives are *possible*, before deciding which are best. In this example, we might imagine that the default $O \Rightarrow R$ doesn't hold, but that $O \wedge H \Rightarrow R$ does: if it is overcast, then the thermostat is set high because I anticipated rain before I left last night. Our agent must use the truth of this controllable atom H to determine the truth of the uncontrollable R, which in turn will influence its decisions.[12] Once accounted for in forming $Cl(KB)$, H can safely be ignored. This leads to the following formulation of goals that account for ability. We assume as usual a QDT-model M reflecting the agent's preferences and defaults.

Definition 4.3 Let M be a QDT-model and KB a propositional belief set. The *uninfluenceable belief set* (given M and KB), denoted $UI(KB)$ is

$$UI(KB) = \{\alpha \in Cl(KB) : \alpha \text{ is uninfluenceable}\}$$

For the remainder of this section we assume that $UI(KB)$ is complete with respect to $\overline{\mathcal{C}}$; that is, for each $P \in \overline{\mathcal{C}}$, either $P \in UI(KB)$ or $\neg P \in UI(KB)$. We use the (complete) uninfluenceable belief set to determine an agent's goals.

Definition 4.4 Let M be a QDT-model and $KB \subseteq \mathbf{L}_{CPL}$. Proposition α is a *complete knowledge (CK-) goal* (with respect to M, KB) iff $M \models I(\alpha | UI(KB))$ and α is controllable.

This definition embodies the intuition that the set of expected uninfluenceable beliefs is fixed and (since beyond the agent's control) unchangeable. Given $UI(KB)$, an agent is charged with ensuring

[12]If a controllable provides some indication of the truth of an uncontrollable or another controllable, (e.g., $H \Rightarrow R$) we should think of this as an *evidential rule* rather than a *causal rule*. Given our assumption about the independence of atoms in $\mathcal{C}$, we must take all such rules to be evidential (e.g., changing the thermostat will not alter the chance of rain). This can be generalized using a more reasonable conditional representation, and ultimately should incorporate causal structure. Note the implicit temporal aspect here; propositions should be thought of as *fluents*. We can avoid an explicit temporal representation by assuming that preference is solely a function of the truth values of fluents.

that all controllable propositions that are ideal in this context are made true.

As with ideal goals, the set of CK-goals is deductively closed and should be viewed as a set of necessary conditions in any rational course of action — if some CK-goal is not true, the agent is in a less-than-ideal $UI(KB)$-world. Of course, goals can only be affected by atomic actions, so we will typically be interested in a set of actions that is guaranteed to achieve each CK-goal. To this end, we define an *(atomic) action set* to be any set of controllable literals. If $\mathcal{A}$ is such a set we use it to denote both the set of actions $do(l)$ for each $l \in \mathcal{A}$, and the proposition formed by the conjoining its elements.

Definition 4.5 Let M be a QDT-model and $KB \subseteq \mathbf{L}_{CPL}$. An *atomic goal set* is any action set $\mathcal{A}$ such that, for each CK-goal α,

$$M \models UI(KB) \wedge \mathcal{A} \supset \alpha$$

Clearly, any such atomic goal set ensures each CK-goal α holds and thus determines a reasonable course of action. Such action sets can be determined by appeal to sufficiency.

Theorem 4.2 *Let $\mathcal{A}$ be some atomic action set. Then $\mathcal{A}$ is a goal set iff $\mathcal{A}$ is sufficient for $UI(KB)$.*

Proof The proof proceeds similarly to that of Proposition 3.3. ■

As a result, the fact that $\mathcal{A}$ is a goal set can be expressed in the language of QDT as follows:

$$M \models \overleftarrow{\Box}_P(\mathcal{A} \supset \overleftarrow{\Box}_P(\mathcal{A} \supset \neg UI(KB)))$$

In our example above, where the robot knows O, it set of uninfluenceable beliefs is $UI(KB) = \{O, R\}$. The set of CK-goals includes any proposition entailed by the set $\{O, R, T, H\}$. However, given $UI(KB)$ all such goals are made true should the agent execute one of the two possible atomic goal sets, $\{T, H\}$ and $\{C, T, H\}$. Typically, we will be interested in minimal goals sets, since these require the fewest actions to achieve ideality. We may impose other metrics and preferences on goals sets as well (e.g., by associating costs with various actions).

Notice that the preference for tea does not prevent the robot from bringing coffee. While a pragmatic filtering mechanism such as minimality would rule out this possibility, such constraints can easily be imposed on the preference ordering if appropriate. Disjunctive goals and "integrity constraints" pose no difficulty. If I prefer exactly one of coffee of tea, the preference $I(C \equiv \neg T)$ can be asserted, preventing the robot from bringing both. In such a case, the generated atomic goal sets will be $\{C, \overline{T}\}$ and $\{\overline{C}, T\}$, ruling out $\{C, T\}$.

With default information and controllability in place, we can briefly return to the alternative interpretation of preference statements described in Section 2. Recall the assertion "I prefer an umbrella when it's raining" was interpreted as an *expected preference* rather than an ideal preference. Formally, such expected preferences can be interpreted as stating that the consequent is preferred given the uninfluenceable belief set induced by the antecedent. In this case, the assertion means $I(U|UI(\{R\}))$. Thus, the logic QDT can be used to express both types of preferences as long as the intended meaning of the preference "assertions" is made clear. Such expected preference statements can also be used to derive new information. For instance, together with the ideal preferences such as $I(D|R)$ and $I(\overline{U}|D)$ (and other background information as before), this statement $I(U|UI(\{R\}))$ allows one to draw conclusions about an user's defaults, for instance, $R \Rightarrow \neg D$. This can be used to further influence the derivation of subsequent goals.

4.3 A Decision-Theoretic Interpretation

In our goal derivation scheme, a certain priority is given to defaults over preferences. Goals are determined by first constructing the default consequences of *KB* and then deciding what to do based on this knowledge as if it were certain. While a seemingly reasonable approach, in a truly decision-theoretic setting acting on the basis of uncertain information is a function not only of its likelihood, but also the consequences of being incorrect. For instance, in our framework we might have the default rule $R \Rightarrow S$, stating that if I run across the freeway I will reach the other side safely. If this allows me to arrive at my destination five minutes sooner (assuming arriving earlier is desirable) than had I crossed at a crosswalk, the default assumption S will ensure that I run across the freeway, the reasoning being that I won't (by default) get hit by a car and I will arrive sooner. In general, however, the (rather drastic) consequences of my default prediction turning out poorly must be traded off against the probability of being right. If the five minutes saved is not worth the risk, then I decide to go to the crosswalk.

To express this tradeoff we must assume that the qualitative scales of preference and normality are calibrated somehow; and nothing in the constraints expressed by the user in our purely qualitative setting allows such an assumption. In the concluding section we discuss potential "qualitative" ways around this problem. However, the scheme presented here has a certain naive appeal, which may be partly due to the observation that defaults are usually expressed with such considerations in mind (2; 2). Furthermore, the scheme is conceptually simple in that it embodies a principle analogous to the *separability* of state estimation and control (2). An agent can calculate what is (probably) true of the world and subsequently and independently base its decisions upon these beliefs. Finally, our scheme is applicable when likelihood and preference information is truly qualitative and explicit calibration of the orderings is not feasible. This is in fact the key motivation for our proposal. Few decision criteria have been proposed for dealing with purely qualitative constraints of this type. In many applications, quantitative or even "calibrated" scales of likelihood and preference may not exist.

Given, the need for qualitative decision criteria, we can describe some conditions under which our assumption of separability is appropriate. In particular, the decisions sanctioned by our scheme, whereby defaults are given precedence over preferences, are approximately sound with respect to the MEU criterion under certain circumstances.

The logic of conditional normality statements can be given a probabilistic interpretation as described in (2). In particular, the purely conditional fragment is equivalent to Adams's (2) system of *ε-semantics*, which has also been applied to the representation of defaults (2). This means that there is a probability assignment that ensures that every default rule $A \Rightarrow B$ corresponds to an assertion of high conditional probability $P(B|A) > 1 - \varepsilon$, for any $\varepsilon > 0$. Thus, we may assume that a user chooses default rules with such a parameter in mind, and that $P(Cl(KB)|KB) > 1 - \varepsilon$. We can also assume that the preference ordering is "constructed" by clustering together worlds that have actual utility within some reasonably small range, and treating distinct clusters as separated by a reasonably large gap in utility. Thus, the user treats certain outcomes as having (more or less) indistinguishable utility, those within the same cluster, while outcomes in different clusters have sufficiently different utilities. To analyze the appropriateness of our goal derivation scheme, we make this assumption precise by assigning a point utility δ_i to each cluster in the preference ordering. Let δ denote the smallest gap $\delta_i - \delta_{i+1}$ between any two adjacent point utilities (the "smallest perceptible change" in utility) and let $\Delta = \delta_0 - \delta_n$ denote the magnitude of the possible range in utility (see Figure 6).

Goals (or decisions) are determined with respect to a given *KB*, which induces a decision problem in the obvious fashion: given the context *UI*(*KB*) what is an optimal course of action (or atomic action set)? Under our assumptions, some action set must have maximal expected utility; let U^* denote the expected utility of any such optimal action. For an arbitrary action set $\mathcal{A}$, we denote by $EU(\mathcal{A})$ the expected utility of $\mathcal{A}$. Since our goal derivation scheme suggests that any goal set $\mathcal{A}$ is a reasonable course of action, we want to compare $EU(\mathcal{A})$ to U^* for any goal set $\mathcal{A}$. The difference between the two indicates the degree to which our qualitative criterion is suboptimal.

Intuitively, an atomic goal set embodies the best decision should the default beliefs of an agent actually be the case. On our interpretation above, we have that $P(Cl(KB)|KB) > 1 - \varepsilon$ and hence that $P(UI(KB)|KB) > 1 - \varepsilon$. The only circumstances under which the goal set potentially reflects a poor decision is when one of the default beliefs is false. This is due to the fact that unexpected propositions (those whose negation is believed by default) play no role in goal derivation. However, the error is bounded by the probability of default violation and the potential loss.

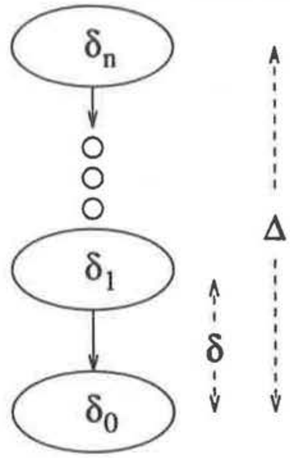


Figure 6: Utilities Assigned to a Preference Ordering

Proposition 4.3 *$U^* - EU(\mathcal{A}) \leq \varepsilon\Delta$ for any goal set $\mathcal{A}$.*

Proof Let $\mathcal{A}$ be some goal set and $\mathcal{B}$ be an optimal action set for the decision problem. Since goal sets determine outcomes with a single ranking of preference (given *UI*(*KB*)), let the outcome(s) associated with $\mathcal{A}$ have utility δ_i. Since $P(UI(KB)|KB) \geq 1-\varepsilon$, we have $EU(\mathcal{A}) \geq (1-\varepsilon)\delta_i + \varepsilon\delta_n$. The outcome(s) of $\mathcal{B}$ given *UI*(*KB*) cannot have utility greater than δ_i (otherwise $\mathcal{B}$ would be a goal set and $\mathcal{A}$ would not); so $EU(\mathcal{B}) \leq (1-\varepsilon)\delta_i + \varepsilon\delta_0$. Thus,
$U^* - EU(\mathcal{A}) \leq \varepsilon(\delta_0 - \delta_n) \leq \varepsilon\Delta$. ∎

Furthermore, goal sets are guaranteed to be better than non-goal actions sets under certain conditions.

Proposition 4.4 *Let $\mathcal{A}$ any atomic goal set and $\mathcal{B}$ any (non-goal) action set. Then $EU(\mathcal{B}) - EU(\mathcal{A}) \leq \varepsilon\Delta - (1-\varepsilon)\delta$.*

Proof As above, $EU(\mathcal{A}) \geq (1-\varepsilon)\delta_i + \varepsilon\delta_n$. Since $\mathcal{B}$ is not a goal set, its outcome(s) given *UI*(*KB*) must have utility less than δ_i, say $\delta i + k$. Then $EU(\mathcal{B}) \leq (1-\varepsilon)\delta_{i+k} + \varepsilon\delta_0$. Thus,
$EU(\mathcal{B}) - EU(\mathcal{A}) \leq \varepsilon(\delta_0 - \delta_n) - (1-\varepsilon)(\delta_i - \delta_{i+k}) \leq \varepsilon\Delta - (1-\varepsilon)\delta$. ∎

Therefore, a goal set $\mathcal{A}$ is guaranteed to be better than any non-goal action set whenever $\delta(1-\varepsilon) \geq \varepsilon\Delta$.

This gives some idea of the circumstances under which the assumption of separability is sound. If the probability of default violation is reasonably small, when "magnified" by the cost associated with error, then our decisions will be acceptable and approximately sound. Of course, it is unreasonable to only reason with qualitative constraints that meet these stringent requirements. But they do suggest useful abstractions for ordinary goal derivation, and the degree to which these conditions are approximated gives reasonable assurance of good decisions. Thus, the separability assumption provides a computationally manageable procedure for finding "satisficing" solutions.

5 Goals in the Presence of Incomplete Knowledge

The goals described above seem reasonable, in accord with the general maxim "do the best thing possible consistent with your knowledge." We dubbed such goals "CK-goals" because they seem correct when an agent has complete knowledge of the world (or at least its uncontrollable component). But CK-goals do not always determine the best course of action if an agent's knowledge is *incomplete*. Consider an agent with the preferences and defaults of the umbrella example, but whose propositional knowledge base is empty. For all the agent knows it could rain or not — it has no indication either way. The definition of CK-goal requires that the agent $do(\overline{U})$, for the best situation consistent with $KB = \emptyset$ is $\overline{RU}$. Leaving its umbrella at home is the best choice should it turn out not to rain; but should it rain, the agent has ensured the *worst* possible outcome. It is not clear that $\overline{U}$ should be a goal. Indeed, one might expect U to be a goal, for no matter how R turns out, the agent has avoided the worst outcome.

5.1 Strategies

CK-goals are appropriate in the case of complete knowledge because the outcome associated with a complete course of action is certain. Nothing is left to chance. However, in the example above there *is* uncertainty regarding the possible outcomes of any course of action: it may rain or not. This uncertainty manifests itself in uncertainty about the *desirability of the outcome*.

In the MEU framework, one can deal with such uncertainty easily; but qualitatively, when trying to do as much as possible with strictly ordinal value information, a different approach is required. In the presence of incomplete knowledge there are various *strategies* for determining goals, corresponding to the attitude an agent adopts toward this uncertainty. In the decision theory literature, a problem of this type is sometimes referred to as a decision problem *under strict uncertainty* (2), alluding to the fact that while there may be several possible outcomes of different utility, the probability of these outcomes is unknown. Several decision criteria have been proposed for such problems, among them minimax regret (2) and Laplace's principle of indifference (2). However, many of these are not applicable in in the case of purely qualitative information, requiring that the desirability of outcomes be associated with (numerically significant) utilities. Two criteria that are applicable are the *maximax* and *maximin* (2) decision rules. It turns out that these criteria can be formulated in QDT.

We first describe some preliminary notions. Let a *complete action set* be any complete truth assignment to the atoms in $\mathcal{C}$. These are the alternative courses of action available to an agent (sometimes referred to simply as *actions*). If *UI*(*KB*) is incomplete, any complete action set $\mathcal{A}$ induces a set of possible outcomes:

Definition 5.1 Let $M = \langle W, \leq_P, \leq_N, \varphi \rangle$ be a QDT-model, and *KB* an objective knowledge base. The *set of outcomes* induced by an action set $\mathcal{A}$ is

$$OUT(KB, \mathcal{A}) = \{w \in W : M \models_w \mathcal{A} \wedge UI(KB)\}$$

The relative desirability of any decision will be a function of its outcome set and those of competing actions.

In the example above, an agent could decide to leave its umbrella at home. This reflects an optimistic strategy for goal derivation: leaving the umbrella *maximizes potential gain*, for it allows the possibility of the agent ending up with the best possible outcome consistent with *KB*, namely, $\overline{R} \wedge \overline{U}$. More generally, for each action define its *best outcome set* as

$$MAX(KB, \mathcal{A}) = \{w \in OUT(KB, \mathcal{A}) : \text{ if } v \in OUT(KB, \mathcal{A}) \text{ then } w \leq_P v\}$$

An optimistic strategy consists of choosing any action with a maximal or most preferred best outcome set. This corresponds to the *maximax* decision rule, for among the alternative courses of action, we choose the action that maximizes the maximum (or preferred) possible outcome.

Definition 5.2 A complete action set $\mathcal{A}$ is an *optimistic best (OB-) action set* given *KB* iff for any complete action set $\mathcal{B}$, if $w \in MAX(KB, \mathcal{A})$ and $v \in MAX(KB, \mathcal{B})$ then $w \leq_P v$.

Definition 5.3 A proposition α is an *optimistic goal* given *KB* iff

$$\bigvee \{\mathcal{A}_i : \mathcal{A}_i \text{ is an optimistic best action set }\} \models \alpha$$

Thus, α is an optimistic goal iff it is required to be achieved when any optimistic best course of action is executed. For example, if $\{A, B\}$ and $\{A, \overline{B}\}$ are OB-actions, A is a goal, but B is not, for the agent is free to choose whether to $do(B)$ or $do(\overline{B})$ (or let it run its own course). This notion of goal has controllability built in (for actions sets only entail controllable propositions) and corresponds to striving for the best possible outcomes consistent with *UI*(*KB*).

Theorem 5.1 *For any KB, α is an optimistic goal iff $M \models I(\alpha|UI(KB))$ and α is controllable.*

Proof We note that if α is an optimistic goal then it is equivalent to a sentence containing only atomic variables from $\mathcal{C}$ (since α is entailed by actions, themselves formed using only variables from $\mathcal{C}$). Thus, every (nontautological) optimistic goal is controllable, and we need only show that $\mathcal{A} \models \alpha$ for each OB-action $\mathcal{A}$ iff $M \models I(\alpha|UI(KB))$ for those α containing only variables from $\mathcal{C}$.

By definition, $\mathcal{A}$ is an OB-action iff $M \models_w \mathcal{A}$ for some $w \in \min(UI(KB), \leq_P)$; and α is a goal iff $M \models_w \alpha$ for each such w. However, since α contains only variables from $\mathcal{C}$, $M \models_w \alpha$ iff $\mathcal{A} \models \alpha$ for the action $\mathcal{A}$ such that $M \models_w \mathcal{A}$. Thus the disjunction of all optimistic best action sets entails exactly the controllable ideal goals α. ■

Furthermore, while any optimistic best action set allows the possibility of a best outcome, any other course of action rules out the possibility:

Theorem 5.2 *Let $\mathcal{A}$ be a complete action set. Then $\mathcal{A}$ is an optimistic best action set iff $M \models T(\mathcal{A}|UI(KB))$.*

Proof This follows immediately from the definition of OB-actions, for $\mathcal{A}$ is an OB-action iff there exists a preferred $UI(KB)$-world satisfying $\mathcal{A}$. ■

This fact suggests a simple "greedy" approach for the construction of optimistic best action sets. Suppose $A_1, \cdots A_n$ is an enumeration of the atoms in $\mathcal{C}$. If $M \models T(A_1|UI(KB))$ then there exists an OB-action set with action A_1; otherwise $\overline{A_1}$ is part of some best action set. Suppose A_1 is chosen; a similar test for A_2 can be undertaken accounting for the choice of A_1: one of $T(A_2|UI(KB) \wedge A_1)$ or $T(\overline{A_2}|UI(KB) \wedge A_1)$ must hold. After n such tests an optimistic best action set has been constructed. In addition, if at any stage $I(A_1 \wedge \cdots A_i|UI(KB))$ is true, this partial action set is also best (in that any completion of it is an OB-action).

In certain domains adopting an optimistic strategy might be a prudent choice (for example, where a cooperative agent determines the outcome of uncontrollables). Of course, another strategy might be the cautious strategy that *minimizes potential loss*. In our example, an agent would take its umbrella to prevent the worst possible outcome $R \wedge \overline{U}$ from occurring. For each action define its *worst outcome set* as

$$MIN(KB, \mathcal{A}) = \{w \in OUT(KB, \mathcal{A}) : \text{ if } v \in OUT(KB, \mathcal{A}) \text{ then } v \leq_P w\}$$

The relative goodness of an action corresponds to the degree of preference of its worst outcomes. A cautious strategy consists of choosing any action with a maximal or most preferred worst outcome set. This corresponds to the *maximin* decision rule, for among the alternative courses of action, we choose the action that maximizes the minimum (or least preferred) possible outcome.

Definition 5.4 Let $\mathcal{A}, \mathcal{B}$ be complete action sets. $\mathcal{A}$ is *as good as* $\mathcal{B}$ ($\mathcal{A} \leq \mathcal{B}$) iff $w \leq_P v$ for any $w \in MIN(KB, \mathcal{A})$, $v \in MIN(KB, \mathcal{B})$. Action $\mathcal{A}$ is an *cautious best (CB-) action set* given *KB* iff for any action $\mathcal{B}$, if $w \in MIN(KB, \mathcal{A})$ and $v \in MIN(KB, \mathcal{B})$ then $w \leq_P v$.

Clearly, $\leq$ imposes a transitive, connected ordering on complete actions sets, and best action sets are those minimal in this ordering. If an agent chooses other than a cautious best action set, it opens the possibility for a worse outcome (this is an immediate consequence of the definition):

Theorem 5.3 *Let $\mathcal{A}$ be a best action set for KB and $\mathcal{B}$ be any complete action set. For any $w \models UI(KB) \wedge \mathcal{A}_i$, there is some $v \models UI(KB) \wedge \mathcal{A}_j$ such that $w \leq_P v$.*

Definition 5.5 A proposition α is a *cautious goal* given *KB* iff

$$\bigvee \{\mathcal{A}_i : \mathcal{A}_i \text{ is a cautious best action set }\} \models \alpha$$

Cautious goals correspond to preventing the worst possible outcomes consistent with *UI*(*KB*) (and as usual have the notion of controllability built in).

The relative goodness of an action sets can also be expressed in QDT.

Proposition 5.4 *$\mathcal{A}$ is as good as $\mathcal{B}$ iff*

$$M \models \overleftrightarrow{\Diamond}_P(\mathcal{B} \wedge UI(KB) \wedge \neg\overleftarrow{\Diamond}_P(\mathcal{A} \wedge UI(KB)))$$

Proof $\mathcal{A}$ is as good as $\mathcal{B}$ iff for every $w \in MIN(KB, \mathcal{A})$ there is some $w \leq_P v$ such that $\mathcal{A} \leq \mathcal{B}$; that is, iff $M \models_v \mathcal{B} \wedge UI(KB)$. For any such v we must have $M \models_w \neg\overleftarrow{\Diamond}_P(\mathcal{A} \wedge UI(KB))$ since $w \in MIN(KB, \mathcal{A})$. Thus $\mathcal{A} \leq \mathcal{B}$ iff $M \models \overleftrightarrow{\Diamond}_P(\mathcal{B} \wedge UI(KB) \wedge \neg\overleftarrow{\Diamond}_P(\mathcal{A} \wedge UI(KB)))$. ■

The best actions sets are those for which this condition is provable for all $\mathcal{B}$. This suggests a (not unexpected) difficulty in choosing cautious best action sets. In the worst case, one may have to "search" through all action sets in order to determine if an action is a cautious best action set. This stands in contrast with optimistic best actions sets, for which a simple test exists (Theorem 5.2) and for which a greedy construction strategy exists.

We cannot expect best action sets, in general, to be sufficient in the same sense that CK-goal sets are. The potential for desirable and undesirable outcomes makes it impossible to ensure that the best outcomes consistent with $UI(KB)$ are forthcoming. However, we can show that if there is some action set that is sufficient for KB then all cautious best action sets will be sufficient.

Proposition 5.5 *If some action set $\mathcal{A}$ sufficient for $UI(KB)$, then every cautious best action set is sufficient.*

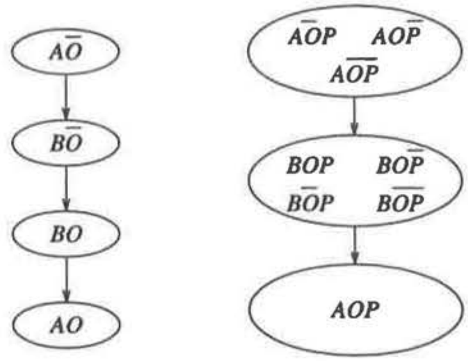
Proof $\mathcal{A}$ is sufficient for $UI(KB)$ iff $M \models \overleftrightarrow{\Box}_P(UI(KB) \supset \overleftrightarrow{\Box}_P(UI(KB) \supset \neg\mathcal{A}))$ iff $OUT(KB, \mathcal{A}) \subseteq \min(UI(KB), \leq_P)$ iff $MIN(KB, \mathcal{A}) \subseteq \min(UI(KB), \leq_P)$. Now if $\mathcal{B}$ is a cautious best action set then $MIN(KB, \mathcal{B}) \subseteq \min(UI(KB), \leq_P)$; and by the reasoning above $\mathcal{B}$ is sufficient for $UI(KB)$. ■

Hence, CK-sufficiency can sometimes be applied even in the case of incomplete knowledge. Its applicability implies that possible outcomes of unknown uncontrollables have no influence on preference: all *relevant* factors are known.

The cautious strategy seems applicable in a situation where one expects the worst possible outcome, for example, in a game against an adversary. Once the agent has performed its action, it expects the worst possible outcome, so there is no advantage to discriminating among the candidate (best) action sets: all have equally good worst outcomes. However, it's not clear that this is the best strategy if the outcome of uncontrollables is essentially "random." If outcomes are simply determined by the natural progression of events, then one should be more selective. We think of nature as neither benevolent (a cooperative agent) or malevolent (an adversary). Therefore, even if we decide to be cautious (choosing among *cautious best* action sets), we should account for the fact that a worst outcome might not occur: we should choose the action sets that take advantage of this fact. Such considerations also apply to games where an opponent might not be able to consistently determine her best moves and an agent wants to exploit this fact. It is easy to distinguish such CN-actions in QDT, choosing those that are "optimistic", or using other means.

5.2 Observations

If an agent can *observe* the truth values of certain unknown propositions before it acts, it can improve its decisions. In many cases, fortuitous observations will preclude the worst possible outcomes and change the actions chosen. To continue the "umbrella" example, suppose R and C are unknown. The agent's cautious goal is U. If it were in the agent's power to determine C (cloudy) or $\overline{C}$ before acting,



(a) (b)

Figure 7: The value of (a) single and (b) multiple observations.

its choice of actions could change. Observing $\overline{C}$ indicates the impossibility of R, and the agent could then decide to exploit this information and $do(\overline{U})$.

To capture this notion, we distinguish two types of uncontrollable atoms, *observables* $\mathcal{O}$ and *unobservables* $\overline{\mathcal{O}}$; formally, we assume $\overline{\mathcal{C}} = \mathcal{O} \cup \overline{\mathcal{O}}$. Suppose *KB* determines a several best action sets of which one, $\mathcal{A}^*$, is chosen by the agent. Intuitively, the observation of some unknown uncontrollable atom O is worthwhile if it can potentially change the agent's choice of best action. Cautious and optimistic goals must be treated differently, for the are influenced by "positive" and "negative" observations respectively.

We first describe the value of observations for an agent that has adopted a cautious goal derivation strategy. Suppose *UI*(*KB*) is the agent's uninfluenceable belief set and that $\mathcal{A}^*$ is the cautious best action set chosen by the agent. Furthermore, assume that variable O is observable but that that both O and $\overline{O}$ are consistent with the agent's beliefs. Since O is observable, it is a "worthwhile" observation to undertake if it has the *potential* to change the agent's decision. Note that we can only use the observation's potential for change, for no observation can be guaranteed to change an agent's decision. This is due to the fact that among the decision's ($\mathcal{A}^*$) worst outcomes must be one that either makes O or $\overline{O}$ true, and the observation may simply "validate" the agent's cautious choice. Consider Figure 7(a), where actions A and B are possible and O is observable. If the value of O is unknown, an agent's cautious choice is B. Should the agent decide to test O, an outcome of O will cause a different cautious choice A. Thus, a test of O has potential impact. Notice that if the test results in $\overline{O}$, the cautious best decision remains B.

Definition 5.6 Let $UI(KB) \not\vdash O$, $UI(KB) \not\vdash \overline{O}$. Observable O has *value* in context *KB* iff, for some

CB-action $\mathcal{A}$ (w.r.t *KB*), $\mathcal{A}$ is not a CB-action for one of $KB \cup \{O\}$ or $KB \cup \{\overline{O}\}$.

In this case, observing O is worthwhile since it *might* (depending on its actual truth value) rule out certain feasible decisions for the agent. In our running example, not knowing whether it will rain, our agent would take its umbrella. However, if a phone call to the weather office could refute the possibility of rain, our agent's decision would differ if it had this information. This is essentially a qualitative analog of *value of information*. Of course, we cannot quantify the potential value of making an observation; but we may compare the relative values of two pieces of information O and P. For simplicity, assume that positive observations O and P are the "improving" outcomes. Let $\mathcal{A}_O$ and $\mathcal{A}_P$ be CB-actions for $KB \cup \{O\}$ and $KB \cup \{P\}$. The value of O is as great as that of P just when

$$M \models \overleftrightarrow{\Diamond}_P(\mathcal{A}_P \wedge UI(KB \cup \{P\}) \wedge \neg \overleftarrow{\Diamond}_P(\mathcal{A}_O \wedge UI(KB \cup \{O\})))$$

Faced with the choice of observing O or P, should O have greater value, we should choose O for it allows (should it turn out positively) greater potential for improvement. Generally speaking, decisions change when an agent observes "positive" instances of observable variables (those that rule out worst outcomes).

It is not hard to see that the value of an observable depends crucially on context. In particular, it may have no value in some context but great value in another — or its value may only be derived by considering its observation jointly with that of another variable. For example, in Figure 7(b), if we take O and P to be unknown observables, we see that neither O nor P alone have value. Observing O or $\overline{O}$, P or $\overline{P}$, cannot affect the cautious choice: B is always the cautious best action. However, should O and P *both* be observed, there is a possibility for a change of decision. If the outcome of this test is $O \wedge P$, then A is the best action. Generally, value of information is a property of sets of observations rather than single observations. The definition above is easily generalized to accommodate this fact.

The considerations above for a cautious decision strategy can be applied to optimistic goals as well. Rather than ruling out undesirable outcomes as above, an observable has value in an optimistic setting just when it rules out certain acceptable outcomes. Thus, it is "negative" observations that have value in this setting. For instance, an optimistic agent will leave its umbrella at home if it doesn't know whether it will rain. A "negative" observation that it will rain (e.g., by phoning the weather office) will change its decision. Figure 7(a) also illustrates this phenomenon — observation $\overline{O}$ changes the optimistic decision from A to B. In this setting it is not hard to see that observation O has value iff one of the following hold for some optimistic best action set $\mathcal{A}$ (relative to *KB*):

$$\neg T(\mathcal{A} | UI(KB \cup \{O\})$$

$$\neg T(\mathcal{A}|UI(KB \cup \{\overline{O}\})$$

If $\mathcal{A}$ is not tolerable given $UI(KB \cup \{O\})$ then it cannot be an OB-action should O be observed; thus it is ruled out as a possible choice. As above, O comes with no guarantees, for one of O or $\overline{O}$ must "confirm" the original action choice. As well, the notion applies more generally to sets of observables, rather than just single variables.

6 Concluding Remarks

6.1 Related Work

Other attempts to define goals using preferences bear some relationship to our system. Doyle and Wellman (2) define goals that exhibit a conditional aspect like ours. Roughly, B is a goal given A just when $A \wedge B$ is preferred to $A \wedge \neg B$ for any *fixed* circumstance. For instance, if such a relationship holds $A \wedge B$ should be preferred given C, given $\neg C$, and so on. Such goals incorporate a *ceteris paribus* assumption: B is preferred to $\neg B$ given A, *all else being equal*. This guarantees that doing B will lead to a better situation *whenever* A holds — this is indeed a very strong, and probably rarely applicable, condition. Our conditional goals are much weaker. No such assurances can be provided. Intuitively, if B is a goal given A, then doing B will lead to a better situation, *all else being normal*. However, this permits defeasible goals, affording greater flexibility and naturalness of expression. Only factors directly relevant to preference need be stated, and others are assumed to be irrelevant. In addition, our goals incorporate elements of controllability.

Pearl (2) has proposed a system using much the same underlying logical apparatus as ours. However, conditional statements are taken to impose specific constraints on utility and probability distributions, allowing expected utility calculations (with "order of magnitude" values) to be performed. While this allows stronger conclusions to be reached in general, it makes stronger demands on the input information as well. Thus, the system cannot be construed as truly qualitative, so in a sense the aim here is different. Tan and Pearl (2) introduce a somewhat more qualitative system. It handles quantified conditional desires (adopting the machinery of qualitative probability (2)). To account for likelihood, they adopt our model of closing under default consequence before consulting preferences. Incompletely specified preferences induce a "compact" model where worlds gravitate toward neutrality, but as noted earlier, this is not an obviously useful strategy. Furthermore, conditional preferences are given a *ceteris paribus* interpretation along the lines of Doyle and Wellman. Aside from the unknown impact on the computation of compact rankings, their particular semantics is of questionable value for representing conditional preferences. For example, a preference for A given

$A \vee B$ requires that $\neg A \wedge \neg B$ be dispreferred. In our semantics, a conditional preference given any α imposes no constraints on the degree of preference of $\neg\alpha$-worlds. Finally, a crucial distinction is that the system of Tan and Pearl fails to incorporate any model of ability. Thus, statements of the form "A is preferred given $\neg A$" are given a direct interpretation. In our model, it is certainly possible to say that A is a *goal* given $\neg A$, *as long as* A is controllable. But this is the careful result of determining exactly which facts can be changed and which ideal preferences hold in the reduced context.

Our representation of preferences draws much from work on deontic logic, where preference may be determined by some legal or moral code. Some work in deontic logic has recently begun to incorporate, as we do here, default information (2; 2). However, much work in deontic logic embodies in some way or another the slogan "do the best consistent with what you know." As our considerations of ability show, this is not always the best manner in which to evaluate or derive obligations; if certain known facts are within an agent's control, its obligations (as its goals) should not be constrained by those facts.

6.2 Summary

We have presented a logic QDT for representing qualitative preference and likelihood information. We have shown how defeasible conditional preferences can be expressed, and described several methods for goal derivation based on the assumption that priority be given to defaults. There are a number of ways in which this work can be extended. Clearly, the account of action and ability is simplistic. An object-level characterization of actions with true causal structure can be added to the conditional framework (2) to make goal derivation more realistic. However, as more sophistication is added to the representation of action and ability for goal derivation, the distinction between goal derivation and planning becomes increasingly blurred.

The assumption of separability and priority of default information must be relaxed in many circumstances. In order to allow reasonable decisions to be made, a logic that allows tradeoffs of likelihood and preference to be expressed in a qualitative fashion is desirable. For instance, if I instruct my robot that it should run across the street (instead of crossing at the crosswalk) to save three minutes while fetching my coffee, it can safely deduce that running across the street is worth the risk if a courier deadline is involved. I have implicitly calibrated parts of its preference and normality rankings. We are currently exploring such mechanisms for reasoning directly with such qualitative tradeoff information. This can be viewed as a mechanism to deal with *imperatives*, or commands to perform an action in a particular circumstance. The knowledge implicit in such commands can then be propagated to other contexts.

Related to this is a fuller investigation of the different forms preference information might take in such a setting. As mentioned earlier, user preferences might be stated independently of typicality information, or might incorporate expected circumstances and controllability information. A well-developed logic for these and other "entangled" constraints is certainly worth pursuing.

Finally, a full investigation of multistage decision making in this qualitative context in currently under way. We take belief states to be ranked models as we do here. Actions transform one such belief state into another. The value of a particular course of action is, as usual, a function of the world states through which an agent passes while executing the actions in the sequence. However, since the true state of the world is only known defeasibly, as are the outcomes of actions, this value must be calculated using some qualitative analog of expected utility, or perhaps using other qualitative decision criteria such as those suggested here. Ultimately, a qualitative form of Markov decision processes (2; 2; 2) may prove feasible.

References

Ernest W. Adams. *The Logic of Conditionals*. D.Reidel, Dordrecht, 1975.

Carlos E. Alchourrón. Philosophical foundations of deontic logic and the logic of defeasible conditionals. In *Workshop on Deontic Logic in Computer Science*, Amsterdam, 1991.

Craig Boutilier. Conditional logics of normality as modal systems. In *Proceedings of the Eighth National Conference on Artificial Intelligence*, pages 594–599, Boston, 1990.

Craig Boutilier. Inaccessible worlds and irrelevance: Preliminary report. In *Proceedings of the Twelfth International Joint Conference on Artificial Intelligence*, pages 413–418, Sydney, 1991.

Craig Boutilier. Normative, subjunctive and autoepistemic defaults: Adopting the Ramsey test. In *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning*, pages 685–696, Cambridge, 1992.

Craig Boutilier. What is a default priority? In *Proceedings of Canadian Society for Computational Studies of Intelligence Conference*, pages 140–147, Vancouver, 1992.

Craig Boutilier. A modal characterization of defeasible deontic conditionals and conditional goals. In *AAAI Spring Symposium on Reasoning about Mental States: Formal Theories and Applications*, pages 30–39, Stanford, 1993.

Craig Boutilier. Conditional logics of normality: A modal approach. *Artificial Intelligence*, 68:87–154, 1994.

Craig Boutilier. Unifying default reasoning and belief revision in a modal framework. *Artificial Intelligence*, 68:33–85, 1994.

Craig Boutilier and Richard Dearden. Using abstractions for decision-theoretic planning with time constraints. In *Proceedings of the Twelfth National Conference on Artificial Intelligence*, pages 1016–1022, Seattle, 1994.

Michael E. Bratman, David J. Israel, and Martha E. Pollack. Plans and resource-bounded practical reasoning. *Computational Intelligence*, 4:349–355, 1988.

Roderick M. Chisholm. Contrary-to-duty imperatives and deontic logic. *Analysis*, 24:33–36, 1963.

Philip R. Cohen and Hector J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42:213–261, 1990.

Thomas Dean and Michael Wellman. *Planning and Control*. Morgan Kaufmann, San Mateo, 1991.

Lisa Dent, Jesus Boticario, John McDermott, Tom Mitchell, and David Zabowski. A personal learning apprentice. In *Proceedings of the Tenth National Conference on Artificial Intelligence*, pages 96–103, San Jose, 1992.

Jon Doyle and Michael P. Wellman. Preferential semantics for goals. In *Proceedings of the Ninth National Conference on Artificial Intelligence*, pages 698–703, Anaheim, 1991.

Richard E. Fikes and Nils J. Nilsson. Strips: A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 2:189–208, 1971.

Simon French. *Decision Theory*. Halsted Press, New York, 1986.

Moisés Goldszmidt. Qualitative probabilities: A normative framework for commonsense reasoning. Technical Report R-190, University of California, Los Angeles, October 1992. Ph.D. thesis.

Moisés Goldszmidt, Paul Morris, and Judea Pearl. A maximum entropy approach to nonmonotonic reasoning. In *Proceedings of the Eighth National Conference on Artificial Intelligence*, pages 646–652, Boston, 1990.

Moisés Goldszmidt and Judea Pearl. On the consistency of defeasible databases. *Artificial Intelligence*, 52:121–149, 1991.

Moisés Goldszmidt and Judea Pearl. Reasoning with qualitative probabilities can be tractable. In *Proceedings of the Eighth Conference on Uncertainty in AI*, pages 112–120, Stanford, 1992.

Peter Haddawy and Steve Hanks. Representations for decision-theoretic planning: Utility functions for deadline goals. In *Proceedings of the Third International Conference on Principles of Knowledge Representation and Reasoning*, pages 71–82, Cambridge, 1992.

Bengt Hansson. An analysis of some deontic logics. *Noûs*, 3:373–398, 1969.

John F. Horty. Moral dilemmas and nonmonotonic logic. *Journal of Philosophical Logic*, 1993. To appear.

Ronald A. Howard. *Dynamic Probabilistic Systems*. Wiley, New York, 1971.

G. E. Hughes and M. J. Cresswell. *An Introduction to Modal Logic*. Methuen, London, 1968.

I. L. Humberstone. Inaccessible worlds. *Notre Dame Journal of Formal Logic*, 24(3):346–352, 1983.

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

Andrew J. I. Jones and Ingmar Pörn. On the logic of deontic conditionals. In *Workshop on Deontic Logic in Computer Science*, Amsterdam, 1991.

Hector J. Levesque. All I know: A study in autoepistemic logic. *Artificial Intelligence*, 42:263–309, 1990.

David Lewis. *Counterfactuals*. Blackwell, Oxford, 1973.

Barry Loewer and Marvin Belzer. Dyadic deontic detachment. *Synthese*, 54:295–318, 1983.

Pattie Maes and Robyn Kozierok. Learning interface agents. In *Proceedings of the Eleventh National Conference on Artificial Intelligence*, pages 459–465, Washington, D.C., 1993.

John McCarthy and P.J. Hayes. Some philosophical problems from the standpoint of artificial intelligence. *Machine Intelligence*, 4:463–502, 1969.

J. Milnor. Games against nature. In R. Thrall, C. Coombs, and R. Davis, editors, *Decisions Processes*, pages 49–59. Wiley, New York, 1954.

Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, San Mateo, 1988.

Judea Pearl. System Z: A natural ordering of defaults with tractable applications to default reasoning. In M. Vardi, editor, *Proceedings of Theoretical Aspects of Reasoning about Knowledge*, pages 121–135. Morgan Kaufmann, San Mateo, 1990.

Judea Pearl. A calculus of pragmatic obligation. In *Proceedings of the Ninth Conference on Uncertainty in Artificial Intelligence*, pages 12–20, Washington, D.C., 1993.

Martha E. Pollack. The uses of plans. *Artificial Intelligence*, 57:43–68, 1992.

David Poole. Decision-theoretic defaults. In *Proceedings of Canadian Society for Computational Studies of Intelligence Conference*, pages 190–197, Vancouver, 1992.

Martin L. Puterman. Markov decision processes. In D. P. Heyman and M. J. Sobel, editors, *Handbooks in Operations Research and Management Science*, pages 331–434. North-Holland, Amsterdam, 1990.

Frank P. Ramsey. Turth and probability. In R. B. Braithwaite, editor, *The Foundations of Mathematics and Other Logical Essays*, pages 156–198. Kegan Paul, London, 1931.

Leonard J. Savage. The theory of statistical decision. *Journal of the American Statistical Association*, 46:55–67, 1986.

Yoav Shoham. *Reasoning About Change: Time and Causation from the Standpoint of Artificial Intelligence*. MIT Press, Cambridge, 1988.

Sek-Wah Tan and Judea Pearl. Specification and evaluation of preferences for planning under uncertainty. In *Proceedings of the Fourth International Conference on Principles of Knowledge Representation and Reasoning*, Bonn, 1994. To appear.

John von Neumann and Oskkar Morgenstern. *Theory of Games and Economic Behavior*. Princeton University Press, Princeton, 1944.

Georg Henrik von Wright. Deontic logic. *Mind*, 60:1–15, 1951.

Georg Henrik von Wright. A new system of deontic logic. In Risto Hilpinen, editor, *Deontic Logic: Introductory and Systematic Readings*, pages 105–120. D.Reidel, Dordecht, 1964. 1981.

A. Wald. *Statistical Decision Functions*. Wiley, New York, 1950.

Acknowledgements

Thanks to Carlos Alchourrón, Jon Doyle, Keiji Kanazawa, Judea Pearl, David Poole and Michael Wellman for very valuable discussion of this topic. This research was supported by NSERC Research Grant OGP0121843.

A Proof of Completeness Theorem

In this appendix we provide the proof of soundness and completeness of the axiomatization of QDT. We note that QDT consists of the axioms describing the bimodal logic CO for each of the pairs of connectives $\Box_P$, $\overleftarrow{\Box}_P$ and $\Box_N$, $\overleftarrow{\Box}_N$, together with the interaction axiom $\overleftrightarrow{\Box}_N A \equiv \overleftrightarrow{\Box}_P A$. Soundness of the initial axioms is proved in exactly the same manner as it is for CO (see (2) for complete details). We simply must show that the new axiom is sound as well.

Lemma A.1 *QDT is sound with respect to the class of QDT-models.*

Proof We simply show that $\overleftrightarrow{\Box}_N A \equiv \overleftrightarrow{\Box}_P A$ holds at all worlds in all models. Let $M = \langle W, \leq_P, \leq_N, \varphi\rangle$ be a QDT-model. $M \models_w \overleftrightarrow{\Box}_N A$ iff $M \models_v A$ for all $v \in W$ iff $M \models_w \overleftrightarrow{\Box}_P A$. ∎

To show completeness we adopt a technique of Humberstone (2). We first note, following an analogous result for CO proved in (2), that the all instances of the following axiom schemata are derivable. Let **HP*** denote the (set of) schemata

HP* $\mathcal{D}(\Box_P A \wedge \overleftarrow{\Box}_P B) \supset \mathcal{B}(A \vee B)$

where $\mathcal{D}$ is any sequence of the connectives $\Diamond_P$ and $\overleftarrow{\Diamond}_P$ having length ≥ 0, and $\mathcal{B}$ is any such sequence of $\Box_P$ and $\overleftarrow{\Box}_P$. Let **HN*** denote the same schemata with $\mathcal{D}$ standing for sequences of $\Diamond_N$ and $\overleftarrow{\Diamond}_N$, and $\mathcal{B}$ standing for sequences of $\Box_N$ and $\overleftarrow{\Box}_N$.

Lemma A.2 *Any instance of the following Humberstone schemata* **HP*** *or* **HN*** *is derivable in QDT.*

We can now prove completeness.

Lemma A.3 *If* $\models_{QDT} A$ *then* $\vdash_{QDT} A$.

Proof To show completeness it is sufficient to show that A is falsifiable for any non-theorem A. Letting Γ be some maximal QDT-consistent set (MCS) which contains $\neg A$, we will construct a model $M = \langle W, \leq_P, \leq_N, \varphi\rangle$ which falsifies A. This technique is employed in (2). The model is constructed with W consisting of MCSs and four relations $\leq_P$, $\overline{\leq}_P$, $\leq_N$, $\overline{\leq}_N$ over W. Relation $\overline{\leq}_P$ is intended to represent the complement of $\leq_P$ (and similarly for $\overline{\leq}_N$). Ultimately, $\leq_P$, $\overline{\leq}_P$ will be mutually exclusive and exhaustive on $W \times W$, as will $\leq_N$, $\overline{\leq}_N$.

The construction proceeds as follows: We start at stage 0 by adding Γ to W, so that $W = \{\Gamma\}$ and $\leq_N = \overline{\leq}_N = \leq_P = \overline{\leq}_P = \emptyset$. At each following stage i, for each set Λ added to W at stage $i-1$ we do the following:

(a) For each formula $\Diamond_P B \in \Lambda$ add a MCS Λ' where $\{B\} \cup \{C : \Box_P C \in \Lambda\} \subseteq \Lambda'$, and add $\langle \Lambda', \Lambda \rangle$ to $\leq_P$.

(b) For each formula $\overleftarrow{\Diamond}_P B \in \Lambda$ add a MCS Λ' where $\{B\} \cup \{C : \overleftarrow{\Box}_P C \in \Lambda\} \subseteq \Lambda'$, and add $\langle \Lambda', \Lambda \rangle$ to $\overline{\leq}_P$.

(c) For each formula $\Diamond_N B \in \Lambda$ add a MCS Λ' where $\{B\} \cup \{C : \Box_N C \in \Lambda\} \subseteq \Lambda'$, and add $\langle \Lambda', \Lambda \rangle$ to $\leq_N$.

(d) For each formula $\overleftarrow{\Diamond}_N B \in \Lambda$ add a MCS Λ' where $\{B\} \cup \{C : \overleftarrow{\Box}_N C \in \Lambda\} \subseteq \Lambda'$, and add $\langle \Lambda', \Lambda \rangle$ to $\overline{\leq}_N$.

That such MCSs exist is claimed without proof (see (2; 2)). Now let M be the totality of this (typically infinite) construction. Evaluating the truth conditions of $\overleftarrow{\Box}_P$ with respect to $\overline{\leq}_P$ (as if $\overline{\leq}_P$ were the complement of $\leq_P$), and those of $\overleftarrow{\Box}_N$ with respect to $\overline{\leq}_N$, we can show the following, assuming $\varphi(A) = \{w : A \in w\}$ for atomic A.

Lemma A.4 $M \models_w B$ *iff* $B \in w$.

Proof We proceed by induction on the structure of B. For atomic B, this follows by the definition of φ. Assuming this for A and B, clearly it holds for both $\neg A$ and $A \supset B$ by standard properties of MCSs. Now suppose $\Box_P B \in w$. By the construction of M, for all $v \leq_P w$, $M \models_v B$, therefore $M \models_w \Box_P B$. If $\Box_P B \notin w$, then $\Diamond_P \neg B \in w$. By the construction of M, there is some $v \leq_P w$ such that $M \models_v \neg B$, therefore $M \not\models_w \Box_P B$. The same argument holds for $\overleftarrow{\Box}_P B$, assuming $\overline{\leq}_P$ to be the complement of $\leq_P$. Similar considerations apply to formulae of the form $\Box_N B$ and $\overleftarrow{\Box}_N B$. ■

Now we have a "structure" which falsifies A, as $\neg A \in \Gamma$ and by the above, $M \models_\Gamma \neg A$. However, M is not a QDT-model, since $\leq_P$ and $\leq_N$ are neither reflexive, transitive, nor connected; and $\overline{\leq}_P$ and $\overline{\leq}_N$ are not the complements of these relations. We now show that $\leq_P$, $\overline{\leq}_P$ and $\leq_N$, $\overline{\leq}_N$ can be extended in such a way that $\leq_P$, $\leq_N$ possess the desired properties and $\overline{\leq}_P$, $\overline{\leq}_N$ are their complements on W, while not changing the fact that $M \models_w B$ iff $B \in w$.

First we show that the MCSs selected in the construction process can be chosen in such a way that the sets added in steps (a) and (b) — which are added to the relation $\leq_P$ and $\overline{\leq}_P$ — can be added to either $\leq_N$ or $\overline{\leq}_N$; and similarly, for steps (c) and (d), the choices can be added to $\leq_P$ or $\overline{\leq}_P$. Suppose at stage i, step (a), we must find a MCS Λ' where $\{B\} \cup \{C : \Box_P C \in \Lambda\} \subseteq \Lambda'$, and add $\langle \Lambda', \Lambda \rangle$ to $\leq_P$. We wish to ensure that Λ' is chosen so that $\langle \Lambda', \Lambda \rangle$ can consistently be added to one of $\leq_N$ or $\overline{\leq}_N$. If no such Λ' exists, then both

$$\{B\} \cup \{C : \Box_P C \in \Lambda\} \cup \{D : \Box_N D \in \Lambda\} \text{ and } \{B\} \cup \{C : \Box_P C \in \Lambda\} \cup \{D : \overleftarrow{\Box}_N D \in \Lambda\}$$

are inconsistent. Thus, $\{B\} \cup \{C : \Box_P C \in \Lambda\} \vdash \neg D_1 \wedge \neg D_2$,
for some $D_1 \in \{D : \Box_N D \in \Lambda\}$, $D_2 \in \{D : \overleftarrow{\Box}_N D \in \Lambda\}$. Thus $\neg \Box_P (D_1 \vee D_2) \in \Lambda$. But clearly $\Box_N (D_1 \vee D_2) \in \Lambda$ and $\overleftarrow{\Box}_N (D_1 \vee D_2) \in \Lambda$, so $\overleftrightarrow{\Box}_N (D_1 \vee D_2) \in \Lambda$. By axiom **PN**, $\overleftrightarrow{\Box}_P (D_1 \vee D_2) \in \Lambda$; so $\Box_P (D_1 \vee D_2) \in \Lambda$, contradicting the inconsistency. So some Λ' can be chosen such that $\langle \Lambda', \Lambda \rangle$ can consistently be added to one of $\leq_N$ or $\overline{\leq}_N$. An analogous argument holds for steps (b), (c) and (d). As a result, every MCS added to W is "connected" to the original MCS Γ by some number of steps through $\leq_P \cup \overline{\leq}_P$ and by some number of steps through $\leq_N \cup \overline{\leq}_N$.

We now show that $\overline{\leq}_P$ can be made the complement of $\leq_P$, and that $\leq_P$ can be made reflexive, transitive and connected. Similar arguments holds for $\leq_N$ and $\overline{\leq}_N$.

Suppose that $\langle v, w \rangle \not\in \leq_P$ and $\langle v, w \rangle \not\in \overline{\leq}_P$, and that it cannot be "consistently" added to either of $\leq_P$ or $\overline{\leq}_P$. Then there must be some $\Box_P B \in w$, $B \not\in v$ and some $\overleftarrow{\Box}_P C \in w$, $C \not\in v$. Both w and v must be some finite "distance" away from out starting point Γ, say m and n "steps", respectively, through $\leq_P \cup \overline{\leq}_P$. Following the "path" which lead to the addition of w to W, we have $M \models_\Gamma \mathcal{D}_1(\Box_P B \wedge \overleftarrow{\Box}_P C)$ where $\mathcal{D}_1$ is a string of m $\Diamond_P$'s and $\overleftarrow{\Diamond}_P$'s (depending on how w was added). Similarly, $M \models_\Gamma \mathcal{D}_2(\neg B \wedge \neg C)$ where $\mathcal{D}_2$ is the string of n $\Diamond_P$'s and $\overleftarrow{\Diamond}_P$'s corresponding to how v was added. But this sentence is equivalent to $\neg \mathcal{B}_2(B \vee C)$, where $\mathcal{B}_2$ is formed by replacing $\Diamond_P$ and $\overleftarrow{\Diamond}_P$ with $\Box_P$ and $\overleftarrow{\Box}_P$ (respectively) in $\mathcal{D}_2$. This means both $\mathcal{D}_1(\Box_P B \wedge \overleftarrow{\Box}_P C) \in \Gamma$ and $\neg \mathcal{B}_2(B \vee C) \in \Gamma$, contradicting the Humberstone schema. Since Γ is consistent, $\langle v, w \rangle$ can be added to either $\leq_P$ or $\overline{\leq}_P$ without affecting the truth of formulae at any world in W, and hence $\leq_P$ and $\overline{\leq}_P$ can be extended to complement one another, making valuation of $\overleftarrow{\Box}_P$ with respect to $\overline{\leq}_P$ the same as valuation with respect to the standard truth conditions.

We can ensure that $\leq_P$ is reflexive, as well. Adding $w \leq_P w$ affects the truth of some sentence only if there is some B such that $\Box_P B \in w$ and $B \not\in w$; but this contradicts the axiom **T** and the fact that w is a MCS.

For transitivity, suppose $v \leq_P w$ and $t \leq_P v$. Adding $t \leq_P w$ can only affect truth if there is some $\Box_P B \in w$ and $B \not\in t$. Since $\Box_P B \in w$, by axiom **4**, $\Box_P \Box_P B \in w$. This means

$\Box_P B \in v$ and $B \in t$, contradicting the assumption.

For completeness, suppose $v \overline{\leq}_P w$. Adding $w \leq_P v$ can affect truth only if there is some $\Box_P B \in v$ and $B \notin w$. If $B \notin w$, then $\neg B \in w$ and by axiom **S**, $\overline{\Box}_P \Diamond_P \neg B \in w$. Now since $v \overline{\leq}_P w$, $\Diamond_P \neg B \in v$, and $\Box_P B \notin v$, contradicting the original assumption.

It is clear that there may be some interaction during these "steps" whereby certain pairs of worlds are moved from the set $\overline{\leq}_P$ to $\leq_P$; but, clearly nothing in principle stops one from constructing a suitable model with the appropriate constraints being fulfilled by the relations. In fact, if we insist that $\leq_P$ be completed maximally before we complete $\overline{\leq}_P$, there need not be any interaction. For instance, at the step where we decide to add each pair of worlds to $\leq_P$ or $\overline{\leq}_P$, we can consider the union of the family of all possible relations $\leq_P$ on $W \times W$ that respect on restrictions on the ordering; we take this set to be $\leq_P$ and let $\overline{\leq}_P$ then be $W \times W - \leq_P$.

Since similar considerations apply to $\leq_N$, we can construct a QDT-model which falsifies the non-theorem A. ■

Theorem 3.1 *The system* QDT *is characterized by the class of QDT-models.*

Proof This follows immediately from Lemmas A.1 and A.3. ■

B System Z

In the following, a copnditional can be taken to be either $I(B|A)$ or $A \Rightarrow B$, and the constraints are imposed on the appropriate ordering. We use $\Rightarrow$ generally in the description of System Z. Pearl (2) describes a natural ordering on default rules named the *Z-ordering*, and uses this to define a nonmonotonic entailment relation, 1-entailment, put forth as an extension of ε-semantics (2). Pearl's default rules r have the form $A \rightarrow B$, where A and B are propositional. As shown in (2), Pearl's system of ε-semantics and our conditional logic agree precisely on their conditional fragments, so we may apply his definitions directly to conditionals in CO. We say a valuation (possible world) w *verifies* the rule $A \Rightarrow B$ iff $w \models A \wedge B$, *falsifies* the rule iff $w \models A \wedge \neg B$, and *satisfies* the rule iff $w \models A \supset B$. For any rule $r = A \Rightarrow B$, we define r^* to be its *material counterpart* $A \supset B$. We assume that T is a finite set of such rules.

Definition 2.1 **(2)** T *tolerates* $A \Rightarrow B$ iff there is some world that verifies $A \Rightarrow B$, and falsifies no rule in T; that is, $\{A \wedge B\} \cup \{C \supset D : C \Rightarrow D \in T\}$ is satisfiable.

Toleration can be used to define a natural ordering on conditionals by partitioning T:

Definition 2.2 (2) For any $i \geq 0$ we define $T_i = \{r : r \text{ is tolerated by } T - T_0 - T_1 - \cdots T_{i-1}\}$,

Assuming T is consistent, this results in an ordered partition $T = T_0 \cup T_1 \cup \cdots T_n$. Now to each rule $r \in T$ we assign a rank (the *Z-ranking*): $Z(r) = i$ whenever $r \in T_i$. Roughly, the idea is that lower ranked rules are more general, or have lower priority. Given this ranking, we can rank possible worlds according to the highest ranked rule they falsify:

$$Z(w) = \min\{n : w \text{ satisfies } r, \text{ for all } r \in T, Z(r) \geq n\}.$$

Lower ranked worlds are to be considered more normal (or more ideal); thus Z-ranking determines a unique preferred structure. Now any proposition A can be ranked according to the lowest ranked world that satisfies it; that is

$$Z(A) = \min\{Z(w) : w \models A\}.$$

Given that lower ranked worlds are considered more normal, we can say that a conditional $A \Rightarrow B$ should hold iff the rank of $A \wedge B$ is lower than that of $A \wedge \neg B$. Hence we have:

Definition 2.3 (2) Formula B is *1-entailed* by A with respect to T (written $A \vdash_1 B$) iff $Z(A \wedge B) < Z(A \wedge \neg B)$ (where Z is determined by T).

For details regarding the types of conclusions sanctioned by 1-entailment, see (2).