THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Preference Change

OPEN ACCESS

Preference Change

**Anaïs Cadilhac · Nicholas Asher ·
Alex Lascarides · Farah Benamara**

Abstract Most models of rational action assume that all possible states and actions are pre-defined and that preferences change only when beliefs do. But several decision and game problems lack these features, calling for a *dynamic* model of preferences: preferences can change when unforeseen possibilities come to light or when there is no specifiable or measurable change in belief. We propose a formally precise dynamic model of preferences that extends an existing static model (Boutilier et al, 2004). Our axioms for updating preferences preserve consistency while minimising change, like Hansson's (1995). But unlike prior models of preference change, ours supports default reasoning with partial preference information, which is essential to handle decision problems where the decision tree isn't surveyable. We also show that our model avoids problems for other models of preference change discussed in Spohn (2009).

Anaïs Cadilhac
IRIT, Université Paul Sabatier, 118 Route de Narbonne, F-31062 Toulouse Cedex 9, France
Tel.: +33 5 61 55 73 43, Fax: +33 5 61 55 62 58, E-mail: cadilhac@irit.fr

Nicholas Asher
IRIT, Université Paul Sabatier, 118 Route de Narbonne, F-31062 Toulouse Cedex 9, France
Tel.: +33 5 61 55 77 05, Fax: +33 5 61 55 62 58, E-mail: asher@irit.fr

Alex Lascarides
School of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh, EH8 9AB, Scotland, UK
Tel.: +44 131 650 4428, Fax: +44 131 650 6626, E-mail: alex@inf.ed.ac.uk

Farah Benamara
IRIT, Université Paul Sabatier, 118 Route de Narbonne, F-31062 Toulouse Cedex 9, France
Tel.: +33 5 61 55 82 98, Fax: +33 5 61 55 62 58, E-mail: benamara@irit.fr

1 Introduction

Standard decision and game theory encode preferences with a static, predefined and complete reward function, assigning each end state of the game or decision problem (and sometimes intermediate states too) a numerical reward (Simon, 1955). In such models, to change these "intrinsic" preferences is simply to change the game or decision problem, and analyzing preference change as a logic of change to one's decision problem involves defining a game or decision problem over games. Spohn (2009) shows that such an approach is flawed, because it cannot distinguish between intuitively different decision problems. It also doesn't address preference change when agents discover unforeseen possibilities or when their beliefs don't change.

This paper develops an axiomatic theory of preference change that addresses problems with prior models and that we hope will be of use in analyzing game and decision problems. Our model of preference change, like Hansson's (1995), resembles an AGM model for belief change (Alchourrón et al, 1985): updating old preferences with new ones preserves consistency while minimising change. But unlike Hansson (1995), we analyze logical interactions between preferences and beliefs, an essential feature for decision and game theory. Our model crucially exploits *default* reasoning with *partial* preference information in contrast to Hansson (1995); Liu (2008, 2011); van Benthem et al (2009). We show that the problems motivating our account require this.

Section 2 uses natural decision problems to motivate a dynamic rather than a static utility function. Section 3 evaluates existing dynamic models of preferences, and uses that to come up with a set of desiderata, as described in section 4. We present our formal model in section 5. Section 5.4 shows how our model avoids Spohn's problems for other models of preference change.

2 Motivation: example decision problems

Strict preference is an asymmetric, transitive and irreflexive relation for an agent over *outcomes*, which include actions (e.g., to buy a car) and states (e.g., to own a car). Preferences can be *intrinsic* or *extrinsic* (Spohn, 2009). Intrinsic preferences are not based on other considerations. So you can prefer states that are inconsistent with reality or impossible to achieve; e.g., you might prefer to be healthy even though you're terminally ill. The reward function of decision theory captures intrinsic preferences: it specifies the extent to which the agent finds a state attractive without regard to their beliefs about whether the state is achievable or what subsequent states are reachable from it.

Extrinsic preferences depend on *beliefs*: if X is preferred to Y because the agent believes that X makes achieving some other desirable outcome Z more likely, then X is extrinsically preferred to Y. For example, you could prefer eating fish over meat not because you like fish better but because you believe avoiding meat will help lower your cholesterol. Decision and game theory capture extrinsic preferences via *expected utility* (Bellman, 1957): the

expected utility of an action sequence $\mathbf{a}$ is the weighted average of the reward at each state s' that is reachable via $\mathbf{a}$, with each weight being determined by a probabilistic belief model $P(s'|\mathbf{a}, \mathbf{e})$ of how likely s' is if $\mathbf{a}$ is performed given the evidence $\mathbf{e}$, possibly complicated by a discounting factor that takes into account how long it takes to achieve the reward. Actions that maximize expected utility reflect a trade off between what agents intrinsically prefer and what they believe they can achieve. Expected utilities are rightly dynamic: as an agent's observations $\mathbf{e}$ change, so do the beliefs $P(s'|\mathbf{a}, \mathbf{e})$ and hence also the expected utilities. But for classical theories, extrinsic preferences change *only* when beliefs do and intrinsic preferences don't change. We'll show this isn't always the right picture.

2.1 Extrinsic preferences aren't always computable

In game theory, players should play an *equilibrium strategy*, which specifies a strategy for each player from which no player would unilaterally deviate—i.e., each player expects at least as good a payoff from their own strategy as any other strategy they could adopt, assuming that all the other players adhere to their specified strategies (Savage, 1954). Equilibrium strategies thus depend on the preferences of every agent.

In many games an agent doesn't know the preferences of others, an uncertainty that game theory models via a probability distribution over the possible *types* of the other players, each type being associated with a complete, static reward function. All algorithms for identifying optimal actions, like *backward induction* (Shoham and Leyton-Brown, 2009, p.119) or approximate solutions like *Monte Carlo Tree Search (MCTS)* (Browne et al, 2012), require agents to know all possible states, actions and player types—any hidden information must be a foreseen possibility. For some games, however, agents may start out unaware of all possible player types. If so, then even if the (intrinsic) rewards over end states are known, agents cannot exploit backward induction or MCTS because these methods require the hypothesis space to be complete and pre-defined. Not knowing what is possible is more serious than not knowing the true value from a known set of possibilities, for which well known solutions exist—we'll call games where the agent doesn't know all possibilities games of *strongly incomplete* information.

The set of possible player types may be unknown because the agent doesn't know all the possible actions other players contemplate. Some win-lose board games that involve negotiations have this feature (e.g., *The Settlers of Catan*; see `catan.com` for its rules). The game tree is non-surveyable because there are an unbounded number of possible trades: agents can promise a particular future move under arbitrary conditions as a part of the trade (e.g., *If you trade clay for wood now, I will give you wheat when I get it but only if you don't block me*). Since negotiations make the game tree non-surveyable, players must optimize on a subpart of it. But they don't know which subparts the other players isolate for performing their calculations. Thus the game is

strongly incomplete: the players lack complete knowledge of the set of possible player types, because they do not know which actions are a part of the opponents' decision problem. Such games call for a *dynamic* function from states to preferences because players must adapt the state space they optimize over when something unforeseen happens (Degremont et al, 2014): in our Settlers example, an agent needs to optimize his reaction to an unforeseen trade offer. So agents must defeasibly infer preferences over intermediate states and have ways of revising them when justified by the evidence.

2.2 Intrinsic preferences can change

Section 2.1 showed that standard algorithms for optimization don't handle all changes to extrinsic preferences exhibited in games of strongly incomplete information. Here, we show such games must support changes to *intrinsic* preferences too (Hansson, 1995). Suppose that an agent has never heard of nor tasted turmeric. And suppose that initially she isn't all that keen on soup and would rather eat fish. One day she has a soup that she finds delicious. She asks about its ingredients and so discovers turmeric, to which she takes an immediate liking. Importantly, her *intrinsic* preferences change: she finds turmeric goes well with soup but not with fish, and she would rather have soup with turmeric than fish (with or without turmeric). Now there are possible end states in her decision problem of which she was initially unaware (i.e., states where she eats soup with turmeric), and these states turn out to be most preferred. Crucially, the new reward function is more than an *extension* of the old one—in our example, all states where the agent eats fish initially had a higher reward than those where she eats soup, but now eating fish has a lower reward than eating soup *unless* turmeric isn't available.

Desires and preferences are parasitic on beliefs in that one can't desire an object if one has no idea of that object; there is no *de re* desire without *de re* beliefs (Asher, 1987; Heim, 1992). Agents must thus formulate their preferences over new concepts as and when they're discovered and not before, and in doing so they may *revise* (and not merely extend) existing preferences over the already known possibilities.

Arguably, this is preference revision triggered by belief change—i.e. the discovery of unforeseen possibilities (unawareness logics (e.g., Modica and Rustichini (1999)) might be useful here). But existing game or decision theoretic models like Markov Decision Processes (MDPs) (Bellman, 1957) don't handle this sort of belief change. The MDP's variables and dependencies don't change over time; but in our example, the discovery of *novel state descriptions* leads the agent to update her belief model to include new random variables and/or new values for existing variables and/or revised dependencies. In other words, the language for describing the decision problem has changed. MDPs also characterize intrinsic preferences with a static reward function. But in our example, discovering novel options prompts the agent to revise her intrinsic

preferences: the initial global preference for fish over soup is retracted. MDPs don't capture this sort of preference dynamics either.

Agents can also change their (intrinsic) preferences in the absence of any new discovery or belief change. Suppose that an agent A prefers to smoke rather than not to smoke. Let the propositional variable s stand for "A smokes" and $\overline{s}$ its negation, and let $\succ$ be the preference relation: so $s \succ \overline{s}$. Moreover, A may prefer a peaceful and healthy life (p) over a non-peaceful life in which her friends nag her ($\overline{p}$). So $p \succ \overline{p}$. But A also believes that s (defeasibly) implies $\overline{p}$—i.e., the two global preferences for s and p cannot be reconciled, given the beliefs. So suppose A decides to that p is more important to her than s. That is, A's preference for p over $\overline{p}$ is unchanged, but she abandons the *global* preference $s \succ \overline{s}$ for a derivative or conditional one: given p, she prefers $\overline{s}$; but if p is false, then she would prefer s—she still craves a cigarette. Now suppose A changes even further. After months of not smoking, A's desire to smoke has changed to an aversion to smoke. That is, A replaces her conditional preference for smoking with a global preference $\overline{s} \succ s$. The change from the conditional preference for s to a global one for $\overline{s}$ is an example of intrinsic preference change without rational deliberation, knowledge discovery or belief change. A change in taste may be responsible for this preference change, borne perhaps from a change in habits. Other examples include: people who become vegetarians for health reasons, who miss bacon at first but who over time lose their taste for it; and children who change their tastes as they mature.

3 Motivation: Existing work on preference change

Previous work has classified preference change according to its etiology— e.g., preference change due to a change in belief or taste or the environment (Bradley, 2007; Lang and van der Torre, 2008). The focus of this paper is not what causes one to adopt a new preference, but rather how one updates existing preferences with a new preference, whatever its cause.

Spohn (2009) describes a general approach to modeling preference change involving *global decision models*, which construct a decision (or game) over a set of decisions (or games). He argues that almost all received models of preference change can be articulated this way, but he also criticizes the approach by showing that these models don't discriminate between certain decision problems where intuitively the optimal actions differ.

A global decision model is like a normal decision tree except that in addition to chance and end nodes and actions that link nodes, it contains also *agent nodes*, with each agent node being the root of a different local decision model. The local decision trees can differ in their set of actions, the actions' likely outcomes, and the rewards of the end states. Varying the rewards corresponds to intrinsic preference change. The outcome of an action can be a chance node, agent node or end node, and so a global model can represent *foreseen* preference change, with the agents choosing actions that affect the type of

agent they will be (it cannot deal with unforeseen preference change of the kind we discussed in section 2).

One of Spohn's minimal pair of decision problems that form the basis of his criticism of global decision models involves intrinsic preference change. In both of his scenarios, the agent forms a preference at an initial time, and he can either decide to act on it (pre-empting future 'agents' with perhaps distinct preferences from performing actions) or he can wait and see whether time changes his preference. In the **holiday** scenario, the player p starts with a preference to go on holiday (b_1) over not going (b_2). But first, p must choose between booking the holiday immediately (a_1), or waiting until the next morning (a_2), in which case there's only a 50% chance that p will still think that b_1 is worth the price. There is no objective way of deciding which of a_1 vs. a_2 is optimal. But intuitively, it seems reasonable for p to mistrust his excitement now (especially given the cost) and to sleep on it; i.e., a_2 is intuitively optimal.

In the second, **market** scenario, p is at a market and believes that the goods on offer are never worth the money that the hawkers demand, nor even the price demanded at the end of a long bargaining process. So initially, the agent prefers not to buy (b_1) over buying (b_2). However, the hawkers are persistent and p must either close his mind to their offers (a_1) and thus stick to his initial preference for b_1, or p listens to them (a_2) and risks, with a 50% chance, being talked into preferring b_2 over b_1. This time, intuitively the optimal action is to ignore the hawkers; i.e., a_1 is optimal over a_2. But the holiday and market scenarios have isomorphic global decision models—see Figure 1. This provides insufficient information for distinguishing their diverging optimal strategies.
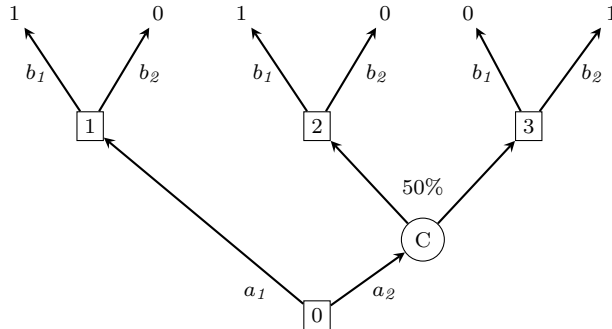
Fig. 1 The global decision model for the holiday and market scenarios. Agent nodes are square and chance nodes are round.

The moral we draw from Spohn's discussion is that global decision models have an inherent problem: they don't offer direct rewards on intermediate nodes, particularly the agent nodes. Thus, they cannot encode the intuition that to act optimally you should be an agent that's not impulsive. In both stories, to avoid impulsive behavior is a *global intrinsic preference*: but this is a property of the *type of agent* one is, not a property of the end states one

might end up in. The preference for a_1 vs. a_2 is dependent on this global preference: for the holiday scenario, if you want to avoid acting on impulse, then prefer a_1 over a_2; for the market, avoiding impulse yields a preference for a_2 over a_1. Global decision models don't express such dependencies, and don't express intrinsic preferences for particular agent types at all. Assigning preferences to intermediate nodes will be a feature of our partial models of preference.

There are several modal characterisations of preference change (Liu, 2008, 2011; van Benthem et al, 2009). They define preference orders over a fixed set of final, independent outcomes, and so they don't address the problem that Spohn points to, although in principle they could. But modeling preference change when an unforeseen possibility is discovered is problematic in these frameworks. Since Liu's dynamic operations change the ordering over a fixed set of outcomes but cannot change the set of outcomes, it can handle preference change due to discovering an unforeseen possibility only through interaction with a model of belief that *doesn't* validate full introspection (i.e., $\neg\mathcal{B}\phi$ doesn't entail $\mathcal{B}\neg\mathcal{B}\phi$). Liu (2008, 2011) provides a such a belief model but its interaction with intrinsic preference change isn't fully explored. Furthermore, the theory puts very high informational demands on the person who is modeling the unaware agent: that person must know all possibilities in advance (even if the agent they are modeling does not), and this isn't always feasible or practical. For instance, section 2 discussed game trees that aren't enumerable—here, this would lead to a non-enumerable set of possible orderings over all possible states.

To address these problems we use *partial models of preference* and a proof theory that supports *defeasible inference*. Our approach allows an agent to make inferences about action without being aware of all possible outcomes, and it naturally allows preferences to order intermediate states in a game or decision. Defeasible inference can handle this level of 'ignorance' because it provides a way of expressing concise preference statements and using them to predict optimal actions even when there are an unbounded (and unknown) set of exceptions to such preference statements.

In section 5, we use defeasible inference to construct and to exploit preference models in decision making. Our method is simpler than Liu's: agents (defeasibly) optimize their behavior with respect to a relatively small and inexpressive preference model that uses only those factors that the agents are aware of and believe are relevant. The agents will add further possible options only when current evidence suggests that they are needed for behaving optimally. Such an approach is inherently defeasible: we saw in section 2 several examples where new observable information is inconsistent with the existing preference model, making retraction necessary to preserve consistency—a consistent preference relation being important for the model to have any predictive power at all (see section 4). In section 5.4 we'll see how partial preference models and defeasible reasoning can distinguish the optimal behaviors in Spohn's holiday and market examples.

Grüne-Yanoff and Hansson (2009) discuss several models of preference change, focussing on whether they preserve *consistency*. Hansson (1995) takes an AGM style approach to revising preferences: one retracts a minimal amount of existing preferences to make it consistent with new information. Hansson's model countenances the discovery of new preferences over previously unknown options, and so his account doesn't fit into the general approach afforded by global decision models. On the other hand, it also doesn't do everything that a global decision model or the other modal approaches such as Liu's (2008) can do: it does not model the interaction between beliefs and preferences at all, and so it cannot model classic decision problems, where extrinsic preferences change because beliefs do. Our approach to preference change follows Hansson in preserving consistency of the preference relation (see section 4) and adopting an AGM style approach to preference change. But we aim to handle both extrinsic and intrinsic preference change.

Andréka et al (2002) offer a complex theory of preference amalgamation and change using the notion of prioritized graphs. They offer a very interesting algebraic treatment for combining preference orders and modeling preferential entailment. But their definitions presuppose a fixed set of outcomes. They also do not investigate the interaction between belief and preference change, though their notion of refinement could be highly relevant to a problem of revision as argued in Liu (2008, 2011).

To summarize, several existing theories advocate preference change in the absence of any belief change. But these theories either do not handle decision problems involving the discovery of unforeseen outcomes (Spohn, 2009; Liu, 2008, 2011; van Benthem et al, 2009) or they don't model any interaction between preference and belief change (Hansson, 1995). Furthermore, none of these models support defeasible reasoning with partial models of preference, which, we've argued, is an irreducible feature of decision making in complex games or games of strongly incomplete information. The next sections provide a formally precise model of preference change that fills these gaps.

4 Modeling Preference Change: Desirable Features

Sections 2 and 3 suggest five desirable features of a symbolic model of preference change. First, it must allow agents to construct representations of other players' preferences from observing what they do, and this representation must be conducive to decision making. We call this desideratum **intelligibility**. Game theory satisfies this desideratum via the posterior probability distribution over player types, which gets updated via observable evidence. But in games of strongly incomplete information an agent may not know the set of possible player types—a required element of the "intelligibility" calculations in game theory. We want a model of preference change that achieves intelligibility even in games of strongly incomplete information.

The second desideratum concerns reasoning about one's own preferences as well as those of others. The logic should support inferences about how pref-

erences may persist, vanish or change as agents acquire or forget information about their situation. Even one's own preference information may be partial when, for instance, one has incomplete information about the possible end states. We therefore need to support default reasoning with **partial** information about preferences, and this partial representation of preferences must evolve as the agent's model of the state space evolves.

Third, following all prior work on preference change, we must guarantee that the preference relation is **consistent**. This is necessary for predicting optimal behavior: an inconsistent preference relation entails any preference ordering and so renders any behavior optimal. Fourth, following Hansson (1995), we want preference change to be **conservative**: one minimally retracts the old preferences required to restore consistency with the new information, so that preferences generally *persist* over time. Finally, conservativeness is balanced by a notion of **entropy**: the smoking example and Spohn's (2009) examples show that preferences can change without conscious reflection.

CP-nets (Boutilier et al, 2004), a compact representation of preferences and their dependencies, offer a suitable starting point for building a theory that meets these desiderata. Cadilhac et al (2011) have shown that representing a partial model of preferences via CP-net descriptions achieves *intelligibility*. They specify axioms for updating one's model of the opponents' preferences on the basis of what they say in conversation. Section 5 deepens these results, providing update rules for revising one's own preferences as well as those of others by observing what they do, not just what they say. By making the update rules similar to those in AGM belief revision (Alchourrón et al, 1985), we make preference revision *consistent* and *conservative*. The *preference dependencies* that are an inherent component of CP-nets provide an elegant way to encode the analogue to entrenchment in AGM belief revision; they adjudicate among alternative minimal changes to preferences Our view of revision also provides a consistent model of *entropy*—i.e., where preferences change without rational deliberation or belief change.

5 A Formal Model of Preference Change

In this section, we describe how we use CP-nets. We supply axioms that update and revise partial preference models for cases where change comes about because of a revision in belief, or from discovering new possible states or options that were not initially a part of the decision problem at all, or simply from changing taste independently of beliefs entirely.

5.1 A brief introduction to CP-nets and partial CP-nets

CP-nets (Boutilier et al, 2004) offer a compact representation of preferences. Similar to Bayes nets (Pearl, 1988), which are graphical models exploiting probabilistic conditional independence to provide a compact representation

of a joint probability distribution, CP-nets are graphical models that exploit *conditional preferential independence* to provide a compact representation of the preference order over all outcomes.

More formally, let V be a finite set of propositional variables, which defines the set of possible outcomes 2^V. Then a *preference relation* $\succeq$ is a reflexive and transitive binary relation on 2^V with strict preference $\succ$ defined in the usual way (i.e., $o \succeq o'$ but $o' \not\succeq o$). An agent is indifferent between two outcomes, written $o \sim o'$, if $o \succeq o'$ and $o' \succeq o$. Definition 1 defines *conditional preference independence* and Definition 2 defines CP-nets: the idea is that the graphical component $\mathcal{G}$ of a CP-net specifies for each variable X its *parent variables* $Pa(X)$ that affect the agent's preferences over the values of X, such that X is conditionally preferentially independent of $V \setminus (\{X\} \cup Pa(X))$ given $Pa(X)$.

Definition 1 *Let V be a set of propositional variables, each variable X_i being linked to a domain $D(X_i)$. Let $\{X, Y, Z\}$ be a partition of V. X is* **conditionally preferentially independent** *of Y given Z if and only if $\forall z \in D(Z)$, $\forall x_1, x_2 \in D(X)$ and $\forall y_1, y_2 \in D(Y)$ we have: $x_1 y_1 z \succeq x_2 y_1 z$ iff $x_1 y_2 z \succeq x_2 y_2 z$.*

Definition 2 *Let V be a set of propositional variables. $\mathcal{N}_V = \langle \mathcal{G}, \mathcal{T} \rangle$ is a* **CP-net** *on V, where $\mathcal{G}$ is a directed graph over V, and $\mathcal{T}$ is a set of Conditional Preference Tables (CPTs) with indifference. That is, $T = \{\mathrm{CPT}(X_j): X_j \in V\}$, where $\mathrm{CPT}(X_j)$ specifies for each combination of values of the parent variables $p \in 2^{Pa(X_j)}$ either $p : x_j \succ \overline{x_j}$, $p : \overline{x_j} \succ x_j$ or $p : x_j \sim \overline{x_j}$ where the $\overline{}$ symbol sets the variable to false.*

Let's illustrate these definitions with a simple example. Suppose an agent prefers to go from Paris to Hong Kong by day rather than overnight. If she takes an overnight trip, she prefers a nonstop flight, but if she goes by day she prefers a flight with a stop. The CP-net is in Figure 2. The variable T stands for the period of travel: $D(T) = \{t_d, t_n\}$, where t_d is a day trip and t_n is a night one. The variable S stands for the stops: $D(S) = \{s, \overline{s}\}$ where s is a trip with stops and $\overline{s}$ is one without.

$$\boxed{T} \qquad \mathrm{CPT(T)} = \quad t_d \succ t_n$$

$$\boxed{S} \qquad \mathrm{CPT(S)} = \quad \begin{array}{l} t_d : s \succ \overline{s} \\ t_n : \overline{s} \succ s \end{array}$$
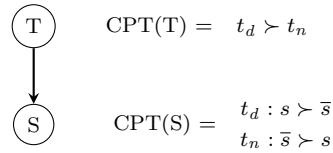
Fig. 2 Travel CP-net

The logic for inferring the preference order over all outcomes from a CP-net consists of two ranked principles. The primary principle is that violating more preference statements is worse than violating fewer of them. The secondary

principle is that violating a preference of something on which your other preferences depend is worse than violating dependent preferences. The preference order over outcomes that follows from Figure 2 is:

$$(t_d \wedge s) \succ (t_d \wedge \overline{s}) \succ (t_n \wedge \overline{s}) \succ (t_n \wedge s) \tag{1}$$

For instance, $t_d \wedge \overline{s}$ is preferred to $t_n \wedge \overline{s}$ because of the logic's secondary principle: they each violate exactly one preference statement ($t_d : s \succ \overline{s} \in CPT(S)$ and $t_d \succ t_n \in CPT(T)$ respectively), but by the secondary principle violating a statement in $CPT(T)$ is worse because T is a parent to S.

The linear *forward sweep* algorithm computes optimal outcomes for acyclic CP-nets. Forward sweep instantiates variables following an order compatible with the graph, choosing for each variable (one of) its preferred values given the value of the parents. An iterated application of the algorithm, where one removes from the sample space the outcome that was identified as preferred in the last iteration, yields the relative preferences over all outcomes (see (1)).

Cadilhac et al (2011) develop a language for (partially) describing CP-nets so as to model partial preferences. Each formula of this language is a partial description of a complete CP-net. For instance, the description language formula $y_1, \ldots y_n : x \succ \overline{x}(CPT(X))$ is satisfied by any complete CP-net $\mathcal{N}$ iff $CPT(X) \in T_{\mathcal{N}}$ contains the entry $p : x \succ \overline{x}$, where $y_1, \ldots, y_n$ are conjuncts in p (so in $\mathcal{N}$, $\{Y_i, 1 \leq i \leq n\} \subseteq Pa(X)$). A CP-net description $\mathcal{DN}$ is a set of such formulas, and $\mathcal{N} \models \mathcal{DN}$ just in case $\mathcal{N}$ satisfies each formula in $\mathcal{DN}$. Thus a partial CP-net $\mathcal{DN}$ defines a partial preference order over outcomes. $\mathcal{DN}$ may entail neither $o \succeq o'$ nor $o' \succeq o$, making o and o' incomparable until $\mathcal{DN}$ is refined into a more specific description. The description language itself has axioms for $\succ$ (transitive and asymmetric) and $\sim$ (transitive, symmetric and reflexive) together with the usual axioms for propositional logic and so delivers a notion of theorem hood $\vdash$ that we use below.

Cadilhac et al (2011) provide rules to update and revise one's partial CP-net model of agents' preferences given what they say in conversation. The type of dialogue act that an agent performs says how to update the old (partial) model of their declared preferences with newly declared preferences. The rules allow an existing partial CP-net to be *revised* rather than extended to reflect the effects of *corrective* speech acts. For instance, in dialogue (1), from the Verbmobil corpus (Muller and Kasper, 2000), the dialogue move (1c) *corrects* (1ab)—the preference stated initially, to meet on Thursday, is replaced with a preference not to meet on this day.

(1)   a. A: how about, Thursday afternoon, at one?

     b. B: that would be good for me,

     c. A: oh wait, that says two thirty, not twelve thirty.
       so, we can not meet that Thursday.

We forego the details of the rules here because modeling preference dynamics via evidence from dialogue isn't sufficient. Cadilhac et al (2011) don't claim such rules affect one's *own* CP-net; even if they did preference change

also happens when information is acquired from *non-verbal* actions like tasting turmeric for the first time. So we now provide a more general model of preference dynamics, which abstracts over the nature of the evidence that prompted the change in belief and/or the acquisition of a new preference.

5.2 CP-net revision when beliefs change

We start by addressing the problem of how belief change can trigger preference change, particularly extrinsic preference change. One way to avoid preference revision, whatever the change in beliefs, would be to make each preference statement conditional on a unique underspecified condition. Thus as beliefs change, different equations in the (partial) CP-net become operative. For instance in dialogue (1), A's declared preference in (1a) for meeting on Thursday is made dependent on an underspecified context $\mathcal{C}$, say, while the declared preference in (1c) to meet on a day other than Thursday is made dependent on a different underspecified context $\mathcal{C}'$. These (conditional) preference declarations are mutually consistent. But a relation between $\mathcal{C}$ and $\mathcal{C}'$ is needed, if we are to use such contexts in inference and decision making. We don't know how to do this in a plausible and efficient way.

Our alternative approach allows belief change that is exogenous to CP-nets to trigger revision to the CP-net. The general idea is this: if the agents' beliefs change, then they may revise *extrinsic* preference statements in their CP-net, that exist only in service to some parent goal. However, while CP-nets have the great advantage of representing dependencies between preferences explicitly and assigning preferences to intermediate states and actions (a desirable feature for handling Spohn's scenarios, for instance), this makes an explicit distinction between intrinsic and extrinsic preferences difficult. The agents can choose to *retain* a preference for an outcome *o even if* they come to believe that o is unlikely or even impossible. After all, retaining intrinsic preferences is rational because what is currently deemed unachievable may later become achievable (especially in dynamic environments with chance moves by nature). Nevertheless, no rational agents should *act* to achieve something that they deem unachievable.

Using CP-nets thus calls for a double interaction between preferences and beliefs. First, agents tentatively adjust their CP-net to reflect their changing beliefs. But since the CP-nets retain intrinsic preferences, even if their fulfilment is unlikely, identifying an optimal action requires the agents to 'filter out' from the preference order entailed by the CP-net alone those outcomes that, according to the agents' exogenous beliefs, are unreachable.

Let's look at the first step: i.e., CP-net change due to belief change. We assume a standard Bayesian belief model and *convert* the qualitative (partial) CP-net into a *numeric* utility function. That is, the current CP-net description $\mathcal{N}_V$, defined over variables V that in turn determines the set of (currently distinguishable) outcomes $\mathcal{O}$, is mapped to a utility function $u : \mathcal{O} \rightarrow \mathbb{R}$ that satisfies the following two constraints for all $o_1, o_2 \in \mathcal{O}$: (i) if $o_1 \sim o_2$ then

$u(o_1) = u(o_2)$, and (ii) if $o_1 \succ o_2$ then $u(o_1) \geq u(o_2)$. These constraints on u suffice to ensure that u is consistent with the logic of the CP-net $\mathcal{N}_V$, but there is an unbounded number of such functions u. We'll return to this issue shortly. But first, we consider how an agent can use u and the updated beliefs $P(\mathcal{O}|\mathbf{e})$ to check whether (extrinsic) preferences should change: the expected utilities are now $u(o)P(o|\mathbf{e})$ rather than $u(o)$, for $o \in \mathcal{O}$.

The rule **Extrinsic Preference Change** below captures the intuition that an existing derivative preference, which is in service to a goal $\mathbf{x}$, normally changes when newly observed evidence $\mathbf{e}$ yields a revised belief about the most likely way of achieving $\mathbf{x}$:

Extrinsic Preference Change: Let the utility of an *outcome description p*, where p is a formula expressed with the vocabulary V of the (partial) CP-net $\mathcal{N}_V$, be defined as the average utility of the specific outcomes o that satisfy p. That is:

$$u(p) =_{def} \frac{\sum \{u(o) : o \in \mathcal{O} \text{ and } o \vdash p\}}{|o : o \in \mathcal{O} \text{ and } o \vdash p|}$$

Suppose $\mathbf{x}$ is a goal, and the agent's partial CP-net $\mathcal{N}_V$ includes the preference statement $\mathbf{x} : a \succ a'$ (so $u(\mathbf{x} \wedge a) > u(\mathbf{x} \wedge a')$). Suppose that given evidence $\mathbf{e}$, the belief model satisfies the following two conditions:
(a) $P(\mathbf{x}|a', \mathbf{e}) > P(\mathbf{x}|\mathbf{e})$ (i.e., the agent believes, given the evidence, that a' contributes to achieving $\mathbf{x}$); and
(b) $u(\mathbf{x} \wedge a')P(\mathbf{x} \wedge a'|\mathbf{e}) > u(\mathbf{x} \wedge a)P(\mathbf{x} \wedge a|\mathbf{e})$ (i.e., the inequality over the two expected utilities has changed).
Then normally, on observing $\mathbf{e}$, the agent updates her CP-net description with a new preference $\mathbf{x} : a' \succ a$, which replaces the old preference $\mathbf{x} : a \succ a'$ (we'll see shortly how our formal definitions of preference revision entail that $\mathbf{x} : a \succ a'$ is retracted when updating with the new preference).

This update principle yields (extrinsic) preference change in a CP-net as a result of belief change. For instance, suppose you want to go shopping (i.e., the goal is *shop*), and to do this you would rather take the *car* than the *bus*. That is, your initial (partial) CP-net includes the conditional preference statement *shop* : *car* $\succ$ *bus*. But before setting out, you observe evidence $\mathbf{e}$ that all the car parks are full. In fact, your relative preference u for using the car over the bus to go shopping, together with your updated beliefs are such that $u(shop \wedge car)P(shop \wedge car|\mathbf{e})$ is less than $u(shop \wedge bus)P(shop \wedge bus|\mathbf{e})$. Then normally, you replace *shop* : *car* $\succ$ *bus* with the new extrinsic preference *shop* : *bus* $\succ$ *car*.

It is important to stress that the above principle says that revision *normally* occurs as opposed to *always* occurs. An agent may decide that even though the expected utilities of a and a' have changed their relative order, a is *still* preferred to a'. This is a sort of induced *intrinsic* preference change—that is, the agent is free to pick an alternative utility function u' that (i) like u, is consistent with the initial partial CP-net $\mathcal{N}_V$ (and in particular, with $\mathbf{x} : a \succ a'$), while (ii) ensuring that condition (b) is no longer satisfied with the

updated beliefs. This alternative, of choosing a utility function to articulate one's initial preferences so as to make revision unnecessary in the context of the new belief, is consistent and compensates both for the fact that CP-nets don't distinguish intrinsic from extrinsic preferences and for the fact that they express which outcomes are more preferred, but not by how much.

This brings us to the second interaction of beliefs with preferences. Since the agents can choose *not* to revise their CP-net when beliefs change—intrinsic preferences needn't change even when they're deemed unachievable—the agents must consider the extraneous belief model as well as the CP-net when deciding how to act. They should not try to achieve goals that they believe to be unachievable. To illustrate this step simply, we add to the description logic $\vdash$ a modal operator $\mathcal{B}$, where the formula $\mathcal{B}_a p$ means that agent a assigns p a sufficiently high probability (Pearl, 1988). We make $\mathcal{B}$ KD45, and we also assume that our logic supports default reasoning, with a nonmonotonic consequence relation $\mathrel{|\!\sim}$; so an agent can come to believe a proposition in the *absence* of evidence to the contrary. An agent uses the logic of CP-nets to order the outcomes, and then uses her beliefs to *filter out* from that ordering any outcomes that are doxastically improbable: that is, whenever all of the agent a's observations Γ support an inference that $\Gamma \mathrel{|\!\sim} \mathcal{B}_a \neg p$, then any outcome o that entails p is removed from the preference ordering. This is captured in the definition of *CP-solution* (Asher and Lascarides, 2013):

Definition 3 *Let G be a game, represented as a joint CP-net (one for each player). Then CP-solution$_a(\phi, G)$ holds iff:*

1. *a is a player in the game G; and*
2. *$o \vdash \phi$ for every belief-compliant optimal outcome o of G: i.e., where Γ includes all the background axioms of belief (including domain-level axioms) and the relevant premises about players in G, $\Gamma \mathrel{|\!\not\sim} \mathcal{B}_a \neg o$ and for any outcome o' that is strictly more optimal for a in G than o, $\Gamma \mathrel{|\!\sim} \mathcal{B}_a \neg o'$.*

The optimal outcome according to the preferences in a CP-net alone might thus be different from the CP-solution that takes extraneous beliefs into account, and we assume agents act on their CP-solutions. This approach, like that of classical decision theory, is rich enough to capture a situation where an agent's most preferred outcome is (currently) deemed by her to be improbably hard to achieve, while at the same time ensuring that the decision making is rational: agents will not attempt to achieve the impossible simply because they wish for it. Moreover, as beliefs change, so does the CP-solution even if the CP-net does not change. Finally, if the exogenous model of beliefs and intentions incorporates an axiom that intentions are dropped once they're achieved (Bratman, 1987), then agents won't plan to achieve preferences that are already believed to be true.

5.3 CP-net consistency when adding new preferences

Cadilhac *et al*'s (2011) rules for updating a CP-net with a new preference don't necessarily preserve consistency. If the original CP-net contains $x_1 \succ x_2$ and

$x_2 \succ x_3$ then $x_1 \succ x_3$ is added to it via transitivity. If this is then updated with the preference $x_3 \succ x_1$, Cadilhac *et al*'s rules do only a *local* revision: only $x_1 \succ x_3$ is retracted. So the result yields $x_1 \succ x_1$ via transitivity, which violates irreflexivity. Local revision is not a problem for their intended domain of application, namely dialogue, because speakers control for contradictions. But updating with a new preference generated from more general (i.e., non-verbal) evidence needs to guarantee the consistency of the CP-net (recall section 4 for motivation).

If the partial preference model is updated with a new preference that is consistent with it, then simply adding the new preference is unproblematic. The challenge is to define update with a new preference that is inconsistent with the existing preferences. To ensure that the updated model is consistent, we need some notion of *preference revision*. Following the AGM approach to belief revision (Alchourrón et al, 1985), we define preference revision as a sequence of two operations: *downdating* the existing preferences to a *maximal subset* that is consistent with the new preference, followed by *adding* the new preference to the result (so new preferences take priority over old ones).

One difficulty for belief revision is how to downdate a belief model when there is more than one maximal subset of old beliefs that are consistent with the new one. To handle this all theories exploit some notion of *entrenchment* (e.g., Gärdenfors and Makinson (1988)): a transitive, binary relation on propositions where "p is more entrenched than q" means that agents are more reluctant to give up their belief in p than in q (all else being equal). Intuitively, more entrenched propositions are more useful in deliberation; e.g., p is a natural law whereas q is a contingent fact. So when there's a choice, agents favour downdating old beliefs to the the maximal subset where fewer entrenched propositions are removed. Defining belief entrenchment involves causality and epistemic explanatory power, both notoriously difficult concepts to analyse.

Since we're making preference revision analogous to AGM belief revision, we also need a concept analogous to entrenchment, to help an agent decide which maximal subset of old preferences to retain. Fortunately, unlike the modal logic of belief, CP-nets have an explicit partial order—the graphical model—that suggests a natural solution for this problem. Recall that CP-nets define which variables influence the preferences over other variables. In effect, the structure of the CP-net defines which preferences are global and which are derivative. So we can regiment the intuition that the fewer factors there are that compel us to have a particular preference, the more entrenched it is and the less prepared we are to give it up (unless failure to do so results in abandoning more preferences overall). This priority for removing derivative preferences over global ones aligns with the secondary principle in the logic of CP-nets for inferring the preference order over all outcomes: i.e., it is worse to violate a preference on variables over which your other preferences depend.

Accordingly, we form a partial order over the outcomes defined by a CP-net that reflects the extent to which agents would be prepared to give up their preference for one outcome as compared to giving up their preference for another. If the agents' preferences for outcomes o_i and o_j are dependent on

each other, then the agent is equally reluctant to give up either of them. On
the other hand, if the preference for o_i depends on a superset of the factors on
which the preference for o_j depends, then the agents are more reluctant to give
up their preferences for o_j. We call this partial order the *preference surrender
value* or PSV (we use the term surrender, an antonym of entrenchment, because
we'll assign numeric PSVs to outcomes where the higher the number, the more
inclined one is to give up a preference for it). Definition 4 defines the partial
order PSV in two steps. First, it detects cyclically dependent outcomes in the
CP-net $\mathcal{N}_V$ and constrains their PSVs to be equal. This forms a partition over
outcomes. Then it assigns elements in each partition a numeric PSV value: the
lower the number the *less* one is inclined to give up the preference.

Definition 4 *Variables $X, Y \in \mathcal{N}_V$ are said to encode cyclically dependent
preferences if X is an ancestor of Y and Y is an ancestor of X in the graphical
component of $\mathcal{N}$. The partial order **preferential surrender value** (or PSV)
over the variables V in a CP-net $\mathcal{N}_V$ is defined as follows:*

1. *For $V' \subseteq V$ such that $\mathcal{N}_V|V'$ describes cyclically dependent preferences
   over V', we say that $\mathrm{PSV}_{\mathcal{N}_V}(o_i) = \mathrm{PSV}_{\mathcal{N}_V}(o_j)$ for all $o_i, o_j \in V'$ (so o_i and
   o_j are state descriptions that assign specific values to each of the variables
   in V').*
2. *With V thus partitioned into equivalence classes of cyclically dependent
   outcomes, we assign each equivalence class a numeric preference surrender
   value or PSV as follows:*
   - *For $V_0 \subseteq V$ such that V_0 is an equivalence class of outcomes such that
     all preference statements in $\mathcal{N}_V$ about outcomes in V_0 depend only on
     elements in V_0 or none at all, we set $\mathrm{PSV}_{\mathcal{N}_V}(o_i) = 0$, for all $o_i \in V_0$*
   - *For any V_n, $n \neq 0$, such that V_n is an equivalence class of outcomes
     in $\mathcal{N}_V$, for all $o_i \in V_n$ we set*

   $$\mathrm{PSV}_{\mathcal{N}_V}(o_i) = 1 + max_{o_j \in V_n, x \in V}\{\mathrm{PSV}(x) : x \text{ is a parent of } o_j \text{ and } x \notin V_n\}$$

Suppose a (perhaps partial) CP-net $\mathcal{N}$ is updated with a new preference
statement $\phi : R(t, t')$, where $R \in \{\prec, \succ, \sim\}$. To maintain consistency, one first
checks whether $\mathcal{N}$'s transitive closure entails $\phi : \overline{R(t, t')}$, where $\phi : \overline{R(t, t')}$ is
equivalent to $t \prec t' \vee t \sim t'$ if $R(t, t')$ is $t \succ t'$. If so, then we must change or
reset formulae in $\mathcal{N}$ so that the result together with $\phi : R(t, t')$ is consistent.
Following earlier discussion, the ranking in Definition 5 favours those resets
with the fewest changes to preferences of any outcomes, and more changes to
outcomes with a larger PSV than a smaller PSV.

Definition 5 *The >-**ranking over Resets** is defined as follows:*
$Reset_n(\mathcal{N}) > Reset_m(\mathcal{N})$ *iff*

- *$Reset_n(\mathcal{N})$ resets fewer equations in $\mathcal{N}$ than $Reset_m(\mathcal{N})$; or*
- *they reset the same number of equations, and:*
  *$min\{PSV(o) : R(o, o') \text{ is reset by } Reset_n(N)\} > min\{PSV(o) : R(o, o') \text{ is
  reset by } Reset_m(N)\}$*

We can now stipulate that during preference revision, any downdating is restricted to Resets of $\mathcal{N}$ that are consistent with the new preference and $>$-maximal. We obtain the following definition where $*$ is the *revision* operator ($\mathcal{N} * \phi$ means that ϕ is added to $\mathcal{N}$ and at the same time other statements are removed if this is needed to ensure that the resulting CP-net is consistent) and $+$ is the *expansion* operator ($\mathcal{N} + \phi$ means that ϕ is added to $\mathcal{N}$ without checking the consistency: nothing is removed).

Definition 6 *The **preference revision** $\mathcal{N} * \phi$ is defined as follows:*

$$\mathcal{N} * \phi = \begin{cases} \bigcap \{Reset_i(\mathcal{N}) : Reset_i(\mathcal{N}) \text{ is } > -maximal \text{ and} \\ \qquad Reset_i(\mathcal{N}) + \phi \text{ consistent}\} + \phi, \text{ if } \phi \text{ is consistent} \\ \bot \text{ otherwise} \end{cases}$$

We illustrate Definition 6 by updating the partial CP-net (2) with $c \succ a$:

$$\begin{aligned} CPT(X) = a &\succ b \\ b &\succ c \end{aligned} \tag{2}$$

All preferences in the CP-net (2) have rank 0 because its only variable X, whose domain is $\{a, b, c\}$ has rank 0. So there are two minimal resets of (2) that are consistent with $c \succ a$, given in (3) and (4):

$$\begin{aligned} CPT(X) = a &\succ b \\ \mathbf{c} &\succ \mathbf{b} \end{aligned} \tag{3} \qquad \begin{aligned} CPT(X) = \mathbf{b} &\succ \mathbf{a} \\ b &\succ c \end{aligned} \tag{4}$$

In words, (3) retains $a \succ b$ and resets $b \succ c$ to $c \succ b$, whereas (4) retains $b \succ c$ and resets $a \succ b$ to $b \succ a$. Since their intersection is empty, updating (2) with $c \succ a$ yields only the new information $c \succ a$.

On the other hand, consider CP-net (5) with two variables X_1 and X_2 whose domains are $D(X_1) = \{a, \overline{a}\}$ and $D(X_2) = \{b, c, d\}$. Updating this CP-net with $\overline{a} : d \prec c$ leaves $a \prec \overline{a}$ and the preferences dependent on a intact.

$$\begin{aligned} CPT(X_1) = & \quad a \prec \overline{a} \\ CPT(X_2) = & \quad a : b \prec c \\ & \quad a : c \prec d \\ & \quad \overline{a} : c \prec b \\ & \quad \overline{a} : b \prec d \end{aligned} \tag{5}$$

Lemma 1, which follows from Definition 6, shows that preference revision is relatively well-behaved.

Lemma 1

1. *Success: $\phi \in \mathcal{N} * \phi$ (and so trivially $\mathcal{N} * \phi \models \phi$).*
2. *Inclusion: $\mathcal{N} * \phi \subseteq \mathcal{N} + \phi$ (that is, the deductive or transitive closure of $\mathcal{N} * \phi$ is contained in that of $\mathcal{N} + \phi$).*
3. *Vacuity: If $\neg\phi \notin \mathcal{N}$ and ϕ is consistent, then $\mathcal{N} * \phi = \mathcal{N} + \phi$.*
4. *Consistency: $\mathcal{N} * \phi$ is consistent if ϕ is consistent.*

5. *Extensionality: If $\vdash \phi \leftrightarrow \psi$, then $\mathcal{N} * \phi = \mathcal{N} * \psi$.*

Our result lacks AGM belief revision constraint of *Closure*, but we don't want this because we reason with *partial* descriptions of preferences. It is a fact of life that preference information is usually incomplete; we are interested in the process of reasoning with incompleteness. Our model for preference change differs from Hansson's (1995) in this respect too: Hansson only considers preferences that refer to complete alternatives (i.e., to elements of a set of mutually exclusive alternatives). CP-nets also have the advantage of providing a notion of importance among the preferences thanks to the dependency structure—we've used this to guide revision to the most preferred ways of minimally changing the old preferences to preserve consistency with the new ones. This sort of guidance is exogenous to Hansson's model.

Definition 6 handles all the types of intrinsic preference change detailed in sections 2 (see section 5.4). It has the *consistency* and *conservativity* properties discussed in section 4. Our model of preferences also has the property of *intelligibility*. Intelligibility relies on (a) updating the model of other agents' preferences based on observing what they do and (b) using this to support decision making. We achieve (a) because Definition 6 provides the means to consistently update one's existing model of another agent's preferences with any new preference that's inferrable from observing his or her latest action. We achieve the second objective because partial CP-nets have a semantics that's defined in terms of complete CP-nets, where the latter support a logic for identifying the optimal action (recall Definition 3 from section 5.2). So we can also predict the agent's optimal actions from our partial model: one simply completes the partial model of preferences by defaulting to indifference for the preference information that is missing entirely, and one uses the resulting complete representation to infer what decision the agent will make next.

Finally, preference revision is relatively simple in this framework assuming that the options over which preferences are defined are logically independent and finite. Most decision problems involve logically independent outcomes, and our induction of partial CP-nets from behavior ensures that the options are finite. Our definition of revision and of ranking depend just on the structure of the CP-net constructed. Testing for consistency over logically independent outcomes involves computing the transitive closure for the relation, which is doable in low polynomial time. The number of consistent revisions to be considered leads in the worst case to an exponential factor in the revision process, but preference revision is thus *decidable*.

5.4 Formal analyses of sample decision problems

Let us now return to the scenarios of preference change from section 2. In the turmeric example, the agent learns of a new possibility, expressed via a new state description: the vocabulary the agent uses to describe the decision problem $\mathcal{N}_V$ changes from V to V' where $V \subset V'$. For example, suppose that

the agent initially prefers to eat fish over soup and prefers coffee to tea. His or her initial vocabulary V consists of two variables: D for the drink with domain $D(D) = \{coffee, tea\}$; and M for the meal with domain $D(M) = \{fish, soup\}$. The (partial) CP-net $\mathcal{N}_V$ is (6):

$$
\begin{aligned}
CPT(D) = \ & coffee \succ tea \\
CPT(M) = \ & fish \succ soup
\end{aligned}
\tag{6}
$$

We will now illustrate different effects on preference revision with different examples of discovery; i.e., different V'.

Suppose first that the preferences over $V' \setminus V$ are *independent* from those over V (see Definition 1): that is, for all $t \in V' \setminus V$, $a, b \in V$ if $t \wedge a \succ t \wedge b$ then $\bar{t} \wedge a \succ \bar{t} \wedge b$. Then, it is immediate from Definition 6 that all prior preferences in $\mathcal{N}_V$ persist in $\mathcal{N}_{V'}$. For instance, if the agent discovers chocolate, extending V to a new vocabulary $V' = V \cup C$ where the domain $D(C) = \{chocolate, \overline{chocolate}\}$, and the agent adopts the new preference is $chocolate \succ \overline{chocolate}$, then by Definition 6 the preferences in (6) persist in the updated preferences (7):

$$
\begin{aligned}
CPT(D) = \ & coffee \succ tea \\
CPT(M) = \ & fish \succ soup \\
CPT(C) = \ & chocolate \succ \overline{chocolate}
\end{aligned}
\tag{7}
$$

Now suppose that there are preferences in V that are *dependent* on the new preferences in $V' \setminus V$. For example, suppose the agent discovers wine and turmeric such that V' contains two new variables: W for the wine with domain $D(W) = \{white\_wine, red\_wine\}$; and T for turmeric with domain $D(T) = \{turmeric, \overline{turmeric}\}$. Suppose that the agent prefers *white_wine* with fish, but *red_wine* with soup, and also now prefers *soup* over *fish* whenever *turmeric* is available, and prefers *turmeric* over $\overline{turmeric}$, whatever the circumstances. Then according to Definition 6, updating the CP-net (6) with this new preference information yields the CP-net (8):

$$
\begin{aligned}
CPT(D) = \ & coffee \succ tea \\
CPT(T) = \ & turmeric \succ \overline{turmeric} \\
CPT(M) = \ & turmeric : soup \succ fish \\
& \overline{turmeric} : fish \succ soup \\
CPT(W) = \ & fish : white\_wine \succ red\_wine \\
& soup : red\_wine \succ white\_wine
\end{aligned}
\tag{8}
$$

Definition 6 entails that old preferences that are independent of new ones (like *coffee* and *tea*) are retained. Likewise, it entails that new variables (e.g., W) that are dependent on old ones don't trigger revision to existing preferences. But if the *old* variables are dependent on the *new* ones (like M, which now depends on T), update is more complex because inconsistency may result between the existing preferences and the new one. Computing the update

involves exploring the recursive structure of the CP-net. The simplest case is where $\mathcal{N}_V$ is without dependencies, binary comparisons only, no indifference and:

$$t \in V' \setminus V, a, b \in V \text{ with } a \succ b \in \mathcal{N}_V \text{ but } t{:}b \succ a \in \mathcal{N}_{V'}$$

The question is: should the preferences in $\mathcal{N}_{V'}$ also shift to $b \succ a$, given $\bar{t}$? Of course they can, but this is equivalent to a shift in preferences among a and b within $\mathcal{N}_V$ itself, because $(\mathbf{x} \wedge t{:}\phi) \wedge (\mathbf{x} \wedge \bar{t}{:}\phi)$ is semantically equivalent to $\mathbf{x}{:}\phi$. In other words, the agent's preference change between a and b would in this case be *independent* of the new vocabulary (or discovery) t, instead being preference change within the "smaller" vocabulary V (and hence holds within a domain of fewer possibilities). But the new preference information in our example does not invoke such an intrinsic preference change on the old vocabulary; it only specifies how the preferences among fish and soup change when the (newly discovered) turmeric is available. Accordingly, Definition 6 yields the CP-net (8), where the old preference for *fish* over *soup* is retained in the context $\overline{turmeric}$.

More generally, for any CP-net $\mathcal{N}_V$ with no indifference, Lemma 2 (which follows immediately from Definition 6) specifies a general effect of updating with new preferences expressed in an extended vocabulary $V' \supset V$:

Lemma 2 *Equations in $\mathcal{N}_V$ persist within $\mathcal{N}_{V'}$ where $V \subseteq V'$ for all pairs of outcomes o_1, o_2 where $\mathrm{PSV}(o_1), \mathrm{PSV}(o_2) > \mathrm{PSV}(t)$, for all $t \in V' \setminus V$. Further, if $\mathbf{x}{:}R(o_1, o_2)$ and $\mathbf{x} \wedge t{:}\overline{R(o_1, o_2)}$, then $\mathbf{x} \wedge \bar{t}{:}R(o_1, o_2)$.*

Introducing indifference complicates matters, because given $t{:}o_1 \succ o_2 \in \mathcal{N}_{V'}$ and $o_1 \prec o_2 \in \mathcal{N}_V$, it is still possible to have $\bar{t}{:}o_1 \sim o_2$ in $\mathcal{N}_{V'}$. But that means that we just slightly weaken Lemma 2 to account for indifference:

$$\text{If } \mathbf{x}{:}R(o_1, o_2) \text{ and } \mathbf{x} \wedge t{:}\overline{R(o_1, o_2)}, \text{ then } \mathbf{x} \wedge \bar{t}{:}R(o_1, o_2) \vee o_1 \sim o_2$$

Next, let's reconsider the smoking example, modeled with two Boolean variables S (smoking) and P (peaceful life). The agent starts with a belief that s (smoking) normally implies $\bar{p}$ (a non-peaceful life) and the CP-net (9):

$$\begin{aligned} CPT(S) &= \ s \succ \bar{s} \\ CPT(P) &= \ p \succ \bar{p} \end{aligned} \tag{9}$$

CP-solutions (see definition 3) captures the agent's moral dilemma: the most preferred outcome $s \wedge p$ is belief-filtered out, and neither s nor p are a part of any CP-solution because neither of them is entailed by *all* optimal belief compliant states (which are $\bar{s} \wedge p$ and $s \wedge \bar{p}$). Now suppose the agent adopts the new preference $p : \bar{s} \succ s$. According to Definition 6, updating (9) with this entails that the preference for s in the context $\bar{p}$ is retained:

$$\begin{aligned} CPT(P) &= \ p \succ \bar{p} \\ CPT(S) &= \ p : \bar{s} \succ s \\ & \quad\ \ \bar{p} : s \succ \bar{s} \end{aligned} \tag{10}$$

The agent then changes preferences again, and updates (10) with $\overline{s} \succ s$ (i.e., a global aversion to smoking). This yields (11) by Definition 6.

$$
\begin{aligned}
CPT(P) &= \ p \succ \overline{p} \\
CPT(S) &= \ \overline{s} \succ s
\end{aligned}
\tag{11}
$$

Finally, we reconsider Spohn's (2009) examples **holiday** and **market** that posed problems for standard decision theory (see section 3). As suggested earlier, implicit in Spohn's formulation is the idea that the agent should not be impulsive: his decision to buy should be the result of deliberation that is as objective as he can make it. While the global decision models cannot make this explicit because you cannot attach rewards directly to the kind of agent you are, we can make it explicit in our model: we introduce a variable R, where $D(R) = \{reason, \overline{reason}\}$. The agent also has (intrinsic) preferences over the variable B, where $D(B) = \{buy, \overline{buy}\}$.

These two stories also involve extrinsic preferences. For the holiday scenario, the variable W takes the value that the agent *wait* before deciding whether to buy the product, or not wait ($\overline{wait}$). For the market scenario, the variable L takes the value that the agent *listen* to the hawker before deciding whether to buy, or not listen ($\overline{listen}$). The holiday scenario is then captured with the *initial* partial CP-net in (12):

$$
\begin{aligned}
CPT(R) &= reason \succ \overline{reason} \\
CPT(B) &= buy \succ \overline{buy} \\
CPT(W) &= reason \wedge buy : wait \succ \overline{wait}
\end{aligned}
\tag{12}
$$

In words, when an agent wants to be reasonable and also wants to (currently) buy the product, he prefers to wait to see if he changes his mind. After waiting, either the preferences over B stay the same, or there is a preference change: the partial CP-net (13) results from updating via Definition 6 the CP-net (12) with the new preference $\overline{buy} \succ buy$:

$$
\begin{aligned}
CPT(R) &= reason \succ \overline{reason} \\
CPT(B) &= \overline{buy} \succ buy \\
CPT(W) &= reason \wedge buy : wait \succ \overline{wait}
\end{aligned}
\tag{13}
$$

In contrast, the market scenario has the following initial partial CP-net:

$$
\begin{aligned}
CPT(R) &= reason \succ \overline{reason} \\
CPT(B) &= \overline{buy} \succ buy \\
CPT(L) &= reason \wedge \overline{buy} : \overline{listen} \succ listen
\end{aligned}
\tag{14}
$$

This makes $\overline{listen}$ optimal, and so the hawker has no opportunity to manipulate these preferences. These analyses show that we can better predict optimal action in the face of foreseen preference change. We don't represent all possible choices in all possible present and future states, but rather only current preferences, which get updated according to subsequent evidence. In contrast, Spohn's model has no downdating or revision of a preference model.

6 Conclusions and Future Work

This paper has described decision and game problems for which a static and pre-defined reward function isn't sufficient for reasoning about rational action. First, there are practical problems in games where all possible ways of getting from the current state to an end state aren't surveyable. Static, pre-defined preferences are also untenable on conceptual grounds, because one can be playing a game while at the same time discovering its hypothesis space, and preference change can stem from a change in taste in the absence of any discovery or belief change. We agreed with Spohn (2009) that global decision models don't always provide sufficiently rich information for identifying the optimal strategies.

We proposed a logically precise model of preference change that extends and refines CP-nets, a compact and qualitative representation of preferences. To handle the agents discovering new possible states and actions during the course of their actions and deliberations, we introduced CP-net *descriptions*: a *partial* model of preferences that can subsequently be added to or updated. Indeed, the vocabulary for describing the game can be extended. We then defined preference update and showed that it preserves consistency. The explicit encoding in CP-nets of dependencies among preferences guides the revision process. The result is a model of preference revision similar to an AGM model of belief revision (Alchourrón et al, 1985). We applied this model to the motivating examples from section 2. We showed how it complements earlier research about preference change and in particular we demonstrated how it overcomes Spohn's (2009) criticisms of global models of dynamic preferences.

This work addresses reasoning about rational action in a context where preferences change in the absence of any specifiable or measurable belief change. In future work, we plan to devise algorithms, both precise and approximate, for computing preference update and the revised optimal strategies that result. We believe that the resulting computational model of dynamic preferences, when integrated with a computational model of dynamic belief, will provide a powerful tool for studying decision and game problems for which standard models of rational action do not provide a satisfactory solution.

References

Afantenos S, Asher N, Benamara F, Cadilhac A, Dégremont C, Denis P, Guhe M, Keizer S, Lascarides A, Lemon O, Muller P, Paul S, Popescu V, Rieser V, Vieu L (2012) Modeling strategic conversation: model, annotation design

and corpus. In: Proceedings of the 16th Workshop on the Semantics and Pragmatics of Dialogue (Seinedial), Paris

Alchourrón C, Gärdenfors P, Makinson D (1985) On the logic of theory change: Partial meeting contraction and revision functions. Journal of Symbolic Logic 50:510–530

Andréka H, Ryan M, Schobbens PY (2002) Operators and laws for combining preference relations. Journal of Logic and Computation 12(1):13–53

Ariely D (2008) Predictably Irrational: The Hidden Forces That Shape Our Decisions. HarperCollins

Asher N (1987) A typology for attitude verbs and their anaphoric properties. Linguistics and Philosophy 10(2):125–197

Asher N, Lascarides A (2013) Strategic conversation. Semantics and Pragmatics 6(2):2:1–:62, DOI dx.doi.org/10.3765/sp.6.2

Baltag A, Moss L, Solecki S (1999) The logic of public announcements, common knowledge and private suspicions. Tech. Rep. SEN-R9922, Centrum voor Wiskunde en Informatica

Bellman R (1957) A markovian decision process. Journal of Mathematics and Mechanics 6:679–684

van Benthem J, Girard P, Roy O (2009) Everything else being equal: A modal logic for *ceteris paribus* preferences. Journal of Philosophical Logic 38:83–125

Boutilier C, Brafman R, Domshlak C, Hoos H, Poole D (2004) Cp-nets: A tool for representing and reasoning with conditional *ceteris paribus* preference statements. Journal of Artificial Intelligence Research 21:135–191

Bradley R (2007) The kinematics of belief and desire. Synthese 156(3):513–535

Bratman M (1987) Intentions, Plans and Practical Reason. Harvard University Press

Browne C, Powley E, Whitehose D, Lucas S, Cowling P, Rohlfshagen P, Tavener S, Perez D, Samothrakis S, Colton S (2012) A survey of monte carlo tree search methods. IEEE Transactions on Computational Intelligence and AI in Games 4(1):1–49

Cadilhac A, Asher N, Benamara F, Lascarides A (2011) Commitments to preferences in dialogue. In: Proceedings of the 12th Annual SIGDIAL Meeting on Discourse and Dialogue, pp 204–215

Degremont C, Paul S, Asher N (2014) A Logic of Sights, to appear in Journal of Logic and Computation

Gärdenfors P, Makinson D (1988) Revisions of knowledge systems using epistemic entrenchment. In: Vardi MY (ed) Proceedings of the Second Conference on Theoretical Aspects of Reasoning about Knowledge, Morgan Kaufmann, San Francisco, pp 83–95

Grüne-Yanoff T, Hansson SO (2009) Preference Change: Approaches from Philosophy, Economics and Psychology. Springer

Hansson S (1995) Changes in preference. Theory and Decision 38(1):1–28

Heim I (1992) Presupposition projection and the semantics of attitude verbs. Journal of Semantics 9:183–221

Kahneman D, Tversky A (1979) Prospect theory: An analysis of decision under risk. Econometrica 47(2):263–291

Lang J, van der Torre L (2008) From belief change to preference change. In: ECAI, pp 351–355

Liu F (2008) Changing for the better: Preference dynamics and agent diversity. PhD thesis, Institute for Logic, Language and Computation (ILLC), University of Amsterdam

Liu F (2011) Reasoning about Preference Dynamics. Springer Verlag

Modica S, Rustichini A (1999) Unawareness and partitional information structures. Games and Economic behavior 27:265–298

Muller S, Kasper W (2000) Hpsg analysis of german. In: Wahlster W (ed) Verbmobil: Foundations of Speech-to-Speech Translation, Springer, Berlin, pp 238–253

Osborne M, Rubinstein A (1990) Bargaining and Markets. Academic Press

Pearl J (1988) Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference. Morgan Kauffmann

Savage L (1954) The Foundations of Statistics. John Wiley

Shoham Y, Leyton-Brown K (2009) Multiagent Systems: Algorithmic, Game-Theoretic and Logical Foundations. Cambridge University Press

Simon H (1955) A behavioral model of rational choice. Quarterly Journal of Economics 69:99–118

Spohn W (2009) Why the received models of considering preference change must fail. In: Grüne T, Hansson S (eds) Preference Change: Approaches from Philosophy, Economics and Psychology, Springer, pp 109–121

Sutton R, Barto A (1998) Reinforcement learning: An introduction. MIT Press

Thomas R (2004) Real-time decision making for adversarial environments using a plan-based heuristic. PhD thesis, Department of Computer Science, Northwestern University

Yong Z, Xinlin W (2012) Intrinsic preferences, revealed preferences and bounded rational decisions. Systems Research and behavioral Science DOI 10.1002/sres.2120