Optimality, Belief and Preference

Dick de Jongh

Institute for Logic, Language and Computation, University of Amsterdam, Netherlands.

dickdj@sccience.uva.nl

Fenrong Liu

Institute for Logic, Language and Computation, University of Amsterdam, Netherlands.

Institute of Philosophy, Chinese Academy of Social Sciences, China.

fenrong@science.uva.nl

Abstract.

We define preference in terms of a constraint sequence, a concept from optimality theory. In case agents only have incomplete information, beliefs are introduced. We propose three definitions to describe different procedures agents may follow to get a preference relation using the incomplete information. Changes of preference are explored w.r.t their sources: changes of the constraint sequence, and changes in beliefs.

1 Motivation

Optimality theory ((PS93)) has been a highly successful approach in linguistics during the last decade. In optimality theory the grammatical or phonological theory does not directly deliver a unique product. First a set of alternative solutions is generated. After that a set of conditions is applied to these alternatives to produce an optimal solution. It is by no means sure that the optimal solution satisfies all the conditions. There may be no such alternative. The conditions, called *constraints*, are strictly ordered according to their importance, and the alternative that satisfies the earlier conditions best (in a way described more precisely below) is considered to be the optimal one. This way of choosing the optimal alternative naturally induces a preference ordering among all the alternatives.

We are interested in formally studying the way the constraints induce the *preference ordering* among the alternatives. This is for us more important than the choice of the optimal alternative. We want to execute our investigations in a wider and maybe somewhat differently directed context. In optimality theory the optimal alternative, like for example the correct grammatical utterance, is chosen unconsciously of course (if that way of speaking is proper at all); we are thinking mostly of applications where conscious choices are made. So it is desirable to present the preference ordering in combination with the reasons underlying it. Also, in optimality theory the application of the constraints to the

Proceedings of the Workshop on Rationality and Knowledge, ESSLLI 2006
Sergei Artemov & Rohit Parikh (editors)

alternatives lead to a *clear* and *unambiguous* result: either the constraint clearly is true of the alternative or it is not, and that is something that is not sensitive to change. We will loosen this condition and consider issues that arise when changes are allowed. It turns out to be very illuminating to introduce the explicit belief operator of doxastic logic into this context.

We think that this point of view gives us a fresh perspective on the issues discussed in previous papers on preference logic and preference change ((Wri63), (Han01) and (BL06)). Of course, there are more ways to obtain a preference order of alternatives from a set of constraints than only the way of optimality theory. A good overview is found in (CMLPM04). We are convinced that our methods can fruitfully be applied in other approaches as well.

We are naturally lead to first consider preference between objects rather than between propositions (compare (DW94)). Consider the following common situation:

Example 1.1 Buying a house is probably the biggest financial decision one will ever make, it is worth taking time to find the right choice. Alice is now in such a situation, for her there are several things to consider: the cost, the quality and the neighborhood, strictly in that order. All these are clear-cut for her, e.g. the cost is good if it is inside her budget, otherwise it is bad. Her decision is then determined by the information whether the alternatives have the desirable properties, and by the given order of importance of the properties.

The following sections are aimed to make this precise. In section 2, we start with a simple language to study the rigid case in which the constraints lead to a clear and unambiguous preference ordering. Section 3 follows with a mathematical exploration on orders and a completeness proof for the simple language. In section 4 we will consider what happens when the (conscious!) agent has incomplete information about the constraints with regard to the alternatives. In section 5 we will look at changes in preference caused by two different sources: changes in beliefs, and changes of the sequence of constraints. Finally, we end up with further discussions about preference in terms of just partially ordered constraint sequences, and our conclusions.

2 From constraints to preference

To discuss preference over objects, we use a first order logic with constants $d_0, d_1 \dots$; variables $x_0, x_1, \dots$; and predicates $P, Q, P_0, P_1, \dots$. In practice, we are thinking of finite domains, monadic predicates, simple formulas usually quantifier free or even variable free. The following definition is directly inspired by optimality theory.

Definition 2.1 *A* ***constraint sequence*** *is a finite ordered sequence of formulas (constraints) written as follows:*

$$C_1 \gg C_2 \cdots \gg C_n \quad (n \in \mathbb{N}),$$

where each of C_m is a formula from the language, and there is exactly one free variable x, which is a common one to each C_m.

The constraint order is read in such a way that the earlier constraints count strictly heavier than the later ones, e.g. $C_1 \wedge \neg C_2 \cdots \wedge \neg C_m$ is preferable over $\neg C_1 \wedge C_2 \cdots \wedge C_m$ and $C_1 \wedge C_2 \wedge C_3 \wedge \neg C_4 \wedge \neg C_5$ is preferable over $C_1 \wedge C_2 \wedge \neg C_3 \wedge C_4 \wedge C_5$. A difference with optimality theory is that we look at *satisfaction* of the constraints whereas in optimality theory *infractions* of the constraints are stressed. This is more a psychological than a formal difference. However, optimality theory knows multiple infractions of the constraints and then counts the number of these infractions. We do not obtain this with our simple objects, but we think that option can be achieved by considering composite objects, like strings. We do not pursue this in the present paper.

Definition 2.2 *Given a constraint sequence of length n,* ***Pref(x,y)*** *is defined as follows:*

$$Pref_1(x,y) ::= C_1(x) \wedge \neg C_1(y),$$
$$Pref_{k+1}(x,y) ::= Pref_k(x,y) \vee (Eq_k(x,y) \wedge C_{k+1}(x) \wedge \neg C_{k+1}(y)), k < n,$$
$$Pref(x,y) ::= Pref_n(x,y),$$

where the auxiliary binary predicate $Eq_k(x,y)$ *stands for* $(C_1(x) \leftrightarrow C_1(y)) \wedge \cdots \wedge (C_k(x) \leftrightarrow C_k(y))$.

In Example 1.1, Alice has the following constraint sequence:

$$C(x) \gg Q(x) \gg N(x),$$

where $C(x)$, $Q(x)$ and $N(x)$ are intended to mean 'x has a low cost', 'x is of good quality' and 'x is in a nice neighborhood', respectively. Consider two houses d_1 and d_2 with the following properties: $P(d_1), P(d_2), \neg Q(d_1), \neg Q(d_2), N(d_1)$ and $\neg N(d_2)$. According to the above definition, Alice prefers d_1 over d_2, i.e. $Pref(d_1, d_2)$.

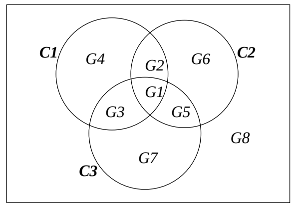
We have chosen a syntactic approach expressing constraints by formulas. Semantically, this comes down to pointing out n sets of worlds in the model. Then Lewis' sphere semantics ((Lew73) p.98-99) comes to mind immediately. Syntactically, a sequence $G_1, \ldots, G_m$ of formulas can be used to express the preferability, i.e. $G_1(x)$ is the most preferable, $G_m(x)$ the least ($G_i(x)$ implies $G_j(x)$ if $i \leq j$). The two approaches are equivalent in the sense that they can be translated into each other.

Theorem 2.3 *A constraint sequence* $C_1 \gg C_2 \cdots \gg C_m$ *gives rise to a G-sequence of length* 2^m. *In the other direction the constraint sequence is logarithmic in the length of the G-sequence.*

Proof. Let us just look at the case that m=3. Assuming that we have the constraint sequence $C_1 \gg C_2 \gg C_3$, we write out the G-sequence in terms of the C_i:

G_1: $C_1 \wedge C_2 \wedge C_3$; G_2: $C_1 \wedge C_2 \wedge \neg C_3$;
G_3: $C_1 \wedge \neg C_2 \wedge C_3$; G_4: $C_1 \wedge \neg C_2 \wedge \neg C_3$;
G_5: $\neg C_1 \wedge C_2 \wedge C_3$; G_6: $\neg C_1 \wedge C_2 \wedge \neg C_3$;
G_7: $\neg C_1 \wedge \neg C_2 \wedge C_3$; G_8: $\neg C_1 \wedge \neg C_2 \wedge \neg C_3$.

and then we get $G_1 \succ G_2 \succ G_3 \succ G_4 \succ G_5 \succ G_6 \succ G_7 \succ G_8$. On can simply read the relation from the following picture:



From constraint sequence to G-sequence in case $m = 3$.

On the other hand, given a G_i-sequence, we can define C_i as follows,

$C_1 = G_4$; $C_2 = G_2 \vee (G_6 \wedge \neg G_4)$;
$C_3 = G_1 \vee (G_3 \wedge \neg G_2) \vee (G_5 \wedge \neg G_4) \vee (G_7 \wedge \neg G_6)$.

And again one can simply read it from a picture of the G-spheres. ■

3 Order and completeness

In this section we will just run through the types of order that we will use. A relation $<$ is a *linear order* if $<$ is irreflexive, transitive and asymmetric, and satisfies *totality*:

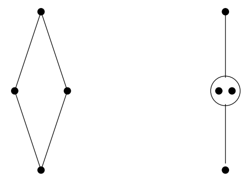
$$x < y \vee x = y \vee y < x$$

More precisely, $<$ is called a *strict* linear order. A *non-strict* linear order $\leq$ is a reflexive, transitive, antisymmetric and total relation. It is for various reasons useful to introduce non-strict variants of orderings as well.

Mathematically, strict and non-strict linear orders can easily be translated into each other:

(1) $x < y \leftrightarrow x \leq y \wedge x \neq y$, or
(2) $x < y \leftrightarrow x \leq y \wedge \neg(y \leq x)$,
(3) $x \leq y \leftrightarrow x < y \vee x = y$, or
(4) $x \leq y \leftrightarrow x < y \vee (\neg(x < y) \wedge \neg(y < x))$.

Optimality theory only considers linearly ordered constraints. These will be seen to lead to a *quasi-linear order* of preferences, i.e. a relation $\preccurlyeq$ that satisfies all the requirements of a non-strict linear order but antisymmetry. A quasi-linear ordering contains *clusters* of elements that are 'equally large'. Such elements are $\leq$ each other. Most naturally one would take for the strict variant $\prec$ an irreflexive, transitive, total relation. If one does that, strict and non-strict orderings can still be translated into each other (only by using alternatives (2) and (4) above though, not (1) and (3)). However, $Pref$ is normally taken to be an asymmetric relation, and we agree with that, so we take the option of $\prec$ as an irreflexive, transitive, asymmetric relation. Then $\prec$ is definable in terms of $\preccurlyeq$ by use of (2), but not $\preccurlyeq$ in terms of $\prec$. That is clear from the picture below, an irreflexive, transitive, asymmetric relation cannot distinguish between the two given orderings.



Incomparability and indifference.

One needs an additional equivalence relation $x \sim y$ to express that x and y are elements in the same cluster; $x \sim y$ can be defined by

(5) $x \sim y \leftrightarrow x \leq y \wedge y \leq x$.

Then, in the other direction, $x \leq y$ can be defined in terms of $<$ and $\sim$:

(6) $x \leq y \leftrightarrow x < y \vee x \sim y$.

It is certainly possible to extend our discussion to partially ordered sets of constraints, and we will make this excursion in section 6. The preference relation will no longer be a quasi-linear order, but a so-called *quasi-order*: in the non-strict case a reflexive and transitive relation, in the strict case an asymmetric, transitive relation. One can still use (2) to obtain a strict quasi-order from a non-strict one and (6) to obtain a non-strict quasi-order from a strict one and $\sim$. However, we will see in section 4 that in some contexts involving beliefs these translations no longer give the intended result. In such a case one has to be satisfied with the fact that (5) still holds and that $\prec$ as well as $\sim$ imply $\preccurlyeq$.

Let us return to the basic case that the constraint sequence is linearly ordered and the preference relation is obtained from the constraint sequence in the manner described in the previous section. We will write $Pref$ for the strict version of preference, $\underline{Pref}$ for the non-strict version, and let Eq correspond to $\sim$. Interestingly, no matter what the constraints are, the non-strict preference relation has the following general properties:

(a) $\underline{Pref}(x, x)$,
(b) $\underline{Pref}(x, y) \vee \underline{Pref}(y, x)$,
(c) $\underline{Pref}(x, y) \wedge \underline{Pref}(y, z) \rightarrow \underline{Pref}(x, z)$.

(a), (b) and (c) express reflexivity, totality and transitivity, respectively. Thus, as explained above, $\underline{Pref}$ is a quasi-linear relation; we lack antisymmetry. It turns out that (a), (b) and (c) are a complete set of principles for preference. We can show that these are the only stable properties that preference has in this context. Let us consider the universal theory **P** obtained from the above three axioms (with the variables interpreted as universally quantified) with modus ponens as the only rule.

Theorem 3.1 Completeness $\vdash_{\mathbf{P}} \varphi$ *iff* φ *is valid in all models obtained from constraint sequences.*

Proof. Assume formula $\varphi(x_1, \ldots, x_n)$ is not derivable in **P**. Then a non-strict quasi-linear ordering exists of some $d_1, \ldots, d_n$ which falsifies $\varphi(d_1, \ldots, d_n)$. Let us just assume that we have a linear order (adaptation to the more general case is simple), and also, w.l.o.g. that the ordering is $d_1 > d_2 > \cdots > d_n$. Then we introduce unary predicates $P_1, \ldots, P_n$ with a constraint sequence $P_1 \gg P_2 \cdots \gg P_n$ and let P_i apply only to d_i. Clearly then the preference order of $d_1, \ldots, d_n$ with respect to the given constraint sequence is from left to right. We have transformed the model into one in which the defined preference has the required properties. [1] ■

It is good to point out here that if one considers the objects as worlds and replaces the monadic predicates by propositional variables, the results so far can be restated in hybrid logic(see e.g. (Bla00)), provided formulas are restricted to be quantifier free. This has advantages and disadvantages. Stating the results in hybrid logic has the advantage that it makes the results more directly comparable to those of other papers that consider preference between worlds or propositions rather than objects. Our approach allows more complex predicate-logical formulas to be used as constraints. Also, it allows the generalizations to the belief contexts of the following sections. At this time we do not see how to obtain these generalizations in hybrid logic.

4 Preference and belief

In the above, we have assumed that objective information about the objects is always available (complete). In this section, we discuss the situation that arises when an agent has only incomplete information, but she needs to decide on her preference. The language will be extended with belief operators $B\varphi$[2] to deal with such uncertainty. Interestingly, the definitions of preference we propose in the following spell out different "procedures" an agent might follow to obtain her preference when processing the incomplete information under uncertainty. Which procedure is taken strongly depends on the domain or the type of agent. In the new language, the definition of constraint remains, i.e. a constraint C_i is a formula from the language *without* belief operators.

Definition 4.1 *Given a constraint sequence of length n,* ***Pref(x,y)*** *is defined as follows*[3]*:*

$$Pref_1(x,y) ::= BC_1(x) \wedge \neg BC_1(y),$$
$$Pref_{k+1}(x,y) ::= Pref_k(x,y) \vee (Eq_k(x,y) \wedge BC_{k+1}(x) \wedge \neg BC_{k+1}(y)), k < n,$$
$$Pref(x,y) ::= Pref_n(x,y), \quad (1)$$

where $Eq_k(x,y)$ stands for $(BC_1(x) \leftrightarrow BC_1(y)) \wedge \cdots \wedge (BC_k(x) \leftrightarrow BC_k(y))$.

[1] Note that, although we used n constraints in the above proof to make the procedure easy to describe, in general ${}^2log(n) + 1$ constraints are sufficient for the purpose.

[2] When we discuss issues in a multi-agent context, we write B_a, $a \in G$, a group of agents.

[3] Preference here is a doxastic notion, as it depends on beliefs, which is different from the notion of preference in (BL06)

To determine the preference relation, one just checks whether one believes the relevant properties of the objects. But at least two other options of defining preference seem reasonable as well.

Definition 4.2 *Given a constraint sequence of length n,* ***Pref(x,y)*** *is defined below:*

$$Pref_1(x,y) ::= BC_1(x) \wedge B\neg C_1(y),$$
$$Pref_{k+1}(x,y) ::= Pref_k(x,y) \vee (Eq_k(x,y) \wedge BC_{k+1}(x) \wedge B\neg C_{k+1}(y)), k < n,$$
$$Pref(x,y) ::= Pref_n(x,y) \quad (2)$$

where $Eq_k(x,y)$ *stands for* $(BC_1(x) \leftrightarrow BC_1(y)) \wedge (B\neg C_1(x) \leftrightarrow B\neg C_1(y)) \wedge \cdots \wedge (BC_k(x) \leftrightarrow BC_k(y)) \wedge (B\neg C_k(x) \leftrightarrow B\neg C_k(y))$.

Definition 4.3 *Given a constraint sequence of length n,* ***Pref(x,y)*** *is defined below:*

$$Supe_1(x,y)^4 ::= C_1(x) \wedge \neg C_1(y),$$
$$Supe_{k+1}(x,y) ::= Supe_k(x,y) \vee (Eq_k(x,y) \wedge C_{k+1}(x) \wedge \neg C_{k+1}(y)), k < n,$$
$$Supe(x,y) ::= Supe_n(x,y),$$
$$Pref(x,y) ::= B(Supe(x,y)), \quad (3)$$

where $Eq_k(x,y)$ *stands for* $(C_1(x) \leftrightarrow C_1(y)) \wedge \cdots \wedge (C_k(x) \leftrightarrow C_k(y))$.

To better understand the difference between the above three definitions, we look at the Example 1.1 again, but in three different variations:

A. Alice favors (1): She looks at what information she can get, she reads that d_1 has a low cost, about d_2 there is no information. This immediately makes her decide for d_1. She will stick with this decision no matter what she may hear about quality or neighborhood.

B. Bob favors (2): He gets the same information. But he has no preference, and that will remain so as long as he hears nothing about the cost of d_2, no matter what he hears about quality or neighborhood.

C. Carol favors (3): She also has the same information. On that basis Carol cannot decide. But some additional information about quality and neighborhood helps her. For instance, when she hears that d_1 is of good quality or is in a good neighborhood, and d_2 is not of good quality and not in a good neighborhood. Then Carol believes that, no matter what, d_1 is superior, so d_1 is her preference.

Speaking more generally in terms of the behaviors of the above agents, it seems that Alice always decides what she prefers on the basis of limited information. In contrast, Bob may choose to wait and request more information. Carol behaves somewhat differently, she first tries to do some reasoning with the available information before making her decision. This suggests a good perspective to think of diversity of agents in general. Apparently, we have the following fact.

Fact 4.4 - *Totality holds for* (1)*, but not for* (2) *or* (3)*;*

[4]Superiority is just defined as preference is in the previous section.

- *Among the above three definitions, we have* (2)$\rightarrow$ (1) *and* (2)$\rightarrow$ (3)*, but* (1) *and* (3) *are incomparable.*

It is striking that, if in definition (3), one plausibly also defines $\underline{Pref}(x, y)$ as $B(\underline{Supe}(x, y))$, then the normal relation between $Pref$ and $\underline{Pref}$ no longer holds: $Pref$ is not definable in terms of $\underline{Pref}$, or even $\underline{Pref}$ in terms of $Pref$ and Eq. This is the possibility we alluded to in section 3 just below (6).

Let us assume the normal principles of $KD45$ for B. For all three definitions, we have the following theorem.

Theorem 4.5 $\underline{Pref}(x, y) \leftrightarrow B\underline{Pref}(x, y)$.

Proof. In fact we prove something more general in $KD45$. Namely, if α is a propositional combination of B-statements, then $\vdash_{KD45} \alpha \leftrightarrow B\alpha$.

¿From left to right, since α is a propositional combination of B-statements, it can be transformed into disjunctive normal form: $\beta_1 \vee ... \vee \beta_k$. It is clear that $\vdash_{KD45} \beta_i \rightarrow B\beta_i$ for each i, because each member γ of the conjunction β_i implies $B\gamma$. If $\alpha = \beta_1 \vee ... \vee \beta_k$ holds then some β_i holds, so $B\beta_i$, so $B\alpha$. Then we immediately have: $\vdash_{KD45} \neg\alpha \rightarrow B\neg\alpha$ (*) as well, since $\neg\alpha$ is also such a statement if α is.

¿From right to left: Suppose $B\alpha$ and $\neg\alpha$. Then $B\neg\alpha$ by(*), so $B\bot$, but this is impossible in $KD45$, therefore α holds.

The theorem follows, as $\underline{Pref}(x, y)$ consists of B-statements. ■

Corollary 4.6 $\neg\underline{Pref}(x, y) \leftrightarrow B\neg\underline{Pref}(x, y)$.

Actually, we think Theorem 4.5 should be the case because we believe that preference describes a state of mind in the same way that belief does. Just as one knows what one believes, one knows what one prefers. So, in fact we think that in a setting where knowledge would be discussed, $\underline{Pref}(x, y) \leftrightarrow K\underline{Pref}(x, y)$ holds just as well.

If we stick to Definition (1), we can generalize the completeness result (Theorem 3.1). Let us consider the language built up from standard propositional letters, plus $\underline{Pref}(x, y)$ (with x, y variables) by the connectives, and belief operators B. Again we have the normal principles of $KD45$ for B.

Definition 4.7 *Consider the B-Pref system including the following valid principles, Modus ponens(MP), as well as Generalization for the operator B.*

(a) $\underline{Pref}(x, x)$,
(b) $\underline{Pref}(x, y) \vee \underline{Pref}(y, x)$,
(c) $\underline{Pref}(x, y) \wedge \underline{Pref}(y, z) \rightarrow \underline{Pref}(x, z)$,
(1.) $\neg B\bot$,
(2.) $B\varphi \rightarrow BB\varphi$,
(3.) $\neg B\varphi \rightarrow B\neg B\varphi$,
(4.) $\underline{Pref}(x, y) \leftrightarrow B\underline{Pref}(x, y)$.

We will write BP for the whole system, and B for $KD45$.

Theorem 4.8 Completeness $\vdash_{\mathsf{BP}} \varphi$ *iff* φ *is valid in all models obtained from constraint sequences.*

Proof. Suppose that $\nvdash_{\mathsf{BP}} \varphi(x_1, ..., x_n, P_1, ..., P_m)$. Also for some sequence of constants $d_1, ..., d_n$ we have $\nvdash_{\mathsf{BP}} \varphi(d_1, ..., d_n, P_1, ..., P_m)$. Consider the set Π consisting of the following formulas:

(1) $Pref(d_i, d_i)$,
(2) $Pref(d_i, d_j) \vee Pref(d_j, d_i)$,
(3) $Pref(d_i, d_j) \wedge Pref(d_j, d_k) \rightarrow Pref(d_i, d_k)$,
(4) $BPref(d_i, d_j) \leftrightarrow Pref(d_i, d_j)$,
(5) $BPref(d_i, d_i)$,
(6) $B(Pref(d_i, d_j) \vee Pref(d_j, d_i))$,
(7) $B(Pref(d_i, d_j) \wedge Pref(d_j, d_k) \rightarrow Pref(d_i, d_k))$,
(8) $B(BPref(d_i, d_j) \leftrightarrow Pref(d_i, d_j))$.

where $i, j, k \in \{1, ..., n\}$. So Π is finite, and $\nvdash_{\mathsf{B}} \Pi \rightarrow \varphi(d_1, ..., d_n, P_1, ..., P_m)$. So, there is a world w in a B-model such that $w \vDash \Pi$ and $w \nvDash \varphi(d_1, ..., d_n, P_1, ..., P_m)$. From $w \vDash \Pi$ by (4) and (8) it follows that the preference relations are the same quasi-linear ordering everywhere in the model[5]. Then $w \nvDash \varphi(d_1, ..., d_n, P_1, ..., P_m)$ boils down to the situation we had in Theorem 3.1. Just as there, we can transform the model into one with the required properties of the constraint sequence. ■

To prove a similar theorem for more agents seems to be more difficult. Ordinary completeness can be obtained, but the problem is to find fitting constraints.

5 Preference Changes

Let us first look at a variation of Example 1.1:

Example 5.1 Alice won a lottery prize of one million dollars. Her situation has changed dramatically. She considers the quality most important.

In other words, the ordering of the constraints has changed. Of course, one can continue to design more variations of that kind. We will focus on the constraint changes, and how they bring about preference change. To discuss such issues, we first make the constraint sequence explicit in the preference. We do this for the simple language. For the language with belief, corresponding changes should be made. Let $\mathfrak{C}$ be a constraint sequence with fixed length n as in Definition 2.3. Then we write $Pref_{\mathfrak{C}}(x, y)$ for the preference defined from that constraint sequence. Let us write $\mathfrak{C}^\frown C$ for adding C to the right of $\mathfrak{C}$, $C^\frown \mathfrak{C}$ for adding C to the left of $\mathfrak{C}$, $\mathfrak{C}^-$ for the sequence $\mathfrak{C}$ with its final element deleted, and finally, $\mathfrak{C}^{i \leftrightarrows i+1}$ for the sequence $\mathfrak{C}$ with its i-th and i+1-th switched. It is clear that we have the following relationships:

[5]Note that at this point an ordinary completeness result without constraints has been established.

$Pref_{\mathfrak{C}^\frown C}(x,y) \leftrightarrow Pref_{\mathfrak{C}}(x,y) \vee (Eq_{\mathfrak{C}}(x,y) \wedge C(x) \wedge \neg C(y))$,
$Pref_{C^\frown \mathfrak{C}}(x,y) \leftrightarrow (C(x) \wedge \neg C(y)) \vee ((C(x) \leftrightarrow C(y)) \wedge Pref_{\mathfrak{C}}(x,y))$,
$Pref_{\mathfrak{C}^-}(x,y) \leftrightarrow Pref_{\mathfrak{C},n-1}(x,y)$,
$Pref_{\mathfrak{C}^{i \leftrightarrows i+1}}(x,y) \leftrightarrow Pref_{\mathfrak{C},i-1}(x,y) \vee (Eq_{\mathfrak{C},i-1}(x,y) \wedge C_{i+1}(x) \wedge \neg C_{i+1}(y)) \vee (Eq_{\mathfrak{C},i-1}(x,y) \wedge$
$(C_{i+1}(x) \leftrightarrow C_{i+1}(y)) \wedge C_i(x) \wedge \neg C_i(y)) \vee (Eq_{\mathfrak{C},i+1}(x,y) \wedge Pref_{\mathfrak{C}}(x,y))$.

Now we can describe preference change due to changes of the constraint sequence. We consider the operations $[^+C]$ of adding C to the right, $[C^+]$ of adding C to the left, $[-]$ of dropping the last element of a constraint sequence of length n, $[i \leftrightarrow i+1]$ of interchanging the i-th and $i+1$-th elements. Then we have the following reduction axioms:

$[^+C]Pref(x,y) \leftrightarrow Pref(x,y) \vee (Eq(x,y) \wedge C(x) \wedge \neg C(y))$,
$[C^+]Pref(x,y) \leftrightarrow ((C(x) \wedge \neg C(y)) \vee ((C(x) \leftrightarrow C(y)) \wedge Pref(x,y)))$,
$[-]Pref(x,y) \leftrightarrow Pref_{n-1}(x,y)$,
$[i \leftrightarrow i+1]Pref(x,y) \leftrightarrow Pref_{i-1}(x,y) \vee (Eq_{i-1}(x,y) \wedge C_{i+1}(x) \wedge \neg C_{i+1}(y)) \vee (Pref_i(x,y) \wedge$
$(C_{i+1}(x) \leftrightarrow C_{i+1}(y))) \vee (Eq_{i+1}(x,y) \wedge Pref(x,y))$.

Of course, the first two are the more satisfactory ones. Now we move to changes in belief, which may cause preference change as well. This occurs often in a multi-agent situation, as say, two agents cooperate to make a better choice. Technically, the update mechanisms of (BMS98) can be immediately applied to our system with belief to calculate preference changes. Consider Example 1.1, but with two agents Alice and Bob:

Example 5.2 This time Alice and Bob only consider the houses' cost (C) and their neighborhood (N), both with $C(x) \gg N(x)$. There are two houses d_1 and d_2 available. The real situation is that $C(d_1), N(d_1), C(d_2)$ and $\neg N(d_2)$. First Alice prefers d_2 over d_1 because she believes $C(d_2)$ and $N(d_1)$. However, Bob's preference is the opposite one because he believes $C(d_1)$. Now Bob tells Alice his belief $C(d_1)$ and Alice accepting this also believes $C(d_1)$, and accordingly changes her preference.

Here we assume that the beliefs of Alice and Bob about these simple statements are very solid (like knowledge) and that is so understood by both. The following diagram shows the situation before Bob's statement,

	$C \wedge N$	$C \wedge \neg N$	$\neg C \wedge N$	$\neg C \wedge \neg N$
a	d_1, d_2	d_2	d_1	$\diagdown$
b	d_1, d_2	d_1, d_2	d_2	d_2

The constraint sequence reads from left to right, and objects are put in the places where they possibly have the properties. After Bob tells Alice that $C(d_1)$, the situation becomes

	$C \wedge N$	$C \wedge \neg N$	$\neg C \wedge N$	$\neg C \wedge \neg N$
a^*	d_1, d_2	d_2	$\diagdown$	$\diagdown$
b	d_1, d_2	d_1, d_2	d_2	d_2

The possibility of $\neg C \wedge N$ of d_1 has been eliminated from the first table: Alice updated her beliefs. Now she prefers d_1 over d_2.

We have assumed that we are using the elimination semantics (e.g. (FHMV03), (Ben06)) in which public announcement of the sentence A leads to the elimination of the $\neg A$ worlds from the model. We have the reduction theorem

$[!A]Pref_{\mathfrak{C}}(x,y) \ \leftrightarrow \ A \rightarrow Pref_{A \rightarrow \mathfrak{C}}(x,y)$,

where $A \rightarrow \mathfrak{C}$ is the constraint sequence obtained by replacing each $C_i(x)$ by $A \rightarrow C_i(x)$.

6 Further Discussions and Conclusions

A new situation occurs when there are several constraints of incomparable strength. In Example 1.1 now, Alice also takes the 'transportation convenience' into account. But for her neighborhood and transportation convenience are really incomparable. Abstractly speaking, it means that the constraint sequence is now *partially ordered*. We show in the following how to define preference based on a partially ordered constraint sequence. We consider a set of constraints $C_1 ..., C_n$ with the relation $\gg$ between them a partial order.

Definition 6.1 *We define $Pref_m(x,y)$, $Eq_m(x,y)$, $Pref_{\{n_1,...,n_k\}}(x,y)$ and $Eq_{\{n_1,...,n_k\}}(x,y)$ in a simultaneous induction, where $\{n_1, ..., n_k\}$ is a set of incomparable nodes:*

$Pref_{\{n_1,...,n_k\}}(x,y) ::= (\underline{Pref_{n_1}}(x,y) \wedge ... \wedge \underline{Pref_{n_k}}(x,y)) \wedge (Pref_{n_1}(x,y) \vee ... \vee Pref_{n_k}(x,y))$
($\top$ *if the set is empty.*)
$Eq_{\{n_1,...,n_k\}}(x,y) ::= Eq_{n_1} \wedge ... \wedge Eq_{n_k}$. ($\top$ *if the set is empty.*)

Let S be the set of immediate predecessors of m, we define

$Pref_m(x,y) ::= Pref_S(x,y) \vee (Eq_S(x,y) \wedge C_m(x) \wedge \neg C_m(y))$.
$Eq_m(x,y) ::= Eq_S(x,y) \wedge (C_m(x) \leftrightarrow C_m(y))$.

For more discussion on the relation between partially ordered constraints and G-spheres, see (Lew81). The situation that the set of constraints is unordered leads to the position advocated by (Kra81).

Conclusions Inspired by optimality theory, the notion of constraint was introduced. We defined preference in terms of a constraint sequence and properties of objects in a simpler language. We then added beliefs explicitly into the language, and proposed three possible ways to define preference. We have proved completeness for both the simple language and the language with beliefs. Preference change was investigated in two possible ways as well: changes in the constraint sequence, and change of beliefs. Reduction axioms were presented. Finally, we discussed how to get preference from a partially ordered constraint sequence. For further study, we would like to extend our framework immediately to a multi-agent context and explore more how the interaction between agents can change their preference. Also, we are aware that a large amount of research on preference has been done in social choice theory and computer science. We would like to compare our approach with such work. As mentioned earlier other types of constraints are used in such research, often with weights. We do think our methods are applicable quite generally. Also, if only for comparison's sake, we will study preference between states (or propositions).

Acknowledgement We would like to thank J. van Benthem, R. Blutner, U. Endriss, T. Mijangos Martínez, F. Roelofsen, T. Yamada, H. Zeevat, and two anonymous reviewers for comments and helpful suggestions.

Bibliography

J. van Benthem. 'one is a lonely number': On the logic of communication. to appear in Z. Chatzidakis, P. Koepke, W.Pohlers, eds., 'Logic Collquium '02', ASL Lecture Notes in Logic 27, AMS Publications, Providence (R.I.), 2006.

J. van Benthem and F. Liu. Dynamic logic of preference upgrade. *Journal of Applied Non-Classical Logic*, 2006. To appear.

P. Blackburn. Representation, reasoning, and relational structures: a hybrid logic manifesto. *Logic Journal of the IGPL*, 8:339–365, 2000.

A. Baltag, L.S. Moss, and S. Solecki. The logic of common knowledge, public announcements, and private suspicions. In I. Gilboa, editor, *Proceedings of the 7th conference on theoretical aspects of rationality and knowledge (TARK 98)*, pages 43–56, 1998.

S. Coste-Marquis, J. Lang, Liberatore P., and P. Marquis. Expressive power and succinctness of propositional languages for preference representation. In *Proc. 9th International Conference on Principles of Knowledge Representation and Reasoning (KR-2004)*. AAAI Press., 2004.

J. Doyle and M.P. Wellman. Representing preferences as ceteris paribus comparatives. *Working Notes of the AAAL Symposium on Decision-Theoretic Planning*, 1994.

R. Fagin, J. Y. Halpern, Y. Moses, and M. Y. Vardi. *Reasoning about Knowledge*. The MIT Press, 2003.

S. Hansson. *Preference logic*, volume 4 of *Handbook of philosophical logic*, chapter 4, pages 319–393. Kluwer, 2001.

A. Kratzer. Partition and revision: the semantics of counterfactuals. *Journal of Philosophical Logic*, 10:201–216, 1981.

D. Lewis. *Counterfactuals*. Harvard University Press, 1973.

D. Lewis. Ordering semantics and premise semantics for counterfactuals. *Journal of Philosophical Logic*, 10:217–234, 1981.

A. Prince and P. Smolensky. *Optimality Theory: Constraint Interaction in Generative Grammar*. Malden, Ma: Blackwell, 1993.

G. H. von Wright. *The Logic of Preference*. Edinburgh, 1963.