Chapter 8
Belief-Based Preference

8.1 Introduction

In this chapter, we resume the issue of "entanglement" of preference and information-based notions like knowledge and belief, that we have studied already in Chapter 5. How does this play when preference comes with richer priority structure?

To plunge right in, let us consider a variation of Example 7.1:

Example 8.1 (buying a house under uncertainty) Alice is going to buy a house. For her there are several things to consider: the cost, the quality and the neighborhood, strictly in that order. Consider two houses d_1 and d_2 that Alice hopes to choose from. Alice only has partial information. Let us assume that she *believes* that d_1 and d_2 have the following properties: $C(d_1)$, $C(d_2)$, $\neg Q(d_1)$, $\neg Q(d_2)$, $N(d_1)$ and $\neg N(d_2)$.

The definition of preference proposed in Chapter 7 does not apply here anymore, as beliefs have entered now. Alice's decision is not determined by her complete information, but by her beliefs under uncertainties. In a more general sense, this allows us to consider more complex scenarios. For instance, do we believe certain properties from the priority base to apply or not? Or even more dramatically, can we form a priority base on the basis of our beliefs? Handling uncertainties of this kind calls for a combination of a doxastic language and a preference language.

For this purpose, the preference language defined in the previous chapter will be extended now with belief operators $B\varphi$. When we do this, it may seem that we are heading into doxastic predicate logic. This is true, but we are not going to be affected by the existing difficult issues in interpreting modal predicate logics (cf. [80]). What we are using in this context is just a very limited part of such a language.[1] We will take the standard modal system **KD45** as the logic for belief, though we are

[1] It would be interesting to consider what more a full doxastic predicate logic language can bring to our preference setting, but we will leave this question to other occasions.

F. Liu, *Reasoning about Preference Dynamics*, Synthese Library 354,
DOI 10.1007/978-94-007-1344-4_8, © Springer Science+Business Media B.V. 2011

aware of the philosophical debates about beliefs and the many options for designing appropriate logical systems.[2]

This chapter is structured as follows. in Section 8.2 we propose three different ways of defining preference in terms of priorities and beliefs. In particular, we will present a doxastic preference logic for the notion of "decisive preference" and prove an extended representation theorem for that case. Section 8.3 will extend our discussions to multi-agent case, in which we will particularly study both cooperative agents and competitive agents, and describe their characteristics in representation theorems. In Section 8.4 we move to preference over propositions, and propose a propositional doxastic preference logic. And also, we will explore the relationship between preference over objects and preference over propositions. Finally, we restate our main points in the conclusions.

8.2 Doxastic Preference Logic

8.2.1 Three Notions of Belief-Based Preference

Working with beliefs, we will first give several definitions of preference in terms of priority sequence in this section. Interestingly, the definitions we consider in the following spell out different "procedures" an agent may follow to decide her preference when processing the incomplete information about the relevant properties. Which procedure is taken strongly depends on the domain or the type of agents. Moreover, we consider a simpler scenario, namely, in the new language, the definition of priority sequence remains the same, i.e., a priority C_i is a formula from the language *without* belief operators.[3]

Definition 8.2 (decisive preference) Given a priority sequence of length n, two objects x and y, $\mathit{Pref}(x, y)$ is defined as follows:

$\mathit{Pref}_1(x, y) := BC_1(x) \wedge \neg BC_1(y)$,
$\mathit{Pref}_{k+1}(x, y) := \mathit{Pref}_k(x, y) \vee (Eq_k(x, y) \wedge BC_{k+1}(x) \wedge \neg BC_{k+1}(y)), k < n$,
$\mathit{Pref}(x, y) := \mathit{Pref}_n(x, y)$,

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

To determine the preference relation, one just runs through the sequence of relevant properties to check whether one believes them of the objects. But at least two other options of defining preference seem reasonable as well.

Definition 8.3 (conservative preference) Given a priority sequence of length n, two objects x and y, $\mathit{Pref}(x, y)$ is defined below:

---

[2] Readers who liked our plausibility models for belief in Chapters 4, 5, may also just continue thinking in these terms when reading what we have to say about the doxastic modality $B\varphi$.

[3] It would be also interesting to look at non-factual priorities containing beliefs of the agents.

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

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 8.4 (deliberate preference) Given a priority sequence of length n, two objects x and y, $Pref(x, y)$ is defined below:

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

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 8.1 again, but in three different variations:

A. Alice favors Definition 8.2: She looks at what information she can get, she reads that d_1 has low cost, about d_2 there is no information. This immediately makes her decide for d_1. This will remains so, no matter what she hears about quality or neighborhood.
B. Bob favors Definition 8.3: The same thing happens to him. But he reacts differently than Alice. 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. Cora favors Definition 8.4: She also has the same information. On that basis Cora cannot decide either. But some more information about quality and neighborhood helps her to decide. For instance, suppose she hears that d_1 has good quality or is in a good neighborhood, and d_2 is not of good quality and not in a good neighborhood. Then Cora believes that, no matter what, d_1 is superior, so d_1 is her preference. Note that such kind of information could not help Bob to decide.

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 the limited information she has. In contrast, Bob chooses to wait and require more information. Cora behaves somewhat differently, she first tries to do some reasoning with all the available information before making her decision. This suggests yet another perspective on diversity of agents than discussed in [132].

Clearly, then, we have the following fact:

[4] Superiority is just defined as preference was in Chapter 7.

Fact 8.5

- *Totality holds for Definition 8.2, but not for Definition 8.3 or 8.4;*
- *Among the above three definitions, Definition 8.3 is the strongest in the sense that if $Pref(x, y)$ holds according to Definition 8.3, then $Pref(x, y)$ holds according to Definition 8.2 and 8.4 as well.*

It is striking that, if in Definition 8.4, 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 with $\underline{Pref}$ any more, or even $\underline{Pref}$ in terms of $Pref$ and Eq.

For all three definitions, we have the following theorem.

Theorem 8.6 $\underline{Pref}(x, y) \leftrightarrow B\underline{Pref}(x, y)$.

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

From left to right, since α is a propositional combination of B-statements, it can be transformed into conjunctive normal form: $\beta_1 \vee \cdots \vee \beta_k$. It is clear that $\vdash_{\mathbf{KD45}} \beta_i \to B\beta_i$ for each i, because each member γ of the conjunction β_i implies $B\gamma$. If $A = \beta_1 \vee \cdots \vee \beta_k$ holds then some β_i holds, so $B\beta_i$, so $B\alpha$. Then we immediately have: $\vdash_{\mathbf{KD45}} \neg\alpha \to B\neg\alpha$ (*) as well, since $\neg\alpha$ is also a propositional combination of B-statements 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 since $\underline{Pref}(x, y)$ is in all three cases indeed a propositional combination of B-statements. □

Corollary 8.7 $\neg\underline{Pref}(x, y) \leftrightarrow B\neg\underline{Pref}(x, y)$.

Actually, we think it is proper that Theorem 8.6 and Corollary 8.7 hold because we believe that preference describes a state of mind in the same way that belief does. Just as one believes what one believes, one believes what one prefers.

8.2.2 *Doxastic Preference Logic*

If we stick to Definition 8.2, we can generalize the representation result from the previous chapter. Let us consider the reduced language built up from standard propositional letters, plus $\underline{Pref}(d_i, d_j)$ by the connectives, and belief operators B. Again we have the normal principles of **KD45** for B.

Theorem 8.8 *The following principles axiomatize exactly the valid ones.*

(1) $\underline{Pref}(d_i, d_i)$.

(2) $\underline{Pref}(d_i, d_j) \vee \underline{Pref}(d_j, d_i)$.

(3) $\underline{Pref}(d_i, d_j) \wedge \underline{Pref}(d_j, d_k) \to \underline{Pref}(d_i, d_k)$.

(4) $\neg B\bot$.

(5) $B\varphi \to BB\varphi$.

(6) $\neg B\varphi \to B\neg B\varphi$.

(7) $\underline{Pref}(d_i, d_j) \leftrightarrow B\underline{Pref}(d_i, d_j)$.

We now consider the **KD45-P** system including the above valid principles, *Modus ponens*(MP), as well as *Generalization* for the operator B.

Definition 8.9 (doxastic preference model) A doxastic preference model of **KD45-P** is a tuple $(S, D, R, \{\preceq_s\}_{s\in S}, V)$, where S is a non-empty set of worlds, D is a set of constants, R is a euclidean, transitive, and serial accessibility relation on S. Namely, it satisfies $\forall xyz((Rxy \wedge Rxz) \to Ryz)$, $\forall xyz((Rxy \wedge Ryz) \to Rxz)$, and $\forall x\exists y Rxy$. For each s, $\preceq_s$ is a quasi-linear order on D, which is the same throughout each euclidean class. V is evaluation function in an ordinary manner.

We remind the reader that in most respects euclidean classes are equivalence classes except that a number of points are irreflexive and have R relations just towards the reflexive members (the *equivalence part*) of the class.

Theorem 8.10 *The* **KD45-P** *system is complete.*

Proof The canonical model of this logic **KD45-P** has the required properties: The belief accessibility relation R is euclidean, transitive, and serial. This means that with regard to R the model falls apart into euclidean classes. In each node $\underline{Pref}$ is a quasi-linear order of the constants. Within a euclidean class the preference order is constant (by $B\underline{Pref} \leftrightarrow \underline{Pref}$). This suffices to prove completeness. □

Theorem 8.11 *The logic* **KD45-P** *has the finite model property.*

Proof By standard methods. □

Theorem 8.12 (representation theorem) $\vdash_{\mathbf{KD45-P}} \varphi$ *iff* φ *is valid in all models obtained from priority sequences.*

Proof Suppose that $\nvdash_{\mathbf{KD45-P}} \varphi(d_1, ..., d_n, p_1, ..., p_m)$. By Theorem 8.10, there is a model with a world w in which φ is falsified. We restrict the model to the euclidean class where w resides. Since the ordering of the constants is the same throughout euclidean classes, the ordering of the constants is now the same throughout the whole model. We can proceed as in Theorem 7.9 defining the predicates $P_1, \ldots, P_n$ in a constant manner throughout the model. □

Remark 8.13 The three definitions above are not the only definitions that might be considered. For instance, we can give a variation (*) of Definition 8.3. For simplicity, we just use one predicate C.

$$Pref(x, y) := \neg B\neg C(x) \wedge B\neg C(y). \quad (*)$$

This means the agent can decide on her preference in a situation in which on the one hand she is not totally ready to believe $C(x)$, but considers it consistent with what she assumes, on the other hand, she distinctly believes $\neg C(y)$. Compared with Definition 8.3, (*) is weaker in the sense that it does not require explicit positive beliefs concerning $C(x)$.

We can even combine Definition 8.2 and (*), obtaining the following:

$$Pref(x, y) := (BC(x) \wedge \neg BC(y)) \vee (\neg B\neg C(x) \wedge B\neg C(y)). \quad (^{*}*)$$

Contrary to (*), this gives a quasi-linear order.

Similarly, for Definition 8.4, if instead of $B(Supe(x, y))$, we use $\neg B\neg(Supe(x, y))$, a weaker preference definition is obtained.

8.3 Extension to the Multi-agent Case

8.3.1 Multi-agent Doxastic Preference Logic

This section extends the results of Section 8.2 to the many agent case. This will generally turn out to be more or less a routine matter. But at the end of the section, we will see that the priority base approach gives us a start of an analysis of cooperation and competition of agents. We consider agents here as cooperative if they have the same goals (priorities), competitive if they have opposite goals. This foreshadows the direction one may take to apply our approach to games. The language we are using is defined as follows:

Definition 8.14 (reduced doxastic preference language) Let Φ be a set of propositional variables, N be a group of agents, and D be a finite domain of objects, the reduced doxastic preference language for many agents is defined in the following:

$$\varphi := p \mid \neg\varphi \mid \varphi \wedge \psi \mid \underline{Pref}^a(d_i, d_j) \mid B^a\varphi \quad \text{where } p \in \Phi, a \in N \text{ and } d_i \in D.$$

Similarly to $\underline{Pref}^a$ expressing non-strict preference, we will use $Pref^a$ to denote the strict version. When we want to use the extended language, we add variables and the statements $P(d_i)$.

Definition 8.15 (priority sequence for agent a) A priority sequence for agent a is a finite ordered sequence of formulas written as follows: $C_1 \gg_a C_2 \cdots \gg_a C_n (n \in \mathbb{N})$, where each C_m $(1 \leq m \leq n)$ is a formula from the language of Definition 8.14, with one single free variable x, but without $\underline{Pref}$ and B.

Here we take decisive preference to define an agent's preference. But the results of this section apply to other definitions just as well. It seems quite reasonable to allow in this definition of $Pref^a$ formulas that contain B^b and $Pref^b$ for agents b other than a. But we leave this for a future occasion.

Definition 8.16 (preference for agent a) Given a priority sequence of length n, two objects x and y, $Pref^a(x, y)$ is defined as follows:

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

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

Definition 8.17 The doxastic preference logic for many agents **KD45-P^G** is consists of the following principles,

(1) $\underline{Pref}^a(d_i, d_i)$.
(2) $\underline{Pref}^a(d_i, d_j) \vee \underline{Pref}^a(d_j, d_i)$.
(3) $\underline{Pref}^a(d_i, d_j) \wedge \underline{Pref}^a(d_j, d_k) \to \underline{Pref}^a(d_i, d_k)$.
(4) $\neg B^a \bot$.
(5) $B^a \varphi \to B^a B^a \varphi$.
(6) $\neg B^a \varphi \to B^a \neg B^a \varphi$.
(7) $\underline{Pref}^a(d_i, d_j) \leftrightarrow B^a \underline{Pref}^a(d_i, d_j)$.

As usual, it also includes *Modus ponens*(MP), as well as *Generalization* for the operators B^a. It is easy to see that the above principles are valid for $\underline{Pref}^a$ extracted from a priority sequence.

Theorem 8.18 *The doxastic preference logic for many agents* **KD45-P^G** *is completely axiomatized by the stated principles.*

Proof The canonical model of this logic **KD45-P^G** has the required properties: The belief accessibility relation R_a is euclidean, transitive, and serial. This means that with regard to R_a the model falls apart into a-euclidean classes. Again, in each node $Pref^a$ is a quasi-linear order of the constants and within an a-euclidean class the a-preference order is constant. This quasi-linearity and constancy are of course the required properties for the preference relation. Same for the other agents. This shows completeness of the logic. □

Theorem 8.19 *The logic* **KD45-P^G** *has the finite model property.*

Proof By standard methods. □

Similarly, a representation theorem can be obtained by showing that the model could have been obtained from priority sequences $C_1 \gg_a C_2 \cdots \gg_a C_m (m \in \mathbb{N})$ for all the agents.

Theorem 8.20 (representation theorem) $\vdash_{\mathbf{KD45-P^G}} \varphi$ *iff* φ *is valid in all models with each* $\underline{Pref}^a$ *obtained from a priority sequence.*

Proof Let there be k agents $a_0, \ldots, a_{k-1}$ and suppose $\varphi(d_1, \ldots, d_n)$. We provide each agent a_j with her own priority sequence $P_{n\times j+1} \gg_{a_j} P_{n\times j+2} \gg_{a_j} \cdots \gg_{a_j} P_{n\times(j+1)}$. It is sufficient to show that any model for **KD45-P$^{\mathbf{G}}$** for the reduced language can be extended by valuations for the $P_j(d_i)$'s in such a way that the preference relations are preserved. For each a_i-euclidean class, we follow the same procedure for $d_1, \ldots, d_n$ w.r.t. $P_{n\times j+1}$, $P_{n\times j+2}$, ..., $P_{n\times(j+1)}$ as in Theorem 7.9 w.r.t $P_1, \ldots, P_n$. The preference orders obtained in this manner are exactly the Pref^{a_j} relations in the model. □

8.3.2 Cooperative and Competitive Agents

In the above case, the priority sequences for different agents are separate, and thus very different. Still stronger representation theorems can be obtained by requiring that the priority sequences for different agents are related, e.g. in the case of *cooperative agents* that they are equal. We will consider the two agent case in the following.

Theorem 8.21 (two cooperative agents) $\vdash_{\mathbf{KD45-P^G}} \varphi$ *iff* φ *is valid in all models obtained from priority sequences shared by two cooperative agents.*

Proof The two agents are a and b. We now have the priority sequence $P_1 \gg_a P_2 \gg_a \cdots \gg_a P_n$, same for b. It is sufficient to show that any model $\mathfrak{M}$ with worlds W for **KD45-P$^{\mathbf{G}}$** for the reduced language can be extended by valuations for the $P_j(d_i)$'s in such a way that the preference relations are preserved. We start by making all $P_j(d_i)$'s true everywhere in the model. Next we extend the model as follows. For each a-euclidean class E in the model carry out the following procedure. Extend $\mathfrak{M}$ with a complete copy $\mathfrak{M}_E$ of $\mathfrak{M}$ for all of the reduced language i.e. without the predicates P_j. Add R_a relations from any of the w in E to the copies v_E such that $w\ R_a\ v$. Now carry out the same procedure as in the proof of Theorem 7.9 in E's copy E_E. What we do in the rest of $\mathfrak{M}_E$ is irrelevant. Now, in w, a will believe in $P_j(d_i)$ exactly as in the model in the previous proof, the overall truth of $P_j(d_i)$ in the a-euclidean class E in the original model has been made irrelevant. The preference orders obtained in this manner are exactly the Pref^a relations in the model. All formulas in the reduced language keep their original valuation because the model $\mathfrak{M}_E$ is bisimilar for the reduced language to the old model $\mathfrak{M}$ as is the union of $\mathfrak{M}$ and $\mathfrak{M}_E$.

Finally do the same thing for b: Add for each b-euclidean class in $\mathfrak{M}$ a whole new copy, and repeat the procedure followed for a. Both a and b will have preferences with regard to the same priority sequence. □

For *competitive agents* we assume that if agent a has a priority sequence $D_1 \gg_a D_2 \gg \cdots \gg_a D_m (m \in \mathbb{N})$, then the opponent b has priority sequence $\neg D_m \gg_b \neg D_{m-1} \gg \cdots \gg_b \neg D_1$.

Theorem 8.22 (two competitive agents) $\vdash_{\mathbf{KD45-P^G}} \varphi$ *iff* φ *is valid in all models obtained from priority sequences for competitive agents.*

Proof Let's assume two agents a and b. For a we take a priority sequence $P_1 \gg_a P_2 \gg_a \cdots \gg_a P_n \gg_a P_{n+1} \gg_a \cdots \gg_a P_{2n}$, and for b, we take $\neg P_{2n} \gg_b \neg P_{2n-1} \gg_b \cdots \gg_b \neg P_n \gg_b \neg P_{n-1} \gg_b \cdots \gg_b \neg P_1$. It is sufficient to show that any model $\mathfrak{M}$ with worlds W for **KD45-P**$^{\mathbf{G}}$ for the reduced language can be extended by valuations for the $P_j(d_i)$'s in such a way that the preference relations are preserved. We start by making all $P_1(d_i) \ldots P_n(d_i)$ true everywhere in the model and $P_{n+1}(d_i) \ldots P_{2n}(d_i)$ all false everywhere in the model. Next we extend the model as follows.

For each a-euclidean class E in the model carry out the following procedure. Extend $\mathfrak{M}$ with a complete copy $\mathfrak{M}_E$ of $\mathfrak{M}$ for all of the reduced language i.e. without the predicates P_j. Add R_a relations from any of the w in E to the copies v_E such that $w\ R_a\ v$. Now define the values of the $P_1(d_i) \ldots P_n(d_i)$ in E_E as in the previous proof and make all $P_m(d_i)$ true everywhere for $m > n$. The preference orders obtained in this manner are exactly the $Pref^a$ relations in the model.

For each b-euclidean class E in the model carry out the following procedure. Extend $\mathfrak{M}$ with a complete copy $\mathfrak{M}_E$ of $\mathfrak{M}$ for all of the reduced language i.e. without the predicates P_j. Add R_b relations from any of the w in E to the copies v_E such that $w\ R_b\ v$. Now define the values of the $\neg P_{2n}(d_i) \ldots \neg P_{n+1}(d_i)$ in E_E as for $P_1(d_i) \ldots P_n(d_i)$ in the previous proof and make all $P_m(d_i)$ true everywhere for $m \leq n$. The preference orders obtained in this manner are exactly the $Pref^b$ relations in the model.

All formulas in the reduced language keep their original valuation because the model $\mathfrak{M}_E$ is bisimilar for the reduced language to the old model $\mathfrak{M}$ as is the union of $\mathfrak{M}$ and all the $\mathfrak{M}_E$. $\square$

Discussion These last representation theorems show that they are as is to be expected not only a strength but also a weakness. The weakness here is that they show that cooperation and competition cannot be differentiated in this language. On the other hand, the theorems are not trivial, one might think for example that if a and b cooperate, $B_a Pref_b(c, d)$ would imply $Pref_a(c, d)$. This is of course completely false, a and b can even when they have the same priorities have quite different beliefs about how the priorities apply to the constants. But the theorems show that no principles can be found that are valid only for cooperative agents. Moreover they show that if one wants to prove that $B_a Pref_b(c, d) \rightarrow Pref_a(c, d)$ is not valid for cooperative agents a counterexample to it in which the agents do not cooperate suffices.

8.4 Preference over Propositions

Most other authors on preference have discussed preference over propositions rather than objects. In this section, we will show that the current approach can be applied to preference over propositions as well. Following the previous section on belief-based preferences, we will propose a propositional system combining preference and beliefs. And we specially take the line that preference is a state of mind and that therefore one prefers one alternative over another if and only if one believes one does. If we take this line, the most obvious way would be to go to second-order logic and consider priority sequence $A_1(\varphi) \gg A_2(\varphi) \gg \ldots, \gg A_n(\varphi)$, where the A_i are properties of propositions. However, we find it close to our intuitions to stay first-order as much as possible. With that in mind, we define the new priority sequence for the propositional case as follows.

Definition 8.23 (propositional priority sequence) A propositional priority sequence is a finite ordered sequence of formulas written as follows:

$$\varphi_1(x) \gg \varphi_2(x) \gg \cdots \gg \varphi_n(x) \quad (n \in \mathbb{N}),$$

where each of $\varphi_m(x)$ is a propositional formula with an additional propositional variable, x, which is a common one to each $\varphi_m(x)$.

Formulas $\varphi(x)$ can express properties of propositions, for instance, applied to ψ, $x \to p_1$ expresses that ψ implies p_1, "ψ has the property p_1".

We apply our approach in previous sections to define preference in terms of beliefs. As we have seen in Section 8.2, there are various ways to do it. We are guided by the definition of decisive preference in formulating the following:

Definition 8.24 (preference over propositions) Given a propositional priority sequence of length n, we define preference over propositions ψ and θ as follows:

$$Pref(\psi, \theta) \quad \text{iff} \quad \text{for some } i, \quad (B\varphi_1(\psi) \leftrightarrow B\varphi_1(\theta)) \wedge \cdots \wedge (B\varphi_{i-1}(\psi) \leftrightarrow B\varphi_{i-1}(\theta)) \wedge (B\varphi_i(\psi) \wedge \neg B\varphi_i(\theta)).$$

Note that preference between propositions is in this case almost a preference between mutually exclusive alternatives: In the general case one can conclude beyond the quasi-linear order that derives directly from our method only that if $B(\psi \leftrightarrow \theta)$, then ψ and θ are equally preferable. Otherwise, any proposition can be preferable over any other.

For some purposes (this will get clearer in the proof of the representation theorem below), we need a further generalization, as in this slightly more complex definition:

Definition 8.25 A propositional priority sequence is a finite ordered sequence of sets of formulas written as follows:

$$\Gamma_1 \gg \Gamma_2 \gg \cdots \gg \Gamma_n,$$

where each set Γ_i consists of propositional formulas that have an additional propositional variable, x, which is a common one to each Γ_i.

A new matching definition of preference is then given by:

Definition 8.26 Given a propositional priority sequence of length n, we define preference over propositions ψ and θ as follows:

$$Pref(\psi, \theta) \quad \text{iff} \quad \exists i (\forall j < i (\exists \varphi \in \Gamma_j B\varphi(\psi) \leftrightarrow \exists \varphi \in \Gamma_j B\varphi(\theta)) \wedge$$
$$(\exists \varphi \in \Gamma_i B\varphi(\psi) \wedge \forall \varphi \in \Gamma_i \neg B\varphi(\theta))).$$

Remark 8.27 In fact, the priority set Γ_m could be expressed by one formula

$$\bigvee_{\varphi \in \Gamma_m} B\varphi.$$

But then we would have to use B in the formulas of the priority sequence, which we prefer not to.

The axiom system **BP** that arises from these considerations combines preference and beliefs in the following manner:

(1) $\underline{Pref}(\varphi, \varphi)$.
(2) $\underline{Pref}(\varphi, \psi) \wedge \underline{Pref}(\psi, \theta) \rightarrow \underline{Pref}(\varphi, \theta)$.
(3) $\underline{Pref}(\varphi, \psi) \vee \underline{Pref}(\psi, \varphi)$.
(4) $B\underline{Pref}(\varphi, \psi) \leftrightarrow \underline{Pref}(\varphi, \psi)$.
(5) $B(\varphi \leftrightarrow \psi) \rightarrow \underline{Pref}(\varphi, \psi) \wedge \underline{Pref}(\psi, \varphi)$.

As usual, it also includes *Modus ponens (MP)*, as well as the Generalization Rule for the operator B. The first three are standard for preference, and we have seen the analogue of (4) in Section 8.2. (5) is new, as a connection between beliefs and preference. It expresses that if two propositions are indistinguishable on the plausible worlds they should be equally preferable. It is easy to see that the above axioms are valid in the models defined as follows.

Definition 8.28 (BP-model) A model of **BP** is a tuple $(S, R, \{\preceq_s\}_{s \in S}, V)$, where S is a non-empty set of worlds, R is a euclidean, transitive, and serial accessibility relation on S. Namely, it satisfies $\forall xyz((Rxy \wedge Rxz) \rightarrow Ryz)$, $\forall xyz((Rxy \wedge Ryz \rightarrow Rxz)$, and $\forall x \exists y Rxy$. Moreover, for each s, $\preceq_s$ is a quasi-linear order on propositions (subsets of S), which is constant throughout each euclidean class and which is determined by the part of the propositions that lies within the 'plausibility part' of the euclidean class. V is an evaluation function in an ordinary manner.

Theorem 8.29 *The* **BP** *system is complete w.r.t the above models.*

Proof Assume $\nvdash_{\mathbf{BP}} \theta$. Take the canonical model $\mathfrak{M} = (S, R, V)$ for the formulas using only the propositional variables of θ. To each world of S a quasi-linear order of all formulas is associated, and it only depends on the extension of the formula (the set of nodes where the formula is true) in the plausible part of the model. This order is constant throughout the euclidean class defined by R. $\neg\theta$ can be extended to a maximal consistent set Γ. We consider the submodel generated by Γ, $\mathfrak{M}' = (S', R, V)$, which naturally is an euclidean class. Since each world in S' has access to the same worlds, each world that satisfies the same atoms satisfies the same formulas. In fact, each formula φ in this model is equivalent to a purely propositional formula, a formula without B or *Pref*. To see this, one just has to realize that $B\psi$ is in the model either equivalent to $\top$ or $\bot$, and the same holds for $Pref(\psi, \theta)$. (Note that this argument only applies because we have just one euclidean class.) Now apply a p-morphism to $\mathfrak{M}'$ which identifies worlds that satisfy the same formula. This gives a finite model consisting of one euclidean class with a constant order that still falsifies θ. Moreover, each world is characterized by a formula $\pm p_1 \wedge \cdots \wedge \pm p_k$ that expresses which atoms are true in it. In consequence, each subset of the model (proposition) is also definable by a purely propositional formula, a disjunction of the formulas $\pm p_1, \wedge \cdots \wedge \pm p_k$ describing its elements. □

Similarly, we have a representation method establishing the next result:

Theorem 8.30 (representation theorem) $\nvdash_{\mathbf{BP}} \varphi$ *iff* φ *is valid in all models obtained from priority sequences.*

Proof The order of the finitely many formulas defining all the subsets of the models can be represented as a sequence

$$\Gamma_1, \ldots, \Gamma_k,$$

where Γ_1 are the best propositions ($\varphi, \psi \in \Gamma_1$ implies $\varphi \trianglelefteq \psi$ and $\psi \trianglelefteq \varphi$, Γ_i are the next best propositions, etc. Then the following is the priority sequence which results in the given order:

$$\{x \leftrightarrow \varphi \mid \varphi \in \Gamma_1\} \gg \cdots \gg \{x \leftrightarrow \varphi \mid \varphi \in \Gamma_k\}.$$

□

So far our discussions on the preference relation over propositions are rather general. We do not presuppose any restriction on such a relation. However, if we think that the preference relation over propositions is a result of lifting a preference relation over possible worlds (as discussed before), we specify its meaning in a more precise way, following the obvious option of choosing different combinations of quantifiers. For example, we can take $\forall\exists$ preference relations over the propositions, i.e., preference relations over propositions lifted from preference relations over worlds in the $\forall\exists$ manner. Regarding the axiomatization, we will then have to add the following two axioms to the above **BP** system, obtaining a new system $\mathbf{BP}^{\forall\exists}$. The latter has two more axioms:

- $B(\varphi \to \psi) \to \underline{Pref}(\psi, \varphi)$.
- $\underline{Pref}(\varphi, \varphi_1) \wedge \underline{Pref}(\varphi, \varphi_2) \to \underline{Pref}(\varphi, \varphi_1 \vee \varphi_2)$.

Theorem 8.31 *The logic* $BP^{\forall\exists}$ *is complete.*

Proof By an adaption of the proof by [94]. The difference is this: [94] uses a combination of preference and the universal modality. Instead, our system is a combination of preference and belief. This means that what is preferred in our system is decided by the plausibility structure of the model. However, this does not affect Halpern's completeness proof much, and we can still use it. □

Remark 8.32 In fact, $[\leq]\varphi$ in Chapter 3 can be defined now as $\underline{Pref}(\varphi, \top)$. Then the preference used in the system $\mathbf{BP}^{\forall\exists}$ is simply the following:

$$\underline{Pref}(\varphi, \psi) \leftrightarrow B(\psi \to \langle\leq\rangle\varphi).$$

Similarly, we get a representation-based result for this special case:

Theorem 8.33 (representation theorem) $\vdash_{\mathbf{BP}^{\forall\exists}} \varphi$ *iff* φ *is valid in all* $\forall\exists$*-models obtained from priority sequences.*

The proof is same as for the basic system.

8.4.1 Preference over Propositions and Preference over Objects

Finally, to conclude this subsection, recall that we had a logic system to discuss preference over objects when beliefs are involved. With our new system just presented, we can talk about preference over propositions. But what is the relation between these two systems? The following theorem provides an answer.

Theorem 8.34 $\vdash_{\mathbf{KD45-P}} \varphi(d_1, \ldots, d_n)$ *iff* $\vdash_{\mathbf{BP}} \varphi(p_1, \ldots, p_n)$ *where the propositional variables* $p_1, \ldots, p_n$ *do not occur in* $\varphi(d_1, \ldots, d_n)$.

Proof In order to prove this theorem, we need to prove the following lemma:

Lemma 8.35 *If* $\nvdash_{\mathbf{KD45-P}} \varphi(d_1, \ldots, d_n)$*, then for each* n *there is a model* $\mathfrak{M} \models \neg\varphi$ *with at least* n *elements.*

Proof Assume that we only have a model $\mathfrak{M} = (S, R, V)$ in which S has m elements, where $m < n$. Take one element of S, say s, and make copies of it, say, s_1, s_2,..., s_k, till we get at least n elements. If sRt, then we make s_iRt, and if tRs, then tRs_i. In this way we get a new model with at least n elements. It is bisimilar to the original model. □

Now we are ready to prove the theorem.
($\Rightarrow$) It is easy to see that all the **KD45-P** axioms and rules are valid in **BP** if one replaces each d_i by p_i.
($\Leftarrow$) It is sufficient to transform any finite **KD45-P** model $\mathfrak{M}$ with only one euclidean class into a **BP** model $\mathfrak{M}'$ with at least n possible worlds in which for each s and each ψ, $\mathfrak{M}', s \models \psi(p_1, \ldots, p_n)$ iff $\mathfrak{M}, s \models \psi(d_1, \ldots, d_n)$. Let $\mathfrak{M} = (S, R, \leq, V)$, then $\mathfrak{M}' = (S', R, \trianglelefteq, V')$, where V' is like V except that for the $p_1, \ldots, p_n$, we assign $V'(p_i) = V'(p_j)$ if $d_i \leq d_j \wedge d_j \leq d_i$, otherwise,

$V'(p_i) \neq V'(p_j)$.[5] According to Lemma 8.35, there are enough subsets to do this. Finally, we set $V'(p_i) \lhd V'(p_j)$ iff $d_i < d_j$ and extend $\lhd$ to other sets in an arbitrary manner. □

If one thinks of propositional variables as representing basic propositions, then this theorem says that reasoning about preference over objects is the same as reasoning about preference over basic propositions. This is not surprising if one thinks of basic propositions as exclusive alternatives, just like objects. Of course, the logic of preference over propositions in general is more expressive. One can look at this latter fact in two different ways: (i) the logic over preference over all propositions as essentially richer than the logic of basic propositions or objects, or (ii) the essence of the logic of propositions is contained in the basic propositions (represented by the propositional variables) and the rest needs to be carried along in the theory to obtain a good logical system–though it may be of little value by itself.[6]

By applying the method of [94] we can again adapt the above proof to obtain:

Theorem 8.36 $\vdash_{\mathbf{KD45-P}} \varphi(d_1, \ldots, d_n)$ *iff* $\vdash_{\mathbf{BP}^{\forall\exists}} \varphi(p_1, \ldots, p_n)$ *where the propositional variables* $p_1, \ldots, p_n$ *do not occur in* $\varphi(d_1, \ldots, d_n)$.

Up to now we have used decisive preference. Another option is to use deliberate preference. Let us look at this in a rather general manner. Assume that $Supe(\varphi, \psi)$ has the property in a model that for each φ, ψ,

$$\models (\varphi \leftrightarrow \varphi') \wedge (\psi \leftrightarrow \psi') \rightarrow (Supe(\varphi, \psi) \leftrightarrow Supe(\varphi', \psi')),$$

we then say "superior" is a *local property* in that model. We can now state the following propositions.

Theorem 8.37 *If we define* $Pref(\varphi, \psi)$ *as* $B(Supe(\varphi, \psi))$ *in any model where* $Supe(\varphi, \psi)$ *is a local partial order, then* $Pref(\varphi, \psi)$ *satisfies the principles of* **BP**, *except possibly connectedness.*

It is to be noted that

$$\varphi \rightarrow \langle \leq \rangle \psi$$

is not a local property even if $\leq$ is a subrelation of R. Nevertheless, in case $\leq$ is a subrelation of R, $B(\varphi \rightarrow \langle \leq \rangle \psi)$ does satisfy the principles of **BP** minus connectedness, and the additional $\mathbf{BP}^{\forall\exists}$ axioms, as we commented in Remark 8.32. For this purpose the following weakening of locality is sufficient:

$$\models (\varphi \leftrightarrow \varphi') \wedge B(\varphi \leftrightarrow \varphi') \wedge (\psi \leftrightarrow \psi') \wedge B(\psi \leftrightarrow \psi')$$
$$\rightarrow (Supe(\varphi, \psi) \leftrightarrow Supe(\varphi', \psi')).$$

[5] Note that the $V'(p_i)$ are only relevant for the ordering $\trianglelefteq$ because the p_i's only occur directly under the *Pref* in $\varphi(p_1, \ldots, p_n)$.

[6] In Chapter 10, we will return to the role of structured propositions in priority graphs, showing how their "internal algebra" can be relevant to preference reasoning after all.

8.5 Conclusion

In this chapter, we have studied preference and priorities together with beliefs, as such entanglements occur naturally in real life scenarios. We constructed a new doxastic preference logic, which extended the standard logic of belief. We proved completeness and representation theorems for it, both in single-agent and multi-agent versions. This led us to consider interesting connections between preference and beliefs. Again, we strengthened the usual completeness results for logics of this kind to representation theorems. In the multi-agent case, these representation theorems were applied to cooperative and competitive agents. Finally, we proposed a new system combining beliefs and preference over propositions. To conclude this chapter, we studied the relationship between preference over objects and preference over propositions. We showed that if we think of propositional variables as representing basic propositions, then reasoning about preference over objects is the same as reasoning about preference between basic propositions.

So far, what we have explored in this part are static properties or aspects of priority-based preference, both pure and belief-entangled. In the next chapter, we will look at our earlier main concern of the dynamics of *changing preferences*, which turns out to go well with our richer modeling of the reasons underlying preference.