Informatik

Stefan Gruner
Bruce Watson
{editors}

Formal Aspects of Computing

Essays dedicated to Derrick Kourie
on the occasion of his 65th Birthday

SHAKER
VERLAG

Chapter 2

Ceteris Paribus Preferences: Prediction via Abduction

Sergei Obiedkov

Department of Applied Mathematics and Information Science
National Research University Higher School of Economics
Russia
E-mail: sergei.obj@gmail.com

Abstract. This chapter presents an approach to preference learning via ceteris paribus preferences (i.e., preferences that hold 'other things being equal') over attribute subsets. I provide semantics for such preferences based on formal concept analysis and show that ceteris paribus preferences valid in a dataset correspond to implications valid in a certain formal context built from this dataset. Preferences computed from a training dataset can then be used to extend the preference relation to new objects based on the attributes they have, but this approach may require exponential time. However, to compute preferences over new objects induced by the preference theory behind the training dataset, it is not necessary to compute this theory explicitly: an abduction algorithm for Horn formulae represented by their characteristic models can be modified to obtain an algorithm for inducing preferences between a pair of new objects that runs in time polynomial in the size of the training dataset.

2.1 Introduction and Motivation

The cross-table on the left-hand side of Figure 2.1 describes five cars, $c_1, \ldots, c_5$, each of which is either a minivan or an SUV, is either red or white from the outside and bright or dark from the inside. This example is adapted from [34] and is also used in [199]. For instance, c_3 corresponds to a white minivan with bright interior. The diagram on the right-hand side shows a subject's preferences over these cars: c_5 is the best, c_4 is the worst, c_1 is better than both c_2 and c_3, which are incomparable. Based on this information, can we guess what red

	minivan	SUV	red exterior	white exterior	bright interior	dark interior
c_1	×			×		×
c_2		×		×		×
c_3	×			×	×	
c_4		×	×			×
c_5	×		×			×

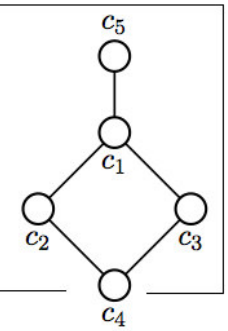

Figure 2.1: *Example of a Preference Context.*

car with bright interior the subject would rather buy: a minivan or an SUV? More generally, given a set of alternatives described by sets of attributes and a preference relation over these alternatives, can we derive preferences over new alternatives?

The approach to solving this problem presented here consists of two steps. We start by generalizing observed preferences over individual alternatives to preferences over their descriptions: from data, we derive statements such as *"I prefer white cars to red cars"*. These derived statements can be used at the second step to predict preferences over new objects. How exactly this is done depends on the semantics of the statements derived at the first step. One way to interpret the statement above is that every white car is preferred to every red car. This corresponds to the totalitarian semantics, which may not be the most appropriate one in this case: indeed, such interpretation forces the subject to disregard all other properties cars might have and even to prefer a hopelessly broken white car to a red car in a perfect state.

In this chapter we subscribe to a variant of *ceteris paribus* semantics which dictates that a subject preferring white cars to red cars must prefer every white car only to every red car that is otherwise similar. Such preferences have been considered, for example, in modal preference logics [237]. There, preferences over formulas ϕ and ψ are determined based on a preference relation defined on possible worlds (nodes of a Kripke frame): ψ is preferred to ϕ if worlds satisfying ψ are preferred to worlds satisfying ϕ. Ceteris paribus semantics puts restrictions on which worlds should be compared. In [237], preferences are parameterized by formulae sets: ψ is preferred to ϕ with respect to a formulae set Γ if every world satisfying ψ is preferred to every world satisfying ϕ that satisfies the same formulae from Γ. The Γ parameter makes it possible to be explicit about which other things must be equal rather than relying on the usual 'all other things being equal' interpretation of the ceteris paribus condition. Because of this, statements like *"I prefer a white car to a red car provided that the speed and the price are the same"* can easily be expressed.

We follow this approach, but make a radical simplification: we consider only non-strict preferences of the form $\phi \preceq_\Gamma \psi$, where ϕ and ψ are atomic conjunctions and Γ is a set of atomic formulae. This simplification allows us to treat ϕ, ψ, and Γ as attribute subsets

rather than as formulae and use formal concept analysis as a framework for preference learning. While obviously less expressive than the language of preference logics, the resulting language is comparable in its expressive power with many of the languages used in artificial intelligence for modeling preferences, including those of conditional preference networks and their extensions [32] [33] [256]. One limitation is that our language only allows for non-strict preferences (i.e., statements of the form *"ψ is at least as good as ϕ"*), whereas AI languages are mostly concerned with strict preferences (i.e., statements of the form *"ψ is better than ϕ"*). Although the approach presented here can be extended to strict preferences, they are not our focus in this chapter.

2.2 Preliminaries

We start with a few definitions from formal concept analysis [127]. A *(formal) context* is a triple $K = (G, M, I)$, where G is called a set of *objects*, M is called a set of *attributes*, and the binary relation $I \subseteq G \times M$ specifies which objects have which attributes. Formal contexts are visualized by cross-tables similar to one on the left-hand side of Figure 2.1. In this example, $G = \{c_1, \ldots, c_5\}$ and M consists of attributes corresponding to the columns of the cross-table. The *derivation operators* $(\cdot)'$ are defined for $A \subseteq G$ and $B \subseteq M$ as follows:

$$A' = \{m \in M \mid \forall g \in A(gIm)\}$$
$$B' = \{g \in G \mid \forall m \in B(gIm)\}$$

A' is the set of attributes shared by objects of A, and B' is the set of objects having all attributes of B. For example, $\{\text{SUV}\}' = \{c_2, c_4\}$ is the set of all SUVs, and $\{c_2, c_4\}' = \{\text{SUV}, \text{dark interior}\}$ is the set of all attributes that they share. For $g \in G$ and $m \in M$, the sets $\{g\}'$ and $\{m\}'$ are called *object intent* and *attribute extent* and sometimes denoted by g' and m', respectively. We will also refer to object intents as object descriptions.

The derivation operators form a *Galois connection* between the power sets of G and M. The double application of $(\cdot)'$ gives rise to two closure operators (one on objects and one on attributes): $(\cdot)''$ is extensive, idempotent, and increasing. For this reason, sets A'' and B'' are said to be *closed*. Thus, the closure of $\{\text{SUV}\}$ equals to $\{\text{SUV}, \text{dark interior}\}$, which means that we have only dark SUVs in our context. This is also captured by the notion of attribute *implication*, which is an expression $A \to B$, where $A, B \subseteq M$ are attribute subsets. It *holds* or *is valid* in a context K (notation: $K \models A \to B$) if $A' \subseteq B'$, i.e., every object of the context with all attributes from A also has all attributes from B. A set $\mathcal{L}$ of implications is *sound* for K if every implication from $\mathcal{L}$ is valid in K; it is *complete* for K if every implication valid in K holds in all contexts for which $\mathcal{L}$ is sound. Note that an implication $A \to B$ corresponds to a conjunction of definite Horn clauses with the same body.

A context is said to be *object-reduced* if none of its object intents can be represented as an intersection of other object intents. It can be shown that *reducing* the context by removing objects whose intents are intersections of other object intents does not affect the validity of attribute implications. The implications valid in a context K are summarized by the Duquenne-Guigues basis [132] which has the minimal number of implications among

implication sets sound and complete for K. For our running example, the basis consists of the following implications (here and below, we omit curly brackets and use shortened attribute names when this does not lead to confusion):

$$\begin{array}{rcl} \text{bright} & \rightarrow & \text{minivan}, \text{white} \\ \text{SUV} & \rightarrow & \text{dark} \\ \text{red} & \rightarrow & \text{dark} \\ \text{minivan}, \text{SUV}, \text{dark} & \rightarrow & M \\ \text{red}, \text{white}, \text{dark} & \rightarrow & M \\ \text{minivan}, \text{white}, \text{bright}, \text{dark} & \rightarrow & M \end{array}$$

Our semantics for preferences will be based on the notion of a preference context introduced in [198] which extends the notion of a formal context.

Definition 2.1 A *preference context* $P = (G, M, I, \leq)$ is a formal context (G, M, I) supplied with a reflexive and transitive preference relation $\leq$ on G (i.e., $\leq$ is a preorder). We write $g < h$, or $h > g$, if $g \leq h$ and $h \not\leq g$.

Figure 2.1 shows an example of a preference context.

2.3 Ceteris Paribus Preferences

Now we will be working with ceteris paribus preferences as they are defined in [197] [199]. This section recalls the definition and a few results from those papers.

Definition 2.2 A set of attributes $B \subseteq M$ is *preferred ceteris paribus* to a set of attributes $A \subseteq M$ with respect to a set of attributes $C \subseteq M$ in a preference context $P = (G, M, I, \leq)$ if $\forall g \in A' \forall h \in B'(\{g\}' \cap C = \{h\}' \cap C \rightarrow g \leq h)$. In this case, we say that the ceteris paribus preference $A \preceq_C B$ is *valid* or *holds* in P, and we denote this by $P \models A \preceq_C B$.

Thus, B is preferred ceteris paribus to A with respect to C if, between two objects that agree on all attributes from C, one with all attributes from B is always at least as good as one with all attributes from A.

Example 2.1 The preference SUV $\preceq_{\text{bright,dark}}$ minivan holds in the context from Figure 2.1, but the stronger preference SUV $\preceq_\emptyset$ minivan does not. For the latter preference to hold, it is necessary that every minivan be at a least as good as every SUV. This is not true for the minivan c_3 and the SUV c_2, which are incomparable. However, they have different interior colour: therefore their existence does not contradict the weaker preference, which can be interpreted as a preference of minivans over SUVs with the same interior colour.

Having defined the language and semantics for preferences, we say that a preference π *follows from* (or *is a semantic consequence of*) a set of preferences Π (notation: $\Pi \models \pi$) if, whenever all preferences from Π are valid in some preference context P (Π is *sound* for P; $P \models \Pi$), the preference π is also valid in P ($P \models \pi$). It is a coNP-complete problem to decide whether

	m_1	s_1	r_1	...	m_2	s_2	r_2	...	m_3	s_3	r_3	...	$\leq$
...													
c_1, c_4	×					×	×						
c_1, c_5	×				×		×		×	×			×
...													

Figure 2.2: *Part of the Ceteris Paribus Translation of the Preference Context from Figure 2.1: m_1 stands for (minivan, 1), s_2 stands for (SUV, 2), etc.*

a given preference is a semantic consequence of a given preference set [199]. A set Π of preferences is said to be *complete* for P if, for all preferences π, we have $P \models \pi$ if and only if $\Pi \models \pi$. As we will see, from a preference context P one can build a formal context K so that preferences valid in P correspond to certain valid implications of K. Because of this, computational tools developed for implications can be used to compute preferences.

Definition 2.3 The *ceteris paribus translation* of $P = (G, M, I, \leq)$ is a formal context $K^P_\sim = (G \times G, (M \times \{1,2,3\}) \cup \{\leq\}, I_\sim)$, where

$$\begin{array}{lcl} (g_1, g_2) I_\sim (m, 1) & \iff & g_1 I m, \\ (g_1, g_2) I_\sim (m, 2) & \iff & g_2 I m, \\ (g_1, g_2) I_\sim (m, 3) & \iff & \{g_1\}' \cap \{m\} = \{g_2\}' \cap \{m\}, \\ (g_1, g_2) I_\sim \leq & \iff & g_1 \leq g_2. \end{array}$$

To avoid confusion with the derivation operators of the preference context P, we denote the derivation operators of $K^P_\sim$ by $(\cdot)^\sim$ instead of $(\cdot)'$.

Example 2.2 In Figure 2.2, we show part of the formal context resulting from the ceteris paribus translation of the preference context from Figure 2.1. The object set of this new context consists of pairs of objects of the original context. Original attributes are replaced by three copies: (g_1, g_2) is associated with the first copy of an attribute m if g_1 has m in the original context; with the second copy, if g_2 has m; and with the third copy, if g_1 and g_2 agree on m, i.e., either they both have m or neither of them does. The additional preference attribute, $\leq$, is associated with (g_1, g_2) if g_2 is at least as good as g_1 according to the preference relation of the preference context.

Definition 2.4 $T_\sim(A \preceq_C B)$, the *translation of a ceteris paribus preference* $A \preceq_C B$, is the implication $(A \times \{1\}) \cup (B \times \{2\}) \cup (C \times \{3\}) \to \{\leq\}$ of the formal context $K^P_\sim$.

Example 2.3 The preference SUV $\preceq_{\text{bright,dark}}$ minivan is translated into the implication $\{(\text{SUV}, 1), (\text{minivan}, 2), (\text{bright}, 3), (\text{dark}, 3)\} \to \{\leq\}$. The preference is valid in the preference context from Figure 2.1 and the implication is valid in the translated context. As it turns out, the validity of ceteris paribus preferences is always preserved under this translation.

Proposition 2.1 $A \preceq_C B$ is valid in a preference context $P = (G, M, I, \leq)$ if and only if its translation is valid in $K^P_\sim$, i.e.: $P \models A \preceq_C B \quad \Longleftrightarrow \quad K^P_\sim \models T_\sim(A \preceq_C B)$.

Proof:

Case '$\Rightarrow$': Suppose that $P \models A \preceq_C B$ and $(A \times \{1\}) \cup (B \times \{2\}) \cup (C \times \{3\}) \subseteq (g_1, g_2)^\sim$ for some $g_1 \in G$ and $g_2 \in G$. Then, $A \subseteq \{g_1\}'$, $B \subseteq \{g_2\}'$, and $\forall c \in C$: $g_1 I c$ if and only if $g_2 I c$. The latter means that $\{g_1\}' \cap C = \{g_2\}' \cap C$. Since $A \preceq_C B$ holds in P, we have $g_1 \leq g_2$ and $(g_1, g_2) I_\sim \leq$ as required. $\surd$

Case '$\Leftarrow$': Conversely, assume $K^P_\sim \models (A \times \{1\}) \cup (B \times \{2\}) \cup (C \times \{3\}) \rightarrow \{\leq\}$. We need to show that $g_1 \leq g_2$ whenever $A \subseteq \{g_1\}'$, $B \subseteq \{g_2\}'$, and $\{g_1\}' \cap C = \{g_2\}' \cap C$. Indeed, in this case, we have $(A \times \{1\}) \cup (B \times \{2\}) \cup (C \times \{3\}) \subseteq \{(g_1, g_2)\}^\sim$ and consequently $(g_1, g_2) I_\sim \leq$, i.e.: $g_1 \leq g_2$. $\surd$

2.4 Preference Prediction

We now come back to the question posed in the introductory section: You are buying a red car with bright interior. Considering the preferences specified in Figure 2.1, will it be a minivan or an SUV?

More generally, the problem is formulated as follows: given a preference context P and two additional objects g and h with descriptions A and B, predict which of the two is better. The purpose of this section is to propose a possible solution to this problem based on ceteris paribus preferences introduced in the previous section.

One possible approach is to find a preference valid in P that forces a particular order for A and B and, consequently, for g and h. If $P \models D \preceq_F E$ with $D \subseteq A$, $E \subseteq B$, and F having no attributes from $A \setminus B$ and from $B \setminus A$ (i.e., $F \cap A = F \cap B$), then we predict $g \leq h$, for otherwise the preference $D \preceq_F E$ would not hold in the preference context P extended with g and h. Similarly, if a preference $E \preceq_F D$ with D, E, and F as above is valid in P, we conclude $h \leq g$. It is, of course, possible to obtain $g \leq h$ and $h \leq g$, in which case we have to postulate indifference between g and h.

Example 2.4 Let P be the preference context from Figure 2.1 and consider determining preferences between cars c_6 and c_7 with descriptions $\{\text{minivan}, \text{red}, \text{bright}\}$ and $\{\text{SUV}, \text{red}, \text{bright}\}$, respectively. As discussed in Example 2.1,

$$P \models \text{SUV} \preceq_{\text{bright,dark}} \text{minivan}. \tag{2.1}$$

From this we conclude $c_7 \leq c_6$, and, indeed, in the example from [34], of which our preference context is only a part, c_6 is preferred to c_7. The problem, however, is that we also have

$$P \models \text{minivan} \preceq_\emptyset \text{SUV}, \text{bright} \tag{2.2}$$

for the trivial reason that there are no SUVs with bright interior in our preference context and, thus, there is no counter-example for this preference. This forces us to state $c_6 \leq c_7$.

However, there is an important difference between preferences (2.1) and (2.2). The first preference is supported by data: the preference context contains four pairs of SUVs and minivans with the same interior colour ($\{c_2, c_4\} \times \{c_1, c_5\}$), and, for each such pair, the minivan is better than the SUV. Although the preference context contains no evidence against the second preference, it does not provide any evidence in favor of the second preference, either: there are no objects in the preference context for which a preference of one over the other could be explained with preference (2.2).

The example of above was an instance of a more general problem: whenever we need to compare two objects with descriptions A and B such that at least one of A and B does not occur in the object intent of any object from the preference context, we will always have both $P \models A \preceq_\emptyset B$ and $P \models B \preceq_\emptyset A$. This motivates the following definition:

Definition 2.5 A preference $A \preceq_C B$ is *supported* by $P = (G, M, I, \leq)$ if $P \models A \preceq_C B$ and $\exists g \in A' \exists h \in B' (g \neq h$ and $g' \cap C = h' \cap C)$.

It seems more appropriate to use only preferences supported by the preference context to predict preferences over objects outside this context: a preference over attribute sets can be used for predicting a preference between two objects only if it explains an observed preference between some objects of the preference context. This would allow us to ignore the preference

$$\text{minivan} \preceq_\emptyset \text{SUV}, \text{bright}$$

for, although it is valid in the preference context, it is not supported by it. Note that this preference is translated into the implication

$$\{(\text{minivan}, 1), (\text{SUV}, 2), (\text{bright}, 2)\} \to \{\leq\}$$

of the context $K^P_\sim$ with a zero support (i.e., $\{(\text{minivan}, 1), (\text{SUV}, 2), (\text{bright}, 2)\}' = \emptyset$), to use the terminology from association rule mining [2]. Preferences supported by P always get translated into implications $X \to \{\leq\}$ of $K^P_\sim$ with $X' \neq \emptyset$. To sum up, we will use the following rule for predicting preferences over new objects:

Definition 2.6 Let $P = \{G, M, I, \leq\}$ be a preference context and $g, h \notin G$ be two objects with descriptions $A \subseteq M$ and $B \subseteq M$ respectively. We say that h is *hypothetically preferred to g with respect to P* if P supports $D \preceq_F E$ such that $D \subseteq A$, $E \subseteq B$, and $F \cap A = F \cap B$.

To determine whether h is hypothetically preferred to g, we need to find a preference $D \preceq_F E$ described in Definition 2.6 or make sure that no such preference is supported by P. One way to achieve this is to generate a semantically complete set of preferences valid in P, ignore those that are not supported by P and see if $A \preceq_{M \setminus (A \triangle B)} B$ follows from the rest (where $A \triangle B$ is the symmetric difference between A and B). Such preference set is described in [197] based on the translation of preferences into implications of $K^P_\sim$. Unfortunately, computing preferences as implications is essentially equivalent to enumerating Horn prime implicates (with a fixed positive literal) of a Boolean formula specified by the set of its satisfying

assignments, and no output-polynomial algorithm is known for this [161]. Furthermore, this preference set can itself be exponential in the size of the preference context. A more efficient method is thus desirable.

In [159] a polynomial-time algorithm is proposed for abduction using Horn theories represented by their characteristic models. A model of a Horn theory is characteristic if it cannot be obtained as the intersection of other models. The representation of a Horn theory by its set of characteristic models is essentially the same as the representation of an implication set by an object-reduced context for which this implication set is sound and complete. The abduction algorithm works fine also on a set of models that contains some (or all) non-characteristic models. In terms of formal concept analysis, the goal of the algorithm is, given a context $K = (G, M, I)$, a set $A \subseteq M$, and an attribute $m \in A'' \setminus A$, to find, if it exists, a minimal (with respect to subset inclusion) *explanation* $E \subseteq A$ such that $K \models E \rightarrow \{m\}$ and $E' \neq \emptyset$. This suggests a strategy for determining if h is hypothetically preferred to g as described in Definition 2.6: assume that h is indeed preferred to g and find an explanation for this in the translated context $K^P_\sim$. If $g \leq h$ were part of the preference context P, the context $K^P_\sim$ would contain an object (g, h) with

$$(g, h)^\sim = (A \times \{1\}) \cup (B \times \{2\}) \cup ((M \setminus (A \triangle B)) \times \{3\}) \cup \{\leq\},$$

where A and B are the attributes of g and h, respectively. This attribute subset is the description of the fact that $g \leq h$. A proof that h is hypothetically preferred to g can be achieved through finding a minimal explanation for $\leq$ among the other attributes of $(g, h)^\sim$. In fact, it is not even necessary to find a minimal explanation; it is enough to make sure that an explanation exists. Note though that an explanation E must be such that $E^\sim$ does not consist only of pairs of the form $(g, g), g \in G$, since Definition 2.5 requires a preference to be supported by two different objects.

This is the rationale behind Procedure 2.1 which is an adaptation of the abduction algorithm from [159] for our problem. We describe the algorithm without resorting to the translation $K^P_\sim$: the algorithm works directly on P. In addition to P, the algorithm receives attribute sets A and B, which should be thought of as descriptions of two objects that must be preferentially ordered with respect to each other. The algorithm returns true if the object described by B is hypothetically preferred to the object described by A and false otherwise.

Procedure 2.1 predict_preference(A, B, P)
Input: Object intents $A, B \subseteq M$, and a preference context $P = (G, M, I, \leq)$.
Output: true if P supports $D \preceq_F E$ for some $D \subseteq A$, $E \subseteq B$, $F \subseteq M$, such that $(F \cap A) = (F \cap B)$, otherwise false.

```
For all g ∈ G do
    D := A ∩ g'
    For all h ∈ G\{g} with g ≤ h do
        E := B ∩ h'
        F := (M\(A△B)) ∩ (M\(g'△h'))
        if P ⊨ D ⪯_F E then return true
return false
```

To find a required preference $D \preceq_F E$ supported by P, Procedure 2.1 iterates through all pairs of different objects (g, h) of P such that $g \leq h$. For each such pair, it forms the weakest preference that would explain both $g \leq h$ and the assumed preference of the object described by B over the object described by A. The left-hand side of this preference must 'match' both g' and A; therefore, the algorithm sets $D := A \cap g'$. Similarly, the right-hand side is $E := B \cap h'$. The attributes in the ceteris paribus condition must 'match' the similarity between g and h, as well as the similarity between A and B. Thus, F is set to contain all attributes on which there is no disagreement between g' and h' and between A and B: each attribute from F belongs to both g' and h' or to none of them and to both A and B or to none of them; (recall that $A \triangle B$ is the symmetric difference between A and B). It is easy to see that, if the resulting preference $D \preceq_F E$ is valid in P, it satisfies the description in Definition 2.6 making it possible to conclude that the object described by B is hypothetically preferred to the object described by A. On the other hand, every preference supported by P must 'match' a pair of objects $g \leq h$ in the sense of Definition 2.5. If, after iterating through all pairs $g \leq h$, we have not built a preference $D \preceq_F E$ explaining why B should be preferred to A, then no such preference is supported by P, and the algorithm answers negatively.

Example 2.5 Consider again cars c_6 and c_7 from Example 2.4. We show how Procedure 2.1, finds a preference explaining $c_7 \leq c_6$ on preference context from Figure 2.1. Here, $A = \{\text{SUV}, \text{red}, \text{bright}\}$ and $B = \{\text{minivan}, \text{red}, \text{bright}\}$. The algorithm starts by assigning $g = c_1$ and $h = c_5$, where $c_1' = \{\text{minivan}, \text{white}, \text{dark}\}$ and $c_5' = \{\text{minivan}, \text{red}, \text{dark}\}$, computes necessary intersections, and considers the resulting preference

$$\emptyset \preceq_{\text{bright,dark}} \text{minivan}, \text{red}$$

stating that a red minivan is always at least as good as any other car with the same interior colour. This preference holds in P; therefore the algorithm returns `true` and terminates. Now, let us try Procedure 2.1 to find a preference suggesting $c_6 \leq c_7$. In this case, $A = \{\text{minivan}, \text{red}, \text{bright}\}$, $B = \{\text{SUV}, \text{red}, \text{bright}\}$, and the first pair (c_1, c_5) results in the preference

$$\text{minivan} \preceq_{\text{bright,dark}} \text{red},$$

which does not hold in P, because P contains cars $c_5 > c_4$ with the same interior colour, c_5 being a minivan and c_4 being red. The next pair to be considered (c_2, c_1) produces the preference

$$\emptyset \preceq_{\text{red,white,bright,dark}} \emptyset,$$

which does not hold due to the same counter-example $c_5 > c_4$. The pair (c_2, c_5) yields

$$\emptyset \preceq_{\text{bright,dark}} \text{red},$$

and the same counter-example works here, too. The algorithm proceeds by deriving candidate preferences from pairs $\{c_3\} \times \{c_1, c_5\}$ and $\{c_4\} \times \{c_1, c_2, c_3, c_5\}$, all of which are falsified by P. Thus, we conclude that $c_6 \not\leq c_7$ and, combining this with $c_7 \leq c_6$, decide that c_6 is strictly preferred to c_7.

It is easy to see that Procedure 2.1 is polynomial in the size of its input. To be more precise, let $n = |G|, m = |M|$, and $k = | \leq |$. Then there are $k - n$ iterations of the inner **for**-loop in total in the worst case, since the two nested loops iterate through all the pairs $g \leq h$ of P for different g and h. At each iteration, we compute E and F applying set-theoretic operations to subsets of M which subsumes the time needed to compute D. Then we test if the resulting preference $D \preceq_F E$ holds in P. For this, we have to search through all pairs $g \not\leq h$ of P trying to find one with $D \subseteq g'$, $E \subseteq h'$, and $F \cap g' = F \cap h'$. Assuming that set-theoretic operations over subsets of M take time $O(m)$, this can be done in time $O((n^2 - k)m)$, resulting in the overall time of $O(m(k - n)(n^2 - k))$.

Another way to check if a preference holds in a preference context is given by Procedure 2.2 below. This algorithm will generally be more efficient if, for each attribute $m \in M$, we precompute its extent m'.

Procedure 2.2 check_preference($D \preceq_F E, P$)
Input: a ceteris paribus preference $D \preceq_F E$ over M, and preference context $P = (G, M, I, \leq)$.
Output: true if $P \models D \preceq_F E$, otherwise false.

```
X := ⋂_{m∈D} m'
Y := ⋂_{m∈E} m'
For all g ∈ X do
    For all h ∈ Y do
        if g ≰ h and (g' ∩ F) = (h' ∩ F) then return false
return true
```

To verify that $P \models D \preceq_F E$, we then need to compute $X = D'$ and $Y = E'$ using $|D| - 1$ and $|E| - 1$ intersections, respectively. After this, we check if there are objects $g \in X$ and $h \in Y$ that agree on all attributes from F, but for which $g \not\leq h$. If we can find such objects, the preference does not hold; otherwise, it does. This optimization does not lead to a better worst-case theoretical complexity, but it may result in a considerable speed-up in practice: instead of spending $O(m)$ time on each of $O(n^2 - k)$ pairs, we would only consider pairs from $D' \times E'$. Further optimizations may include precomputing, for each $m \in M$, all pairs of objects that agree on m: $\{(g, h) \in G^2 \mid g' \cap \{m\} = h' \cap \{m\}\}$.

2.5 Related Work

Preferences have been studied in many fields as diverse as philosophy, psychology, decision theory, and economics, to give a few examples [34]. They are of fundamental interest for many applications of artificial intelligence, and several preference representation languages, as well as various approaches to learning preferences from data, have been proposed by AI researchers [103] [124]. On the other hand, a number of approaches to modeling various notions of preferences have been developed within preference logics — notably modal preference logics. The key principle here is to extend a given preference relation on individual outcomes to sets of outcomes in one of several reasonable ways and, based on that, derive

preferences between propositions about these outcomes. Paper [27] provides a discussion on the relation between preference logics and AI preference languages.

In this chapter, we focused on logical preference theories restricted to preferences between conjunctions of atomic propositions. We are interested in *ceteris paribus* preferences, i.e.: preferences that hold between propositions other things being equal. In the *classical* preference logic of von Wright [246] ceteris paribus preferences between propositions are required to hold only when *all* other things —those not mentioned in the propositions being compared— are equal. Similarly, in CP-nets, one of the best-known preference modeling AI formalisms [32], a preference of one attribute value over another is conditioned by values of all other variables being identical.

In [237] a more general logic is described where it is possible to explicitly specify which other things must be equal for a preference to hold. Similar extensions have been developed for CP-nets [256], and a unifying framework was proposed in [27] in the form of what they call a 'prototypical' preference logic, expressive enough to encode CP-nets and many of their extensions.

The language we use here can be regarded as a syntactic fragment of this prototypical logic, but our semantics follows that of the logic from [237] and, thus, is slightly different from the approach taken in [27]. In particular, we interpret preferences on arbitrary preorders, whereas [27] requires the preference relation on outcomes to be a total preorder. Another difference is that the semantics based on preference contexts (being essentially a variant of possible-world semantics of modal logics) allows for objects with identical descriptions to be treated as different entities, thus, taking into account a possibility that the language used for their description might not cover all aspects relevant to determining preferences. This is not allowed in [27], as well as in many other AI preference languages, which makes them incapable of handling cases when two identically described outcomes are preferred to different sets of outcomes — a situation that very well may happen when building a preference model from real-life data.

2.6 Conclusion and Outlook

In this chapter we presented semantics for ceteris paribus preferences based on preference contexts, which extend formal contexts as defined in formal concept analysis. Based on this, we proposed a two-step method for predicting preferences over objects described by attribute sets. The first step consists in extracting ceteris paribus preferences over attribute subsets from a preference relation defined on objects in the training set. The second step is to predict preferences over new objects based on preferences over attribute sets extracted at the previous step.

The first step is computationally hard, and it can result in an exponentially large set of preferences, thus making the second step hard, too. However, we have shown that explicit computation of preferences over attribute sets is not necessary for predicting preferences over new objects: exactly the same prediction can be achieved by a modification of an abduction algorithm developed for Horn theories represented by their characteristic models. Presented

with a couple of objects, the algorithm will compute, in polynomial time (in the size of the training dataset), a ceteris paribus preference forcing this or that preferential order between the two objects if such a preference is supported by the training dataset. This shortcut is made possible by representing ceteris paribus preferences as implications of a special formal context whose size is quadratic in the size the original dataset.

It remains to see how well the approach to preference learning presented in this chapter works in practice. It may be possible that, in application to real-life data, it will be necessary to take into account statistical considerations by using for prediction preferences that admit exceptions, but only those that are supported by a large volume of data. This is easy with the proposed approach: instead of translating preferences into implications, they could be translated into association rules satisfying certain thresholds for confidence and support [2]. Another possibility (also interesting from the scalability point of view) is to randomize the polynomial-time prediction algorithm in such a way that it guarantees certain accuracy in prediction.

Acknowledgments: The author is grateful to the editors of this volume for giving him the opportunity to be part of it. It is indeed a great pleasure and honour to be able to contribute to a book dedicated to *Derrick Kourie*, a fine academic and a wonderful person.

About the Author: *Sergei Obiedkov* is associate professor at the National Research University, Higher School of Economics in Moscow, Russia. Before taking this position he spent a year as a postdoctoral researcher at the University of Pretoria, South Africa, under the supervision of Derrick Kourie. Sergei Obiedkov holds a PhD in theoretical computer science. His research interests include artificial intelligence, formal concept analysis, knowledge representation, data analysis, logic, algorithms, and applications.