Group Decis Negot (2012) 21:99–127
DOI 10.1007/s10726-011-9274-2

# Argumentation-Based Qualitative Preference Modelling with Incomplete and Uncertain Information

**Wietske Visser · Koen V. Hindriks · Catholijn M. Jonker**

Published online: 15 November 2011


**Abstract** This paper presents an argumentation-based framework for the modelling of, and automated reasoning about multi-attribute preferences of a qualitative nature. The framework presents preferences according to the lexicographic ordering that is well-understood by humans. Preferences are derived in part from knowledge. Knowledge, however, may be incomplete or uncertain. The main contribution of the paper is that it shows how to reason about preferences when only incomplete or uncertain information is available. We propose a strategy that allows reasoning with incomplete information and discuss a number of strategies to handle uncertain information. It is shown how to extend the basic framework for modelling preferences to incorporate these strategies.

**Keywords** Qualitative multi-attribute preferences · Argumentation · Incomplete information · Uncertain information

## 1 Introduction

Our introduction of an argumentation-based framework for modelling qualitative multi-attribute preferences under incomplete or uncertain information is motivated by research into negotiation support systems. In this context, we are faced with the

W. Visser (✉) · K. V. Hindriks · C. M. Jonker
Man Machine Interaction Group, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands
e-mail: Wietske.Visser@tudelft.nl

K. V. Hindriks
e-mail: K.V.Hindriks@tudelft.nl

C. M. Jonker
e-mail: C.M.Jonker@tudelft.nl

need to express a user's preferences. A necessary (but not sufficient) condition for an offer to become an agreement is that both parties feel that it satisfies their preferences well enough. Unfortunately, eliciting and representing a user's preferences is not unproblematic. Existing negotiation support systems are based on quantitative models of preferences. These kinds of models are based on utilities; a utility function determines for each outcome a numerical value of utility. However, it is difficult to elicit such models from users, since humans generally express their preferences in a more qualitative way. We say we like something more than something else, but it seems strange to express liking something exactly twice as much as an alternative. In this respect, *qualitative* preference models will provide a better correspondence with the way preferences are expressed by humans. We also think that qualitative models will allow a human user to interact more naturally with an agent negotiating on his behalf or supporting him in his negotations, and will investigate this in future. There are, however, several challenges that need to be met before qualitative models can be usefully applied. Doyle and Thomason (1999) provide an overview including among others the challenge to deal with partial information (information-limited rationality) and, more generally, the challenge to formalize various reasoning-related tasks (knowledge representation, reasons, and preference revision).

For any real-life application it is important to be able to handle *multi-attribute* preferences. It is a natural approach to derive object preferences from general preferences over properties or attributes. For example, it is quite natural to say that you prefer one house over another because it is bigger and generally you prefer larger houses over smaller ones. This might still be so if the first house is more expensive and you generally prefer cheaper options. So there is an interplay between attributes and the preferences a user holds over them in determining object preferences. This means that object preferences can be quite complex. One approach to obtain preferences about objects is to start with a set of properties of these objects and derive preferences from a ranking of these properties that indicates the relative importance or priority of each of these properties. This approach to obtain preferences is typical in multi-attribute decision theory (Keeney and Raiffa 1993), a quantitative theory that derives object preferences from utility values assigned to outcomes which are derived from numeric weights associated with properties or attributes of objects. On the other hand, also several qualitative approaches have been proposed (Brewka 2004; Brewka et al. 2004; Coste-Marquis et al. 2004; Liu 2008).

Next, a user's preferences and knowledge about the world may be incomplete, uncertain, inconsistent and/or changing. For example, a user may lack some information regarding the objects he has to choose between, or he might have contradictory information from different sources. Preferences may change for various reasons, e.g. new information becoming available, experience, changing goals, or interaction with persuasive others. For now, we focus on the situation in which information about objects is *incomplete* or *uncertain*, but we will address other types of incompleteness, uncertainty, inconsistency and change in future.

The topic is related to decision making under uncertainty (e.g. Dubois et al. 2003; Boutilier 1994). In DMU, the aim is to find the best decision in case of uncertainty about the current state of the world, and hence about the outcomes of decisions. Our approach is more general and can be applied in different contexts; we compare the

preference between abstract 'objects', which could be states of the world (as in decision making), but also e.g. products, contracts, holiday arrangements, or houses. Also, the best option may not always be available (e.g. in negotiation, you typically have to find a compromise) so that also the preference between non-optimal solutions is important.

One of the challenges of reasoning about preferences is their multi-attribute nature. There are several distinct notions: importance of attributes, degree of satisfaction of attributes, and degree of belief of facts. In some approaches, (some of) these measures are assumed to be commensurate (e.g. Amgoud and Prade 2009 and classical utility theory Keeney and Raiffa 1993), others (including this paper) suppose *non-commensurability*. In this paper we focus on the case where it is not completely certain which attributes the objects have (there are different degrees of belief), combined with relative importance of attributes. We leave the degree of satisfaction of attributes for future work. Dubois et al. (2008) present several multi-attribute preference ordering rules, but do not take uncertainty into account. Bonet and Geffner (1996) present a qualitative model for decision making with plausibility measures of input situations, but they treat plausible and likely beliefs equally. Amgoud and Prade (2009) present an argument-based approach to multi-criteria decision making, but assume that the knowledge base is consistent, fully certain and complete.

The approach we take is based on argumentation. In recent years, argumentation has evolved to be a core study within artificial intelligence and has been applied in a range of different topics (Bench-Capon and Dunne 2007). We incorporate some of the ideas introduced in existing qualitative approaches but also go beyond these approaches by introducing a framework that is able to reason about preferences also when only incomplete information is available or when the available information is not certain. Because of its non-monotonic nature, argumentation is useful for handling inconsistent, incomplete and uncertain information. Although a lot of work has been done on argumentation-based negotiation (for a comprehensive review, see Rahwan et al. 2004), most of this work considers only the bidding phase in which offers are exchanged. For preparation, the preferences of a user have to be made clear (both to the user himself and to the agent supporting him), hence we need to express and reason with them. We focus here on the modelling of a single user's preferences by means of an argumentation process. The idea is that a user weighs his preferences, which gives him better insight into his own preferences, and so this weighing is part of the preference elicitation process. The weighing of arguments maps nicely onto argumentation. For example, 'I like to travel by car because it is faster than going by bike' is countered by 'But cycling is healthier than driving the car and that is more important to me, so I prefer to take the bike'. This possibility to construct arguments that are attacked by counterarguments is another advantage of argumentation, since it is a very natural way of reasoning for humans and fits in with a user's own reasoning processes. This is a general feature of argumentation and we will make extensive use of it: arguments like those above form the basis of our system. We believe that this way of reasoning will also be very useful in the preference elicitation process since the user's insight into his preferences grows piece by piece as he is expressing them. The introduction of an argumentation-based framework for reasoning about preferences even when only incomplete information is available seems particularly suitable for

such a step-by-step process. It allows the user to extend and refine the system representation of his preferences gradually and as the user sees fit. Another motivation to use argumentation is the link with multi-agent dialogues (Amgoud et al. 2000), which will be very interesting in our further work on negotiation.

In this paper we present an argumentation-based framework for reasoning with qualitative multi-attribute preferences. In Sect. 2, we introduce qualitative multi-attribute preferences, in particular the lexicographic preference ordering. In Sect. 3 we start by modelling this ordering for reasoning with complete and certain information in an argumentation framework. Then we proceed and extend this framework in such a way that it can also handle incomplete information. In Sect. 4, we propose a strategy (based on the lexicographic ordering) with some desired properties to derive object preferences in the case of incomplete information. In Sect. 5 this strategy is subsequently incorporated into the argumentation framework. In Sect. 6 we discuss the situation where information about objects is uncertain and introduce an epistemic argumentation framework to reason with such uncertain information. Section 7 presents concrete, qualitative preference strategies that provide different ways for handling uncertain information. Section 8 concludes the paper.

## 2 Qualitative Multi-Attribute Preferences

Qualitative multi-attribute preferences over objects are based on a set of relevant attributes or goals, which are ranked according to their importance or priority. Without loss of generality, we only consider binary (Boolean) attributes (cf. Brewka 2004). Moreover, it is assumed that the presence of an attribute is preferred over its absence. For example, given that *garden* is an attribute, a house that has a garden is preferred over one that does not have one. The importance ranking of attributes is defined by a total preorder (a total, reflexive and transitive relation), which we will denote by $\trianglerighteq$. This relation is not required to be antisymmetric, so two or more attributes can have the same importance. The relation $\trianglerighteq$ yields a stratification of the set of attributes into importance levels. Each importance level consists of attributes that are deemed equally important. Together with factual information about which objects have which attributes, the attribute ranking forms the basis on which various object preference orderings can be defined. One of the most well-known preference orderings is the lexicographic ordering, which we will use here. Brewka (2004) and Coste-Marquis et al. (2004) define more multi-attribute preference orderings, such as the discrimin and best-out orderings. In this paper we focus on the lexicographic ordering because it defines a total preference relation (contrary to the discrimin ordering) and it is more discriminating than the best-out ordering. Furthermore, the experimental research of Bonnefon and Fargier (2006) shows that among several qualitative approaches to order options based on their positive and negative aspects, cardinality-based approaches such as the lexicographic ordering best predict the actual choices made by humans. Since the other orderings are structurally similar to the lexicographic ordering, a similar argumentation framework could be defined for them if desired.

The lexicographic preference ordering first considers the highest importance level. If some object has more attributes on that level than another, the first is preferred. If

**Table 1** An example of objects and attributes

| | *large* | ≜ | *garden* | ≜ | *closeToWork* | ▷ | *nearShops* | ≜ | *quiet* | ▷ | *detached* |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *villa* | ✓ | | ✓ | | ✗ | | ✗ | | ✗ | | ✓ |
| *apartment* | ✓ | | ✗ | | ✓ | | ✓ | | ✗ | | ✗ |
| *cottage* | ✗ | | ✓ | | ✗ | | ✓ | | ✓ | | ✓ |

both objects have the same number of attributes on this level, the next importance level is considered, and so on. Two objects are equally preferred if they have the same number of attributes on every importance level. We illustrate the lexicographic preference ordering by means of an example.

*Example 1* Paul wants to buy a house. According to him, the most important attributes are *large* (e.g. minimally 100 $m^2$), *garden* and *closeToWork*, which among themselves are equally important. The next most important attributes are *nearShops* and *quiet*. Being *detached* is the least important. Paul can choose between three options: a *villa*, an *apartment* and a *cottage*. The attributes of these objects are displayed in Table 1. In this table, the attributes are ordered in decreasing importance from left to right. $\triangleq$ between attributes indicates equal importance, $\triangleright$ a transistion to a lower importance level. A ✓ indicates that an object has the attribute, a ✗ means that the attribute is absent. Which house should Paul choose? He first considers the highest importance level, which in this case comprises *large*, *garden* and *closeToWork*. The *villa* and the *apartment* both have two of these attributes, while the *cottage* only has one. So at this moment Paul concludes that both the *villa* and the *apartment* are preferred to the *cottage*. For the preference between the *villa* and the *apartment* he has to look further. At the next importance level, the *apartment* has one attribute and the *villa* has none. So the *apartment* is preferred over the *villa*. Note that although the *cottage* has the most attributes in total, it is still the least preferred option because of its bad score at the more important attributes.

**Definition 1** (*Lexicographic preference ordering*) Let $\mathcal{P}$ be a set of attributes or goals, and $\trianglerighteq$ a total preorder on $\mathcal{P}$ representing the relative importance among attributes. We write $P \triangleright Q$ for $P \trianglerighteq Q$ and $Q \not\trianglerighteq P$, and $P \triangleq Q$ for $P \trianglerighteq Q$ and $Q \trianglerighteq P$. We use $|\cdot|$ to denote the cardinality of a set. Object $a$ is *strictly preferred* over object $b$ according to the lexicographic ordering if there exists an attribute $P$ such that $|\{P' \mid a \text{ has } P' \text{ and } P \triangleq P'\}| > |\{P' \mid b \text{ has } P' \text{ and } P \triangleq P'\}|$ and for all $Q \triangleright P$: $|\{Q' \mid a \text{ has } Q' \text{ and } Q \triangleq Q'\}| = |\{Q' \mid b \text{ has } Q' \text{ and } Q \triangleq Q'\}|$. Object $a$ is *equally preferred* as object $b$ according to the lexicographic ordering if for all $P$: $|\{P' \mid a \text{ has } P' \text{ and } P \triangleq P'\}| = |\{P' \mid b \text{ has } P' \text{ and } P \triangleq P'\}|$.

## 3 Basic Argumentation Framework for Preferences

In this section we present an argumentation framework for deriving preferences according to the lexicographic ordering, based on complete and certain information.

In later sections we extend this basic framework in order to deal with incomplete and uncertain information.

### 3.1 Abstract AF and Semantics

In order to formally model and reason with preferences we define an argumentation framework (AF). We use as our starting point the well-known argumentation theory of Dung (1995). An abstract argumentation framework (Dung 1995) is a pair $\langle \mathcal{A}, \rightarrow \rangle$ where $\mathcal{A}$ is a set of arguments, and $\rightarrow$ a binary defeat relation (informally, a counter-argument relation) on $\mathcal{A}$.

To define which arguments are justified, we use Dung's (1995) preferred semantics.

**Definition 2** (*Preferred semantics*) A *preferred extension* of an AF $\langle \mathcal{A}, \rightarrow \rangle$ is a maximal (w.r.t. $\subseteq$) set $S \subseteq \mathcal{A}$ such that: $\forall A, B \in S : A \not\rightarrow B$ and $\forall A \in S$: if $B \rightarrow A$ then $\exists C \in S : C \rightarrow B$. An argument is credulously (sceptically) *justified* w.r.t. preferred semantics if it is in some (all) preferred extension(s).

Informally, a preferred extension is a coherent point of view that can be defended against all its attackers. In case of contradictory information there will be multiple preferred extensions, each advocating one point of view. The contradictory conclusions will be credulously, but not sceptically justified.

An AF is abstract in the sense that both the set of arguments and the defeat relation are assumed to be given, and the construction and internal structure of arguments is not taken into account. If we want to reason with argumentation, we have to instantiate an abstract AF by specifying the structure of arguments and the defeat relation.

### 3.2 Arguments

Arguments are built from formulas of a logical language (see Sect. 3.4), that are chained together using inference steps (see Sect. 3.5). Every inference step consists of premises and a conclusion. Inferences can be chained by using the conclusion of one inference step as a premise in the following step. Thus a tree of chained inferences is created, which we use as the formal definition of an argument (similar to e.g. Vreeswijk 1997).

**Definition 3** (*Argument*) An *argument* is a tree, where the nodes are inferences, and an inference can be connected to a parent node if its conclusion is a premise of that node. Leaf nodes only have a conclusion (a formula from the knowledge base), and no premises. A subtree of an argument is also called a *subargument*. `inf` returns the last inference of an argument (the root node), and `conc` returns the conclusion of an argument (the conclusion of its last inference).

Some example arguments will be given in Example 3 after the presentation of the specific language and inference schemes that are used to build them.

### 3.3 Defeat

This section provides the formal definition of defeat that we will use. The most common type of defeat is rebuttal. An argument rebuts another argument if its conclusion is the negation of the conclusion of the other argument. Rebuttal is always mutual. Another type of defeat is undercut. An undercutter is an argument for the inapplicability of an inference used in another argument (for the specific undercutters used in our framework, see Sect. 3.5). Undercut works only one way. Defeat is defined recursively, which means that rebuttal can attack an argument on all its premises and (intermediate) conclusions, and undercut can attack it on all its inferences.

**Definition 4** (*Defeat*) An argument $A$ *defeats* an argument $B$ if

- $\texttt{conc}(A) = \varphi$ and $\texttt{conc}(B) = \neg\varphi$ (*rebuttal*), or
- $\texttt{conc}(A) =$'$\texttt{inf}(B)$ is inapplicable' (*undercut*), or
- $A$ defeats a subargument of $B$.

### 3.4 Language

The language has to allow us to express everything we want to talk about when reasoning about preferences. To start, we need to be able to state the facts about objects: which attributes they do and do not have. We also have to express the importance ranking of attributes, so we need to be able to say that one attribute is more important than another, or that two attributes are equally important. Of course, we want to say that one object is preferred over another, and that two objects are equally preferred. Finally, we need to be able to express how many attributes of equal importance a certain object has, since the lexicographic preference ordering is based on counting these. To this end, we introduce a special predicate $has(a, [P], n)$ which expresses that object $a$ has $n$ attributes with equal importance as attribute $P$. Since we have no names for importance levels, we denote them by any attribute of that level, placed between square brackets. It is not necessary that the attribute used is among the attributes that the object has; in our example, $has(apartment, [quiet], 1)$ is true even though the *apartment* is not *quiet*. All of the things described can be expressed in the following language.

**Definition 5** (*Language*) Let $\mathcal{P}$ be a set of attribute names with typical elements $P$, $Q$, and $\mathcal{O}$ a set of object names with typical elements $a$, $b$, and let $n$ be a non-negative integer. The *input language* $\mathcal{L}^{in}$ and full *language* $\mathcal{L}$ are defined as follows.

$$\varphi \in \mathcal{L}^{in} ::= P(a) \mid \neg P(a) \mid P \rhd Q \mid P \triangleq Q$$
$$\psi \in \mathcal{L} ::= \varphi \in \mathcal{L}^{in} \mid pref(a, b) \mid eqpref(a, b) \mid has(a, [P], n)$$

Formulas of this language have the following informal meaning:

| | |
|---|---|
| $P(a)$ | object $a$ has attribute $P$ |
| $\neg P(a)$ | object $a$ does not have attribute $P$ |

| | |
|---|---|
| $P \rhd Q$ | attribute $P$ is more important than attribute $Q$ |
| $P \triangleq Q$ | attribute $P$ is equally important as attribute $Q$ |
| $pref(a, b)$ | object $a$ is strictly preferred over object $b$ |
| $eqpref(a, b)$ | object $a$ is equally preferred as object $b$ |
| $has(a, [P], n)$ | object $a$ has $n$ attributes equally important as attribute $P$ (not necessarily including $P$ itself) |

The idea is that preferences over objects are derived from facts about which objects have which attributes, and the importance order among attributes. These facts are contained in a *knowledge base*, which is a set of formulas from $\mathcal{L}^{in}$. A knowledge base is complete if, given a set of objects to compare and a set of attributes to compare them on, it contains for every object $a$ and for every attribute $P$, either $P(a)$ or $\neg P(a)$, and for all attributes $P$, $Q$, either $P \rhd Q$, $Q \rhd P$ or $P \triangleq Q$.

*Example 2* The information from Example 1 can be expressed in the form of the following knowledge base that is based on the language $\mathcal{L}^{in}$.

$large \triangleq garden \triangleq closeToWork \rhd nearShops \triangleq quiet \rhd detached$

| | | |
|---|---|---|
| $large(villa)$ | $large(apartment)$ | $\neg large(cottage)$ |
| $garden(villa)$ | $\neg garden(apartment)$ | $garden(cottage)$ |
| $\neg closeToWork(villa)$ | $closeToWork(apartment)$ | $\neg closeToWork(cottage)$ |
| $\neg nearShops(villa)$ | $nearShops(apartment)$ | $nearShops(cottage)$ |
| $\neg quiet(villa)$ | $\neg quiet(apartment)$ | $quiet(cottage)$ |
| $detached(villa)$ | $\neg detached(apartment)$ | $detached(cottage)$ |

### 3.5 Inferences

An argument is a derivation of a conclusion from a set of premises. Such a derivation is built from multiple steps called inferences. Every inference step consists of premises and a conclusion, and has a label. The inferences that can be made are defined by inference schemes. The inference schemes of our framework are listed in Table 2. The first and second inference schemes are used to count the number of attributes of equal importance as some attribute $P$ that object $a$ has. This type of inference is inspired by *accrual* (Prakken 2005), which combines multiple arguments with the same conclusion into one accrued argument for the same conclusion. Although our application is different, we use a similar mechanism. We want all attributes that are present to be counted. Otherwise we would conclude incorrect preferences (e.g. if the *large* attribute of the *apartment* were not counted, we would incorrectly derive that the *villa* were preferred over the *apartment*). Inference scheme 1, which counts 0, can always be applied since it has no premises. Inference scheme 2 can be applied on any subset of the set of attributes of some importance level that an object $a$ has. This means that it is possible to construct an argument that does not count all attributes that are present (a so-called non-maximal count). To ensure that only maximal counts are used, we provide an inference scheme to make arguments that defeat non-maximal counts (inference scheme 3). An argument of this type says that any count which is not maximal is not applicable. This type of defeat is called undercut (see below). Inference scheme 4 says

**Table 2** Inference schemes for the basic argumentation framework (complete and certain information)

| | | |
|---|---|---|
| 1 | $\dfrac{}{has(a, [P], 0)}$ | $count(a, [P], \varnothing)$ |
| 2 | $\dfrac{P_1(a), \ldots, P_n(a) \quad P_1 \triangleq \cdots \triangleq P_n}{has(a, [P_1], n)}$ | $count(a, [P_1], \{P_1, \ldots, P_n\})$ |
| 3 | $\dfrac{P_1(a), \ldots, P_n(a) \quad P_1 \triangleq \cdots \triangleq P_n \triangleq P}{count(a, [P], S \subset \{P_1, \ldots, P_n\}) \text{ is inapplicable}}$ | $count(a, [P], S)uc$ |
| 4 | $\dfrac{has(a, [P], n) \quad has(b, [P'], m) \quad P \triangleq P' \quad n > m}{pref(a, b)}$ | $prefinf(a, b, [P])$ |
| 5 | $\dfrac{has(a, [Q], n) \quad has(b, [Q'], m) \quad Q \triangleq Q' \rhd P \quad n \neq m}{prefinf(a, b, [P]) \text{ is inapplicable}}$ | $prefinf(a, b, [P])uc$ |
| 6 | $\dfrac{has(a, [P], n) \quad has(b, [P'], m) \quad P \triangleq P' \quad n = m}{eqpref(a, b)}$ | $eqprefinf(a, b, [P])$ |
| 7 | $\dfrac{has(a, [Q], n) \quad has(b, [Q'], m) \quad Q \triangleq Q' \not\triangleq P \quad n \neq m}{eqprefinf(a, b, [P]) \text{ is inapplicable}}$ | $eqprefinf(a, b, [P])uc$ |

that an object $a$ is preferred over an object $b$ if the number of attributes of a certain importance level that $a$ has is higher than the number of attributes on that same level that $b$ has. For the lexicographic ordering, it is also required that $a$ and $b$ have the same number of attributes on any level higher than that of $P$. We model this by defining an inference scheme 5 that undercuts scheme 4 if there is a more important level than that of $P$ on which $a$ and $b$ do not have the same number of attributes. Finally, inference schemes 6 and 7 do the same as 4 and 5, but for equal preference. We need these because equal preference cannot be expressed in terms of strict preference.

*Example 3* We now illustrate the inference schemes with some arguments that can be made from the knowledge base in Example 2. The example arguments are listed in Table 3 (for space reasons, the inference labels are left out). Argument $A$ illustrates the general working; a preference for the apartment over the cottage is derived, based on the fact that there is an importance level where the apartment has two attributes and the cottage only one. Argument $B$ illustrates a zero count. Here a preference for the apartment over the villa is derived, based on the fact that there is an importance level where the apartment has one attribute and the villa zero. In argument $C$ a non-maximal count is used (stating that the apartment has zero attributes of the level of *nearShops*), which leads to another conclusion, namely that the villa and the apartment are equally preferred. However, there are undercutters to attack such arguments (argument $D$).

**Table 3** Example arguments

| | |
|---|---|
| $A$: | $\dfrac{\dfrac{large(apartment) \quad closeToWork(apartment) \quad large \triangleq closeToWork}{has(apartment, [large], 2)} \quad \dfrac{garden(cottage)}{has(cottage, [garden], 1)} \quad large \triangleq garden \quad 2 > 1}{pref(apartment, cottage)}$ |
| $B$: | $\dfrac{\dfrac{nearShops(apartment)}{has(apartment, [nearShops], 1)} \quad \dfrac{}{has(villa, [nearShops], 0)} \quad nearShops \triangleq nearShops \quad 1 > 0}{pref(apartment, villa)}$ |
| $C$: | $\dfrac{\dfrac{}{has(villa, [nearShops], 0)} \quad \dfrac{}{has(apartment, [nearShops], 0)}\ * \quad nearShops \triangleq nearShops \quad 0 = 0}{eqpref(villa, apartment)}$ |
| $D$: | $\dfrac{nearShops(apartment)}{*\text{ is inapplicable}}$ |

## 3.6 Validity

The argumentation framework defined in previous sections indeed models lexicographic preference, assuming a complete and consistent knowledge base.

**Proposition 1** *Let $\mathcal{A}(KB)$ denote all arguments that can be built from a knowledge base $KB$. Then there is an argument $A \in \mathcal{A}(KB)$ such that the conclusion of $A$ is $pref(a, b)$ and $A$ is sceptically justified under preferred semantics iff $a$ is preferred over $b$ according to the lexicographic preference ordering (Definition 1) given $KB$.*

*Proof* Suppose $a$ is preferred over $b$. This means that there exists an attribute $P$ such that $|\{P' \mid a \text{ has } P' \text{ and } P \triangleq P'\}| > |\{P' \mid b \text{ has } P' \text{ and } P \triangleq P'\}|$ and for all $Q \rhd P$: $|\{Q' \mid a \text{ has } Q' \text{ and } Q \triangleq Q'\}| = |\{Q' \mid b \text{ has } Q' \text{ and } Q \triangleq Q'\}|$. Let $P_1, \ldots, P_n$ denote all attributes of equal importance as $P$ such that $a$ has $P_i$ and let $P'_1, \ldots, P'_m$ denote all attributes of equal importance as $P$ such that $b$ has $P_i$. Note that $n > m$. Then the knowledge base is as follows: $P_1 \triangleq \cdots \triangleq P_n \triangleq P'_1 \triangleq \ldots P'_m$ and $P_1(a), \ldots, P_n(a)$ and $P'_1(b), \ldots, P'_m(b)$. The following argument ($A$) can be built (note that this argument can also be built if $m$ is equal to 0, by using the empty set count):

$$\dfrac{\dfrac{P_1(a), \ldots, P_n(a) \quad P_1 \triangleq \cdots \triangleq P_n}{has(a, [P_1], n)} \quad \dfrac{\dfrac{P'_1(b), \ldots, P'_m(b) \quad P'_1 \triangleq \cdots \triangleq P'_m}{has(b, [P'_1], m)} \quad P_1 \triangleq P'_1 \quad n > m}{pref(a, b)}$$

We will now play devil's advocate and try to defeat this argument. We can try rebuttal and undercut of the argument and its subarguments. Rebuttal of premises is not applicable, since the knowledge base is consistent. Rebuttal of (intermediate) conclusions is not possible either, since there is no way to derive a negation. Then there are three inferences we can try to undercut (the last inference of the argument and the last inferences of two subarguments). For the left-hand count, this can only be done if there is another $P_j$ such that $P_j \triangleq P$ and $P_j \notin \{P_1, \ldots, P_n\}$ and $P_j(a)$ is the case. However, $P_1, \ldots, P_n$ encompass all such attributes, so count undercut is not possible. The same argument holds for the other count. At this point it is useful to note that these two counts are the only ones that are undefeated. Any lesser

count will be undercut by the count undercutter that takes all of $P_1, \ldots, P_n$ (resp. $P'_1, \ldots, P'_m$) into account. Such an undercutter has no defeaters, so any non-maximal count is not justified. The final thing that is left to try is undercut of $prefinf(a, b, [P_1])$. The undercutter of $prefinf(a, b, [P_1])$ is based on two counts. We have seen that any non-maximal count will be undercut. If the maximal counts are used, we have $n = m$, since we have for all $Q \rhd P$: $|\{Q' \mid a \text{ has } Q' \text{ and } Q \triangleq Q'\}| = |\{Q' \mid b \text{ has } Q' \text{ and } Q \triangleq Q'\}|$. So the undercutter inference rule cannot be applied since $n \neq m$ is not true. This means that for every possible type of defeat, either the defeat is inapplicable or the defeater of $A$ is itself defeated by undefeated arguments. This means that $A$ is in every preferred extension and hence sceptically justified according to preferred semantics.

Suppose $a$ is not preferred over $b$. This means that for all attributes $P$, either $|\{P' \mid a \text{ has } P' \text{ and } P \triangleq P'\}| \leq |\{P' \mid b \text{ has } P' \text{ and } P \triangleq P'\}|$ or there exists an attribute $Q \rhd P$ such that $|\{Q' \mid a \text{ has } Q' \text{ and } Q \triangleq Q'\}| \neq |\{Q' \mid b \text{ has } Q' \text{ and } Q \triangleq Q'\}|$. This means that any argument with conclusion $pref(a, b)$ (which has to be of the form above) is either undercut by $count(b, [P], S)uc$ because it uses a non-maximal count, or by $prefinf(a, b, [P])uc$ because there is a more important level where a preference can be derived. This means that any such argument will not be in any preferred extension and hence not sceptically justified under preferred semantics.

The same line of argument can be followed for $eqpref$. □

## 4 Incomplete Information

So far, we have defined an argumentation system that can reason about preferences according to the lexicographic preference ordering. Above, we have assumed that the information about the objects that are compared is complete. But, as stated in the introduction, this is not always the case. In this section we investigate how incomplete information can best be handled when reasoning about preferences.

Suppose it is not known whether an object has a specific attribute, e.g. we know that $P(a)$ but we do not know whether $P(b)$ or $\neg P(b)$. This might not be a problem. If the preference between $a$ and $b$ can be decided based on attributes that are more important than $P$, the knowledge whether $P(b)$ or $\neg P(b)$ is the case is irrelevant. But otherwise this information is necessary to decide a lexicographic preference. In that case, different approaches or strategies for drawing conclusions are possible. However, not all strategies give desired results. In the following, we will discuss some naive strategies and their shortcomings, from which we will derive some desired properties of strategies, and define and model a strategy that gives intuitive results.

### 4.1 Naive Strategies

*Optimistic, resp. Pessimistic, Strategy* This strategy always assumes that an object has, resp. does not have, the attribute that is not known. This strategy can always derive some preference between two objects, since it completes the knowledge by making particular assumptions, and can then derive a complete preference ordering over objects. But there is no guarantee that the inferences made are correct. In fact,

**Table 4** Example of intransitive preference with the disregard attribute strategy

| | $P$ | $\triangleq$ | $Q$ | $\triangleq$ | $R$ |
|---|---|---|---|---|---|
| $a$ | ✓ | | ? | | ✗ |
| $b$ | ✗ | | ✓ | | ? |
| $c$ | ? | | ✗ | | ✓ |

any inferred preference can only be correct if all the assumptions it is based on are either correct or irrelevant. Since we do not know whether assumptions are correct and the strategy does not check for relevance, the inference can only be correct by chance. For example, suppose it is not known whether the *villa* has a *garden* and whether it is *closeToWork*. The optimistic strategy would assume that it has both attributes, in which case an incorrect preference of the *villa* over the *apartment* would be derived. The pessimistic strategy on the other hand would assume the *villa* has neither of the attributes, and would derive an incorrect preference of the *cottage* over the *villa*.

Note that using the framework defined in Sect. 3 without adaptation would boil down to using a pessimistic strategy: if it is not known whether an object has a certain attribute, the attribute is (implicitly) assumed to be absent. This is due to the fact that only attributes for which it is known that an object has them are counted. Attributes that an object does not have and attributes for which this information is unavailable are treated the same way (i.e. not taken into account when counting).

*Disregard Attribute Strategy* This strategy does not take into account the attributes for which information about the objects to be compared is incomplete. It can always derive some preference between two objects, since the information regarding the remaining attributes is complete, so a complete preference ordering over objects can be derived. But the inference might not be correct, since the attributes that are disregarded might be relevant in defining a preference order. For example, suppose it is not known whether the *cottage* is *large*. In that case, the attribute *large* will not be taken into account when comparing the *cottage* to another object. This leaves only the attributes *garden* and *closeToWork* on the highest importance level, of which all attributes have exactly one. Since the *cottage* has the most attributes on the next importance level, a preference of the *cottage* over the *villa* as well as the *apartment* will be derived, even though in the original example the *cottage* was the least preferred object.

This strategy has another undesired effect. Consider the situation in Table 4. When comparing $a$ and $b$, this strategy only takes attribute $P$ into account, and concludes a preference of $a$ over $b$. Similarly, preferences of $b$ over $c$, and of $c$ over $a$ can be derived. So with this strategy, intransitive preferences can be derived, which is undesired.

*Cautious Strategy* In order to prevent the derivation of preferences that are only correct by chance, a natural alternative is to use a cautious strategy that prevents such inferences. This strategy infers nothing unless all information about the objects under comparison is available. It never makes incorrect preference inferences, but it lacks in decisiveness. Even if the unknown information is irrelevant to make an inference, nothing is inferred.

### 4.2 Desired Properties for Strategies

Given the limitations of the strategies discussed above, it is clear that we need a more balanced strategy that takes two main concerns into account, which we call decisiveness and safety.

*Decisiveness* We call a strategy *decisive* if it does not infer too little. As mentioned above, an unknown attribute might be irrelevant for deciding a preference. This is the case if the preference is already determined by more important attributes. For example, suppose that we do not know whether the *apartment* has attribute *nearShops*. Then we can still conclude that the *apartment* is preferred over the *cottage*, based on the attributes *large*, *garden*, and *closeToWork*. It is not required that a preference is derived in every case, since the missing information might be essential, but all preferences that are certain (for which no essential information is missing) should be derived. The cautious strategy is not decisive.

*Safety* We call a strategy *safe* if it does not infer too much. Suppose again that we do not know whether the *apartment* has attribute *nearShops*. Whereas this is irrelevant for deciding a preference between *apartment* and *cottage*, we do need this information for deciding the preference between the *villa* and the *apartment*. A strategy that makes assumptions about the missing information, or that disregards the attribute in question, will make unfounded inferences, and hence be unsafe. The optimistic, pessimistic and disregard attribute strategies are not safe.

### 4.3 A Decisive and Safe Strategy

We have seen above what may go wrong when a naive strategy is used to deal with incomplete information. In this section we define an alternative strategy that does satisfy the properties of decisiveness and safety identified above. A preference inference should never be based on an unfounded assumption for a strategy to be safe. But to be decisive, a strategy needs to be able to distinguish relevant from irrelevant information. Our approach is based on the following intuition. When comparing two objects under incomplete information, multiple situations are possible. That is, whenever it is not known whether an object has an attribute, there is a possibility that it does and a possibility that it does not. If a preference can be inferred in every possible situation, then apparently the missing information is not relevant, and it is safe to infer that preference. It is not necessary to check every possible situation, but it suffices to look at extreme cases. For every object, we can construct a best- and worst-case scenario, or best and worst possible situation. A possible situation is a *completion* of an object in the sense that all missing information is filled in.

**Definition 6** (*Completion*) A *completion* of an object $a$ is an extension of the knowledge base with (previously missing) facts about $a$ such that for every attribute $P$, either $P(a)$ or $\neg P(a)$ is in the extended knowledge base. So if $a$ has $n$ unspecified attributes, there are $2^n$ possible completions of $a$.

**Table 5** Examples of objects and attributes with incomplete information

| | $P$ | $\triangleright$ | $Q$ | $\triangleright$ | $R$ |
|---|---|---|---|---|---|
| $a$ | ✓ | | ✓ | | ? |
| $b$ | ? | | ✗ | | ✓ |

**a**

| | $P$ | $\triangleright$ | $Q$ |
|---|---|---|---|
| $a$ | ✓ | | ? |
| $b$ | ? | | ✓ |

**b**

| | $P$ | $\triangleq$ | $Q$ |
|---|---|---|---|
| $a$ | ✓ | | ? |
| $b$ | ✗ | | ✓ |

**c**

Since we assumed that presence of an attribute is preferred over absence, the most preferred completion assumes presence of all unknown attributes, and the least preferred completion assumes absence. If even the least preferred completion of $a$ is preferred over the most preferred completion of $b$, then $a$ must always be preferred over $b$, since $a$ could not be worse and $b$ could not be better. For example, consider the objects and attributes in Table 5a. Recall our assumption that presence of attributes is preferred over absence. So in the worst case for $a$, $a$ does not have attribute $R$. And in the best case for $b$, $b$ has attribute $P$. But even in this situation, $a$ will be preferred over $b$, based on attribute $Q$. There is no way that this situation can improve for $b$ or deteriorate for $a$, so it is safe to infer a preference for $a$ over $b$. The strategy's power to make such inferences makes it decisive.

The next example illustrates that this approach does not infer a preference when the missing information is relevant. Consider Table 5b. In the situation that is worst for $a$ and best for $b$, $b$ will be preferred because it has both attributes, while $a$ only has $P$. But in the other extreme situation, that is best for $a$ and worst for $b$, $a$ is preferred. This means that in reality, anything is possible, and it is not safe to infer a preference.

We have seen when a preference for $a$ over $b$ can be inferred, and in which case no preference can be inferred. There are, however, two more possibilities. One is the case in which a preference of the most preferred completion of $a$ over the least preferred completion of $b$ can be derived, but only equal preference between the least preferred completion of $a$ and the most preferred completion of $b$. This is illustrated in Table 5c. In this case, we would like to derive at least a weak preference of $a$ over $b$. This is important, because in many cases a weak preference is strong enough to base a decision on, even if a strict preference cannot be derived. When having to decide between $a$ and $b$, choosing $a$ cannot be wrong when $a$ is weakly preferred over $b$. Failing to derive a weak preference makes a strategy less decisive.

The last possibility is equal preference. We only want to derive an equal preference between two objects $a$ and $b$ if all possible completions of $a$ are equally preferred as all possible completions of $b$. This also means that the most and least preferred completions of $a$ and $b$ have to be equally preferred. This can only be the case if all information about $a$ and $b$ is known, for as soon as some information is missing, there will be multiple possible completions which are not equally preferred.

## 5 Argumentation Framework for Preferences with Incomplete Information

This section presents how our framework is extended to incorporate the decisive and safe strategy for incomplete information as presented in Sect. 4.3. We first present

the changes to the language and then the changes to the inference rules. The defeat definition does not have to change.

### 5.1 Language

To distinguish between the different completions of an object, we introduce a completion label. We use the object name without label to denote the object in general, that is, the object with any completion. The superscript $^+$ is used for the most preferred completion of an object, $^-$ for the least preferred completion. For example, consider object $a$ in Table 5a. The most preferred completion of $a$ has attribute $R$, and is denoted $a^+$. The least preferred completion of $a$ does not have attribute $R$, and is denoted $a^-$.

Reasoning with completions as discussed above can be viewed as a kind of assumption-based reasoning. To be able to support such reasoning, we extend the language and introduce weak negation, denoted by $\sim$, which is also used in Prakken and Sartor (1997). This is used to formalize a kind of assumption-based reasoning. A formula $\sim \varphi$ can always be assumed, but is defeated by $\varphi$ (see the next section for the details). So the statement $\sim \varphi$ should be interpreted as '$\varphi$ cannot be derived'.

Finally, we add formulas of the type $wpref(a, b)$ which express weak preference, just as $pref(a, b)$ and $eqpref(a, b)$ express strict and equal preference, respectively. We use weak preference in the sense that an object $a$ is weakly preferred over an object $b$ if any completion of $a$ is either preferred over or equally preferred as any completion of $b$, but no strict or equal preference can be derived.

This leads to the following redefinition of the language.

**Definition 7** (*Language*) Let $\mathcal{P}$ be a set of attribute names with typical elements $P$, $Q$, and $\mathcal{O}$ a set of object names with typical elements $a, b$, and let $n$ be a non-negative integer, and $x, y \in \{+, -, \{\}\}$ a label for objects (where $\{\}$ means no label). The *input language* $\mathcal{L}^{in}$ and full *language* $\mathcal{L}$ are defined as follows.

$$\varphi \in \mathcal{L}^{in} ::= P(a) \mid \neg P(a) \mid P \rhd Q \mid P \triangleq Q$$
$$\psi \in \mathcal{L} ::= \varphi \in \mathcal{L}^{in} \mid pref(a^x, b^y) \mid eqpref(a^x, b^y) \mid wpref(a^x, b^b) \mid has(a^x, [P], n) \mid \sim \psi$$

### 5.2 Inferences

The inference rules of the extended framework are listed in Table 6. Two inference rules are added that define the meaning of the weak negation $\sim$. According to inference rule 8, a formula $\sim \varphi$ can always be inferred, but such an argument will be defeated by an undercutter built with inference rule 9 if $\varphi$ is the case.

$P$ is supposed to be among the attributes of the least preferred completion of $a$ ($a^-$) only if it is known that $a$ has $P$. This is modelled by inference rule 2b in Table 6. For the most preferred completion of $a$, it is only required that it is not known that $a$ does not have $P$; if this is not known, $a^+$ will be assumed to have $P$. This is modeled by using premises of the form $\sim \neg P(a)$ instead of $P(a)$. This can be seen in inference rule 2a. Inference rules 4 through 7 remain unchanged, except that completion labels are added.

**Table 6** Inference schemes for incomplete information

| | | |
|---|---|---|
| 1 | $\dfrac{}{has(a^x, [P], 0)}\ count(a^x, [P], \varnothing)$ | |
| 2a | $\dfrac{\sim \neg P_1(a), \ldots, \sim \neg P_n(a) \quad P_1 \triangleq \cdots \triangleq P_n}{has(a^+, [P_1], n)}\ count(a^+, [P_1], \{P_1, \ldots, P_n\})$ | |
| 2b | $\dfrac{P_1(a), \ldots, P_n(a) \quad P_1 \triangleq \cdots \triangleq P_n}{has(a^-, [P_1], n)}\ count(a^-, [P_1], \{P_1, \ldots, P_n\})$ | |
| 3a | $\dfrac{\sim \neg P_1(a), \ldots, \sim \neg P_n(a) \quad P_1 \triangleq \cdots \triangleq P_n}{count(a^+, [P_1], S \subset \{P_1, \ldots, P_n\}) \text{ is inapplicable}}\ count(a^+, [P_1], S)uc$ | |
| 3b | $\dfrac{P_1(a), \ldots, P_n(a) \quad P_1 \triangleq \cdots \triangleq P_n}{count(a^-, [P_1], S \subset \{P_1, \ldots, P_n\}) \text{ is inapplicable}}\ count(a^-, [P_1], S)uc$ | |
| 4 | $\dfrac{has(a^x, [P], n) \quad has(b^y, [P'], m) \quad P \triangleq P' \quad n > m}{pref(a^x, b^y)}\ prefinf(a^x, b^y, [P])$ | |
| 5 | $\dfrac{has(a^x, [Q], n) \quad has(b^y, [Q'], m) \quad Q \triangleq Q' \rhd P \quad n \neq m}{prefinf(a^x, b^y, [P]) \text{ is inapplicable}}\ prefinf(a^x, b^y, [P])uc$ | |
| 6 | $\dfrac{has(a^x, [P], n) \quad has(b^y, [P'], m) \quad P \triangleq P' \quad n = m}{eqpref(a^x, b^y)}\ eqprefinf(a^x, b^y, [P])$ | |
| 7 | $\dfrac{has(a^x, [Q], n) \quad has(b^y, [Q'], m) \quad Q \triangleq Q' \not\triangleq P \quad n \neq m}{eqprefinf(a^x, b^y, [P]) \text{ is inapplicable}}\ eqprefinf(a^x, b^y, [P])uc$ | |
| 8 | $\dfrac{}{\sim \varphi}\ asm(\sim \varphi)$ | 9 $\dfrac{\varphi}{asm(\sim \varphi) \text{ is inapplicable}}\ asm(\sim \varphi)uc$ |
| 10 | $\dfrac{pref(a^-, b^+)}{pref(a, b)}$ | 11 $\dfrac{eqpref(a^-, b^+) \quad pref(a^+, b^-)}{wpref(a, b)}$ |
| 12 | $\dfrac{eqpref(a^+, b^-) \quad eqpref(a^-, b^+)}{eqpref(a, b)}$ | |

To infer overall preferences from the preferences over certain completions, three more inference rules are defined. Inference rule 10 states that if (even) $a^-$ is preferred over $b^+$, then $a$ must be preferred over $b$, as we saw above. When $a^+$ is preferred over $b^-$, but $a^-$ is only equally preferred as $b^+$, this is not strong enough to infer a strict preference of $a$ over $b$, but we can infer a weak preference of $a$ over $b$ using inference rule 11. Rule 12 states that in order to infer equal preference between $a$ and $b$, both the most preferred completion of $a$ and the least preferred completion of $b$, and the least preferred completion of $a$ and the most preferred completion of $b$ must be equally preferred.

*Example 4* In the case of Table 5a, argument $A$ in Table 7 can be built. Argument $B$ shows that a weak preference can be inferred in the situation of Table 5c.

**Table 7** Example arguments

$$A: \dfrac{\dfrac{Q(a)}{has(a^-,[Q],1)} \quad \dfrac{}{has(b^+,[Q],0)} \quad Q \triangleq Q \quad 1>0}{\dfrac{pref(a^-,b^+)}{pref(a,b)}}$$

$$B: \dfrac{\dfrac{\dfrac{P(a)}{has(a^-,[P],1)} \quad \dfrac{\sim\overline{\neg Q(b)}}{has(b^+,[Q],1)} \quad P \triangleq Q \quad 1=1}{eqpref(a^-,b^+)} \quad \dfrac{\dfrac{\sim\overline{\neg P(a)} \quad \sim\overline{\neg Q(a)} \quad P \triangleq Q}{has(a^+,[P],2)} \quad \dfrac{Q(b)}{has(b^-,[Q],1)} \quad P \triangleq Q \quad 2>1}{pref(a^+,b^-)}}{wpref(a,b)}$$

## 6 Uncertain Information

In the last two sections we focused on the situation where some information regarding the presence or absence of attributes for a given outcome is lacking. With the proposed safe and decisive strategy however, it may still be the case that no preference can be inferred. What should we do in such a case? One approach is to ask the user for the missing information. But the user might not have this information, and might not have the time or resources to look it up. Still, in many situations there is other information available on the basis of which the missing facts can be derived. For example, if the destination country of a certain holiday is not given, but it is specified that the trip will be to Rome, we can infer that the country will be Italy. In this case, the derived fact is completely certain since there is no doubt that Rome is in Italy. In other cases, the derived information may be less than fully certain, e.g. because the applied rule only holds by default and there are some exceptions to it, or because the used facts or applied rule are not certain themselves, e.g. because the source of the information is unreliable. For example, for a holiday to Rome in July we can infer that it will be sunny because that is usually the case. But this conclusion is not completely certain because it may be an exceptionally rainy July in Rome this year. Or we may conclude that the hotel we will stay in will be clean based on the reviews we have read online, but again this conclusion is not certain because we cannot trust the source completely (the hotel itself may have posted fake reviews).

The main point we want to make here is that even if missing information can be derived, the acquired facts may have different certainty levels. Such 'degrees of belief' play an important role in the deriviation of preferences. Consider a simple case in which preference is determined by a single attribute $P$ (like before, we assume that presence of an attribute is preferred over absence). If both for outcome $a$ and for outcome $b$ it can be derived that $P$ is present, but for $a$ this conclusion is more certain than for $b$, it would be rational to prefer $a$. This is because there is a bigger chance that the information about $b$ is incorrect. The situation gets more complicated in case of multiple attributes. For example, are two certainly true attributes and one certainly false attribute better or worse than three attributes whose truth is not completely certain? In Sect. 7 we present different qualitative strategies to derive preferences in case of more or less certain information. But first we will formalise the concept of certainty levels.

**Table 8** Inference schemes for epistemic formulas

| | | |
|---|---|---|
| 1 | $\dfrac{L_1, \ldots, L_m, \sim L_p, \ldots, \sim L_q \Rightarrow L : l \quad L_1 : l_1, \ldots, L_m : l_m \quad \sim L_p, \ldots, \sim L_q}{L : \min(l, l_1, \ldots, l_m)}$ | $DMP$ |
| 2 | $\dfrac{}{\sim L}$ | $asm(\sim L)$ |
| 3 | $\dfrac{L}{asm(\sim L) \text{ is inapplicable}}$ | $asm(\sim L)uc$ |
| 4 | $\dfrac{\sim L \quad \sim \neg L}{L : 0}$ | |
| 5 | $\dfrac{L : l \quad \neg L : l}{L : 0}$ | |
| 6 | $\dfrac{\neg L : l}{L : -l}$ | |

## 6.1 Certainty Levels

Different approaches to model degrees of belief can be found in the literature. Among the best-known ones are subjective probability, Dempster-Shafer belief functions, and possibility theory (see Huber 2009 for an overview and references). In this paper we take a qualitative approach in which the knowledge base is stratified according to the certainty of the formulas (see also Amgoud et al. 2005, who use similar certainty levels but apply them to decision making in a different way). Each stratum in the knowledge base corresponds to one level of certainty. Note that in the literature, the notion of certainty levels (or similar notions) is sometimes referred to as preference or priority between formulas. In this paper, we use the term preference only to refer to preference between objects or outcomes, and priority to refer to the relative importance of attributes.

Essentially, a certainty level is the qualitative counterpart of the (subjective) probability of an attribute being true. In the case of the highest certainty level, denoted $N$, this probability is 1; certainly true information is always true. Similarly, the probability is 0 for the negation of formulas with certainty level $N$. Also, as is common in the literature on subjective probability, we assume that the subjective probability of truth of a completely uncertain formula (with certainty level 0) is 0.5 (the principle of indifference, Huber 2009). For intermediate certainty levels, an exact probability can not always be given due to the qualitative nature of certainty levels, but intuitively the probability is higher for higher certainty levels.

A knowledge base is $\mathcal{K} = \mathcal{K}_1 \cup \cdots \cup \mathcal{K}_N$ where $\mathcal{K}_1$ contains the knowledge with the lowest certainty, and knowledge in $\mathcal{K}_N$ is fully certain. We assume that $\mathcal{K}_N$ is consistent, but $\mathcal{K}$ may not be. Note that there is no subset $\mathcal{K}_0$; a formula with certainty level 0 is completely uncertain and does not belong in a knowledge base. Every epistemic formula $\varphi$ in the knowledge base has an associated certainty level $l$, denoted $\varphi : l$ (i.e. $\varphi \in \mathcal{K}_l$). We will use the notation $\varphi : -l$ to denote that the negation of $\varphi$ has certainty level $l$ (see inference scheme 6 in Table 8). This notation is convenient as it provides a uniform way of expressing the certainty that some attribute is present or absent in an outcome. In this paper we assume that non-epistemic information about the relative importance of desired attributes is fully certain; we leave the situations where this is not the case for future work.

Note that the concept of certainty levels is very general and can also be applied in the cases discussed in the previous sections. If the information is complete and

fully certain, there are two certainty levels (true and false): $N = 1$ and level 0 does not occur. The situation with incomplete information is the same except that some information may have certainty level 0 (complete uncertainty). In the generic situation, we have the same three levels, 0 for complete uncertainty and the scale ends $N$ and $-N$ for complete certainty, plus any number of certainty levels in between.

The argumentation framework for deriving preferences based on uncertain information can be considered as consisting of two separate parts: an epistemic part for reasoning about the (uncertain) attributes of outcomes and a preferential part, which contains several approaches to derive preferences from uncertain information. In this paper the main focus is on the preferential part, which will be discussed in Sect. 7. For this part, only the 'output' (justified conclusions) of the epistemic part matters, i.e. which outcomes have which attributes with what certainty. The way in which this information is derived is not important for the preferential part of the framework. However, the preferential part does assume that for every attribute $P$ and every outcome $a$, there is a single certainty associated with $P(a)$. This means that the epistemic part must resolve conflicts and incompleteness. There are several ways in which to do this in a reasonable way. However, in the remainder of the current section we will give only one possible specification of the epistemic part, and will not discuss all possible alternatives as it is outside the scope of this paper.

### 6.2 Epistemic Argumentation Framework

In order to derive new facts form other facts in the knowledge base, we introduce a new kind of formula to the input language: rules. A rule is of the form $L_1, \ldots, L_k, \sim L_l, \ldots, \sim L_m \Rightarrow L_n$ where $L_i = P(a)$ or $\neg P(a)$. Its informal reading is: if all of $L_1, \ldots, L_k$ hold, then typically $L_n$ holds, except if one of $L_l, \ldots, L_m$ holds. The same kind of rules were used by Prakken and Sartor (1997). If there are no exceptions, and the rule is fully certain, then it is called a strict rule. Otherwise it is defeasible. Defeasible rules describe what is 'normally' the case. Using this kind of rules can add some information to an incomplete knowledge base. This can be beneficial in situations where a user does not have certain information, and does not have the time or resources to verify information.[1]

**Definition 8** (*Epistemic language*) As before, we distinguish a subset of the full language called the input language. A knowledge base can only contain formulas of the input language; other formulas have to be derived by inference. The epistemic input language $\mathcal{L}_e^{in}$ is defined as follows (where $L$ is a literal ($P(a)$ or $\neg P(a)$, where $P$ is an attribute and $a$ an outcome) and $l$ is a certainty level such that $0 < l \leq N$).

$$\varphi \in \mathcal{L}_e^{in} ::= L : l \mid L, \ldots, L, \sim L, \ldots, \sim L \Rightarrow L : l$$

[1] For rules to be fully applicable, it would also be required to extend the language with literals that refer to other knowledge than just which outcomes have which attributes, and to specify explicitly which attributes are desired and influence preference. A full discussion of this issue is outside the scope of this paper.

The full epistemic language $\mathcal{L}_e$ is defined as follows (where $L$ is a literal ($P(a)$ or $\neg P(a)$) and $l$ is a certainty level such that $-N \leq l \leq 0$).

$$\varphi \in \mathcal{L}_e ::= \varphi \in \mathcal{L}_e^{in} \mid \sim L \mid L : l$$

To apply a rule, inference scheme 1 in Table 8, called defeasible modus ponens, is introduced. The level of certainty of the conclusion is the same as the level of the least certain premise, this is called the weakest link principle (see e.g. Pollock 2001 for a motivation). The inferences 2 and 3 for weak negation are the same as before.

Now, for any atom $\varphi$, we can have any of the following situations.

- Either $\varphi$ or $\neg\varphi$ (but not both) is in $\mathcal{K}$ or can be derived, with only one level of certainty.
- Neither $\varphi$ nor $\neg\varphi$ is in $\mathcal{K}$ or can be derived.
- The formula $\varphi$ occurs in $\mathcal{K}$ or can be derived multiple times with different levels of certainty.
- Both $\varphi$ and $\neg\varphi$ are in $\mathcal{K}$ or can be derived.

The first situation is the most straightforward case, and this information can be used directly by the preferential part of the framework. The incompleteness in the second situation is a case of complete uncertainty with respect to $\varphi$. We use inference scheme 4 in Table 8 to derive a certainty level 0 for an atom $\varphi$ if neither $\varphi$ nor $\neg\varphi$ can be derived. The other two situations are more complicated, and there are multiple possibilities for handling these cases. For example, if there are multiple arguments concluding $\varphi$ and/or $\neg\varphi$, one could aggregate these arguments such that arguments with the same conclusion strengthen each other, but such a conclusion is weakened by counter-arguments. This approach is not trivial; see Prakken (2005) for a discussion of the issues concerning accrual of arguments. For the sake of simplicity, the approach we take here is to make sure that only the conclusion with the highest certainty level will be justified. To this end, the definition of rebuttal would have to be adapted such that an argument only rebuts another argument if their conclusions are each other's negation and the first argument has a higher certainty level than the second. Also, a more certain argument for $\varphi$ would have to defeat a less certain argument for $\varphi$. One remaining issue is what to do with arguments for $\varphi$ and $\neg\varphi$ with equal certainty. If such arguments rebut each other, there will be multiple preferred extensions, which means less sceptically justified conclusions. One could also argue that this situation is equivalent to the case of complete uncertainty. This can be modelled with inference scheme 5 in Table 8.

In order to get the desired results in all cases described above, the definition of defeat has to be slightly changed. First of all, for an argument $A$ to rebut another argument $B$, the conclusion of $A$ should not only be the negation of the conclusion of $B$, but it should also be at least as certain. This definition of rebuttal is similar to the ones used in preference-based argumentation (Amgoud and Cayrol 2002) and argumentation with defeasible priorities (Prakken and Sartor 1997). Next, since we want only the most certain argument for some conclusion to be justified, we introduce a new kind of defeat such that an argument $A$ defeats an argument $B$ if their conclusions

are the same but $A$'s conclusion is more certain. The definition of undercut remains unchanged.

**Definition 9** (*Defeat*) An argument $A$ *defeats* an argument $B$ if

- $\texttt{conc}(A) = \varphi : l$ and $\texttt{conc}(B) = \neg\varphi : l'$ (*rebuttal*) and $l > l' > 0$, or
- $\texttt{conc}(A) = \varphi : l$ and $\texttt{conc}(B) = \varphi : l'$ and $l > l' > 0$, or
- $\texttt{conc}(A) =$'$\texttt{inf}(B)$ is inapplicable' (*undercut*), or
- $A$ defeats a subargument of $B$.

In the following, we will assume that the epistemic part of the argumentation framework will resolve conflicts between formulas and their negations with possibly different certainty levels.

## 7 Argumentation Framework for Preferences with Uncertain Information

Now that we have introduced a framework for epistemic reasoning with uncertain information, we turn to the question how to derive preferences, if any, from such uncertain information. Different approaches to infer preferences from information with varying degrees of certainty are possible. We discuss several different ones. The strategies we present in this section all apply the lexicographic ordering in the sense that a preference between two objects is determined at the highest importance level of attributes where a preference can be derived. They differ in the way a preference is determined within one importance level.

In the lexicographic ordering, objects are compared w.r.t. their attributes on an importance level. The highest importance level where a preference can be derived determines the overall preference. In the Boolean case, preference within an importance level is determined solely by the number of true attributes of both objects (the number of false attributes can be ignored because it can be computed when the number of true attributes is known). This comparison is relatively easy, since we only have to compare two numbers. In the case of uncertainty, instead of two possible values for an attribute, we have multiple ($2N + 1$), one for each level of certainty. So in this case, we compare two tuples of numbers. Such a comparison can be done in different ways, resulting in different strategies. Abstractly, we can say that the triples are compared by a 'beats' relation $B$. If on some importance level object $a$ has $m_N$ certainly true, …, and $m_{-N}$ certainly false attributes, object $b$ has $m'_N$ certainly true, …, and $m'_{-N}$ certainly false attributes, and $\langle m_N, \ldots, m_{-N}\rangle >_B \langle m'_N, \ldots, m'_{-N}\rangle$, then object $a$ is preferred over object $b$ on that importance level. The key issue is how to define $B$.

### 7.1 Purely Qualitative Strategies

We briefly mention some extreme cases. First, it would be possible to reduce the number of certainty levels to two ($N$ and $-N$) by treating the levels in between either the same way as $N$ (optimistic approach) or the same way as $-N$ (pessimistic approach). This is a generalisation of the optimistic and pessimistic strategies discussed in Sect. 4.1, and the same objections apply. A third option is to treat all positive

**Table 9** Examples of objects and attributes with uncertain information

| | $P \triangleq$ | $Q \triangleq$ | $R$ |
|---|---|---|---|
| $a$ | 2 | 1 | 1 |
| $b$ | 1 | 1 | -1 |
| $c$ | 1 | 2 | 0 |

**a**

| | $P \triangleq$ | $Q \triangleq$ | $R \triangleq$ | $S \triangleq$ | ... |
|---|---|---|---|---|---|
| $a$ | 2 | -2 | -2 | -2 | ... |
| $b$ | 1 | 1 | 1 | 1 | ... |

**b**

| | $P \triangleq$ | $Q \triangleright$ | $R \triangleq$ | $S \triangleq$ | $T$ |
|---|---|---|---|---|---|
| $a$ | -1 | 2 | 2 | 2 | -1 |
| $b$ | 2 | -1 | 1 | 1 | 1 |

**c**

| | $P \triangleq$ | $Q$ |
|---|---|---|
| $a$ | 1 | 0 |
| $b$ | -1 | 1 |

**d**

certainty levels the same way as $N$ and all negative certainty levels the same way as $-N$. This shows great confidence in the correctness of information, but ignores the differences in certainty. Finally, it is possible to treat all certainty levels between $-N$ and $N$ the same way as 0 (this essentially reduces the problem to the case with incomplete information discussed in Sect. 4). This focuses on the uncertainty of the information, but does not take into account that there may be different degrees of uncertainty. Any non-extreme approach should distinguish between different certainty levels.

An obvious strategy is dominance: an outcome $a$ is (weakly) preferred to an outcome $b$ if for all attributes, it is at least as certain that $a$ has it as that $b$ has it. For example, in the situation in Table 9a, object $a$ is preferred to object $b$ since its certainty level is at least as high for every attribute. This strategy is clear-cut and quite safe. On the other hand, it is not very decisive: the resulting preference relation is far from complete. Also, it does not take into account what it means for two attributes to be equally important, namely that they are interchangeable. If $P$ and $Q$ have equal importance, it does not matter for preference whether an object has $P$ but not $Q$, or $Q$ but not $P$. For example, in the situation in Table 9a, objects $a$ and $c$ are incomparable according to dominance, while it would be intuitive to prefer $a$. To solve this issue, the definition of dominance can be straightforwardly adapted to the following definition of 'ordered dominance'. The attributes within an importance level can be rank-ordered according to the certainty that an object has the attribute (in the case of ties, i.e. multiple attributes having the same certainty for an outcome, consecutive ranks are assigned to them at random). Such an ordering may be different for every object. Now object $a$ is (weakly) preferred to object $b$ if for every rank, it is at least as certain that $a$ has the attribute with that rank in $a$'s ordering as that $b$ has the attribute with that rank in $b$'s ordering. This definition results in a preference of $a$ over $c$ in Table 9a. This definition captures the intuition behind equal importance of attributes, and again, the derived preferences are intuitive. However, it still lacks in decisiveness, since many objects will be incomparable even though it may be reasonable to prefer one over the other.

In order to define a more decisive strategy, one could consider using the lexicographic ordering on certainty levels. That is, within one importance level, we first count all attributes with certainty $N$. If one object has more of those than another, the first object is preferred over the second (within this level). If both objects have

the same number, we go on to count the attributes with certainty $N-1$, and so on. When both objects have the same number of attributes on every certainty level, we go on to consider the next importance level. The advantage of this strategy is that it results in a complete preference relation, i.e. two objects are never incomparable. On the other hand, some derived preferences may not be intuitive, since no number of less certain attributes can be valuated higher than a single more certain attribute. Consider for example the situation in Table 9b, where $N = 2$. It is known for certain that object $a$ has attribute $P$ and none of the other attributes. Object $b$ has all attributes with certainty 1. The lexicographic strategy will always prefer object $a$, no matter how many attributes there are in the importance level, even though it would be more intuitive to prefer $b$ when the odds are taken into account. In the next section we propose a strategy that does take the odds into account and allows compensation.

### 7.2 Compensatory Strategy

The reason that it would be more intuitive to prefer object $b$ to object $a$ in Table 9b as the number of attributes in the importance level increases, is that it becomes more likely that object $b$ will have more attributes than object $a$. In other words, the expected number of attributes of $b$, which increases with every attribute with certainty 1 that is added, will be higher than the expected number of attributes of $a$, which stays 1. In this section we present a strategy for determining preference that is based on the expected number of attributes of each object. As said before, the strategy applies the lexicographic ordering in the sense that a preference between two objects is determined at the highest importance level of attributes where a preference can be derived. Within an importance level, it prefers one object to another if the first has a higher expected number of attributes on that level than the second.

In order to calculate the expected number of attributes that an outcome has on some importance level, we need to know the subjective probability $pr$ associated with each certainty level. Some probabilities may be known, such as $pr(N) = 1$ and $pr(-N) = 0$, others have to be estimated. The probability function has to be monotonic, i.e. $pr(l) > pr(l')$ iff $l > l'$, and $pr(l) = 1 - pr(-l)$. Here we only treat the case for unconditional probabilities. Extending this with conditional probabilities is a straightforward application of well-known techniques and would unnecessarily complicate the presentation of the argumentation framework here. If the probabilities of an outcome having different attributes are independent, given $P_1(a) : l_1, \ldots, P_n(a) : l_n$, the expected number of attributes among $P_1, \ldots, P_n$ that object $a$ has is given by $\sum_{i=1}^{n} pr(l_i)$.

To incorporate this strategy into the argumentation framework, the interpretation of the formula $has(a, [P], n)$ is changed slightly from 'object $a$ has $n$ attributes equally important as attribute $P$' to 'object $a^x$ *expectedly* has $n$ attributes equally important as attribute $P$'. Inference scheme 1 in Table 10 takes as premises the attributes of a certain importance level and the certainty levels of an object $a$ having these attributes. It concludes that the expected number of attributes of the object on that importance level is the sum of the probabilities of the certainty levels. Inference scheme 2 is an

**Table 10** Inference schemes for the compensatory strategy for uncertain information

| | | |
|---|---|---|
| 1 | $\frac{P_1(a) : l_1, \ldots, P_n(a) : l_n \quad P_1 \triangleq \cdots \triangleq P_n}{has(a, [P_1], \sum_{i=1}^{n} pr(l_i))}$ | $count(a, [P_1], \{P_1, \ldots, P_n\})$ |
| 2 | $\frac{P_1(a) : l_1, \ldots, P_n(a) : l_n \quad P_1 \triangleq \cdots \triangleq P_n}{count(a, [P_1], S \subset \{P_1, \ldots, P_n\}) \text{ is inapplicable}}$ | $count(a, [P_1], S)uc$ |
| 3 | $\frac{has(a, [P], n) \quad has(b, [P], m) \quad n > m}{pref(a, b)}$ | $p(a, b, [P])$ |
| 4 | $\frac{has(a, [Q], n) \quad has(b, [Q], m) \quad Q \rhd P \quad n \neq m}{p(a, b, [P]) \text{ is inapplicable}}$ | $p(a, b, [P])uc$ |
| 5 | $\frac{has(a, [P], n) \quad has(b, [P], m) \quad n = m}{eqpref(a, b)}$ | $eqp(a, b, [P])$ |
| 6 | $\frac{has(a, [Q], n) \quad has(b, [Q], m) \quad Q \rhd P \quad n \neq m}{eqp(a, b, [P]) \text{ is inapplicable}}$ | $eqp(a, b, [P])uc$ |

undercutter that defeats an argument built with scheme 1 if not all attributes at the importance level in question are considered. Note that *all* attributes on the importance level should be included; if some attribute is not present, the certainty level will just be $-N$. Inference scheme 3 infers that an object $a$ is preferred over an object $b$ if the expected number of attributes of $a$ is higher than the expected number of attributes of $b$ on a certain importance level. This scheme is undercut by inference scheme 4 if there is a higher importance level where $a$ and $b$ do not have the same expected number of attributes. Schemes 5 and 6 do the same for equal preference.

*Example 5* Consider the situation in Table 9c. Since both objects have one attribute with certainty 2 and one attribute with certainty $-1$ on the highest importance level, the preference is determined on the second importance level. Which object is preferred depends on the probabilities of the different certainty levels. Since $N = 2$, we have $pr(2) = 1$ and $pr(-2) = 0$, but the probabilities for the other certainty levels can be subjectively estimated. For example, if we take $pr(1) = 0.75$ and $pr(-1) = 0.25$, the expected number of attributes of object $a$ on the second importance level will be $2^* pr(2) + pr(-1) = 2^*1 + 0.25 = 2.25$ and the expected number of attributes of object $b$ on the second importance level will be $3^* pr(1) = 3^*0.75 = 2.25$, and both objects will be equally preferred. This is illustrated with argument A in Table 11. Other probability estimates lead to other preferences. In argument $B$, we take $pr(1) = 0.8$ and $pr(-1) = 0.2$, and $b$ is preferred over $a$. In argument $C$, we take $pr(1) = 0.7$ and $pr(-1) = 0.3$, and $a$ is preferred over $b$.

For this strategy we have been specific in assigning probabilities to certainty levels. As can be seen in Example 5, small differences in the estimated probabilities can lead to completely different preferences. This makes the strategy decisive (it always infers a preference), but not very safe, since subjective probabilities are not always exactly known and may not be estimated accurately. In the next section we present a safer version of the compensatory strategy that generalises both the compensatory strategy

**Table 11** Some example arguments in the compensatory strategy

| | | | |
|---|---|---|---|
| A: | $\dfrac{\dfrac{R(a):2 \quad S(a):2 \quad T(a):-1 \quad R \triangleq S \triangleq T}{has(a,[R],2.25)} \quad \dfrac{R(b):1 \quad S(b):1 \quad T(b):1 \quad R \triangleq S \triangleq T}{has(b,[R],2.25)}}{eqpref(a,b)}$ | | $2.25 = 2.25$ |
| B: | $\dfrac{\dfrac{R(a):2 \quad S(a):2 \quad T(a):-1 \quad R \triangleq S \triangleq T}{has(a,[R],2.2)} \quad \dfrac{R(b):1 \quad S(b):1 \quad T(b):1 \quad R \triangleq S \triangleq T}{has(b,[R],2.4)}}{pref(b,a)}$ | | $2.2 < 2.4$ |
| C: | $\dfrac{\dfrac{R(a):2 \quad S(a):2 \quad T(a):-1 \quad R \triangleq S \triangleq T}{has(a,[R],2.3)} \quad \dfrac{R(b):1 \quad S(b):1 \quad T(b):1 \quad R \triangleq S \triangleq T}{has(b,[R],2.1)}}{pref(a,b)}$ | | $2.3 > 2.1$ |

of this section and the safe and decisive strategy for incomplete information presented in Sect. 4.3.

### 7.3 A Safer Compensatory Strategy

The strategy presented here is a generalisation of the compensatory strategy of the previous section, inspired by the strategy for incomplete information presented in Sect. 5. The idea is as follows. Instead of assigning a single probability $pr(l)$ to every certainty level $l$, we specify a range of probability with a lower bound $pr^-(l)$ and an upper bound $pr^+(l)$. Now we can use the same intuition as before. If the worst case for object $a$ is still preferred over the best case for object $b$, then $a$ has to be preferred over $b$ (on some importance level). Note that fully certain information still gets a single probability value: $pr^-(-N) = pr^+(-N) = 0$ and $pr^-(N) = pr^+(N) = 1$. The strategy is less decisive than the compensatory strategy presented in the previous section, since it is not always able to derive a preference between two objects. However, it is still more decisive than the dominance-based strategy.

Inference schemes 1–6 in Table 12 are the same as those for the compensatory strategy in Table 10, except that labels ($-$ and $+$) are added to objects and $pr^-$ and $pr^+$ are used in inference scheme 1. These inference schemes can be used to infer preferences between best and worst cases of objects, similar to best and worst completions in Sect. 5. The inference schemes to infer preferences between objects are exactly the same as before: inference schemes 7–9 in Table 12 are the same as inference schemes 10–12 in Table 6.

*Example 6* Consider the same situation in Table 9c again. If we take $pr^+(1) = 0.9$, $pr^-(1) = 0.8$, $pr^+(-1) = 0.2$ and $pr^-(-1) = 0.1$, we can build argument $A$ in Table 13, concluding that $b$ is preferred over $a$. If we take $pr^+(1) = 0.8, pr^-(1) = 0.7$, $pr^+(-1) = 0.3$ and $pr^-(-1) = 0.2$, no justified arguments concluding a preference between $a$ and $b$ can be constructed.

**Proposition 2** *The safer compensatory strategy presented here generalises the compensatory strategy in Sect. 7.2.*

**Table 12** Inference schemes for the safer compensatory strategy for uncertain information

| | | |
|---|---|---|
| 1 | $\dfrac{P_1(a) : l_1, \ldots, P_n(a) : l_n \quad P_1 \triangleq \cdots \triangleq P_n}{has(a^x, [P_1], \sum_{i=1}^{n} pr^x(l_i))}$ | $count(a^x, [P_1], \{P_1, \ldots, P_n\})$ |
| 2 | $\dfrac{P_1(a) : l_1, \ldots, P_n(a) : l_n \quad P_1 \triangleq \cdots \triangleq P_n}{count(a^x, [P_1], S \subset \{P_1, \ldots, P_n\}) \text{ is inapplicable}}$ | $count(a^x, [P_1], S)uc$ |
| 3 | $\dfrac{has(a^x, [P], n) \quad has(b^y, [P], m) \quad n > m}{pref(a^x, b^y)}$ | $p(a^x, b^y, [P])$ |
| 4 | $\dfrac{has(a^x, [Q], n) \quad has(b^y, [Q], m) \quad Q \rhd P \quad n \neq m}{p(a^x, b^y, [P]) \text{ is inapplicable}}$ | $p(a^x, b^y, [P])uc$ |
| 5 | $\dfrac{has(a^x, [P], n) \quad has(b^y, [P], m) \quad n = m}{eqpref(a^x, b^y)}$ | $eqp(a^x, b^y, [P])$ |
| 6 | $\dfrac{has(a^x, [Q], n) \quad has(b^y, [Q], m) \quad Q \rhd P \quad n \neq m}{eqp(a^x, b^y, [P]) \text{ is inapplicable}}$ | $eqp(a^x, b^y, [P])uc$ |
| 7 | $\dfrac{pref(a^-, b^+)}{pref(a, b)}$ | |
| 8 | $\dfrac{eqpref(a^-, b^+) \quad pref(a^+, b^-)}{wpref(a, b)}$ | |
| 9 | $\dfrac{eqpref(a^+, b^-) \quad eqpref(a^-, b^+)}{eqpref(a, b)}$ | |

Let $pr^-(l) = pr^+(l)$ for every certainty level $l$, i.e. the probability range is actually a single probability value. Then for every object $a$, the expected number of attributes is the same for $a^+$ and $a^-$ and this strategy coincides with the compensatory strategy in Sect. 7.2.

*Example 7* Consider the same situation in Table 9c again. With $pr^-(1) = pr^+(1) = 0.75$ and $pr^-(-1) = pr^+(-1) = 0.25$, we can construct argument $B$ in Table 13, which is analogous to argument $A$ in Table 11.

**Proposition 3** *The safer compensatory strategy presented here generalises the safe and decisive strategy in Sect. 4.3.*

As said before, the case with incomplete information corresponds to the case with three certainty levels: $N$ and $-N$ (certain presence/absence of attributes) and 0 (unknown). If we take $pr^-(0) = 0$ and $pr^+(0) = 1$, then for every object $a$, the strategy proposed here counts the number of present and unknown attributes for $a^+$ and only the certainly present attributes for $a^-$, and hence coincides with the strategy presented in Sect. 7.

*Example 8* Consider the situation in Table 5c. If we translate this to certainty levels, taking $N = 1$, we get the situation in Table 9d. In this case, argument $C$ in Table 13 can be built (for reasons of space, this argument's two subarguments are displayed

**Table 13** Some example arguments in the safer compensatory strategy

$$\dfrac{\dfrac{\dfrac{R(a):2 \quad S(a):2 \quad T(a):-1 \quad R \triangleq S \triangleq T}{has(a^+,[R],2.2)} \quad \dfrac{R(b):1 \quad S(b):1 \quad T(b):1 \quad R \triangleq S \triangleq T}{has(b^-,[R],2.4)} \quad 2.2 < 2.4}{pref(b^-,a^+)}}{pref(b,a)}\ \text{A:}$$

$$\dfrac{\dfrac{\dfrac{R(a):2 \quad S(a):2 \quad T(a):-1 \quad R \triangleq S \triangleq T}{has(a^+,[R],2.25)} \quad \dfrac{R(b):1 \quad S(b):1 \quad T(b):1 \quad R \triangleq S \triangleq T}{has(b^-,[R],2.25)} \quad 2.25 = 2.25}{eqpref(a^+,b^-)} \quad \dfrac{\vdots}{eqpref(a^-,b^+)}}{eqpref(a,b)}\ \text{B:}$$

$$\dfrac{\dfrac{P(a):1 \quad Q(a):0 \quad P \triangleq Q}{has(a^-,[P],1)} \quad \dfrac{P(b):-1 \quad Q(b):1 \quad P \triangleq Q}{has(b^+,[P],1)} \quad 1 = 1}{eqpref(a^-,b^+)}\ \text{C}_1\text{:}$$

$$\dfrac{\dfrac{P(a):1 \quad Q(a):0 \quad P \triangleq Q}{has(a^+,[P],2)} \quad \dfrac{P(b):-1 \quad Q(b):1 \quad P \triangleq Q}{has(b^-,[P],1)} \quad 2 > 1}{pref(a^+,b^-)}\ \text{C}_2\text{:}$$

$$\dfrac{\text{C}_1 \quad \text{C}_2}{wpref(a,b)}\ \text{C:}$$

separately). When we compare this argument to argument $B$ in Table 7 we see that the conclusions are indeed the same.

## 8 Conclusion

In this paper we have made the following contributions. Approaches based on argumentation can be used to model qualitative multi-attribute preferences such as the lexicographic ordering. The advantage of argumentation over other approaches emerges most clearly in the case of incomplete or uncertain information. Our approach to the incomplete information case allows to reason about preferences from best- and worst-case perspectives (called completions here), and the consequences for overall preferences. In addition we proposed different ways to reason about preferences in case of uncertain information.

In our future work we would like to distinguish more explicitly between mental attitudes such as beliefs, goals, desires and preferences. This will also allow us to reason about these attitudes, for example that a certain preference we have is based on some specific beliefs. We hope to gain insight from modal preference languages with belief operators such as the one presented in Liu (2008).

In the current paper we have focused on the case where we have incompleteness of uncertainty in the epistemic part of the knowledge base (i.e. about the attributes that objects do or do not have). It would be interesting to explore the case where also information about what attributes influence preference and the importance order between them is incomplete or uncertain. This is especially useful when modelling preferences of others, where it is not realistic that all relevant information is available.

Other interesting cases are inconsistency in and change of the knowledge that is used to determine preferences.

Other interesting areas for future work include the representation of dependent preferences (e.g. 'I only want a balcony if the house does not have a garden, otherwise I do not care'), different degrees of satisfaction of attributes, and preferences based on underlying interests or values. We would also like to look into the relation with e.g. CP-nets (Boutilier et al. 2004) and value-based argumentation (Kaci and van der Torre 2008).

Finally, we believe that the argumentation-based framework for preferences presented here can be usefully applied in the preference elicitation process. It allows the user to extend and refine the system representation of his preferences gradually and as the user sees fit. To facilitate this elicitation process more research is needed on how our framework can support a user e.g. by indicating which information is still missing.

**Acknowledgments** This research is supported by the Dutch Technology Foundation STW, applied science division of NWO and the Technology Program of the Ministry of Economic Affairs. It is part of the Pocket Negotiator project with grant number VICI-project 08075.

## References

Amgoud L, Cayrol C (2002) Inferring from inconsistency in preference-based argumentation frameworks. J Autom Reason 29:125–169

Amgoud L, Prade H (2009) Using arguments for making and explaining decisions. Artif Intell 173(3–4):413–436

Amgoud L, Maudet N, Parsons S (2000) Modelling dialogues using argumentation. In: Proceedings of the fourth international conference on multi-agent systems, pp 31–38

Amgoud L, Bonnefon JF, Prade H (2005) An argumentation-based approach to multiple criteria decision. In: Proceedings of the eighth European conference on symbolic and quantitative approaches to reasoning with uncertainty (ECSQARU 2005), pp 269–280

Bench-Capon TJM, Dunne PE (2007) Argumentation in artificial intelligence. Artif Intell 171:619–641

Bonet B, Geffner H (1996) Arguing for decisions: a qualitative model of decision making. In: Proceedings of the 12th conference on uncertainty in artificial intelligence (UAI 1996), pp 98–105

Bonnefon JF, Fargier H (2006) Comparing sets of positive and negative arguments: empirical assessment of seven qualitative rules. In: Proceedings of the 17th European conference on artificial intelligence (ECAI 2006), IOS Press, pp 16–20

Boutilier C (1994) Toward a logic for qualitative decision theory. In: Proceedings of the fourth international conference on principles of knowledge representation and reasoning (KR 1994), pp 75–86

Boutilier C, Brafman RI, Domshlak C, Hoos HH, Poole D (2004) CP-nets: a tool for representing and reasoning with conditional ceteris paribus preference statements. J Artif Intell Res 21:135–191

Brewka G (2004) A rank based description language for qualitative preferences. In: Proceedings of the sixteenth European conference on artificial intelligence (ECAI 2004), pp 303–307

Brewka G, Benferhat S, Le Berre D (2004) Qualitative choice logic. Artif Intell 157(1–2):203–237

Coste-Marquis S, Lang J, Liberatore P, Marquis P (2004) Expressive power and succinctness of propositional languages for preference representation. In: Proceedings of the ninth international conference on principles of knowledge representation and reasoning (KR 2004), pp 203–212

Doyle J, Thomason RH (1999) Background to qualitative decision theory. AI Mag 20(2):55–68

Dubois D, Fargier H, Perny P (2003) Qualitative decision theory with preference relations and comparative uncertainty: an axiomatic approach. Artif Intell 148:219–260

Dubois D, Fargier H, Bonnefon JF (2008) On the qualitative comparison of decisions having positive and negative features. J Artif Intell Res 32:385–417

Dung PM (1995) On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and $n$-person games. Artif Intell 77:321–357

Huber F (2009) Belief and degrees of belief. In: Huber F, Schmidt-Petri C (eds) Degrees of belief, Synthese Library, vol 342. Springer, Netherlands, pp 1–33
Kaci S, van der Torre L (2008) Preference-based argumentation: arguments supporting multiple values. Int J Approx Reason 48(3):730–751
Keeney RL, Raiffa H (1993) Decisions with multiple objectives: preferences and value trade-offs. Cambridge University Press, Cambridge
Liu F (2008) Changing for the better: preference dynamics and agent diversity. PhD thesis, Universiteit van Amsterdam
Pollock JL (2001) Defeasible reasoning with variable degrees of justification. Artif Intell 133(1-2):233–282
Prakken H (2005) A study of accrual of arguments, with applications to evidential reasoning. In: Proceedings of the tenth international conference on artificial intelligence and law (ICAIL 2005), pp 85–94
Prakken H, Sartor G (1997) Argument-based extended logic programming with defeasible priorities. J Appl Non-Class Logics 7:25–75
Rahwan I, Ramchurn SD, Jennings NR, McBurney P, Parsons S, Sonenberg L (2004) Argumentation-based negotiation. Knowl Eng Rev 18(4):343–375
Vreeswijk GAW (1997) Abstract argumentation systems. Artif Intell 90(1–2):225–279