Interpolative and extrapolative reasoning in propositional theories using qualitative knowledge about conceptual spaces

Steven Schockaert[a], Henri Prade[b]

[a] *Cardiff University, School of Computer Science & Informatics, 5 The Parade, Cardiff CF24 3AA, UK*
[b] *Toulouse University, Université Paul Sabatier, IRIT, CNRS, 118 Route de Narbonne, 31062 Toulouse Cedex 09, France*

Abstract

Many logical theories are incomplete, in the sense that non-trivial conclusions about particular situations cannot be derived from them using classical deduction. In this paper, we show how the ideas of interpolation and extrapolation, which are of crucial importance in many numerical domains, can be applied in symbolic settings to alleviate this issue in the case of propositional categorization rules. Our method is based on (mainly) qualitative descriptions of how different properties are conceptually related, where we identify conceptual relations between properties with spatial relations between regions in Gärdenfors conceptual spaces. The approach is centered around the view that categorization rules can often be seen as approximations of linear (or at least monotonic) mappings between conceptual spaces. We use this assumption to justify that whenever the antecedents of a number of rules stand in a relationship that is invariant under linear (or monotonic) transformations, their consequents should also stand in that relationship. A form of interpolative and extrapolative reasoning can then be obtained by applying this idea to the relations of betweenness and parallelism respectively. After discussing these ideas at the semantic level, we introduce a number of inference rules to characterize interpolative and extrapolative reasoning at the syntactic level, and show their soundness and completeness w.r.t. the proposed semantics. Finally, we show that the considered inference problems

Email addresses: s.schockaert@cs.cardiff.ac.be (Steven Schockaert), prade@irit.fr (Henri Prade)

Preprint submitted to Elsevier *April 24, 2013*

are PSPACE-hard in general, while implementations in polynomial time are possible under some relatively mild assumptions.

Keywords: Commonsense reasoning, Conceptual spaces, Interpolative reasoning, Analogical reasoning

1. Introduction

Symbolic approaches to knowledge representation typically start from a finite set of natural language labels, which are associated to atomic propositions (in propositional settings), to predicates (in first-order settings), or to atomic concepts (in description logics). The meaning of these labels is then expressed implicitly by encoding how the corresponding propositions are related to each other using a logical theory. Clearly, such a logical theory can only capture a small fraction of the actual meaning of the labels at the cognitive level. Formalizing commonsense reasoning then boils down to developing principled approaches to extend or refine logical inference, such that the conclusions that can be derived from a logical theory become, in some way, closer to what we infer at the cognitive level. Some approaches to non-monotonic reasoning [1], for instance, deal with exceptions by assuming that rules only apply to *typical* instances of the concepts involved or are only valid in *normal* situations, even though these notions of typicality and normality are not explicitly expressed at the symbolic level. Essentially, such approaches are non-monotonic because the factual knowledge they work on is incomplete: we may know that Tweety is a bird, but not that it is a penguin. Later observations may enrich our factual knowledge base, necessitating a revision of some of the assumptions that were made (e.g. Tweety can fly). As another form of commonsense reasoning, in this paper we look at techniques for dealing with a lack of generic knowledge. For the ease of presentation, we will assume that generic knowledge is expressed as a set of propositional rules. We are then interested in situations where factual knowlege is, in principle, complete, but where none of the given rules applies to the situation at hand. For instance, we may know that it is advisable to rest (i) when feeling nauseated and having a high fever, and (ii) when feeling nauseated without any fever, but not have any information about what is advisable in case of nausea with a mild fever.

Similarity-based reasoning. Humans can often cope with such a lack of knowledge by drawing analogies, or by resorting to knowledge about similar situations [2, 3]. This observation has led to a number of theories of approximate reasoning, which are mainly based on the premise that from similar conditions we can draw similar conclusions. Most notably, a large number of fuzzy set based approaches have been proposed that build on the idea that the more a situation is compatible with the fuzzy labels in the antecedent of a rule, the more it should be compatible with the fuzzy labels in its consequent. Such rules, called gradual rules in [4], are based on a measure of similarity that is implicit in the definition of the membership functions of the fuzzy labels. Related to gradual rules, certainty rules [4] rather encode that the more similar a situation is to the fuzzy labels in the antecedent of a rule, the more certain that the consequent holds. Some related approaches avoid the use of fuzzy sets and rather encode similarity assessments in an explicit way (e.g. [5, 6, 7]). Given that a_1 and $a \rightarrow b$ hold, such methods allow to derive b with a certainty that depends on the degree of similarity $sim(a, a_1)$.

Despite the intuitive appeal of similarity-based approaches, and their popularity in the context of control and classification problems, they face a number of difficulties when used for commonsense reasoning. First, quantitative similarity degrees can be hard to obtain in practice, a problem which is aggravated by the observation that similarity judgements are context-dependent [8, 9]. The quantitative nature of similarity degrees also makes it difficult to encode rules, e.g. exactly how similar should a given premise a_1 be to the antecedent of the rule $a \rightarrow b$ to derive b with a given certainty? Finally, similarity based methods tend to lack a principled way of dealing with conjunctions. For instance, assume that the rule $a \wedge b \rightarrow c$ and the facts a_1 and b_1 are known to hold, and moreover $sim(a, a_1) = 0.6$ and $sim(b, b_1) = 0.4$. To assess whether we can plausibly derive $a_1 \wedge b_1 \rightarrow c$ using similarity based reasoning, we would need to assess to what extent $a_1 \wedge b_1$ is similar to $a \wedge b$. Usually, a truth-functional approach is assumed, assuming e.g. $sim(a \wedge b, a_1 \wedge b_1) = \min(sim(a, a_1), sim(b, b_1))$ or $sim(a \wedge b, a_1 \wedge b_1) = sim(a, a_1) \cdot sim(b, b_1)$. Such views, however, are hard to justify from a cognitive point of view.

Betweenness. The aforementioned shortcomings of similarity-based reasoning seem closely related to the use of degrees. A key observation is that in practice, similarity-based approaches are often used to implement a form of *interpolation* of symbolic rules: given the rules $a_1 \rightarrow b_1$ and $a_3 \rightarrow b_3$, and a premise a_2 which is known to be between a_1 and a_3, a conclusion is obtained between b_1 and b_3. Interpolative inference, however, can also be implemented in a qualitative way, taking *betweenness* as primitive rather than similarity.

Indeed, it suffices to know which propositions are conceptually between a_1 and a_3 and which propositions are between b_1 and b_3. For instance, since a mild fever is conceptually between high fever and no fever, in the earlier example we conclude that resting is advised in case of nausea with a mild fever. This basic form of interpolative inference can then be further refined, depending on the kind of background information that is available about the conceptual relationships between the propositions (labels). For example, if we know that a_2 is closer to a_1 than to a_3, we may insist that the conclusion should also be closer to b_1 than to b_3. The idea of interpolating symbolic knowledge can also be extended to various forms of *extrapolative* inference, as is illustrated in the following example.

Example 1. *Consider the following knowledge base, containing observations about the comfort level of different housing options:*

$$\text{mansion} \rightarrow \text{exclusive} \tag{1}$$

$$\text{villa} \wedge \text{suburbs} \rightarrow \text{luxurious} \tag{2}$$

$$\text{apartment} \wedge \text{suburbs} \rightarrow \text{basic} \tag{3}$$

$$\text{apartment} \wedge \text{centre} \rightarrow \text{very-comfortable} \tag{4}$$

Clearly, this knowledge base is incomplete. For instance, we have no information at all about the comfort level of a villa in the centre. However, from the rules that are provided, it seems reasonable to assume that villas are more comfortable than apartments (by comparing (2) *and* (3)*) and that housing in the centre is more comfortable than housing in the suburbs (by comparing* (3) *and* (4)*). As a form of extrapolative inference, this leads us to conclude that a villa in the centre would at least be as comfortable as a villa in the suburbs, i.e. either* luxurious *or* exclusive*. We may also wonder about apartments in the outskirts of the city. As living in the outskirts is conceptually between living in the centre and living in the suburbs, from* (3) *and* (4) *we may reasonably assume, as a form of interpolative inference, that the comfort level of an apartment in the outskirts would be between* basic *and* very-comfortable*.*

Objectives of the paper. The aim of this paper is to develop a principled approach to interpolative and extrapolative reasoning, as a general way to avoid the use of degrees when dealing with incomplete generic knowledge. In particular, we address the following research questions:

1. How can interpolative and extrapolative inference be formalized? What are its computational properties and how can automated inference procedures be implemented?
2. What is the nature of the background knowledge that is needed to support interpolative and extrapolative reasoning?
3. What is the semantic justification for interpolation and extrapolation? While most approaches to non-monotonic reasoning have a principled semantic foundation, typically based on the notion of preferred worlds [10, 1] or the idea of stable models [11, 12], this is to a much lesser extent the case for current methods that deal with incompleteness of rule bases.

The underlying idea is that among all possible refinements of a given knowledge base we favour those which are most *regular*, an intuition which will be formalized using the theory of conceptual spaces [13]. This theory posits that natural language labels can be identified with a convex region in a particular geometric space — called a conceptual space — whose dimensions correspond to cognitively meaningful qualities. Using conceptual spaces, notions such as betweenness can be given a clear geometric interpretation, which allows us to derive a semantic characterization of various interpolative and extrapolative inference relations. It is important to note, however, that although conceptual spaces are crucial to justify our approach, in practical applications, we do not actually require that the conceptual space representations of properties are available. In particular, the inference mechanism itself will only require qualitative knowledge about how the conceptual representations of labels are spatially related. For instance, to support a basic interpolative inference relation, it is only required that we know which labels are conceptually between which other labels. Furthermore note that this form of commonsense reasoning will actually be monotonic: increasing the rule base may allow us to refine earlier conclusions, but will never violate them. In contrast, in the setting of non-monotonic reasoning, increasing the factual knowledge may lead us to consider different rules to be applicable, as more specific rules may override more general ones.

Depending on the considered application, the required qualitative knowledge can be provided by an expert or it can be derived automatically using data-driven techniques. In the first case, an expert may choose to manually encode some rules, and rely on interpolation or extrapolation to avoid the need for a complete specification of the considered domain (e.g. explicitly enumerating the comfort level for all housing types and all location types). The resulting inference relation would be guaranteed to provide sound conclusions, although for those parts of the domain that were not explicitly modeled by the expert, available information may be less precise (but not trivial). In the second case, where data-driven techniques are used to obtain background information about the conceptual relationship of different labels (e.g. by analysing documents from the web), the aim is rather to generate plausible conclusions from imperfect conceptual background knowledge. In this way, we can combine the rigour of a logic-based framework with the flexibility of data-driven methods. The commonsense aspect of the approach thus lies in the possibility to go beyond classical deduction by taking advantage of structural domain knowledge that has been induced from data, without resorting to purely statistical techniques as in [14].

Organization. The paper is structured as follows. After discussing related work in Section 2, we present a high-level overview of our approach in Section 3. Section 4 subsequently discusses in more detail how qualitative spatial descriptions of conceptual spaces can be used to encode the conceptual relationship of different atomic properties. Next, in Section 5 we investigate how conceptual relations between atomic properties can be leveraged to conceptual relations between unions and intersections of properties (represented as sets of vectors of properties), and how the latter conceptual relations can be used to refine a given rule base at the semantic level. In Section 6 we then focus on the syntactic level: we introduce a set of inference rules and show that they are sound and complete w.r.t. the semantics from Section 5. Section 7 analyzes the computational complexity of interpolative and extrapolative reasoning and presents an implementation method. In Section 8 we present some further thoughts on how to apply our method in practice, and in particular on the question of how to handle inconsistencies that are introduced by our method. We present our conclusions and a number of directions for future work in Section 9. Finally, note that this paper forms a substantially extended and revised version of [15]. Among others, we now provide a complete characterization of extrapolative inference and include the idea of comparative distance (whereas only the interpolative inference relation was characterized in [15]). We moreover present an implementation method, as well as the proofs.

2. Related work

Our work is clearly related to existing approaches in cognition and knowledge representation that are based on a spatial representation of knowledge. However, our approach also touches upon several other domains, including non-monotonic reasoning, similarity-based reasoning, regularization in machine learning, and qualitative physics. We briefly clarify the relationship with each of these domains.

Spatial representations of meaning

One of the central motivations of this paper is to approach commonsense reasoning by abandoning the idea that atomic propositions are independent from each other, in favour of a view which allows them to be conceptually related in a way that cannot be fully expressed at the logical level. Although atomic propositions are traditionally assumed to be independent, several 20th century philosophers have argued against such a view. Wittgenstein [16] was among the first to realize that sometimes we need more than a purely syntactic approach to logic, considering that atomic logical formulas may exclude each other while they are not contradictory. A statement such as *place P is green at time T and place P it is blue at time T* is treated as nonsensical by Wittgenstein rather than false, where he writes "It is, of course, a deficiency of our notation that it does not prevent the formation of such nonsensical constructions, and a perfect notation will have to exclude such structures by definite rules of syntax" [16]. In the same spirit, Carnap [17] uses the notion of an attribute space to group predicates of the same type. An attribute space is an abstract representation of a certain domain. For instance, the attribute space of colours consists of all (infinitely many) colour instances. In practice, these attribute spaces are usually described using a finite set of labels, which correspond to predicates at the logical level. By thus partitioning the predicates into separate attribute spaces, one can restrict interpretations to those that make exactly one predicate true from each attribute space, for any given individual. Quine [18] uses the related notion of a quality space to characterize similarity, putting forward the view that similarity cannot be defined in logical terms, and thus requires a deeper representation of atomic propositions. These works have led to the more recent development of conceptual spaces [13] by Gärdenfors, in an attempt to use the idea of a spatial representation of meaning to tackle problems in artificial intelligence (AI), among others. Conceptual spaces will be discussed in more detail in Section 4.

Apart from the work on conceptual spaces, the idea of assuming a spatial representation to reason about concepts also underlies [19], where an approach to integrate heterogeneous databases is proposed based on spatial relations between concepts. This approach starts from the observation that types from one database may not have an exact counterpart in another database. Conceptual relations between types are therefore considered which express e.g. that all typical instances of type A belong to type B but some instances of type A may be outside B. Such relations can formally be modelled as egg/yolk relations, which are a form of qualitative spatial relations between ill-defined spatial regions. Somewhat related, in [20] we presented a general method for merging conflicting propositional knowledge bases coming from different sources, based on the view that different sources may have a slightly different understanding of a given label. The different ways in which such a label may be understood are encoded in terms of four primitive relations that essentially correspond to qualitative spatial relations between (unknown) geometric representations of the labels. Although the kind of spatial relations encountered in these existing works are mainly mereotopological, qualitative spatial reasoning about betweenness is an active topic of research [21, 22].

Nonmonotonic reasoning

In general, several facets of commonsense reasoning have been extensively studied within the field of AI. Of particular interest is the work on System P for reasoning about rules with exceptions [1]. In this approach, a non-monotonic consequence relation is defined by encoding axiomatically how new rules may be derived from existing rules. In particular, the nonmonotonic consequence relation $\mid\!\sim$ is defined by the following inference rules:

Reflexivity $\alpha \mid\!\sim \alpha$

Left logical equivalence If $\alpha \equiv \alpha'$ and $\alpha \mid\!\sim \beta$ then $\alpha' \mid\!\sim \beta$

Right weakening If $\beta \models \beta'$ and $\alpha \mid\!\sim \beta$ then $\alpha \mid\!\sim \beta'$

OR If $\alpha \mid\!\sim \gamma$ and $\beta \mid\!\sim \gamma$ then $\alpha \vee \beta \mid\!\sim \gamma$

Cautious monotony If $\alpha \mid\!\sim \beta$ and $\alpha \mid\!\sim \gamma$ then $\alpha \wedge \beta \mid\!\sim \gamma$

Cut If $\alpha \wedge \beta \mid\!\sim \gamma$ and $\alpha \mid\!\sim \beta$ then $\alpha \mid\!\sim \gamma$

where $\equiv$ and $\models$ denote equivalence and entailment in classical logic, respectively. Intuitively, $\alpha \mathrel{|\!\sim} \beta$ means that in normal situations where α holds, it is also the case that β holds. The normative approach by System P about how a non-monotonic consequence relation should behave has been very influential in the field of non-monotonic reasoning. While the purpose of our paper is not to study non-monotonic consequence relations, our approach does resemble System P in that our goal is also to produce new rules, which are appropriate to a given situation. However, whereas System P is concerned with finding the most specific rules that are compatible with our (incomplete) knowledge about the situation at hand, in interpolative and extrapolative reasoning there is no genuine issue of incompleteness at the factual level. Rather, we are interested in situations where the given situation is not explicitly covered by a rule base, but is intermediate between, or analogous to situations that are covered.

Similarity-based reasoning

Somewhat related, several authors have studied similarity-based consequence relations which are based on the intuition that α approximately entails β iff every model of α is similar to some model of β [23, 6, 24]. In [25], for instance, a similarity-based consequence relation is contrasted with the consequence relation from System P, revealing that similarity-based reasoning satisfies monotonicity and most of the axioms of System P, but not the cut rule. More generally, a large number of authors have proposed systems for approximate, similarity-based reasoning within the field of fuzzy set theory. Most of these works are based on Zadeh's generalized modus ponens [26] (but see [27] for an early example of a more qualitative approach), which allows us to derive a fuzzy restriction on the value of variable Y from the knowledge that *if* X *is* A *then* Y *is* B and X *is* A' with A, A' and B fuzzy sets. The basic idea is that the more A is similar to A', the more the inferred restriction on Y will be close to B. When this idea is applied to a set of parallel rules, such that the fuzzy sets in the antecedents of the rules overlap, it leads to a form of interpolative reasoning. Furthermore, several authors have proposed methods to interpolate fuzzy rule bases in general, i.e. without requiring overlap of the fuzzy sets; we refer to [28] for a recent overview. While these techniques are also about interpolating rules, they differ from our approach in a number of ways. First, they are mostly restricted to unidimensional, numerical domains, and they require that quantitative representations of symbolic labels be available in the form of fuzzy sets. Furthermore, they treat logical connectives, such as conjunctions in the antecedent, in a truth-functional (and therefore heuristic) way.

In [29], a logic called CSL is introduced which has a construct $A \leftleftarrows B$ denoting all objects that are more similar to instances of concept A than to instances of concept B. The qualitative nature of this logic brings it closer to the approach we present in this paper. As it is based exclusively on closeness, and not on other aspects of spatial localization such as being in between, CSL is not directly suitable as a basis for interpolative or extrapolative reasoning. Interestingly, however, as a result of this restriction, CSL can be described using a preferential semantics [30].

Regularity

In the propositional setting, the idea of interpolation and extrapolation has been studied in [31], but from a rather different angle. In particular, the paper discusses how the belief that certain propositions hold at certain moments in time can be extended to beliefs about other moments in time, using persistence assumptions as a starting point. Nonetheless, as in our paper, the main idea is to use general meta-principles to find those completion(s) of a knowledge base that are most regular in some sense.

The idea of regularity can also be found in work on analogical proportions. An analogical proportion is an expression of the form $a : b :: c : d$ which reads as *a is to b as c is to d*. If a, b, c and d are binary propositions, this can be formalized as $(a \rightarrow b \equiv c \rightarrow d) \wedge (b \rightarrow a \equiv d \rightarrow c)$ (see [32]). In [33], an approach to classification is outlined which uses the view that, as a form of regularity, the more of the condition attributes of three training items form an analogical proportion with the condition attributes of the item to be classified, the more it becomes likely that also the decision attribute should form an analogical proportion. Using connectives from multi-valued logic, analogical proportions can be defined for graded propositions, which allows us to extend this idea to numerical attributes. In [34], an extrapolative inference mechanism has been proposed which is based on such analogical proportions between graded proportions. The latter technique can be seen as a special case of the approach we develop in this paper. In this paper, however, we also consider forms of interpolative and extrapolative reasoning that are not based on analogical proportions, and the proposed techniques are moreover not restricted to linearly ordered domains.

More generally, the idea of regularity appears in various forms in learning settings. In graph regularization [35], for instance, the desire for regularity is even made explicit in the form of a graph which connects instances that should receive a similar classification. In other approaches, the idea of regularity is implicit in the choice of the underlying classification functions that are allowed (e.g. being restricted to hyperplanes in the case of support vector machines), and is thus imposed to avoid overfitting. As an example of another domain where the idea of regularity surfaces, [36] presents an approach to derive a preference ordering, starting from a set of generic preferences. To choose a specific ordering among all those satisfying the constraints, the principle of minimal specificity from possibility theory is adopted as a way to avoid introducing any irregularities that have not been explicitly specified as constraints.

The notion of matrix abduction, proposed in [37] is also related to our work in its use of regularity as a criterion to complete missing values, although it operates at a lower-level representation. Specifically, consider a matrix whose rows correspond to objects and whose columns correspond to binary features, such that exactly one entry of the matrix is '?', corresponding to a missing value, and all the other entries are 0 or 1. Then [37] proposes to choose the missing value such that the regularity of the matrix is maximized. Specifically, a partial order relation is induced from both of the possible completions of the matrix, and the completion which is favoured is the one whose associated partial order relation is most natural in some sense. Note that, somewhat related, in abductive reasoning for causal diagnosis, it is also common to favour the simplest explanations (e.g. preferring single fault diagnoses to explain observed symptoms).

Qualitative physics

Finally, there is some resemblance between the inference procedure presented in this paper and the early work on qualitative reasoning about physical systems [38, 39], which deals with monotonicity constraints such as "if the value of x increases, then (all things being equal) the value of y decreases". Our inference procedure differs from these approaches as the domains we reason about do not need to be linearly ordered. Moreover, in the special case of linearly ordered domains, we assume no prior information about which partial mappings are increasing and which are decreasing.

3. Overview of the approach

In this section, we introduce some notations and basic concepts that will be used throughout the paper. We also present the main intuitions of our approach at an informal level.

Let $A_1, ..., A_n$ be finite sets of labels, where each set A_i corresponds to a certain type of properties[1] (e.g. colors), and the labels of A_i correspond to particular properties of the corresponding type (e.g. red, green, orange). The labels in A_i are assumed to correspond to jointly exhaustive and pairwise disjoint (JEPD) properties. Note that each element $(a_1, ..., a_n)$ from the Cartesian product $\mathcal{A} = A_1 \times ... \times A_n$ then corresponds to a maximally descriptive specification of the properties that some object or situation may satisfy. We furthermore assume that $A_i \cap A_j = \emptyset$ for $i \neq j$. We will refer to the sets A_i as attribute domains, and to their elements as attributes or, when used in a propositional language, as atoms.

We consider propositional rules of the form $\beta \rightarrow \gamma$, where β and γ are propositional formulas, built in the usual way from the set of atoms $A_1 \cup ... \cup A_n$ and the connectives $\wedge$ and $\vee$. Note that we do not need to explicitly consider negation, as the negation of an atom $a \in A_i$ corresponds to the disjunction of the atoms in $A_i \setminus \{a\}$. We say that an element $(a_1, ..., a_n) \in \mathcal{A}$ is a model of a formula (or a rule) α, written $(a_1, ..., a_n) \models_{\mathcal{A}} \alpha$ if the corresponding propositional interpretation $\{a_1, ..., a_n\}$ is a model of α in the usual sense, where we see propositional interpretations as sets containing all atoms that are interpreted as true. For formulas (or rules or sets of rules) α_1 and α_2, we say that α_1 entails α_2, written $\alpha_1 \models_{\mathcal{A}} \alpha_2$ if for every $\omega \in \mathcal{A}$, $\omega \models_{\mathcal{A}} \alpha_1$ implies $\omega \models_{\mathcal{A}} \alpha_2$. Note that the notion of entailment we consider is classical entailment, modulo the assumption that the propositions in each set A_i are JEPD.

Example 2. *Consider the following attribute domains:*

$$\begin{aligned}
A_1 = \{&\textit{row-house}_1, \textit{semi-detached}_1, \textit{bungalow}, \textit{villa}, \textit{mansion}, \textit{bedsit}, \textit{studio},\\
&\textit{one-bed-ap}, \textit{two-bed-ap}, \textit{three-bed-ap}, \textit{loft}, \textit{penthouse}\}\\
A_2 = \{&\textit{row-house}_2, \textit{semi-detached}_2, \textit{detached}, \textit{apartment}\}\\
A_3 = \{&\textit{very-small}, \textit{small}, \textit{medium}, \textit{large}, \textit{very-large}\}\\
A_4 = \{&\textit{basic}, \textit{comfortable}, \textit{very-comfortable}, \textit{luxurious}, \textit{exclusive}\}
\end{aligned}$$

[1]Throughout this paper, we use the terms *concept* and *property* interchangeably.

where A_1 lists the housing types that are possible in the given context, A_2 provides a coarser description of some of these housing types, and A_3 and A_4 contain the labels that are used to describe housing sizes and comfort levels respectively. Note that subscripts are used for the housing options row-house and semi-detached to ensure that different attribute domains are disjoint. When there is no cause for confusion, we will omit these subscripts. The following set of rules R provides a partial specification of how these attribute domains are related to each other:

$$\begin{aligned}
& \textit{bungalow} \rightarrow \textit{medium} && \textit{bungalow} \rightarrow \textit{detached} && (5)\\
& \textit{mansion} \rightarrow \textit{very-large} && \textit{mansion} \rightarrow \textit{detached} && (6)\\
& \textit{large} \wedge \textit{detached} \rightarrow \textit{lux} && \textit{large} \wedge \textit{row-house} \rightarrow \textit{comf} && (7)\\
& \textit{small} \wedge \textit{detached} \rightarrow \textit{bas} \vee \textit{comf} && \textit{mansion} \rightarrow \textit{excl} && (8)
\end{aligned}$$

where some labels are abbreviated for the ease of presentation. For example ($villa$, $detached$, $large$, lux) is a model of each of these rules.Note that the only conclusions that can be derived from R are more or less trivial, e.g.

$$R \models_{\mathcal{A}} \textit{villa} \rightarrow \textit{very-small} \vee \textit{small} \vee \textit{medium} \vee \textit{large} \vee \textit{very-large} \tag{9}$$

$$R \models_{\mathcal{A}} (\textit{small} \vee \textit{large}) \wedge \textit{detached} \rightarrow \textit{basic} \vee \textit{comfortable} \vee \textit{luxurious} \tag{10}$$

Note that (9) *follows from our assumption that the labels in an attribute domain are exhaustive.*

3.1. Commonsense inference

A *rule base* R over the atoms in $\mathcal{A}$ usually only provides an incomplete specification of how the given attribute domains are related to each other. We are interested in refining the available knowledge in R using a number of generic meta-principles. To this end, we will make use of background knowledge about the conceptual relationship of different formulas, which we assume to be encoded in a set of assertions Σ (to be formalized in Section 6). We are then interested in defining a consequence relation $\vdash$ that extends the entailment relation $\models_{\mathcal{A}}$ (i.e. supraclassicality). Specifically, we assume that the following inference rule is valid:

$$\frac{R \quad \models_{\mathcal{A}} \quad \beta \rightarrow \gamma}{(R, \Sigma) \quad \vdash \quad \beta \rightarrow \gamma} \tag{S}$$

The first meta-principle we consider is that intermediate conditions should lead to intermediate conclusions.

For instance, given that both large and small detached houses have a comfort level that is between *basic* and *luxurious*, we derive that also medium detached houses should have a comfort level between these bounds. More generally, if a propositional formula β is conceptually between the formulas β_1 and β_2, the idea of interpolative inference is that whatever we can derive from β should be conceptually between what we can derive from β_1 and what we can derive from β_2. The exact nature of this conceptual *betweenness* will be formalized in the following sections, but intuitively β is conceptually between β_1 and β_2 if β has all the features that β_1 and β_2 have in common. We could say, for instance, that a bistro is conceptually between a bar and a restaurant, or that a studio is conceptually between a bedsit and an apartment. In ecology, we may consider that *taiga* is between *tundra* and *temperate-forest*. We could take the view that the painting style of Renoir is conceptually between the painting styles of Monet and Manet.

To encode information about betweenness at the syntactic level, we use a modality $\bowtie$, i.e. the formula $\beta_1 \bowtie \beta_2$ is true whenever a situation holds which is conceptually between a situation satisfying β_1 and a situation satisfying β_2 (or alternatively, when talking about concepts, $\beta_1 \bowtie \beta_2$ is true for instances that are between β_1 and β_2). Typically, it will not be possible to have a precise definition of $\beta_1 \bowtie \beta_2$, as, in fact, our logical language may not be rich enough to precisely capture exactly those situations. However, in practice, we may obtain knowledge about upper and lower approximations of $\beta_1 \bowtie \beta_2$. We write $\Sigma \vdash \gamma \rightarrow \beta_1 \bowtie \beta_2$ to denote that everything which satisfies γ is conceptually between β_1 and β_2, and $\Sigma \vdash \beta_1 \bowtie \beta_2 \rightarrow \gamma$ to denote that anything which is conceptually between β_1 and β_2 should definitely satisfy γ.

Example 3. *It is not the case that all lofts are conceptually between a three-bedroom apartment and a penthouse (e.g. a small loft with only one bedroom), so loft $\rightarrow$ three-bed-ap $\bowtie$ penthouse does not hold. However, we do have trivially that*

$$\textit{three-bed-ap} \vee \textit{penthouse} \rightarrow \textit{three-bed-ap} \bowtie \textit{penthouse}$$

Conversely, however, some lofts are between a three-bedroom apartment and a penthouse, so we cannot remove the disjunct loft in the consequent of the *following implication*

$$\textit{three-bed-ap} \bowtie \textit{penthouse} \rightarrow \textit{three-bed-ap} \vee \textit{loft} \vee \textit{penthouse}$$

Note that in the considered domain, there are no apartments with more than three bedrooms (with the possible exception of penthouses), hence three-bed-ap, loft and penthouse exhaustively cover all situations that are between a three-bedroom apartment and a penthouse.

On the other hand, we may consider that all studios are between bedsits and one-bedroom apartments. Under this view, we should be able to derive the following rules from Σ*:*

$$\textit{bedsit} \vee \textit{studio} \vee \textit{one-bed-ap} \rightarrow \textit{bedsit} \bowtie \textit{one-bed-ap}$$
$$\textit{bedsit} \bowtie \textit{one-bed-ap} \rightarrow \textit{bedsit} \vee \textit{studio} \vee \textit{one-bed-ap}$$

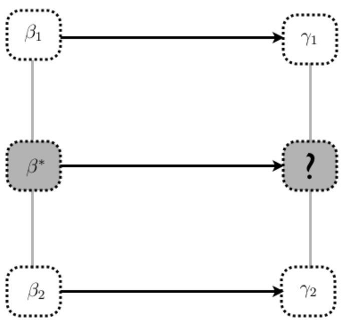
Using the binary modality $\bowtie$, we can define the following interpolative inference rule:

$$\frac{\begin{array}{rcl} (R,\Sigma) & \vdash & \beta_1 \rightarrow \gamma_1 \\ (R,\Sigma) & \vdash & \beta_2 \rightarrow \gamma_2 \\ \Sigma & \vdash & \beta^* \rightarrow \beta_1 \bowtie \beta_2 \\ \Sigma & \vdash & \gamma_1 \bowtie \gamma_2 \rightarrow \gamma^* \end{array}}{(R,\Sigma) \vdash \beta^* \rightarrow \gamma^*} \qquad \textbf{(I)}$$

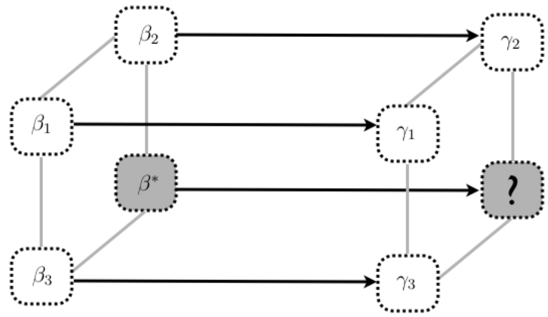
A diagrammatic representation of this interpolation principle is shown in Figure 1(a): given two rules $\beta_1 \rightarrow \gamma_1$ and $\beta_2 \rightarrow \gamma_2$, we derive a rule which applies to a situation β^* which is intermediate between β_1 and β_2. The conclusion γ^* of that rule is required to exhaustively cover all situations that are intermediate between γ_1 and γ_2.

The second meta-principle states that analogous changes in the condition of a rule should lead to analogous changes in the conclusion.

Let us write $\langle \beta_1, \beta_2 \rangle$ to denote the change that is needed to convert a specification compatible with β_1 into a specification compatible with β_2. The intuition will be that $\langle \beta_1, \beta_2 \rangle$ determines a *direction-of-change*. In contrast to betweenness, this notion of direction is not symmetric, e.g. while $\langle \textit{two-bed-ap}, \textit{three-bed-ap} \rangle$ denotes the direction of an increasing number of bedrooms, $\langle \textit{three-bed-ap}, \textit{two-bed-ap} \rangle$ denotes a decreasing number. Given



(a) Interpolative inference



(b) Extrapolative inference

Figure 1: Modelling betweenness and parallelism between regions.

a third formula β_3, we are then interested in those situations that can be obtained by changing a situation compatible with β_3 in the direction specified by $\langle\beta_1, \beta_2\rangle$. In particular, we will write $\beta_3 \triangleright \langle\beta_1, \beta_2\rangle$ for the formula that covers all such situations. For example, we could consider that progressive rock differs from hard rock by having less standard song structures and arrangements, while keeping the same instruments. Then $\textit{heavy-metal} \triangleright \langle\textit{hard-rock}, \textit{prog-rock}\rangle$ would cover all music genres that use heavy metal instruments and timbres, but less standard song structures and arrangements. This would include all progressive metal, as well as some instances of avant-garde metal, among others. In biology, we may consider that the difference between *dog* to *coyote* is analogous to the difference between *cat* and *leopard*, or to the difference between *cat* and *lynx*, which we could encode as $\textit{coyote} \rightarrow \textit{dog} \triangleright \langle\textit{cat}, \textit{leopard}\rangle$ and $\textit{coyote} \rightarrow \textit{dog} \triangleright \langle\textit{cat}, \textit{lynx}\rangle$ respectively. Note in particular that we do not take into account the amount of change: while we may consider that the change from *cat* to *leopard* is bigger than the change from *dog* to *coyote*, the direction of change is the same.

As for betweenness, we will mainly be interested in approximating $\beta_3 \triangleright \langle\beta_1, \beta_2\rangle$ rather than finding an exact definition, i.e. we will be looking for propositional formulas that imply, and that are implied by $\beta_3 \triangleright \langle\beta_1, \beta_2\rangle$.

Example 4. *The change from a bedsit to a studio essentially corresponds to an increase in size and comfort. In this sense, such a change is similar to the change from a two-bedroom apartment to a three-bedroom apartment, or even to a penthouse. We may consider, for instance:*

$$\Sigma \vdash \textit{two-bed-ap} \vee \textit{three-bed-ap} \vee \textit{penthouse} \rightarrow \textit{two-bed-ap} \triangleright \langle\textit{bedsit}, \textit{studio}\rangle$$

$$\Sigma \vdash \textit{two-bed-ap} \rhd \langle \textit{bedsit}, \textit{studio} \rangle \rightarrow \textit{two-bed-ap} \vee \textit{three-bed-ap} \vee \textit{loft} \vee \textit{penthouse}$$

Only the direction of the change is taken into account here, and not the amount of change. For instance, we might have

$$\Sigma \vdash \textit{one-bed-ap} \vee \textit{two-bed-ap} \vee \textit{three-bed-ap} \vee \textit{penthouse}$$
$$\rightarrow \textit{one-bed-ap} \rhd \langle \textit{one-bed-ap}, \textit{two-bed-ap} \rangle$$

Note that the meaning of $\langle ., . \rangle$ by itself cannot be expressed at the syntactic level, i.e. $. \rhd \langle ., . \rangle$ is treated as a ternary modality. Extrapolative inference can then be formalized as follows:

$$\frac{\begin{array}{rcl} (R,\Sigma) & \vdash & \beta_1 \rightarrow \gamma_1 \\ (R,\Sigma) & \vdash & \beta_2 \rightarrow \gamma_2 \\ (R,\Sigma) & \vdash & \beta_3 \rightarrow \gamma_3 \\ \Sigma & \vdash & \beta^* \rightarrow \beta_1 \rhd \langle \beta_2, \beta_3 \rangle \\ \Sigma & \vdash & \gamma_1 \rhd \langle \gamma_2, \gamma_3 \rangle \rightarrow \gamma^* \end{array}}{(R,\Sigma) \vdash \beta^* \rightarrow \gamma^*} \qquad \textbf{(E)}$$

A diagrammatic representation is given in Figure 1(b). In this case, three rules $\beta_1 \rightarrow \gamma_1$, $\beta_2 \rightarrow \gamma_2$ and $\beta_3 \rightarrow \gamma_3$ are available, and we are interested in deriving conclusions about a fourth situation β^* which differs from β_1 as β_3 differs from β_2. The conclusion γ^* which is derived exhaustively covers all situations that differ from γ_1 as γ_3 differs from γ_2.

Finally, we assume that the consequence relation $\vdash$ is deductively closed:

$$\frac{\begin{array}{rcl} (R,\Sigma) & \vdash & \beta_1 \rightarrow \gamma_1 \\ (R,\Sigma) & \vdash & \beta_2 \rightarrow \gamma_2 \\ \{\beta_1 \rightarrow \gamma_1, \beta_2 \rightarrow \gamma_2\} & \models_{\mathcal{A}} & \beta_3 \rightarrow \gamma_3 \end{array}}{(R,\Sigma) \vdash \beta_3 \rightarrow \gamma_3} \qquad \textbf{(D)}$$

Example 5. *Consider the rules from Example 2. Applying* ***(S)****, we immediately have*

$$(R,\Sigma) \vdash \textit{bungalow} \rightarrow \textit{medium} \qquad (R,\Sigma) \vdash \textit{mansion} \rightarrow \textit{very-large}$$

Assuming

$$\Sigma \vdash \textit{bungalow} \vee \textit{villa} \vee \textit{mansion} \rightarrow \textit{bungalow} \bowtie \textit{mansion}$$
$$\Sigma \vdash \textit{medium} \bowtie \textit{very-large} \rightarrow \textit{medium} \vee \textit{large} \vee \textit{very-large}$$

*we find using **(I)** that*

$$(R, \Sigma) \vdash \textit{bungalow} \vee \textit{villa} \vee \textit{mansion} \rightarrow \textit{medium} \vee \textit{large} \vee \textit{very-large}$$

*and using **(D)** that* $(R, \Sigma) \vdash \textit{villa} \rightarrow \textit{medium} \vee \textit{large} \vee \textit{very-large}$ *which refines the trivial conclusion that we obtained in Example 2. Similarly, we find using **(S)** that*

$$\begin{aligned}
(R, \Sigma) &\vdash \textit{large} \wedge \textit{detached} \rightarrow \textit{lux} \\
(R, \Sigma) &\vdash \textit{large} \wedge \textit{row-house} \rightarrow \textit{comf} \\
(R, \Sigma) &\vdash \textit{small} \wedge \textit{detached} \rightarrow \textit{bas} \vee \textit{comf}
\end{aligned}$$

If we now assume that (writing det *for* detached*)*

$$\begin{aligned}
\Sigma \vdash\ &(\textit{large} \vee \textit{medium} \vee \textit{small} \vee \textit{very-small}) \wedge \textit{row-house} \\
&\quad \rightarrow (\textit{large} \wedge \textit{row-house}) \triangleright \langle \textit{large} \wedge \textit{det}, \textit{small} \wedge \textit{det} \rangle \\
\Sigma \vdash\ &\textit{comf} \triangleright \langle \textit{lux}, \textit{bas} \vee \textit{comf} \rangle \rightarrow \textit{bas} \vee \textit{comf}
\end{aligned}$$

*then **(E)** yields*

$$(R, \Sigma) \vdash (\textit{large} \vee \textit{medium} \vee \textit{small} \vee \textit{very-small}) \wedge \textit{row-house} \rightarrow \textit{bas} \vee \textit{comf} \tag{11}$$

*from which we obtain using **(D)** that* $(R, \Sigma) \vdash \textit{medium} \wedge \textit{row-house} \rightarrow \textit{bas} \vee \textit{comf}$*. Indeed, the rule in* (11) *plays the role of* $\beta_1 \rightarrow \gamma_1$ *from the definition of **(D)**, whereas the rule* $\beta_2 \rightarrow \gamma_2$ *is trivial.*

At this point it may not be clear whether the inference relation defined by **(S)**, **(I)**, **(E)** and **(D)** always behaves according to intuition, nor what the implications are at the semantic level of adopting **(I)** and **(E)**. To this end, in the following sections, we will develop a semantic counterpart of the inference relation $\vdash$, which will clarify the nature of the modalities $\bowtie$ and $\triangleright\langle , \rangle$ and will allow us to implement decision procedures for reasoning tasks of interest. A crucial issue that will be discussed in detail is the interaction between the aforemetioned modalities on the one hand, and logical conjunction on the other hand, e.g. discussing under which conditions $(\alpha_1 \wedge \beta_1) \bowtie (\alpha_2 \wedge \beta_2)$ is equivalent to $(\alpha_1 \bowtie \alpha_2) \wedge (\beta_1 \bowtie \beta_2)$.

The principles **(I)** and **(E)** allow us to refine a rule base R by exploiting background knowledge about the conceptual relationship of labels from the same attribute domain. This background knowledge is of a qualitative nature. Extrapolative reasoning, for instance, is based on the idea of direction of change, but does not take the amount of change into account. An analogical proportion such as $a : b :: c : d$, on the other hand, not only expresses that the change from a to b goes in the same direction as the change from c to d, but also that the amount of change between a and b is the same as the amount of change between c and d [40]. To take information about the amount of change into account, and thus generalize analogical reasoning — making it also more cautious when necessary — we will use expressions such as $\beta_1 \rhd_{[\lambda,\mu]} \langle \beta_2, \beta_3 \rangle$ where $0 \leq \lambda \leq \mu < +\infty$. Intuitively, this formula covers all situations that can be obtained by changing β_1 in the same direction as the change from β_2 to β_3, such that the amount of change is between λ and μ times as large as the amount of change between β_2 and β_3. This allows us to make inferences which are more precise in cases where suitable information about the amount of change is available. For example, knowing that the amount of change between *dog* and *coyote* is approximately the same as the amount of change between *cat* and *lynx*, the latter relationship is more useful than the relationship between *cat* and *leopard* if we want to derive knowledge about coyotes from knowledge about dogs.

The most straightforward use of this generalization is to express whether the amount of change from β_1 should be smaller, equal, or larger than the amount of change from β_2 to β_3. In particular, $\beta_1 \rhd_{[1,1]} \langle \beta_2, \beta_3 \rangle$ corresponds to the solution X that makes $\beta_1 : X :: \beta_2 : \beta_3$ a perfect analogical proportion, while $\beta_1 \rhd_{[0,1[} \langle \beta_2, \beta_3 \rangle$ and $\beta_1 \rhd_{]1,+\infty[} \langle \beta_2, \beta_3 \rangle$ express amounts of change that are smaller and larger, respectively, than the amount of change from β_2 to β_3. Note that $\beta_1 \rhd_{[0,+\infty[} \langle \beta_2, \beta_3 \rangle$ corresponds to $\beta_1 \rhd \langle \beta_2, \beta_3 \rangle$. Also note that in the aforementioned cases, the approach remains entirely qualitative. Other choices for the intervals $[\lambda, \mu]$ may give the approach a more numerical flavour, and would mainly be useful in scenarios where the conceptual relationships are obtained using data-driven techniques.

For τ a non-empty subset of $[0, +\infty[$, Principle **(E)** can be refined to

$$\frac{\begin{array}{rcl} (R, \Sigma) & \vdash & \beta_1 \rightarrow \gamma_1 \\ (R, \Sigma) & \vdash & \beta_2 \rightarrow \gamma_2 \\ (R, \Sigma) & \vdash & \beta_3 \rightarrow \gamma_3 \\ \Sigma & \vdash & \beta^* \rightarrow \beta_1 \rhd_\tau \langle \beta_2, \beta_3 \rangle \\ \Sigma & \vdash & \gamma_1 \rhd_\tau \langle \gamma_2, \gamma_3 \rangle \rightarrow \gamma^* \end{array}}{(R, \Sigma) \vdash \beta^* \rightarrow \gamma^*} \tag{E'}$$

Along similar lines, for σ a non-empty subset of $[0, 1]$, we consider expressions of the form $\beta_1 \bowtie_\sigma \beta_2$ to put constraints on the relative closeness to β_1 and β_2. Specifically, $\beta_1 \bowtie_{[0,0.5[} \beta_2$ corresponds to those situations between β_1 and β_2 that are closer to β_1 than to β_2. Note that $\beta_1 \bowtie_{[0,0.5[} \beta_2$ is a refinement of the construct $\beta_1 \leftleftarrows \beta_2$ from CSL [29], as betweenness is not required in CSL, only comparative closeness. Similarly, $\beta_1 \bowtie_{]0.5,1]} \beta_2$ corresponds to situations that are closer to β_2 than to β_1, and $\beta_1 \bowtie_{[0.5,0.5]} \beta_2$ to situations that are exactly halfway. When data-driven techniques are used, other intervals may again be useful. In scenarios where labels can be assumed to be equidistant, we may also know e.g. that *small* $\rightarrow$ *very-small* $\bowtie_{[0.25,0.25]}$ *very-large* and *large* $\rightarrow$ *very-small* $\bowtie_{[0.75,0.75]}$ *very-large*, where the idea is that *medium* is halfway between *very-small* and *very-large* and *small* is halfway between *very-small* and *medium*.

Principle **(I)** can be refined to

$$\frac{\begin{array}{rcl} (R, \Sigma) & \vdash & \beta_1 \rightarrow \gamma_1 \\ (R, \Sigma) & \vdash & \beta_2 \rightarrow \gamma_2 \\ \Sigma & \vdash & \beta^* \rightarrow \beta_1 \bowtie_\sigma \beta_2 \\ \Sigma & \vdash & \gamma_1 \bowtie_\sigma \gamma_2 \rightarrow \gamma^* \end{array}}{(R, \Sigma) \vdash \beta^* \rightarrow \gamma^*} \tag{I'}$$

Example 6. *Consider again the setting of Example 5, and assume that a villa is conceptually halfway between a bungalow and a mansion, and that a large house is conceptually halfway between a medium house and a very-large house. We then get*

$$\Sigma \vdash \textit{villa} \rightarrow \textit{bungalow} \bowtie_{[0.5,0.5]} \textit{mansion} \tag{12}$$

$$\Sigma \vdash \textit{medium} \bowtie_{[0.5,0.5]} \textit{very-large} \rightarrow \textit{large} \tag{13}$$

which allows us the obtain the refined conclusion $(R, \Sigma) \vdash$ *villa* $\rightarrow$ *large using* ***(I')*** *and* ***(D)****. Alternatively, we could assume that a villa is conceptually closer to a bungalow than to a mansion, by assuming that*

$$\Sigma \vdash \textit{bungalow} \vee \textit{villa} \rightarrow \textit{bungalow} \bowtie_{[0,0.5[} \textit{mansion}$$

Together with

$$\Sigma \vdash \textit{medium} \bowtie_{[0,0.5[} \textit{very-large} \rightarrow \textit{medium} \vee \textit{large}$$

we would then find $R \vdash$ *villa* $\rightarrow$ *medium* $\vee$ *large.*

3.2. Mappings between attribute domains

At the semantic level, a rule base R can be seen as a mapping between sets of vectors of attributes. In particular, assume that all the labels in the antecedents of the rules in R belong to the attribute domains $B_1, ..., B_s$ and that the labels in the consequents belong to $C_1, ..., C_k$. The rule base R can then equivalently be expressed as a function f_R from subsets of $\mathcal{B} = B_1 \times ... \times B_s$ to subsets of $\mathcal{C} = C_1 \times ... \times C_k$, defined for $X \subseteq \mathcal{B}$ as

$$f_R(X) = \bigcap \{Y \in 2^{\mathcal{C}} \mid R \models_{\mathcal{A}} (\bigvee_{(x_1,...,x_s) \in X} \bigwedge_{i=1}^{s} x_i) \rightarrow (\bigvee_{(y_1,...,y_k) \in Y} \bigwedge_{i=1}^{k} y_i)\}$$

where $\mathcal{A} = A_1 \times ... \times A_n$ and $\{A_1, ..., A_n\} = \{B_1, ..., B_s\} \cup \{C_1, ..., C_k\}$ as before. It is not hard to see that this function indeed expresses the same knowledge as the rule base R. Furthermore, there exists a single $Y^* \in 2^{\mathcal{C}}$ such that $f_R(X) = Y^*$ and

$$R \models_{\mathcal{A}} (\bigvee_{(x_1,...,x_s) \in X} \bigwedge_{i=1}^{s} x_i) \rightarrow (\bigvee_{(y_1,...,y_k) \in Y^*} \bigwedge_{i=1}^{k} y_i)$$

Example 7. *Consider again the rules base from Example 2. We have that* $\mathcal{B} = A_1 \times A_2 \times A_3$*, as no rule refers to comfort levels in its antecedent, and* $\mathcal{C} = A_2 \times A_3 \times A_4$*, as only the coarser housing types of* A_2 *are referred to in the consequent of rules. We find that a small detached villa is either basic or comfortable:*

$$f_R(\{(\text{villa}, \text{det}, \text{small})\}) = \{(\text{det}, \text{small}, \text{bas}), (\text{det}, \text{small}, \text{comf})\}$$

Indeed, from

$$small \wedge detached \rightarrow bas \vee comf$$

It follows that

$$small \wedge detached \wedge villa \rightarrow (bas \vee comf) \wedge small \wedge detached$$

from which we can already conclude

$$f_R(\{(villa, det, small)\}) \subseteq \{(det, small, bas), (det, small, comf)\}$$

It is furthermore clear that there are no rules in R *which could be used to further refine* $f_R(\{(villa, det, small)\})$*.*

Similarly, we find that a bungalow is detached and medium-sized, while we find no restrictions on the possible comfort levels:

$$f_R(\{(bun, x, y) \mid x \in A_2, y \in A_3\}) = \{(det, medium, z) \mid z \in A_4\}$$

We may see f_R as an approximate (i.e. incomplete) model of a given domain, which may be refined as soon as new information becomes available. In particular, for two $2^{\mathcal{B}} \rightarrow 2^{\mathcal{C}}$ functions f and f' which are monotone w.r.t. set inclusion, we say that f is a refinement of f', written $f \leq f'$, iff

$$\forall X \subseteq \mathcal{B} \,.\, f(X) \subseteq f'(X) \tag{14}$$

This idea of using monotone set-valued functions to describe approximate models is closely linked to the theory of Scott domains; see e.g. [41] for an elaboration of this idea. At the semantic level, completing the rule base R boils down to refining the corresponding function f_R. In our approach, such refinements will be based on meta-knowledge about the nature of the relationship between $\mathcal{B}$ and $\mathcal{C}$. This will lead us to replace f_R by the largest refinement, w.r.t. $\leq$, which is compatible with the imposed meta-knowledge. In particular, as will become clear below, Principles **(I)** and **(E)** amount to refine f_R by imposing some form of *monotonicity*, whereas **(I')** and **(E')** amount to refine f_R by imposing a form of *linearity*. As could be expected, the meta-knowledge underlying **(I')** and **(E')** is stronger than the meta-knowledge underlying **(I)** and **(E)**.

4. Formalization using conceptual spaces

The approach sketched in Section 3 requires information about how the labels of an attribute domain are conceptually related to each other. Provided that this information is available, inference can be carried out purely at the symbolic level. However, to justify the inference procedure, and to provide an adequate semantics for it, we need to be precise on how relationships such as betweenness should be interpreted. To this end, we assume that the cognitive meaning of the attributes can be represented geometrically, as convex regions in a conceptual space. Conceptual relationships can then be given a clear spatial interpretation, as will be discussed in Section 4.1. Taking advantage of this link with conceptual spaces, Section 4.2 subsequently reviews some opportunities for the acquisition of conceptual relations. Finally, Section 4.3 presents the idea of regular mappings between conceptual spaces as a basis for interpolative and extrapolative reasoning.

4.1. Geometric modelling of attribute domains

The theory of conceptual spaces [13] is centered around the assumption that the meaning of a natural property can be adequately modelled as a convex region in some geometric space. Formally, a conceptual space is the Cartesian product $Q_1 \times ... \times Q_m$ of quality dimensions, each of which corresponds to an atomic, cognitively meaningful feature, called a quality. A standard example is the conceptual space of colors, which can be described using the quality dimensions hue, saturation and intensity. Labels to describe colors, in some natural language, are then posited to correspond to convex regions in this conceptual space, a view which is closely related to the ideas of prototype theory [42]. The label *red* for instance will be represented by the set of points whose hue is in the spectrum normally associated with red, whose saturation is sufficiently high, and whose intensity is neither too high nor too low. Note that while e.g. *red* may be an atomic property at the symbolic level, at the cognitive level it is defined in terms of more primitive notions. Quality dimensions may be continuous or discrete, and can even be finite. In practice, however, it is common to identify conceptual spaces with Euclidean spaces [42], and to define cognitive similarity in terms of Euclidean distance. We will also adopt this simplifying view throughout the paper, although part of the discussion readily generalizes to more general spaces[2].

Now consider again the example of housing types. A conceptual space to represent housing types would have a large number of dimensions, relating to shape, size, colour, texture, etc. Each house that exists in the world will correspond to one specific point in this conceptual space. Conversely, however, there may be points in that conceptual space which do not correspond to structures that can be physically realized, or that would not be recognized as houses (e.g. a building structure of 20 km long and 1 cm wide). Each attribute from the domain A_1 corresponds to a convex region in the conceptual space, where intuitively, e.g. the region corresponding to *villa* corresponds to those building structures that are more similar to prototypical villas than to prototypical instances of the other attributes in A_1.

In general, each attribute domain A_i thus corresponds to a partition of some conceptual space in convex regions, where each attribute of A_i corresponds to a partition class. As some regions of conceptual spaces may correspond to types of instances that do not exist in the real world, the number of partition classes may, in principle be higher than the number of attributes in A_i. We write $reg(x)$ to denote the convex region that corresponds to label x. For example, reg(*three-bed-ap*) represents all sections of building structures that could be classified as three-bedroom apartments. This set will contain both points that correspond to actual apartments (which exist somewhere in the world) and possible apartments (which may in principle be built one day).

This conceptual space representation of the attributes is considerably richer than what can be described at the symbolic level, and therefore also allows for richer forms of inference. However, in most application domains, it is not reasonable to assume that such representations are available. Moreover, the precise representation of a property in a conceptual space strongly depends on the considered context and may be subjective. Usually, however, it is assumed that a particular conceptual space representation can be obtained from a generic representation by appropriately rescaling the quality dimensions [44]. Note that this observation implies that also similarity judgements may differ across contexts and people, as rescaling the dimensions may

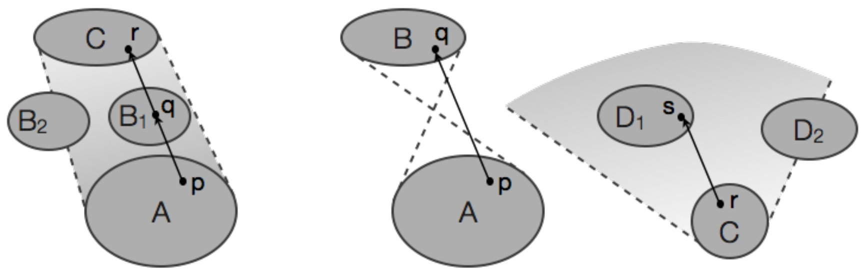
[2] In particular, interpolative reasoning can be carried out w.r.t. any space for which betweenness can meaningfully be defined. See e.g. [43] for a formalization of conceptual spaces in terms of a primitive betweenness relation.

influence the relative Euclidean distance between points. For instance, while apples are usually considered to be closer to tomatoes than to chocolate, in the context of desserts, they may be closer to chocolate (e.g. because both can be used in cakes). To alleviate these issues, we will rely on (mainly) qualitative knowledge about the spatial relationship of different properties. Such qualitative knowledge may be easier to obtain, and because the spatial relations that will be considered are invariant under affine transformations (such as rescaling the quality dimensions), they are more robust against changes in context and person. In particular, such relations are not affected by rescaling of the quality dimensions, although they would still depend on context changes that introduce additional quality dimensions.

For each attribute domain A_i, we assume that information is available about *betweenness* and *parallelism* of the conceptual space representation of its attributes. For example, we may intuitively think of a *studio* to be between a *bedsit* and a *one-bedroom apartment*. In the domain of music genres, we may consider that the change from *hard-rock* to *progressive-rock* is parallel to the change from *heavy-metal* to *progressive-metal*, and that *progressive-rock* is between *hard-rock* and *avant-garde*. The notions of betweenness and parallelism are straightforwardly defined for points in Euclidean spaces. In particular, let us write $bet(p, q, r)$ to denote that q lies between p and r (on the same line), and $par(p, q, r, s)$ to denote that the vectors $\overrightarrow{pq}$ and $\overrightarrow{rs}$ point in the same direction. The fact that point q is between points p and r means that for every point x it holds that $d(q, x) \leq \max(d(p, x), d(r, x))$, and in particular, that whenever p and r are close to a prototype of some concept, then q is close to it as well. In this sense, we may see $bet(p, q, r)$ as a way to express that whatever natural properties p and r have in common, p and q have them in common as well (identifying points in a conceptual space with instances). On the other hand, $par(p, q, r, s)$ intuitively means that to arrive at s, r needs to be changed in the same way as p needs to be changed to arrive at q, i.e. at the qualitative level, p is to q as r is to s (although the amount of change may be different).

The notions of betweenness and parallelism, which are defined for points, need to be extended to regions, in order to describe relationships between attributes. As is well known, this can be done in different ways [45]. We will consider the following two notions of betweenness for regions A, B, and C (in a given Euclidean space):

$$\overline{bet}(A, B, C) \quad \text{iff} \quad \exists q \in B \,.\, \exists p \in A \,.\, \exists r \in C \,.\, bet(p, q, r) \tag{15}$$



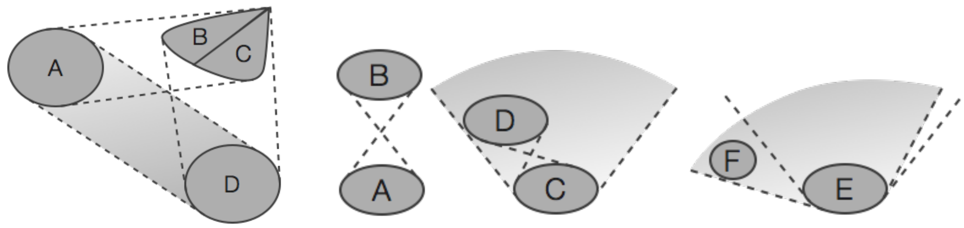
(a) Betweenness (b) Parallelism

Figure 2: Modelling betweenness and parallelism between regions.

$$\underline{bet}(A, B, C) \quad \text{iff} \quad \forall q \in B \,.\, \exists p \in A \,.\, \exists r \in C \,.\, bet(p, q, r) \tag{16}$$

In particular, if A and C are convex regions, $\overline{bet}(A, B, C)$ holds if B overlaps with the convex hull of $A \cup C$, whereas $\underline{bet}(A, B, C)$ holds if B is included in this convex hull. These two notions of betweenness are illustrated in Figure 2(a), where $\underline{bet}(A, B_1, C)$, $\overline{bet}(A, B_1, C)$ and $\overline{bet}(A, B_2, C)$ hold, but not $\underline{bet}(A, B_2, C)$. Note in particular that both relations are reflexive w.r.t. the first two arguments, in the sense that e.g. $\overline{bet}(A, A, C)$ holds, as well as symmetric, in the sense that e.g. $\overline{bet}(A, B, C) \equiv \overline{bet}(C, B, A)$. However, transitivity does not necessarily hold for regions, e.g. from $\underline{bet}(A, B, C)$ and $\underline{bet}(B, C, D)$ we cannot infer that $\underline{bet}(A, B, D)$; a counterexample is depicted in Figure 3(a)[3]. In the terminology of rough set theory [46], $\overline{bet}$ and $\underline{bet}$ correspond to upper and lower approximations of betweenness. The underlying idea is that, given our finite set of labels, we may not be able to exactly describe the convex hull of two regions A and B. All we can do, then, is to list all labels which are completely included in the convex hull (i.e. define the lower approximation of the convex hull), and all labels which have a non-empty intersection with the convex hull (i.e. define the upper approximation

[3]This counterexample also illustrates a technical subtlety of the considered framework. If we want to represent the meaning of each label as a topologically closed set, then regions will inevitably share their boundary with other regions, which is not compatible with the view that different labels (of the same attribute domain) refer to pairwise disjoint properties. One solution to this problem is to associate with each label a topologically open region, and introduce topologically closed regions that correspond to borderline instances, i.e. instances for which it is hard to tell whether they belong to one concept or to another. In Figure 3(a), nothing prevents us from taking B and C to be topologically open regions.



(a) We have $\underline{bet}(A, B, C)$ and $\underline{bet}(B, C, D)$ but not $\underline{bet}(A, B, D)$.

(b) We have $\underline{par}(A, B, C; D)$ and $\underline{par}(C, D, E; F)$ but not $\underline{par}(A, B, E; F)$.

Figure 3: The relations $\underline{bet}$ and $\underline{par}$ are not transitive.

of the convex hull). In the following, for the ease of presentation we will often identify labels with the corresponding regions, writing e.g. $\underline{bet}(a, b, c)$ for $\underline{bet}(reg(a), reg(b), reg(c))$.

Example 8. *In the domain of housing types, we may consider that we have* $\overline{bet}(\textit{three-bed-ap}, \textit{loft}, \textit{penthouse})$ *but not* $\underline{bet}(\textit{three-bed-ap}, \textit{loft}, \textit{penthouse})$*. Note that this corresponds to what was expressed at the syntactic level in Example 3.*

For regions A, B, C and D (in a given Euclidean space), two notions of parallelism will be considered:

$$\overline{par}(A, B, C; D) \quad \text{iff} \quad \exists s \in D \,.\, \exists p \in A, \exists q \in B, \exists r \in C \,.\, \exists \lambda \geq 0 \,.\, \overrightarrow{rs} = \lambda \cdot \overrightarrow{pq}$$
$$\underline{par}(A, B, C; D) \quad \text{iff} \quad \forall s \in D \,.\, \exists p \in A, \exists q \in B, \exists r \in C \,.\, \exists \lambda \geq 0 \,.\, \overrightarrow{rs} = \lambda \cdot \overrightarrow{pq}$$

For parallelism, the role of the convex hull is replaced by a notion of conical extension, which is illustrated in Figure 2(b). In particular, if $\overline{par}(A, B, C; D)$ holds, some point in D differs from some point in C in the same direction that some point in B differs from some point in A. In the scenario of Figure 2(b), this means that D overlaps with the shaded area. Likewise, when $\underline{par}(A, B, C; D)$ holds, D is included in the shaded area of Figure 2(b). For all regions A, B and C, we have that $\underline{par}(A, B, C; C)$ holds (and thus also $\overline{par}(A, B, C; C)$), as well as $\underline{par}(A, B, A; B)$. However, as for betweenness, transitivity does not hold, as is illustrated in Figure 3(b).

Example 9. *We may consider that the following relations hold (cf. Example*

4)

$$\overline{par}(\textit{bedsit}, \textit{studio}, \textit{two-bed-ap}; \textit{loft})$$
$$\underline{par}(\textit{bedsit}, \textit{studio}, \textit{two-bed-ap}; \textit{three-bed-ap})$$

In other words, for every point s in reg(three-bed) we can find points $p \in$ reg(bedsit), $q \in$ reg(studio) and $r \in$ reg(two-bed-ap) such that $\overrightarrow{pq}$ is parallel to $\overrightarrow{rs}$. In other words, every three-bedroom apartment differs from some two-bedroom apartment in (qualitatively) the same way as some studio differs from some bedsit. Similarly, we have that some (but not all) lofts differ from some two-bedroom apartment in the same way that some studio differs from some bedsit.

As a refinement of $bet(p, q, r)$, we may put constraints on the distance ratio $\frac{d(p,q)}{d(p,r)}$, where d denotes the Euclidean distance. Note that distance ratios for points on the same line are preserved under affine transformations. In particular, for σ a subset of $[0, 1]$, we consider the following relations:

$$\overline{bet_\sigma}(A, B, C) \quad \text{iff} \quad \exists q \in B \,.\, \exists p \in A \,.\, \exists r \in C \,.\, bet(p, q, r) \wedge (\exists \lambda \in \sigma \,.\, d(p, q) = \lambda \cdot d(p, r)) \tag{17}$$

$$\underline{bet_\sigma}(A, B, C) \quad \text{iff} \quad \forall q \in B \,.\, \exists p \in A \,.\, \exists r \in C \,.\, bet(p, q, r) \wedge (\exists \lambda \in \sigma \,.\, d(p, q) = \lambda \cdot d(p, r)) \tag{18}$$

Clearly, $\overline{bet} = \overline{bet_{[0,1]}}$ and $\underline{bet} = \underline{bet_{[0,1]}}$. Other notable cases are $\sigma = [0.5, 0.5]$, when (17)–(18) express that B is halfway between A and C, $\sigma =]0.5, 1]$, when (17)–(18) express that B is closer to C than to A, and $\sigma = [0, 0.5[$, when (17)–(18) express that B is closer to A than to C. Note that we can only have $\underline{bet_{[0,0]}}(reg(a), reg(b), reg(c))$ if $a = b$ since labels from the same attribute domain are assumed to be disjoint. A similar refinement of parallelism is possible, where for τ a subset of $[0, +\infty[$, we define

$$\overline{par_\tau}(A, B, C; D) \quad \text{iff} \quad \exists s \in D \,.\, \exists p \in A, \exists q \in B, \exists r \in C \,.\, \exists \lambda \in \tau \,.\, \overrightarrow{rs} = \lambda \cdot \overrightarrow{pq} \tag{19}$$

$$\underline{par_\tau}(A, B, C; D) \quad \text{iff} \quad \forall s \in D \,.\, \exists p \in A, \exists q \in B, \exists r \in C \,.\, \exists \lambda \in \tau \,.\, \overrightarrow{rs} = \lambda \cdot \overrightarrow{pq} \tag{20}$$

In the particular case where $\tau = [1, 1]$, (19)-(20) extend the idea of a parallelogram, whose relationship to the idea of analogical proportion is well known.

4.2. Acquiring conceptual relations

Regarding the applicability of our approach, an important question is how the required relational knowledge about conceptual spaces can be obtained, e.g. how do we find out that a *studio* is conceptually between a *bedsit* and a *one-bedroom apartment*? Depending on the specific application, different options may be available.

Manual encoding

In some domains, it is feasible to manually encode a complete qualitative description of a conceptual space. Most notably, this is the case for conceptual spaces that are unidimensional, for which it suffices to provide a ranking of the labels of interest. For instance, a conceptual space of housing sizes may be described by encoding that

$$\textit{very-small} < \textit{small} < \textit{medium} < \textit{large} < \textit{very-large}$$

From this description, we immediately obtain that, for example, the relations $\underline{bet}(\textit{very-small}, \textit{medium}, \textit{large})$ and $\underline{par}(\textit{very-small}, \textit{small}, \textit{medium}; \textit{very-large})$ hold. Note that in unidimensional domains, the relations $\underline{bet}$ and $\overline{bet}$ coincide when applied to labels of the same attribute domain (as these correspond to disjoint intervals), as do the relations $\underline{par}$ and $\overline{par}$. In multi-dimensional domains, it may still be the case that providing a qualitative description is mainly a matter of ranking. The qualitative description of such simple multi-dimensional conceptual spaces can easily be modelled using a diagram. Figure 4 provides an example of such a diagram, where lines define tuples that satisfy $\underline{bet}$ and parallel lines define tuples that satisfy $\underline{par}$. For example, in the case of Figure 4, we have that

$$\underline{par}(\textit{2-bed-ap}, \textit{3-bed-ap}, \textit{2-bed-rowhouse}; \textit{3-bed-rowhouse})$$

As the formalization of such diagrams is straightforward, we will not discuss it in detail. However, it should be clear that a diagrammatic representation can only be obtained when some simplifying conditions are met. For example, in Figure 4, it is tacitly assumed that $\underline{bet} = \overline{bet}$ and $\underline{par} = \overline{par}$, and moreover that many transitivity and mixed transitivity properties are assumed for $\underline{bet}$ and $\underline{par}$ that do not hold in general (e.g. $\underline{bet}(a, b, c) \wedge \underline{par}(b, c, d; e) \Rightarrow \underline{par}(a, c, d; e)$). While such diagrams can therefore not account for the full generality offered by the relations $\underline{bet}$, $\overline{bet}$, $\underline{par}$ and $\overline{par}$, in simple domains, as the housing example of Figure 4, they offer a convenient way of modelling our intuitions about the conceptual relationships that hold.

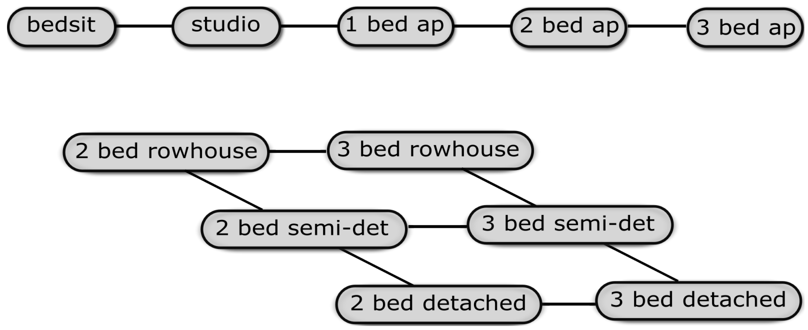


Figure 4: Qualitative description of a conceptual space of housing types.

Natural language processing

A second possibility is to extract conceptual relations from natural language. In [47], for instance, the idea of latent relational analysis was introduced, with the aim of identifying analogical proportions. As instances of $\underline{par_{[1,1]}}$ correspond to analogical proportions, conceptual relations of this particular type can be obtained in the same way. The main idea is that two pairs of words are likely to be related analogously, i.e. form an analogical proportion, when the lexical contexts in which they co-occur are similar. For example, the words *kitten* and *cat* are found in sentences such as "a kitten is a young cat", while the words *chick* and *chicken* are found in sentences such as "a chick is a young chicken". From such observations, the analogical proportion *kitten* : *cat* :: *chick* : *chicken* can be discovered. Another technique for discovering analogical proportions from the web was proposed in [48], estimating the strength of analogical proportions by converting co-occurrence statistics using Kolmogorov information theory. Instances of $\underline{bet_{[0.5,0.5]}}(a, b, c)$ could be identified with analogical proportions of the form $a : b :: b : c$. In principle, instances of $\underline{bet}$ and $\overline{bet}$ can be discovered by applying general methods for extracting ternary relations from text [49]. However, it is clear that the use of information extraction techniques to define the relations $\underline{bet}$, $\overline{bet}$, $\underline{par}$ and $\overline{par}$ will necessarily by highly heuristic, due to the inherent imperfection of such methods.

Data-driven techniques

If sufficient information is available about instances of concepts or properties, several data-driven approaches can be used, which directly take advantage of the geometric nature of the relations of interest. For instance, [42] suggests to start from pairwise similarity judgements between instances, and use multi-dimensional scaling to obtain coordinates for them in a Euclidean space. Representations of concepts can then be obtained by determining the corresponding Voronoi tessellation, after which the conceptual relations of interest can be evaluated by straightforward geometric calculations. The number of dimensions of the resulting space can be freely chosen, where fewer dimensions lead to more conceptual relationships, as only the most prominent dimensions of the conceptual space are then taken into account. Thus, the lower the number of dimensions, the less cautious the resulting inference mechanism will be. In [50], the feasibility of such an approach was demonstrated in the domain of music genres, using similarity judgements that were obtained indirectly using user-contributed meta-data from the website last.fm[4].

Rather than starting from similarity judgements, [14] suggests an approach based on singular value decomposition (SVD), which is a form of dimensionality reduction. Translated to our setting, the approach would start from a matrix where rows correspond to instances and columns correspond to binary features that these instances may or may not have. Instances are then represented in a high-dimensional space with one dimension for each feature, and coordinates are either 0 or 1, depending on whether the instance has the corresponding feature. Using SVD, a linear transformation is then determined which maps this high-dimensional space onto a space of lower dimension, with real-valued coordinates. As before, the resulting representations of the instance can be used to generate geometric representations of the concepts. Again we have that the chosen number of dimensions determines to what extent all quality dimensions, or only the most prominent ones are taken into account. Note that this latter approach offers an interesting connection between representations of concepts as sets of features and geometric representations, which also allows us to make the relationship between analogical proportions and parallelism explicit. In particular, assume that the analogical proportion $a : b :: c : d$ holds for the instances a, b, c and d, and let their sets of features be denoted by A, B, C and D respectively. According to the formal definition of analogical proportion [40, 32], we have that $A \setminus B = C \setminus D$ and $B \setminus A = D \setminus C$ (where we write $\setminus$ for set difference). From this observation, it is easy to see that the representations of a, b, c and d form

[4] `http://www.last.fm`

a parallelogram in the initial, high-dimensional $\{0,1\}$-valued space. Since, parallelograms are preserved under linear transformations, a, b, c and d will still form a parallelogram in the resulting conceptual spaces representation. Note that in principle, due to the lower number of dimensions, the conceptual space may contain more quadruples of instances that form a parallelogram than the initial space. These parallelograms intuitively correspond to pairs of instances that are analogical in all *relevant* aspects.

Alternatively, the SVD approach can also be applied when feature representations of concepts, rather than instances, are available. The concepts are then represented as points, and betweenness and parallelism of these points may be taken to be indicative of betweenness and parallelism of the unknown representations of the concepts as regions.

Example 10. *Consider a rule base about red wines. From a web page about red wines*[5] *we learn the following knowledge*

$$\begin{aligned}
\textit{beaujolais} &\rightarrow \textit{low-tan} \wedge \textit{light-body} \\
\textit{inexpensive} \wedge \textit{burgundy} &\rightarrow \textit{low-tan} \wedge \textit{light-body} \\
\textit{bardolino} &\rightarrow \textit{low-tan} \wedge \textit{light-body} \\
\textit{valpolicella} &\rightarrow \textit{low-tan} \wedge \textit{light-body} \\
\textit{inexpensive} \wedge \textit{bordeaux} &\rightarrow \textit{low-tan} \wedge \textit{med-body} \\
\textit{chianti} &\rightarrow \textit{low-tan} \wedge \textit{med-body} \\
\textit{rioja} &\rightarrow \textit{low-tan} \wedge \textit{med-body} \\
\textit{merlot} &\rightarrow (\textit{low-tan} \vee \textit{mid-tan}) \wedge \textit{med-body} \\
\textit{mid-range} \wedge \textit{burgundy} &\rightarrow \textit{low-tan} \wedge \textit{med-body} \\
\textit{inexpensive} \wedge \textit{cabernet-sauvignon} &\rightarrow \textit{low-tan} \wedge \textit{med-body} \\
\textit{mid-range} \wedge \textit{bordeaux} &\rightarrow \textit{mid-tan} \wedge \textit{med-body} \\
\textit{above-average} \wedge \textit{cabernet-sauvignon} &\rightarrow \textit{mid-tan} \wedge \textit{med-body} \\
\textit{zinfandel} &\rightarrow (\textit{mid-tan} \vee \textit{high-tan}) \wedge (\textit{med-body} \vee \textit{full-body}) \\
\textit{shiraz} &\rightarrow (\textit{mid-tan} \vee \textit{high-tan}) \wedge (\textit{med-body} \vee \textit{full-body}) \\
\textit{high-end} \wedge \textit{bordeaux} &\rightarrow \textit{high-tan} \wedge \textit{full-body} \\
\textit{high-end} \wedge \textit{burgundy} &\rightarrow \textit{high-tan} \wedge \textit{full-body} \\
\textit{high-end} \wedge \textit{cabernet-sauvignon} &\rightarrow \textit{high-tan} \wedge \textit{full-body}
\end{aligned}$$

[5] `http://randomnetstuff.com/wineplace/redwine.php`, accessed July 28th, 2011.

$$
\begin{aligned}
\textit{barolo} &\rightarrow \textit{high-tan} \wedge \textit{full-body} \\
\textit{barbaresco} &\rightarrow \textit{high-tan} \wedge \textit{full-body} \\
\textit{brunello-di-montalcino} &\rightarrow \textit{high-tan} \wedge \textit{full-body} \\
\textit{low-tan} \wedge \textit{light-body} &\rightarrow \textit{light-red} \\
\textit{low-tan} \wedge \textit{med-body} &\rightarrow \textit{med-red} \\
\textit{mid-tan} \wedge \textit{med-body} &\rightarrow \textit{dark-red} \vee \textit{opaque} \\
\textit{high-tan} \wedge \textit{full-body} &\rightarrow \textit{opaque}
\end{aligned}
$$

Given the large number of wine types that exist, inevitably some types are not contained in this rule base. Let us consider, as an example, the wines Barbera and Bandol. From another web page about wine[6]*, we extract information about wine–food pairings. For each listed wine type, we create a vector with one component for every listed food type. This component is set to 1 if the wine is mentioned as appropriate for the corresponding food type, and to 0 otherwise. Assuming that wine–food pairings are based on the taste of the wine, they may provide a valuable source of information about how different wines are related to each other. Using the SVD technique, we reduce the dimension of these vectors to a specified (typically somewhat arbitrary) number of abstract dimensions; in this example, we take 5 dimensions. As every wine type is then represented as a point in a 5-dimensional space, we can calculate which wines are between which other ones. In particular, if the cosine of the angle between* $\overrightarrow{\mathbf{ab}}$ *and* $\overrightarrow{\mathbf{ac}}$ *is sufficiently close to 1, for* $\mathbf{a}$, $\mathbf{b}$ *and* $\mathbf{c}$ *the vector representations of three wine types* a, b *and* c, *and if* $\|\overrightarrow{\mathbf{ab}}\| < \|\overrightarrow{\mathbf{ac}}\|$, *we assume that the taste of wine* b *is between that of wines* a *and* c. *Similarly, if the cosine of the angle between* $\overrightarrow{\mathbf{ab}}$ *and* $\overrightarrow{\mathbf{cd}}$ *is sufficiently close to 1, we assume that the change from wine* a *to wine* b *goes in the same direction as the change from wine* c *to wine* d. *In this way, we obtained that Barbera is between Chianti and Merlot (requiring the cosine to be at least 0.975). Given that these latter two wines, have either low or medium tannins, we can derive using interpolative inference that the same should hold for Barbera. Bandol, on the other hand, is not found to be between any two other types of wine. However, we do find that the following pairs* $\langle (a, b), (c, d) \rangle$ *correspond to par-*

[6] `http://www.theworldwidewine.com/Wine_and_Food/This_wine_which_foods.php`, accessed July 18th, 2011.

allel changes in wine type (again requiring the cosine to be at least 0.975):

$$\langle (\textit{bandol}, \textit{cabernet-sauvignon}), (\textit{bordeaux}, \textit{barolo}) \rangle$$
$$\langle (\textit{bandol}, \textit{zinfandel}), (\textit{barolo}, \textit{barbera}) \rangle$$
$$\langle (\textit{bandol}, \textit{zinfandel}), (\textit{bordeaux}, \textit{barbera}) \rangle$$

Given that the change from both Barolo and Bordeaux to Barbera is towards lower tannins, using extrapolative reasoning we derive that the amount of tannins in Bandol should not be smaller than the amount found in Zinfandel, i.e. we find that Bandol either has medium or high tannins.

Note that these data-driven approaches essentially use quantitative information to obtain a qualitative representation. One reason for not using a purely quantitative approach is that the available data is not likely to be sufficiently informative to build accurate conceptual space representations, but still allows us to discover information about qualitative relations between regions. In the case of Example 10, for instance, the quantitative representation we start with represents each concept as a point rather than a convex region. A second reason is that geometric calculations, such as determining convex hulls or Voronoi tessellations, are computationally expensive in high-dimensional spaces. When all we are interested in are spatial relations such as betweenness and parallelism, we can avoid to actually build the conceptual space, using a linear programming approach that was proposed in [50]. Finally, as mentioned before, qualitative representations are invariant w.r.t. changes in the relative importance of the different quality dimensions.

4.3. Regular mappings between conceptual spaces

In this section, we show how the idea of interpolation and extrapolation can be formalized, and explicate the assumptions that warrant such forms of inference. We start from the functional view on rule bases suggested in Section 3.2. In particular, given the view of attribute domains as granular descriptions of conceptual spaces, we may look at mappings between attribute domains as granular descriptions of mappings between conceptual spaces. Let $\mathcal{A} = A_1 \times ... \times A_n$, $\mathcal{B} = B_1 \times ... \times B_s$ and $\mathcal{C} = C_1 \times ... \times C_k$ be defined as before. As each A_i, B_j and C_l correspond to a conceptual space, also $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ correspond to conceptual spaces, which we will denote by $\mathfrak{A}$, $\mathfrak{B}$ and $\mathfrak{C}$ respectively. Note that the set of quality dimensions underlying $\mathfrak{A}$ is the union of the quality dimensions underlying the conceptual space representations of $A_1, ..., A_n$ and similar for $\mathfrak{B}$ and $\mathfrak{C}$.

The mapping f_R, induced by the knowledge base R, can then be seen as a mapping from subsets of $\mathfrak{B}$ to subsets of $\mathfrak{C}$. The nature of this mapping will strongly depend on the nature of the rules involved. For example, let us consider the following rules

$$\begin{aligned} \textit{studio} &\rightarrow \textit{small} \\ \textit{high-tannins} \wedge \textit{full-body} &\rightarrow \textit{opaque} \\ \textit{museum} \wedge \textit{has-live-animals} &\rightarrow \textit{zoo} \\ \textit{large} \wedge \textit{orchestra} &\rightarrow \textit{symphony} \end{aligned}$$

In these rules, the consequent is implied (only) by the meaning of the terms that appear in the antecedent, e.g. a museum which has live animals is called a zoo, by definition of the word *zoo*. As a result, we can make the assumption that the quality dimensions in $\mathfrak{C}$ form a subset of the quality dimensions in $\mathfrak{B}$. This in turn means that there is a linear mapping m (viz. a projection) from $\mathfrak{B}$ to the lower-dimensional space $\mathfrak{C}$. The (known) mapping f_R can then be seen as an approximation of the (unknown) mapping m. Note that this means that there is an underlying functional dependency, e.g. every housing option will have a specific size, and every wine will have a specific degree of transparency. We will refer to rules of this kind as *categorization rules*.

Categorization rules can be contrasted with *phenomenological rules*, which encode observations about the world, e.g.:

$$\begin{aligned} \textit{morning} &\rightarrow \textit{heavy-traffic} \\ \textit{autumn} \wedge \textit{UK} &\rightarrow \textit{rainy} \end{aligned}$$

While in the four previous rules, the conclusion is the consequence of a categorical definition, in case of phemenological rules, there is usually no underlying functional dependency: knowing the time of the day does not allow us to precisely know the amount of traffic, while knowing the date and location does not allow us to precisely know the amount of rain. Interestingly, a similar link between commonsense reasoning and functional dependencies was already pointed out in [51].

The method we propose in this paper is based on the assumption that a rule base approximates a linear mapping, and thus only applies to categorization rules. However, it is worth noting that even for phenomenological rules, our method can often produce plausible conclusions when it is applied locally, e.g. while we should not conclude that there is heavy traffic at mid-day from the observations that there is heavy traffic in the morning and in the evening, it makes sense to conclude that there is heavy-traffic mid-morning if we know that there is heavy traffic in the early morning and late morning. Moreover, whether a rule classifies as phenomenological or not may depend on the underlying conceptual space representations that are assumed. For instance, the conceptual space representation of *morning* may in fact include a quality dimension for the amount of traffic, when mornings (in cities) and rush hour are considered to be so tied together that traffic jams become a characteristic feature of mornings (in a similar way that flying has become a characteristic feature of birds, even if it is not a defining feature). As a result of these considerations, when applying our method in practice, we would by default assume that it contains categorization rules. When this does not introduce any inconsistencies, the conclusions it produces are in some sense the most plausible ones, given the information that it available. When inconsistencies are introduced in the process, however, we may need to partially revise the assumption that the rules are categorization rules and/or that the conceptual relations we have are appropriate in the given context (i.e. take into account all the relevant features). While interpolative and extrapolative reasoning are monotonic forms of reasoning in principle, in practice they would be applied in a non-monotonic fashion, in the sense that adding more rules to a knowledge base may introduce inconsistencies, which would then lead us to apply interpolation and extrapolation more cautiously (or not at all). In Section 8, we will come back to the issue of how to deal with inconsistencies that are introduced by interpolative or extrapolative reasoning.

By taking the view that f_R is the approximation of an unknown mapping m from points of $\mathfrak{B}$ to points of $\mathfrak{C}$, we have for $X \subseteq \mathcal{B}$ that $\mathit{reg}(f_R(X)) \supseteq \{m(p) \,|\, p \in \mathit{reg}(X)\}$, where we write $\mathit{reg}(X)$ for the geometric representation of X as a region in $\mathfrak{B}$ or $\mathfrak{C}$ as before. More precisely, we take the view that the rule base R is *compatible* with some underlying mapping m in the sense that for each $X \subseteq \mathcal{B}$ and $Y \subseteq \mathcal{C}$ it holds that

$$\Big(R \models_{\mathcal{A}} \big(\bigvee_{\mathbf{x}\in X} \bigwedge_{i=1}^{s} x_i\big) \rightarrow \big(\bigvee_{\mathbf{y}\in Y} \bigwedge_{i=1}^{k} y_i\big)\Big) \Rightarrow \Big(m^*\big(\bigcup_{\mathbf{x}\in X} \mathit{reg}(\mathbf{x})\big) \subseteq \bigcup_{\mathbf{y}\in Y} \mathit{reg}(\mathbf{y})\Big) \tag{21}$$

where the $2^{\mathfrak{B}} \rightarrow 2^{\mathfrak{C}}$ mapping m^* is defined as the pointwise extension of m.

The actual conceptual spaces $\mathfrak{B}$ and $\mathfrak{C}$, and a fortiori the mapping m, are inaccessible in most applications. For instance, we cannot assume that a

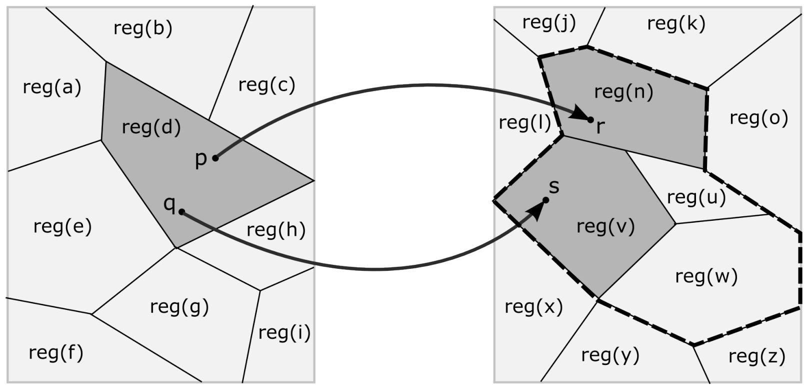


Figure 5: Mappings from $\mathfrak{B}$ (left) to $\mathfrak{C}$ (right).

precise definition of a loft is available, or even an exhaustive enumeration of the qualities on which such a definition would depend. Moreover, using our finite vocabulary, we can only encode approximations of the mapping m, even in the face of complete knowledge. Let us write $\widehat{f}$ for the most informative approximation that can be described using the available labels, i.e. $\widehat{f}$ is the $2^{\mathcal{B}} \to 2^{\mathcal{C}}$ mapping defined for $X \in 2^{\mathcal{B}}$ by

$$\widehat{f}(X) = \{y \in \mathcal{C} \mid x \in X, m^*(reg(x)) \cap reg(y) \neq \emptyset\}$$

In other words, f_R corresponds to the knowledge we actually have, while $\widehat{f}$ corresponds to the maximal knowledge about the mapping m that we could hope to obtain.

Example 11. *Figure 5 displays a setting where $\mathcal{B}$ and $\mathcal{C}$ consist of only one attribute domain, where $\mathcal{B} = B_1 = \{a, b, c, d, e, f, g, h, i\}$ and $\mathcal{C} = C_1 = \{j, k, l, n, o, u, v, w, x, y, z\}$. The mapping m maps each point from $\mathfrak{B}$ to a point of $\mathfrak{C}$. The only knowledge that we have about m is in the forms of rules in the rule base R, which act as constraints on the mapping m. Assume, for example, that R contains the rule $d \to n \vee u \vee v \vee w$, and no other information about d. Then we have that $f_R(\{d\}) = \{n, u, v, w\}$. This mapping f_R is an approximation of the mapping m, for which e.g. $m(p) = r$ and $m(q) = s$ holds. Now suppose that for every point x in $reg(d)$ it holds that $m(x) \in reg(n) \cup reg(v)$, i.e. $m^*(reg(d)) = \{m(x) \mid x \in reg(d)\} \subseteq reg(n) \cup reg(v)$. Then the most precise mapping $\widehat{f}$ that can be described using the vocabulary offered by $\mathcal{B}$ and $\mathcal{C}$ is such that $\widehat{f}(\{d\}) = \{n, v\}$. This means that the rule*

$d \rightarrow n \vee u \vee v \vee w$ *could be refined to* $d \rightarrow n \vee v$*, but this knowledge is missing from* R*.*

The mapping $\widehat{f}$ is not available to us either; it corresponds to the semantic counterpart of a complete rule base, i.e. a rule base which entails all rules that are compatible with m. All we know about $\widehat{f}$ is that it is a refinement of f_R.

Proposition 1. *If* R *is compatible with* m *in the sense of* (21)*, it holds that* $\widehat{f} \leq f_R$*.*

Hence the goal of refining the knowledge base R corresponds, at the semantic level, to finding a mapping $\widehat{f_R}$ for which it is known that $\widehat{f} \leq \widehat{f_R} \leq f_R$. As all our domain knowledge is already encoded in R, a suitable $\widehat{f_R} \neq f_R$ can only be obtained from meta-knowledge about the mapping $\widehat{f}$, or indirectly, from meta-knowledge about the mapping m. Here our restriction to categorization rules plays a key role, as it suggests that m should satisfy the properties of a linear transformation. This leads to the following postulates about the mapping m ($p, q, r, s \in \mathfrak{B}$, $\lambda \geq 0$):

(bet1) $bet(p, q, r) \Rightarrow bet(m(p), m(q), m(r))$

(bet2) $bet(p, q, r) \wedge d(p, q) = \lambda \cdot d(r, q) \Rightarrow d(m(p), m(q)) = \lambda \cdot d(m(r), m(q))$

(par1) $par(p, q, r, s) \Rightarrow par(m(p), m(q), m(r), m(s))$

(par2) $par(p, q, r, s) \wedge d(p, q) = \lambda \cdot d(r, s) \Rightarrow d(m(p), m(q)) = \lambda \cdot d(m(r), m(s))$

In the following section, these postulates will be related to the commonsense inference principles that were introduced in Section 3.1. Specifically, it will become clear that the validity of Principle **(I)** is tied to **(bet1)**, the validity of **(E)** is tied to **(par1)**, the validity of **(I')** is tied to **(bet1)** and **(bet2)**, and the validity of **(E')** is tied to **(par1)** and **(par2)**.

5. Semantic characterization

To characterize interpolative and extrapolative inference, it is useful to note that knowledge can be described on three levels, in the given setting. First, there is the syntactic level, where new rules are produced from given sets of rules. This level will form the topic of Section 6. Second, there is the conceptual spaces level, where labels are represented as geometric regions, and knowledge takes the form of relations between geometric representations of properties and concepts, as was described in Section 4. Finally, there is an intermediate level which we will refer to as the semantic level, as it describes knowledge in terms of a standard propositional logic semantics. In particular, as explained in the beginning of Section 3, the elements of $\mathcal{A}$ take the role of interpretations of the propositional rules and formulas that we consider. This intermediate, semantic level will form the topic of the present section. In particular, at the semantic level, we are interested in approximating the mapping m between the conceptual spaces $\mathfrak{B}$ and $\mathfrak{C}$, as well as relations such as betweenness and parallelism, using the vocabulary at hand. Typically, for a given conceptual space, a number of attribute domains $D_1, ..., D_l$ will be available, containing labels to refer to designated regions of the conceptual space. For the conceptual spaces $\mathfrak{B}$ and $\mathfrak{C}$, these are respectively the attribute domains $B_1, ..., B_s$ and $C_1, ..., C_k$. In general, the available attribute domains determine the level of granularity with which we can describe the underlying conceptual space.

The considered setting of three levels requires us to take into account some subtleties regarding the notion of logical consistency. A logical formula α was defined to be consistent w.r.t. $D_1 \times ... \times D_l$ if there exists an element $\mathbf{d} \in D_1 \times ... \times D_l$ that corresponds to a propositional model of that formula. However, in general, we are not guaranteed that $\mathbf{d}$ actually corresponds to a non-empty region of the underlying conceptual space $\mathfrak{D}$. That is because some of the attribute domains may refer to the same quality dimensions. For example, while $\mathit{blue} \wedge \mathit{wine}$ may be consistent at the propositional level, it corresponds to an empty set in the underlying conceptual space.

Let us call two attribute domains orthogonal if the sets of quality dimensions to which they refer are disjoint. In other words, two attribute domains are orthogonal if their elements refer to different features of a given domain. In particular, if two attribute domains A and B are orthogonal, for every $a \in A$ and $b \in B$, $a \wedge b$ will correspond to a non-empty set in the underlying conceptual space, which can be seen as the Cartesian product of $\mathit{reg}(a)$ and $\mathit{reg}(b)$. In other words, fixing the value of an attribute does not further restrict the possible values of the other attributes. For example, consider the rule $\mathit{large} \wedge \mathit{detached} \rightarrow \mathit{lux}$ from Example 2. The attribute domains A_2 and A_3 (corresponding to *large* and *detached*) are orthogonal to each other, but neither is orthogonal to the attribute domain A_4 (corresponding to *lux*). The concept of orthogonality is closely related to the idea of logical independence.

We call a set $X \subseteq D_1 \times ... \times D_l$ realizable if we are guaranteed that X

contains at least one element that corresponds to a non-empty region. If all of the attribute domains $D_1, ..., D_l$ are orthogonal, this is simply when $X \neq \emptyset$. In general, it may happen that some attribute domains are not orthogonal, but that these attribute domains are irrelevant w.r.t. X. In particular, we define a set $X^{\downarrow i}$ which essentially contains the elements from X that do not depend on the i^{th} attribute domain:

$$(a_1, ..., a_{i-1}, a_{i+1}, ..., a_l) \in X^{\downarrow i} \quad \text{iff} \quad \forall x \in D_i \, . \, (a_1, ..., a_{i-1}, x, a_{i+1}, ..., a_l) \in X \tag{22}$$

Now we can recursively define the notion of realizability. In particular, we say that X is realizable if (i) all attribute domains underlying X are orthogonal and $X \neq \emptyset$ or (ii) there exists an i such that $X^{\downarrow i}$ is realizable.

In this section, we will first analyze how conceptual relations can be defined between subsets of $D_1 \times ... \times D_l$. In Sections 5.1 and 5.2 we will look at betweenness and parallelism, before also taking comparative distance into account in Section 5.3. Finally, in Section 5.4, we will show how these relations can be used to refine the function f_R, using the postulates (**bet1**), (**bet2**), (**par1**) and (**par2**) that were introduced in Section 4.3. Throughout the section, we will assume that the relations $\overline{bet}$, $\underline{bet}$, $\overline{par}$ and $\underline{par}$ are completely specified on the level of individual labels.

5.1. Betweenness

As explained in Section 4, our approach will start from available qualitative knowledge about the conceptual spaces underlying the attribute domains $A_1, ..., A_n$. For three labels $x, y, z \in A_i$, we may know for example that y is fully between x and z, i.e. $\underline{bet}(reg(x), reg(y), reg(z))$ holds, or that y is partially between x and z, i.e. $\overline{bet}(reg(x), reg(y), reg(z))$ holds. Such betweenness information is sufficient for interpolative reasoning in situations where the antecedent of each rule consists of a single atom, taken from some fixed attribute domain B, and where the consequent of each rule consists of a single atom, taken from some fixed attribute domain C. To handle general rule bases, however, we need to lift the betweenness information we have about labels (i.e. atoms) to betweenness information about complex formulas. At the semantic level this means that we need information about the betweenness of subsets of Cartesian products $D_1 \times ... \times D_l$ of attribute domains. Indeed, each element of $D_1 \times ... \times D_l$ corresponds to a conjunction of atoms at the syntactic level, hence subsets of $D_1 \times ... \times D_l$ correspond to arbitrary formulas. Note that each subset of $D_1 \times ... \times D_l$ indeed corresponds to a region in some underlying conceptual space $\mathfrak{D}$.

First we focus on characterizing betweenness for elements of $D_1 \times ... \times D_l$. A central observation is that betweenness for a vector of labels cannot be reduced to betweenness for the labels in the respective components. In particular, notice that when $\underline{bet}(a_1, b_1, c_1)$ and $\underline{bet}(a_2, b_2, c_2)$ hold, we do not necessarily have that (b_1, b_2) is between (a_1, a_2) and (c_1, c_2). Indeed, even for points in a Euclidean space of dimension two or more, betweenness in each dimension does not entail collinearity. Here the intuition of $\overline{bet}$ and $\underline{bet}$ as upper and lower approximations of betweenness becomes important. In particular, while betweenness in each component is not sufficient, it is a necessary condition, hence the upper approximation of betweenness can still be defined as:

$$\overline{bet}(\mathbf{a}, \mathbf{b}, \mathbf{c}) \quad \text{iff} \quad \forall j \,.\, \overline{bet}(a_j, b_j, c_j)$$

where we write e.g. a_i for the i^{th} component of vector $\mathbf{a}$. To extend $\underline{bet}$ to vectors of labels, first note that when the attribute domains $D_1, ..., D_l$ are not orthogonal, we cannot provide any non-trivial guarantees as nothing is between the regions corresponding to $\mathbf{a}$ and $\mathbf{c}$ when one of these regions is empty. Hence $\underline{bet}(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is false unless $\mathbf{a} = \mathbf{b} = \mathbf{c}$. On the other hand, if the attribute domains $D_1, ..., D_l$ are all orthogonal, we define:

$$\begin{aligned} \underline{bet}(\mathbf{a}, \mathbf{b}, \mathbf{c}) \quad \text{iff} \quad & (\mathbf{a} = \mathbf{b}) \vee (\mathbf{b} = \mathbf{c}) \\ & \vee \big(\exists j \,.\, (\forall i \neq j \,.\, a_i = b_i = c_i) \wedge \underline{bet}(a_j, b_j, c_j)\big) \end{aligned} \tag{23}$$

Indeed, in the cases where $\mathbf{a} = \mathbf{b}$ or $\mathbf{b} = \mathbf{c}$, we trivially have that $\underline{bet}(\mathbf{a}, \mathbf{b}, \mathbf{c})$ holds due to the fact that $bet(p, p, q)$ holds for all points p and q. However, note that even in these trivial cases, we still need to require the orthogonality of the attribute domains $D_1, ..., D_l$. In the case where $\mathbf{a} = \mathbf{b} = \mathbf{c}$, on the other hand, we do have that $\underline{bet}(\mathbf{a}, \mathbf{b}, \mathbf{c})$ is trivially satisfied, even if all three regions are empty (as can easily be seen from the definition in (16)). The last disjunct in the right-hand side of (23) covers the general case, where we can only guarantee betweenness for a vector of labels if the vectors coincide in all but one component.

Example 12. *The quality dimensions underlying attribute domains A_1 and A_2 from Example 2 clearly overlap. For example, it is not possible for a* bungalow *to also be an* apartment*, or for a* loft *to be a* row-house*. On the* *other hand, we may consider that attribute domains A_2 and A_3 are orthogonal. Note that this orthogonality holds irrespective of whether there actually exist apartments that are* very-large*. What is important is that nothing in the definition of an* apartment *prevents it from possibly being* very-large*. As a result, we can derive e.g. that*

$$\underline{bet}\big((\text{apartment}, \text{small}), (\text{apartment}, \text{large}), (\text{apartment}, \text{very-large})\big)$$

holds but not e.g.

$$\underline{bet}\big((\text{bungalow}, \text{detached}), (\text{bungalow}, \text{semi-detached}), (\text{bungalow}, \text{row-house})\big)$$

Now we move to betweenness of subsets of $D_1 \times ... \times D_l$. In particular, we will define sets $\overline{bet}(X_1, X_2)$ and $\underline{bet}(X_1, X_2)$ containing respectively those elements from $D_1 \times ... \times D_l$ that are possibly partially between elements of X_1 and elements of X_2, and those elements that are guaranteed to be completely between elements of X_1 and elements of X_2. For $\overline{bet}$ this is straightforward:

$$\overline{bet}(X_1, X_2) = \{\mathbf{b} \,|\, \mathbf{a} \in X_1, \mathbf{c} \in X_2, \overline{bet}(\mathbf{a}, \mathbf{b}, \mathbf{c})\} \tag{24}$$

The following proposition shows that this definition is indeed correct in the sense that it is compatible with the geometric notion of betweenness:

Proposition 2. *Let X_1, X_2 and Y be subsets of $D_1 \times ... \times D_l$. We have that*

$$\overline{bet}(reg(X_1), reg(Y), reg(X_2)) \Rightarrow Y \cap \overline{bet}(X_1, X_2) \neq \emptyset$$

Note that we are slightly abusing notation here, writing $\overline{bet}$ both for a predicate which takes three regions as argument (as defined in (15)), and for a set which takes two subsets of $D_1 \times ... \times D_l$ as argument (as defined in (24)).

To define $\underline{bet}$, the realizability of the arguments X_1 and X_2 again comes into play. We define:

$$\begin{aligned} \underline{bet}(X_1, X_2) = & Z \cup \{\mathbf{b} \,|\, \mathbf{a} \in X_1, \mathbf{c} \in X_2, \underline{bet}(\mathbf{a}, \mathbf{b}, \mathbf{c})\} \\ & \cup \bigcup_{\substack{1 \leq i \leq l \\ x \in D_i}} \{(b_1, ..., b_{i-1}, x, b_{i+1}, ..., b_l) \,|\, \mathbf{b}^{(i)} \in \underline{bet}(X_1^{\downarrow i}, X_2^{\downarrow i})\} \end{aligned} \tag{25}$$

with

$$Z = \begin{cases} X_1 \cup X_2 & \text{if } X_1 \text{ and } X_2 \text{ are realizable} \\ X_1 & \text{if only } X_2 \text{ is realizable} \\ X_2 & \text{if only } X_1 \text{ is realizable} \\ X_1 \cap X_2 & \text{otherwise} \end{cases}$$

$$\mathbf{b}^{(i)} = (b_1, ..., b_{i-1}, b_{i+1}, ..., b_l)$$

The set Z intuitively corresponds to the trivial situations of betweenness $\mathbf{a} = \mathbf{b}$ and $\mathbf{b} = \mathbf{c}$ in (23). If X_1 and X_2 are realizable, this means that all elements from X_1 and X_2 should be considered to be between X_1 and X_2. If X_2 is realizable but not X_1, we know that X_2 corresponds to a non-empty region in the corresponding conceptual space. The set X_1, on the other hand, may or may not correspond to a non-empty region, the lack of realizability merely means that we can not guarantee that $\mathit{reg}(X_1) \neq \emptyset$. In such a case, we cannot guarantee that the elements of X_2 are between X_1 and X_2 as this depends on whether or not X_1 corresponds to an empty region. However, it is not hard to see from (16) that $\underline{bet}(\mathit{reg}(X_1), \mathit{reg}(X_1), \mathit{reg}(X_2))$ holds regardless of whether $\mathit{reg}(X_1) = \emptyset$. The case where only X_1 is realizable is similar. If neither of X_1 and X_2 are realizable, betweenness can only be guaranteed for elements in $X_1 \cap X_2$, noting that $\mathit{reg}(X_1 \cap X_2) = \emptyset$ as soon as one of $\mathit{reg}(X_1) = \emptyset$ or $\mathit{reg}(X_2) = \emptyset$ holds.

The second component in the right-hand side of (25) expresses that $\mathbf{b}$ is in $\underline{bet}(X_1, X_2)$ if it is between some element from X_1 and some element from X_2. The third component is needed to correctly address the case where some of the attribute domains $D_1, ..., D_l$ are not orthogonal (in which case the second component is the empty set). In such a case, we will only find elements that are guaranteed to be between X_1 and X_2 if the attribute domains which are not orthogonal to the others are in some sense irrelevant. The following proposition shows the correctness of the definition of $\underline{bet}$ in (25).

Proposition 3. *Let X_1, X_2 and Y be subsets of $D_1 \times ... \times D_l$. We have that*

$$Y \subseteq \underline{bet}(X_1, X_2) \Rightarrow \underline{bet}(\mathit{reg}(X_1), \mathit{reg}(Y), \mathit{reg}(X_2))$$

Example 13. *Let $X_1 = \{(x, \text{det}, \text{small}) \mid x \in A_1\}$ and $X_2 = \{(x, \text{det}, \text{large}) \mid x \in A_1\}$, and assume that A_2 and A_3 are orthogonal (considering that the size of a house is irrelevant in deciding whether it is e.g. detached or not), while A_1*

and A_2 are not. Noting that

$$\underline{bet}(\{(\mathrm{det},\mathrm{small})\},\{(\mathrm{det},\mathrm{large})\}) = \{(\mathrm{det},\mathrm{small}),(\mathrm{det},\mathrm{medium}),(\mathrm{det},\mathrm{large})\}$$

we find $\underline{bet}(X_1, X_2) = \{(x, \mathrm{det}, y) \,|\, x \in A_1, y \in \{\mathrm{small}, \mathrm{medium}, \mathrm{large}\}\}$

5.2. Parallelism

The treatment of parallelism is largely analogous to the treatment of betweenness. If the attribute domains $D_1, ..., D_l$ are all orthogonal, the relation $\underline{par}$ can be defined for elements of $D_1 \times ... \times D_l$ as follows:

$$\begin{aligned}\underline{par}(\mathbf{a},\mathbf{b},\mathbf{c};\mathbf{d}) \quad \text{iff} \quad & (\mathbf{a}=\mathbf{c} \wedge \mathbf{b}=\mathbf{d}) \vee (\mathbf{c}=\mathbf{d}) \\ & \vee (\exists j\,.\,(\forall i \neq j\,.\,(a_i = b_i) \wedge (c_i = d_i)) \wedge \underline{par}(a_j, b_j, c_j; d_j))\end{aligned} \tag{26}$$

As for betweenness, we find that a component-wise assessment of parallelism is not usually sufficient to provide guarantees on the parallelism of the vectors. The exceptions covered by the definition of $\underline{par}$ are the trivial case where $\mathbf{a} = \mathbf{c}$ and $\mathbf{b} = \mathbf{d}$, the trivial case where $\mathbf{c} = \mathbf{d}$, and the case where the transition from $\mathbf{a}$ to $\mathbf{b}$, and from $\mathbf{c}$ to $\mathbf{d}$ only affects one component. If the attribute domains $D_1, ..., D_l$ are not orthogonal, we cannot guarantee anything about the parallelism of elements from $D_1 \times ... \times D_l$, unless in the entirely trivial case where $\mathbf{a} = \mathbf{b} = \mathbf{c} = \mathbf{d}$.

For any Cartesian product $D_1 \times ... \times D_l$ of attribute domains, $\overline{par}$ is defined as

$$\overline{par}(\mathbf{a},\mathbf{b},\mathbf{c};\mathbf{d}) \quad \text{iff} \quad \forall j\,.\,\overline{par}(a_j, b_j, c_j; d_j)$$

To assess parallelism w.r.t. subsets X_1, X_2 and X_3 of $D_1 \times ... \times D_l$, we define:

$$\begin{aligned}&\underline{par}(X_1,X_2,X_3) = Z_2 \cup Z_3 \cup \{\mathbf{d} \,|\, \mathbf{a} \in X_1, \mathbf{b} \in X_2, \mathbf{c} \in X_3, \underline{par}(\mathbf{a},\mathbf{b},\mathbf{c};\mathbf{d})\} \\ &\cup \{(d_1,...,d_{i-1},x,d_{i+1},...,d_l) \,|\, \mathbf{d} \in \underline{par}(X_1^{\downarrow i}, X_2^{\downarrow i}, X_3^{\downarrow i}), 1 \leq i \leq l, x \in D_i\} \\ &\cup \{(d_1,...,d_{i-1},x,d_{i+1},...,d_l) \,|\, \mathbf{d} \in \underline{par}(X_1^{\downarrow x}, X_2^{\downarrow y}, X_3^{\downarrow y}), 1 \leq i \leq l, x, y \in D_i, \\ &\qquad\qquad D_i \text{ orthogonal to } D_1,...,D_{i-1},D_{i+1},...,D_l\}\end{aligned} \tag{27}$$

with

$$Z_2 = \begin{cases} X_2 & \text{if } X_1 \cap X_3 \text{ is realizable} \\ X_1 \cap X_2 \cap X_3 & \text{otherwise} \end{cases}$$

$$Z_3 = \begin{cases} X_3 & \text{if } X_1 \text{ and } X_2 \text{ are realizable} \\ X_2 \cap X_3 & \text{if } X_1 \text{ is realizable but not } X_2 \\ X_1 \cap X_3 & \text{if } X_2 \text{ is realizable but not } X_1 \\ X_1 \cap X_2 \cap X_3 & \text{otherwise} \end{cases}$$

and for $y \in D_i$ and $X \subseteq D_1 \times ... \times D_l$ we define

$$(a_1, ..., a_{i-1}, a_{i+1}, ..., a_l) \in X^{\downarrow y} \quad \text{iff} \quad (a_1, ..., a_{i-1}, y, a_{i+1}, ..., a_l) \in X$$

The set Z_2 corresponds to those elements that can be added to $\underline{par}(X_1, X_2, X_3)$ due to the trivial case where $\mathbf{a} = \mathbf{c} \wedge \mathbf{b} = \mathbf{d}$ in (26), whereas $\overline{Z_3}$ corresponds to the trivial case where $\mathbf{c} = \mathbf{d}$. In particular, if $X_1 \cap X_3$ is realizable we know that for every point in X_2 we can find a point which belongs to both X_1 and X_3, hence $(X_1 \cap X_3, X_2, X_1 \cap X_3, X_2)$ trivially defines two parallel directions, hence so does (X_1, X_2, X_3, X_2). The definition of Z_3 is based on the intuition that (X_1, X_2, X_3, X_3) trivially defines two parallel directions, but only if X_1, X_2 and X_3 correspond to non-empty regions. As in the case of the set Z in the definition of $\underline{bet}$, we need to take realizability into account. The third argument of the union in the right-hand-side of (27) expresses the basic case: $\mathbf{d}$ is in $\underline{par}(X_1, X_2, X_3)$ if there are elements in X_1, X_2 and X_3 such that the corresponding directions are parallel. The fourth argument allows us to ignore irrelevant attribute domains, similar as in the definition of $\underline{bet}$ in (25). For parallelism, however, there is another case where situations of parallism in $D_1 \times ... \times D_{i-1} \times D_{i+1} \times ... \times D_l$ may be extended to $D_1 \times ... \times D_l$, which is covered by the last argument of the union. The central idea is that when D_i is orthogonal to the other attribute domains, we can be more tolerant; note that $X_2^{\downarrow i} \times X_3^{\downarrow i} \subseteq \bigcup_{y \in D_i} X_2^{\downarrow y} \times X_3^{\downarrow y}$. Indeed, all that is required is that we can extend vectors $\mathbf{a}$, $\mathbf{b}$, $\mathbf{c}$ and $\mathbf{d}$ in $D_1 \times ... \times D_{i-1} \times D_{i+1} \times ... \times D_l$, by choosing the same value x from D_i to extend $\mathbf{a}$ and $\mathbf{b}$ and the same value y to extend $\mathbf{c}$ and $\mathbf{d}$, without there being a requirement for $x = y$ to hold.

On the other hand, regardless of the realizability of X_1, X_2 and X_3, we always define:

$$\overline{par}(X_1, X_2, X_3) = \{\mathbf{d} \,|\, \mathbf{a} \in X_1, \mathbf{b} \in X_2, \mathbf{c} \in X_3, \overline{par}(\mathbf{a}, \mathbf{b}, \mathbf{c}, \mathbf{d})\}$$

The correctness of the proposed definitions is demonstrated by the following propositions.

Proposition 4. *Let X_1, X_2, X_3 and Y be subsets of $D_1 \times ... \times D_l$. We have that*

$$Y \subseteq \underline{par}(X_1, X_2, X_3) \Rightarrow \underline{par}(reg(X_1), reg(X_2), reg(X_3); reg(Y))$$

Proposition 5. *Let X_1, X_2, X_3 and Y be subsets of $D_1 \times ... \times D_l$. We have that*

$$\overline{par}(reg(X_1), reg(X_2), reg(X_3); reg(Y)) \Rightarrow Y \cap \overline{par}(X_1, X_2, X_3) \neq \emptyset$$

5.3. Comparative distance

The relations $\underline{bet_\sigma}$, $\overline{bet_\sigma}$, $\underline{par_\tau}$ and $\overline{par_\tau}$ behave in general similar to their counterparts $\underline{bet}$, $\overline{bet}$, $\underline{par}$ and $\overline{par}$. The main exception is when $\sigma = [\lambda, \lambda]$ or $\tau = [\mu, \mu]$ is a degenerate interval, i.e. a singleton, in which case $\underline{bet_\sigma}$ and $\underline{par_\tau}$ hold as soon as these relations hold for each of the components, i.e. we have for each $\lambda \in [0, 1]$ and $\mu \in [0, +\infty[$ that

$$\begin{aligned}\underline{bet_{[\lambda,\lambda]}}(\mathbf{a}, \mathbf{b}, \mathbf{c}) \quad &\text{iff} \quad \forall i \,.\, \underline{bet_{[\lambda,\lambda]}}(a_i, b_i, c_i)\\ \underline{par_{[\mu,\mu]}}(\mathbf{a}, \mathbf{b}, \mathbf{c}; \mathbf{d}) \quad &\text{iff} \quad \forall i \,.\, \underline{par_{[\mu,\mu]}}(a_i, b_i, c_i; d_i)\end{aligned}$$

provided that the attribute domains $D_1, ..., D_l$ are orthogonal; otherwise, we still cannot provide any guarantees on betweenness or parallelism beyond the trivial cases where $\mathbf{a} = \mathbf{b} = \mathbf{c}$ and $\mathbf{a} = \mathbf{b} = \mathbf{c} = \mathbf{d}$ respectively. The underlying reason that the definition of $\underline{bet_\sigma}$ becomes more tolerant when σ is a singleton is due to the fact that for points $p = (p_1, p_2, ..., p_n)$, $q = (q_1, q_2, ..., q_n)$, $r = (r_1, r_2, ..., r_n)$ in a Euclidean space, it holds that whenever we have $q_i = p_i + \lambda(r_i - p_i)$ for all i, we also have $\overrightarrow{pq} = \lambda \cdot \overrightarrow{pr}$; and similar for $\underline{par_\tau}$.

If σ and τ are not singletons, $\underline{bet_\sigma}$ and $\underline{par_\tau}$ are defined as:

$$\begin{aligned}\underline{bet_\sigma}(\mathbf{a}, \mathbf{b}, \mathbf{c}) \quad &\text{iff} \quad \alpha_1 \vee \alpha_2 \vee (\mathbf{a} = \mathbf{b} = \mathbf{c})\\ &\qquad \vee \big(\exists j \,.\, (\forall i \neq j \,.\, a_i = b_i = c_i) \wedge \underline{bet_\sigma}(a_j, b_j, c_j)\big)\\ \underline{par_\tau}(\mathbf{a}, \mathbf{b}, \mathbf{c}; \mathbf{d}) \quad &\text{iff} \quad \beta_1 \vee \beta_2 \vee ((\mathbf{a} = \mathbf{b}) \wedge (\mathbf{c} = \mathbf{d}))\\ &\qquad \vee (\exists j \,.\, (\forall i \neq j \,.\, (a_i = b_i) \wedge (c_i = d_i)) \wedge \underline{par_\tau}(a_j, b_j, c_j; d_j))\end{aligned}$$

with

$$\alpha_1 \equiv \begin{cases} \mathbf{a} = \mathbf{b} & \text{if } 0 \in \sigma \\ \bot & \text{otherwise} \end{cases} \qquad \alpha_2 \equiv \begin{cases} \mathbf{b} = \mathbf{c} & \text{if } 1 \in \sigma \\ \bot & \text{otherwise} \end{cases}$$

$$\beta_1 \equiv \begin{cases} (\mathbf{a} = \mathbf{c}) \wedge (\mathbf{b} = \mathbf{d}) & \text{if } 1 \in \tau \\ \bot & \text{otherwise} \end{cases} \qquad \beta_2 \equiv \begin{cases} \mathbf{c} = \mathbf{d} & \text{if } 0 \in \tau \\ \bot & \text{otherwise} \end{cases}$$

and where again we assume that $D_1, ..., D_l$ are orthogonal. Note that $\underline{bet_\sigma}$ and $\underline{par_\tau}$ are essentially defined as $\underline{bet}$ and $\underline{par}$ when σ and τ are not singletons, although some of the trivial cases only hold when they contain 0 or 1.

Moreover, in all cases $\overline{bet_\sigma}$ and $\overline{par_\tau}$ are defined like $\underline{bet}$ and $\underline{par}$:

$$\begin{aligned} \overline{bet_\sigma}(\mathbf{a}, \mathbf{b}, \mathbf{c}) \quad &\text{iff} \quad \forall j \,.\, \overline{bet_\sigma}(a_j, b_j, c_j) \\ \overline{par_\tau}(\mathbf{a}, \mathbf{b}, \mathbf{c}; \mathbf{d}) \quad &\text{iff} \quad \forall j \,.\, \overline{par_\tau}(a_j, b_j, c_j; d_j) \end{aligned}$$

Now we extend these relations to subsets of $D_1 \times ... \times D_l$. Let X_1 and X_2 be subsets of $D_1 \times ... \times D_l$; we define

$$\begin{aligned} \underline{bet_\sigma}(X_1, X_2) = Z_1 \cup Z_2 \cup \{\mathbf{b} \,|\, \mathbf{a} \in X_1, \mathbf{c} \in X_2, \underline{bet_\sigma}(\mathbf{a}, \mathbf{b}, \mathbf{c})\} \\ \cup \{(b_1, ..., b_{i-1}, x, b_{i+1}, ..., b_l) \,|\, \mathbf{b} \in \underline{bet_\sigma}(X_1^{\downarrow i}, X_2^{\downarrow i}), 1 \leq i \leq l, x \in D_i\} \end{aligned} \tag{28}$$

with

$$Z_1 = \begin{cases} X_1 & \text{if } X_2 \text{ is realizable and } 0 \in \sigma \\ X_1 \cap X_2 & \text{otherwise} \end{cases}$$

$$Z_2 = \begin{cases} X_2 & \text{if } X_1 \text{ is realizable and } 1 \in \sigma \\ X_1 \cap X_2 & \text{otherwise} \end{cases}$$

The situation for $\underline{par_\tau}$ is mostly similar as for $\underline{par}$:

$$\begin{aligned} \underline{par_\tau}(X_1, X_2, X_3) = Z_2 \cup Z_3 \cup \{\mathbf{d} \,|\, \mathbf{a} \in X_1, \mathbf{b} \in X_2, \mathbf{c} \in X_3, \underline{par_\tau}(\mathbf{a}, \mathbf{b}, \mathbf{c}; \mathbf{d})\} \\ \cup \{(d_1, ..., d_{i-1}, x, d_{i+1}, ..., d_l) \,|\, \mathbf{d} \in \underline{par_\tau}(X_1^{\downarrow i}, X_2^{\downarrow i}, X_3^{\downarrow i}), 1 \leq i \leq l, x \in D_i\} \\ \cup \{(d_1, ..., d_{i-1}, x, d_{i+1}, ..., d_l) \,|\, \mathbf{d} \in \underline{par_\tau}(X_1^{\downarrow x}, X_2^{\downarrow y}, X_3^{\downarrow y}), 1 \leq i \leq l, x, y \in D_i, \\ D_i \text{ orthogonal to } D_1, ..., D_{i-1}, D_{i+1}, ..., D_l\} \end{aligned} \tag{29}$$

with

$$Z_2 = \begin{cases} X_2 & \text{if } X_1 \cap X_3 \text{ is realizable and } 1 \in \tau \\ X_1 \cap X_2 \cap X_3 & \text{otherwise} \end{cases}$$

$$Z_3 = \begin{cases} X_3 & \text{if } X_1 \text{ and } X_2 \text{ are realizable and } 0 \in \tau \\ & \quad \text{or if } X_1 \cap X_2 \text{ is realizable} \\ X_2 \cap X_3 & \text{if } X_1 \text{ is realizable but not } X_2 \text{ and } 0 \in \tau \\ X_1 \cap X_3 & \text{if } X_2 \text{ is realizable but not } X_1 \text{ and } 0 \in \tau \\ X_1 \cap X_2 \cap X_3 & \text{otherwise} \end{cases}$$

Proposition 6. *Let X_1, X_2, X_3, and Y be subsets of $D_1 \times ... \times D_l$, let σ be a subset of $[0,1]$ and τ a subset of $[0,+\infty[$. We have that*

$$Y \subseteq \underline{bet_\sigma}(X_1, X_2) \Rightarrow \underline{bet_\sigma}(reg(X_1), reg(Y), reg(X_2))$$
$$Y \subseteq \underline{par_\tau}(X_1, X_2, X_3) \Rightarrow \underline{par_\tau}(reg(X_1), reg(X_2), reg(X_3); reg(Y))$$

Proposition 7. *Let X_1, X_2, X_3, and Y be subsets of $D_1 \times ... \times D_l$, let σ be a subset of $[0,1]$ and τ a subset of $[0,+\infty[$. We have that*

$$\overline{bet_\sigma}(reg(X_1), reg(Y), reg(X_2)) \Rightarrow Y \cap \overline{bet_\sigma}(X_1, X_2) \neq \emptyset$$
$$\overline{par_\tau}(reg(X_1), reg(X_2), reg(X_3); reg(Y)) \Rightarrow Y \cap \overline{par_\tau}(X_1, X_2, X_3) \neq \emptyset$$

5.4. Refining the rule base

From Postulates (**bet1**), (**bet2**), (**par1**) and (**par2**), we know that for all regions A_1, A_2, A_3 in $\mathfrak{B}$, it holds that:

$$m^*(bet_\sigma(A_1, A_2)) \subseteq bet_\sigma(m^*(A_1), m^*(A_2)) \tag{30}$$
$$m^*(par_\sigma(A_1, A_2, A_3)) \subseteq par_\sigma(m^*(A_1), m^*(A_2), m^*(A_3)) \tag{31}$$

where m^* is the pointwise extension of m as before, and ($\sigma \subseteq [0,1]$, $\tau \subseteq [0,+\infty[$)

$$bet_\sigma(A_1, A_2) = \{q \mid \exists p \in A_1, r \in A_2, \lambda \in \sigma \,.\, bet(p,q,r) \wedge \overrightarrow{pq} = \lambda \cdot \overrightarrow{pr}\}$$
$$par_\tau(A_1, A_2, A_3) = \{s \mid \exists p \in A_1, q \in A_2, r \in A_3, \lambda \in \tau \,.\, \overrightarrow{rs} = \lambda \cdot \overrightarrow{pq}\}$$

In other words, we have that $bet_\sigma(A_1, A_2)$ is the true set of points that are located between A_1 and A_2 with a relative distance from A_1 contained in σ, and similar for $par_\tau(A_1, A_2, A_3)$. In particular, we have that

$$\underline{bet_\sigma}(A_1, A_2) \subseteq bet_\sigma(A_1, A_2) \tag{32}$$
$$bet_\sigma(m^*(A_1), m^*(A_2)) \subseteq \overline{bet_\sigma}(m^*(A_1), m^*(A_2)) \tag{33}$$
$$\underline{par_\tau}(A_1, A_2, A_3) \subseteq par_\tau(A_1, A_2, A_3) \tag{34}$$
$$par_\tau(m^*(A_1), m^*(A_2), m^*(A_3)) \subseteq \overline{par_\tau}(m^*(A_1), m^*(A_2), m^*(A_3)) \tag{35}$$

Together with (30)–(31) and the monotonicity of $\widehat{f}$ w.r.t. set inclusion, this yields

$$\widehat{f}(\underline{bet_\sigma}(X_1, X_2)) \subseteq \overline{bet_\sigma}(\widehat{f}(X_1), \widehat{f}(X_2)) \tag{36}$$
$$\widehat{f}(\underline{par_\tau}(X_1, X_2, X_3)) \subseteq \overline{par_\tau}(\widehat{f}(X_1), \widehat{f}(X_2), \widehat{f}(X_3)) \tag{37}$$

for all $X_1, X_2, X_3 \in 2^{\mathcal{B}}$.

Recall that the mapping $\widehat{f}$ is unknown to us, and corresponds to the semantic counterpart of the 'ideal' rule base, containing all knowledge about m that can be encoded using the given vocabulary. However, given that $\widehat{f} \leq f_R$ is assumed to hold (i.e. all knowledge encoded in R is correct), (36) and (37) allow us to refine the mapping f_R, which is at our disposal, to the most conservative refinement $\widehat{f_R}$ that satisfies these two constraints, i.e. we define $\widehat{f_R}$ to be the largest fixpoint, w.r.t. the ordering $\leq$ defined in (14), of

$$\begin{aligned}\widehat{f_R}(\{\mathbf{x}\}) = f_R(\{\mathbf{x}\}) &\cap \bigcap_{\sigma\subseteq[0,1]} \{\overline{bet_\sigma}(\widehat{f_R}(Y), \widehat{f_R}(Z)) | \mathbf{x} \in \underline{bet_\sigma}(Y, Z)\} \\ &\cap \bigcap_{\tau\subseteq[0,+\infty[} \{\overline{par_\tau}(\widehat{f_R}(X), \widehat{f_R}(Y), \widehat{f_R}(Z)) | \mathbf{x} \in \underline{par_\tau}(X, Y, Z)\}\end{aligned}$$

and $\widehat{f_R}(X) = \bigcup_{\mathbf{x}\in X} \widehat{f_R}(\{\mathbf{x}\})$. From the well-known Knaster-Tarski theorem [52], we know that this largest fixpoint exists, and can be found in an iterative way as follows. Let $\widehat{f_R}^{(0)} = f_R$ and

$$\begin{aligned}\widehat{f_R}^{(i+1)}(\{\mathbf{x}\}) = f_R(\{\mathbf{x}\}) &\cap \bigcap_{\sigma\subseteq[0,1]} \{\overline{bet_\sigma}(\widehat{f_R}^{(i)}(Y), \widehat{f_R}^{(i)}(Z)) | \mathbf{x} \in \underline{bet_\sigma}(Y, Z)\} \\ &\cap \bigcap_{\tau\subseteq[0,+\infty[} \{\overline{par_\tau}(\widehat{f_R}^{(i)}(X), \widehat{f_R}^{(i)}(Y), \widehat{f_R}^{(i)}(Z)) | \mathbf{x} \in \underline{par_\tau}(X, Y, Z)\}\end{aligned} \tag{38}$$

and $\widehat{f_R}^{(i+1)}(X) = \bigcup_{\mathbf{x}\in X} \widehat{f_R}^{(i+1)}(\{\mathbf{x}\})$. It is clear that $\widehat{f_R}^{(i+1)} \leq \widehat{f_R}^{(i)}$, i.e. this definition allows us to repeatedly refine the initial function $\widehat{f_R}^{(0)}$. From the finiteness of the attribute domains $\mathcal{B}$ and $\mathcal{C}$, it follows that this process must end, i.e. that there is an $i_0 \in \mathbb{N}$ such that $\widehat{f_R}^{(i_0+1)} = \widehat{f_R}^{(i_0)} = \widehat{f_R}$.

Proposition 8. *Let f_R, $\widehat{f}$ and $\widehat{f_R}$ be defined as before. It holds that*

$$\widehat{f_R} \leq f_R$$

Moreover, if R is compatible with m in the sense of (21)*, it follows from Postulates* (**bet1**)*,* (**bet2**)*,* (**par1**) *and* (**par2**) *that*

$$\widehat{f} \leq \widehat{f_R}$$

Example 14. *Let us again consider the rules from Example 2, and let us determine the comfort level of a medium-sized detached villa:*

$$\begin{aligned}
&\widehat{f_R}(\{(\text{villa}, \text{det}, \text{med})\}) \\
&\subseteq \widehat{f_R}(\underline{bet}(X_1, X_2)) \\
&\subseteq \overline{bet}(\widehat{f_R}(X_1), \widehat{f_R}(X_2)) \\
&\subseteq \overline{bet}(f_R(X_1), f_R(X_2)) \\
&= \overline{bet}(\{(\text{det}, \text{small}, \text{bas}), (\text{det}, \text{small}, \text{comf})\}, \{(\text{det}, \text{large}, \text{lux})\}) \\
&= \{(\text{det}, x, y) \,|\, x \in \{\text{small}, \text{med}, \text{large}\}, y \in \{\text{bas}, \text{comf}, \text{lux}\}\}
\end{aligned}$$

where X_1 and X_2 are as defined in Example 13. Furthermore, we also have

$$\widehat{f_R}(\{(\text{villa}, \text{det}, \text{med})\}) \subseteq f_R(\{(\text{villa}, \text{det}, \text{med})\}) = \{(\text{det}, \text{med}, x) \,|\, x \in A_4\}$$

Together we thus find

$$\widehat{f_R}(\{(\text{villa}, \text{det}, \text{med})\}) \subseteq \{(\text{det}, \text{med}, \text{bas}), (\text{det}, \text{med}, \text{comf}), (\text{det}, \text{med}, \text{lux})\}$$

6. Syntactic characterization

In this section, we will analyze how conceptual relations behave at the syntactic level, i.e. how available background information on betweenness and parallelism of atoms can be lifted to relations between arbitrary propositional formulas. Together with the meta-principles **(S)**, **(I)**, **(I')**, **(E)**, **(E')** and

(D) introduced in Section 3.1, we will then be able to fully characterize a commonsense inference relation, and show its soundness and completeness w.r.t. the semantics from Section 5.

As elements from $\mathcal{A}$ can be identified with interpretations, any propositional formula α naturally corresponds to a set $\mathit{set}(\alpha)$ of elements S from $\mathcal{A}$, which are the elements that correspond to the models of α, i.e. $\mathit{set}(\alpha) = \{\omega \in \mathcal{A} \,|\, \omega \models_{\mathcal{A}} \alpha\}$. Conversely, every element $\mathbf{a} = (a_1, ..., a_n)$ from $\mathcal{A}$ corresponds to an interpretation, which can syntactically be characterized as a conjunction of atoms $a_1 \wedge ... \wedge a_n$; let us write $\mathit{conj}(\mathbf{a})$ for the conjunction corresponding to $\mathbf{a} \in \mathcal{A}$.

As input to the inference relation, in addition to the rule base R, we assume that a set of assertions Σ is given, which specifies the conceptual relations at the atom level. We define the closure Σ^* of Σ to be the (infinite) set of all assertions that directly follow from assertions in Σ. In particular, we define Σ^* as the smallest set of assertions which satisfies $\Sigma \subseteq \Sigma^*$, and moreover for all a, b, c, and d:

$$
\begin{aligned}
\underline{bet}_{\sigma}(a,b,c) \in \Sigma^* \quad &\Rightarrow \quad \underline{bet}_{1-\sigma}(c,b,a) \in \Sigma^* \\
\overline{bet}_{\sigma}(a,b,c) \in \Sigma^* \quad &\Rightarrow \quad \overline{bet}_{1-\sigma}(c,b,a) \in \Sigma^* \\
\underline{bet}_{\sigma}(a,b,c) \in \Sigma^* \wedge \sigma \subseteq \sigma' \quad &\Rightarrow \quad \underline{bet}_{\sigma'}(c,b,a) \in \Sigma^* \\
\overline{bet}_{\sigma}(a,b,c) \in \Sigma^* \wedge \sigma \subseteq \sigma' \quad &\Rightarrow \quad \overline{bet}_{\sigma'}(c,b,a) \in \Sigma^* \\
\underline{par}_{\tau}(a,b,c;d) \in \Sigma^* \wedge \tau \subseteq \tau' \quad &\Rightarrow \quad \underline{par}_{\tau'}(a,b,c;d) \in \Sigma^* \\
\overline{par}_{\tau}(a,b,c;d) \in \Sigma^* \wedge \tau \subseteq \tau' \quad &\Rightarrow \quad \overline{par}_{\tau'}(a,b,c;d) \in \Sigma^* \\
\underline{bet}_{\sigma}(a,a,c) \in \Sigma^* \quad &\Rightarrow \quad \overline{bet}_{\sigma}(a,a,c) \in \Sigma^* \\
\underline{par}_{\tau}(a,b,c;d) \in \Sigma^* \quad &\Rightarrow \quad \overline{par}_{\tau}(a,b,c;d) \in \Sigma^* \\
0 \in \sigma \quad &\Rightarrow \quad \underline{bet}_{\sigma}(a,a,c) \in \Sigma^* \\
\sigma \neq \emptyset \quad &\Rightarrow \quad \underline{bet}_{\sigma}(a,a,a) \in \Sigma^* \\
1 \in \tau \quad &\Rightarrow \quad \underline{par}_{\tau}(a,b,a;b) \in \Sigma^* \\
0 \in \tau \quad &\Rightarrow \quad \underline{par}_{\tau}(a,b,c;c) \in \Sigma^* \\
\tau \neq \emptyset \quad &\Rightarrow \quad \underline{par}_{\tau}(a,a,b;b) \in \Sigma^*
\end{aligned}
$$

where we write $1 - \sigma$ for the set $\{1 - x \,|\, x \in \sigma\}$. Note that even though the set Σ^* is infinite, it is straightforward to decide whether a given assertion such as $\underline{bet}_{\sigma}(a,b,c)$ is in Σ^* from the definition of Σ. In the following, we will use the assumption that the assertions in Σ are sound in the following sense:

$$\underline{bet}_\sigma(a,b,c) \in \Sigma \;\Rightarrow\; \underline{bet}_\sigma(reg(a), reg(b), reg(c)) \tag{39}$$
$$\overline{bet}_\sigma(reg(a), reg(b), reg(c)) \;\Rightarrow\; \overline{bet}_\sigma(a,b,c) \in \Sigma^* \tag{40}$$
$$\underline{par}_\tau(a,b,c;d) \in \Sigma \;\Rightarrow\; \underline{par}_\tau(reg(a), reg(b), reg(c); reg(d)) \tag{41}$$
$$\overline{par}_\tau(reg(a), reg(b), reg(c); reg(d)) \;\Rightarrow\; \overline{par}_\tau(a,b,c;d) \in \Sigma^* \tag{42}$$

for all labels a, b, c, d from the same attribute domain A_i. In addition, the following assumptions are needed to show completeness, although they can be abandoned when completeness is not required (or feasible):

$$\underline{bet}_\sigma(reg(a), reg(b), reg(c)) \;\Rightarrow\; \underline{bet}_\sigma(a,b,c) \in \Sigma^* \tag{43}$$
$$\overline{bet}_\sigma(a,b,c) \in \Sigma \;\Rightarrow\; \overline{bet}_\sigma(reg(a), reg(b), reg(c)) \tag{44}$$
$$\underline{par}_\tau(reg(a), reg(b), reg(c); reg(d)) \;\Rightarrow\; \underline{par}_\tau(a,b,c;d) \in \Sigma^* \tag{45}$$
$$\overline{par}_\tau(a,b,c;d) \in \Sigma \;\Rightarrow\; \overline{par}_\tau(reg(a), reg(b), reg(c); reg(d)) \tag{46}$$

In the following, we do not explicitly consider $\underline{bet}$, $\underline{par}$, $\overline{bet}$ and $\overline{par}$, seeing these relations merely as abbreviations for $\underline{bet_{[0,1]}}$, $\underline{par_{[0,+\infty[}}$, $\overline{bet_{[0,1]}}$ and $\overline{par_{[0,+\infty[}}$.

Before we can characterize the conceptual relations, we need a syntactic counterpart for the notion of realizability, which played a key role in Section 5. As it turns out, at the syntactic level, realizability corresponds to a strong notion of consistency. More precisely, a formula α is called consistent if it has at least one model. Given a formula α, let us write *domains*(α) for the set of attribute domains to which α refers, i.e. for each atom a occurring in α, the attribute domain A_i which contains a will be in *domains*(α). We then say that α is strongly consistent if there exists a consistent formula β such that $\beta \models \alpha$ and all attribute domains in *domains*(α) are orthogonal to each other. The correspondence with realizability is demonstrated by the following lemma.

Lemma 1. *It holds that a formula α is strongly consistent iff* $\mathrm{set}(\alpha)$ *is realizable.*

To formalize what can be inferred from Σ, we again follow a normative approach. In particular, we consider that an expression of the form $\alpha \rightarrow \beta \bowtie_\sigma \gamma$ is supported by Σ if it can be derived using the following rules.

($\underline{\textbf{bet}}_\mathbf{1}$) If $\underline{bet}_\sigma(a, x, b) \in \Sigma$ then $\Sigma \vdash x \to a \bowtie_\sigma b$.

($\underline{\textbf{bet}}_\mathbf{2}$) If $\alpha_2 \models_\mathcal{A} \alpha_1$, $\beta_1 \models_\mathcal{A} \beta_2$, $\gamma_1 \models_\mathcal{A} \gamma_2$ and $\sigma_1 \subseteq \sigma_2$, then whenever we have $\Sigma \vdash \alpha_1 \to \beta_1 \bowtie_{\sigma_1} \gamma_1$ we also have $\Sigma \vdash \alpha_2 \to \beta_2 \bowtie_{\sigma_2} \gamma_2$.

($\underline{\textbf{bet}}_\mathbf{3}$) If $\Sigma \vdash \alpha \to \beta \bowtie_\sigma \gamma$ then $\Sigma \vdash \alpha \to \gamma \bowtie_{(1-\sigma)} \beta$.

($\underline{\textbf{bet}}_\mathbf{4}$) If β is strongly consistent and $0 \in \sigma$, we have $\Sigma \vdash \alpha \to \alpha \bowtie_\sigma \beta$.

($\underline{\textbf{bet}}_\mathbf{5}$) It always[7] holds that $\Sigma \vdash \alpha \to \alpha \bowtie_\sigma \alpha$.

($\underline{\textbf{bet}}_\mathbf{6}$) If $\Sigma \vdash \alpha_1 \to \beta_1 \bowtie_\sigma \gamma_1$ and $\Sigma \vdash \alpha_2 \to \beta_2 \bowtie_\sigma \gamma_2$ then $\Sigma \vdash \alpha_1 \vee \alpha_2 \to (\beta_1 \vee \beta_2) \bowtie_\sigma (\gamma_1 \vee \gamma_2)$.

($\underline{\textbf{bet}}_\mathbf{7}$) If $\Sigma \vdash \alpha \to \beta \bowtie_\sigma \gamma$, δ is strongly consistent and the attribute domains in *domains*$(\alpha) \cup$ *domains*$(\beta) \cup$ *domains*(γ) are orthogonal to those in *domains*(δ), then $\Sigma \vdash \alpha \wedge \delta \to (\beta \wedge \delta) \bowtie_\sigma (\gamma \wedge \delta)$.

($\underline{\textbf{bet}}_\mathbf{8}$) If $\Sigma \vdash \alpha_1 \to \beta_1 \bowtie_{[\lambda,\lambda]} \gamma_1$, $\Sigma \vdash \alpha_2 \to \beta_2 \bowtie_{[\lambda,\lambda]} \gamma_2$ and the attribute domains in *domains*$(\alpha_1) \cup$ *domains*$(\beta_1) \cup$ *domains*(γ_1) are orthogonal to those in *domains*$(\alpha_2) \cup$ *domains*$(\beta_2) \cup$ *domains*(γ_2), then $\Sigma \vdash \alpha_1 \wedge \alpha_2 \to (\beta_1 \wedge \beta_2) \bowtie_{[\lambda,\lambda]} (\gamma_1 \wedge \gamma_2)$.

($\underline{\textbf{bet}}_\mathbf{9}$) It holds that $\Sigma \vdash \bot \to \beta \bowtie_\sigma \gamma$.

where we write $1 - \sigma$ as a shorthand for $\{1 - \lambda \mid \lambda \in \sigma\}$. The soundness and completeness of these inference rules w.r.t. the semantics from Section 5 is shown by the following proposition.

Proposition 9. *Assume that Σ satisfies* (39) *and* (43). *It holds that $\Sigma \vdash \alpha \to \beta \bowtie_\sigma \gamma$ can be derived from the inference rules* ($\underline{\textbf{bet}}_\mathbf{1}$)–($\underline{\textbf{bet}}_\mathbf{9}$) *iff set*$(\alpha) \subseteq \underline{\text{bet}}_\sigma($*set*$(\beta),$ *set*$(\gamma))$.

If Σ only satisfies (39) we still have that whenever $\Sigma \vdash \alpha \to \beta \bowtie_\sigma \gamma$ can be derived, we also have *set*$(\alpha) \subseteq \underline{bet}_\sigma($*set*$(\beta),$ *set*$(\gamma))$, i.e. inference is still sound. Conversely, to show that $\Sigma \vdash \alpha \to \beta \bowtie_\sigma \gamma$ can be derived whenever *set*$(\alpha) \subseteq \underline{bet}_\sigma($*set*$(\beta),$ *set*$(\gamma))$ holds, we only need (43).

[7] We implicitly do assume that $\sigma \neq \emptyset$ however.

Example 15. *Consider the rules about wines from Example 10, and assume that the attribute domains are given by*

$$
\begin{aligned}
A_1 &= \{\textit{beaujolais}, \textit{bardolino}, \textit{valpolicalla}, \dots\}\\
A_2 &= \{\textit{inexpensive}, \textit{mid-range}, \textit{above-average}, \textit{high-end}\}\\
A_3 &= \{\textit{low-tannins}, \textit{mid-tannins}, \textit{high-tannins}\}\\
A_4 &= \{\textit{light-body}, \textit{medium-body}, \textit{full-body}\}\\
A_5 &= \{\textit{light-red}, \textit{medium-red}, \textit{dark-red}, \textit{opaque}\}
\end{aligned}
$$

The assertions in Σ about A_1 are assumed to have been obtained using data-driven techniques, as was explained in Section 4.2. The assertions about A_2, A_3, A_4 and A_5 are obtained from the natural ranking of the labels, by assuming that they are all represented in a uni-dimensional domain. Recall that in uni-dimensional domains, the relations $\overline{bet_\sigma}$ and $\underline{bet_\sigma}$ coincide, as do $\overline{par_\sigma}$ and $\underline{par_\sigma}$. In the following, we will assume that A_1 and A_2 are orthogonal, and that A_3 and A_4 are orthogonal.

Using ($\underline{\mathbf{bet_1}}$), we find:

$$\Sigma \vdash \textit{barbera} \to \textit{chianti} \bowtie_{[0,1]} \textit{merlot} \tag{47}$$
$$\Sigma \vdash \textit{medium-body} \to \textit{light-body} \bowtie_{[0,1]} \textit{full-body} \tag{48}$$

and using ($\underline{\mathbf{bet_7}}$):

$$\Sigma \vdash (\textit{mid-tan} \wedge \textit{med-body}) \to (\textit{mid-tan} \wedge \textit{light-body}) \bowtie_{[0,1]} (\textit{mid-tan} \wedge \textit{full-body}) \tag{49}$$

Next, we consider formulas of the form $\beta \bowtie_\sigma \gamma \to \alpha$

($\overline{\mathbf{bet_1}}$) We always have $\Sigma \vdash a \bowtie_\sigma b \to \bigvee\{x \mid \overline{bet_\sigma}(a,x,b) \in \Sigma^*\}$.

($\overline{\mathbf{bet_2}}$) If $\alpha_1 \models_{\mathcal{A}} \alpha_2$, $\beta_2 \models_{\mathcal{A}} \beta_1$, $\gamma_2 \models_{\mathcal{A}} \gamma_1$ and $\sigma_2 \subseteq \sigma_1$, then whenever we have $\Sigma \vdash \beta_1 \bowtie_{\sigma_1} \gamma_1 \to \alpha_1$ we also have $\Sigma \vdash \beta_2 \bowtie_{\sigma_2} \gamma_2 \to \alpha_2$.

($\overline{\mathbf{bet_3}}$) If $\Sigma \vdash \beta \bowtie_\sigma \gamma \to \alpha$ then $\Sigma \vdash \gamma \bowtie_{1-\sigma} \beta \to \alpha$.

($\overline{\mathbf{bet_4}}$) If $\Sigma \vdash \beta_1 \bowtie_\sigma \gamma \to \alpha_1$, $\Sigma \vdash \beta_2 \bowtie_\sigma \gamma \to \alpha_2$ then $\Sigma \vdash (\beta_1 \vee \beta_2) \bowtie_\sigma \gamma \to \alpha_1 \vee \alpha_2$.

($\overline{\mathbf{bet_5}}$) If $\Sigma \vdash \beta_1 \bowtie_\sigma \gamma_1 \to \alpha_1$ and $\Sigma \vdash \beta_2 \bowtie_\sigma \gamma_2 \to \alpha_2$ then $\Sigma \vdash (\beta_1 \wedge \beta_2) \bowtie_\sigma (\gamma_1 \wedge \gamma_2) \to \alpha_1 \wedge \alpha_2$.

($\overline{\mathbf{bet_6}}$) It holds that $\Sigma \vdash \bot \bowtie_\sigma \gamma \rightarrow \alpha$.

Again we can show soundness and completeness.

Proposition 10. *Assume that Σ satisfies (40) and (44). It holds that $\Sigma \vdash \beta \bowtie_\sigma \gamma \rightarrow \alpha$ can be derived from the inference rules ($\overline{\mathbf{bet_1}}$)–($\overline{\mathbf{bet_6}}$) iff $\overline{\mathrm{bet}}_\sigma(set(\beta), set(\gamma)) \subseteq set(\alpha)$.*

Example 16. *Continuing the wine example, we obtain from ($\overline{\mathbf{bet_1}}$) that*

$$\Sigma \vdash \textit{mid-tannin} \bowtie_{[0,1]} \textit{low-tannin} \rightarrow \textit{low-tannin} \vee \textit{mid-tannin} \tag{50}$$
$$\Sigma \vdash \textit{mid-tannin} \bowtie_{[0,1]} \textit{mid-tannin} \rightarrow \textit{mid-tannin} \tag{51}$$
$$\Sigma \vdash \textit{medium-body} \bowtie_{[0,1]} \textit{medium-body} \rightarrow \textit{medium-body} \tag{52}$$

Combining (50) and (51) using ($\overline{\mathbf{bet_3}}$) and ($\overline{\mathbf{bet_4}}$), we find

$$\Sigma \vdash (\textit{low-tannin} \vee \textit{mid-tannin}) \bowtie_{[0,1]} \textit{low-tannin} \rightarrow (\textit{low-tannin} \vee \textit{mid-tannin}) \tag{53}$$

Combining (52) and (53), we find

$$\begin{aligned}\Sigma \vdash ((\textit{low-tan} \vee \textit{mid-tan}) \wedge \textit{med-bod}) \bowtie_{[0,1]} (\textit{low-tan} \wedge \textit{med-bod})\\ \rightarrow ((\textit{low-tan} \vee \textit{mid-tan}) \wedge \textit{med-bod})\end{aligned} \tag{54}$$

In a similar way, we can also consider expressions of the form $\delta \rightarrow \gamma \rhd_\tau \langle \alpha, \beta \rangle$:

($\underline{\mathbf{par}}_\mathbf{1}$) If $\underline{par_\tau}(a, b, c; d) \in \Sigma$ then $\Sigma \vdash d \rightarrow c \rhd_\tau \langle a, b \rangle$.

($\underline{\mathbf{par}}_\mathbf{2}$) If $\alpha_1 \models_\mathcal{A} \alpha_2$, $\beta_1 \models_\mathcal{A} \beta_2$, $\gamma_1 \models_\mathcal{A} \gamma_2$, $\delta_2 \models_\mathcal{A} \delta_1$ and $\tau_1 \subseteq \tau_2$, then whenever we have $\Sigma \vdash \delta_1 \rightarrow \gamma_1 \rhd_{\tau_1} \langle \alpha_1, \beta_1 \rangle$ we also have $\Sigma \vdash \delta_2 \rightarrow \gamma_2 \rhd_{\tau_2} \langle \alpha_2, \beta_2 \rangle$.

($\underline{\mathbf{par}}_\mathbf{3}$) If α is strongly consistent and $1 \in \tau$, we always have $\Sigma \vdash \beta \rightarrow \alpha \rhd_\tau \langle \alpha, \beta \rangle$.

($\underline{\mathbf{par}}_\mathbf{4}$) We always[8] have that $\Sigma \vdash \alpha \rightarrow \alpha \rhd_\tau \langle \alpha, \alpha \rangle$.

($\underline{\mathbf{par}}_\mathbf{5}$) If α is strongly consistent, we always have $\Sigma \vdash \beta \rightarrow \beta \rhd_\tau \langle \alpha, \alpha \rangle$.

[8] We implicitly do assume that $\tau \neq \emptyset$ however.

($\underline{\mathbf{par}}_6$) If α and β are strongly consistent and $0 \in \tau$, we always have $\Sigma \vdash \gamma \rightarrow \gamma \rhd_\tau \langle \alpha, \beta \rangle$.

($\underline{\mathbf{par}}_7$) If β is strongly consistent and $0 \in \tau$, we always have $\Sigma \vdash \alpha \rightarrow \alpha \rhd_\tau \langle \alpha, \beta \rangle$.

($\underline{\mathbf{par}}_8$) If α is strongly consistent and $0 \in \tau$, we always have $\Sigma \vdash \beta \rightarrow \beta \rhd_\tau \langle \alpha, \beta \rangle$.

($\underline{\mathbf{par}}_9$) If $\Sigma \vdash \delta_1 \rightarrow \gamma_1 \rhd_\tau \langle \alpha_1, \beta_1 \rangle$ and $\Sigma \vdash \delta_2 \rightarrow \gamma_2 \rhd_\tau \langle \alpha_2, \beta_2 \rangle$ then $\Sigma \vdash \delta_1 \vee \delta_2 \rightarrow (\gamma_1 \vee \gamma_2) \rhd_\tau \langle \alpha_1 \vee \alpha_2, \beta_1 \vee \beta_2 \rangle$.

($\underline{\mathbf{par}}_{10}$) If $\Sigma \vdash \delta \rightarrow \gamma \rhd_\tau \langle \alpha, \beta \rangle$, ϕ and ψ are strongly consistent, and the attribute domains in $\mathit{domains}(\alpha) \cup \mathit{domains}(\beta) \cup \mathit{domains}(\gamma) \cup \mathit{domains}(\delta)$ are orthogonal to those in $\mathit{domains}(\phi) \cup \mathit{domains}(\psi)$, then $\Sigma \vdash \delta \wedge \psi \rightarrow \gamma \wedge \psi \rhd_\tau \langle \alpha \wedge \phi, \beta \wedge \phi \rangle$.

($\underline{\mathbf{par}}_{11}$) If $\Sigma \vdash \delta_1 \rightarrow \gamma_1 \rhd_{[\lambda,\lambda]} \langle \alpha_1, \beta_1 \rangle$, $\Sigma \vdash \delta_2 \rightarrow \gamma_2 \rhd_{[\lambda,\lambda]} \langle \alpha_2, \beta_2 \rangle$, and the attribute domains in $\mathit{domains}(\alpha_1) \cup \mathit{domains}(\beta_1) \cup \mathit{domains}(\gamma_1) \cup \mathit{domains}(\delta_1)$ are orthogonal to those in $\mathit{domains}(\alpha_2) \cup \mathit{domains}(\beta_2) \cup \mathit{domains}(\gamma_2) \cup \mathit{domains}(\delta_2)$, then $\Sigma \vdash \delta_1 \wedge \delta_2 \rightarrow (\gamma_1 \wedge \gamma_2) \rhd_{[\lambda,\lambda]} \langle \alpha_1 \wedge \alpha_2, \beta_1 \wedge \beta_2 \rangle$.

($\underline{\mathbf{par}}_{12}$) It holds that $\Sigma \vdash \bot \rightarrow \gamma \rhd_\tau \langle \alpha, \beta \rangle$.

Proposition 11. *Assume that Σ satisfies (41) and (45). It holds that $\Sigma \vdash \delta \rightarrow \gamma \rhd_\tau \langle \alpha, \beta \rangle$ can be derived from the inference rules ($\underline{\mathbf{par}}_1$)–($\underline{\mathbf{par}}_{12}$) iff $\mathit{set}(\delta) \subseteq \underline{\mathrm{par}}_\tau(\mathit{set}(\alpha), \mathit{set}(\beta), \mathit{set}(\gamma))$.*

Example 17. *Using ($\underline{\mathbf{par}}_1$), we find*

$$\Sigma \vdash \mathit{bandol} \rightarrow \mathit{zinfandel} \rhd_{[0,+\infty[} \langle \mathit{barbera}, \mathit{barolo} \rangle \tag{55}$$
$$\Sigma \vdash \mathit{high\text{-}tannins} \rightarrow \mathit{mid\text{-}tannins} \rhd_{[0,+\infty[} \langle \mathit{low\text{-}tannins}, \mathit{mid\text{-}tannins} \rangle \tag{56}$$
$$\Sigma \vdash \mathit{full\text{-}body} \rightarrow \mathit{medium\text{-}body} \rhd_{[0,+\infty[} \langle \mathit{light\text{-}body}, \mathit{medium\text{-}body} \rangle \tag{57}$$

Using ($\underline{\mathbf{par}}_{10}$), we then obtain

$$\Sigma \vdash (\mathit{high\text{-}tan} \wedge \mathit{med\text{-}bod}) \rightarrow (\mathit{mid\text{-}tan} \wedge \mathit{med\text{-}bod}) \rhd_{[0,+\infty[} \langle (\mathit{low\text{-}tan} \wedge \mathit{med\text{-}bod}), (\mathit{mid\text{-}tan} \wedge \mathit{med\text{-}bod}) \rangle \tag{58}$$
$$\Sigma \vdash (\mathit{mid\text{-}tan} \wedge \mathit{full\text{-}bod}) \rightarrow (\mathit{mid\text{-}tan} \wedge \mathit{med\text{-}bod}) \rhd_{[0,+\infty[} \langle (\mathit{low\text{-}tan} \wedge \mathit{light\text{-}bod}), (\mathit{low\text{-}tan} \wedge \mathit{med\text{-}bod}) \rangle \tag{59}$$

Finally, we specify how formulas of the form $\gamma \rhd_\tau \langle \alpha, \beta \rangle \to \delta$ can be derived:

($\overline{\mathbf{par}}_\mathbf{1}$) We always have $\Sigma \vdash c \rhd_\tau \langle a, b\rangle \to \bigvee\{d \mid \overline{par_\tau}(a, b, c; d) \in \Sigma^*\}$.

($\overline{\mathbf{par}}_\mathbf{2}$) If $\alpha_2 \models_\mathcal{A} \alpha_1$, $\beta_2 \models_\mathcal{A} \beta_1$, $\gamma_2 \models_\mathcal{A} \gamma_1$, $\delta_1 \models_\mathcal{A} \delta_2$ and $\tau_2 \subseteq \tau_1$, then whenever we have $\Sigma \vdash \gamma_1 \rhd_{\tau_1} \langle \alpha_1, \beta_1\rangle \to \delta_1$ we also have $\Sigma \vdash \gamma_2 \rhd_{\tau_2} \langle \alpha_2, \beta_2\rangle \to \delta_2$.

($\overline{\mathbf{par}}_\mathbf{3}$) If $\Sigma \vdash \gamma \rhd_\tau \langle \alpha_1, \beta\rangle \to \delta_1$ and $\Sigma \vdash \gamma \rhd_\tau \langle \alpha_2, \beta\rangle \to \delta_2$ then $\Sigma \vdash \gamma \rhd_\tau \langle \alpha_1 \vee \alpha_2, \beta\rangle \to \delta_1 \vee \delta_2$.

($\overline{\mathbf{par}}_\mathbf{4}$) If $\Sigma \vdash \gamma \rhd_\tau \langle \alpha, \beta_1\rangle \to \delta_1$ and $\Sigma \vdash \gamma \rhd_\tau \langle \alpha, \beta_2\rangle \to \delta_2$ then $\Sigma \vdash \gamma \rhd_\tau \langle \alpha, \beta_1 \vee \beta_2\rangle \to \delta_1 \vee \delta_2$.

($\overline{\mathbf{par}}_\mathbf{5}$) If $\Sigma \vdash \gamma_1 \rhd_\tau \langle \alpha, \beta\rangle \to \delta_1$ and $\Sigma \vdash \gamma_2 \rhd_\tau \langle \alpha, \beta\rangle \to \delta_2$ then $\Sigma \vdash (\gamma_1 \vee \gamma_2) \rhd_\tau \langle \alpha, \beta\rangle \to \delta_1 \vee \delta_2$.

($\overline{\mathbf{par}}_\mathbf{6}$) If $\Sigma \vdash \gamma_1 \rhd_\tau \langle \alpha_1, \beta_1\rangle \to \delta_1$, $\Sigma \vdash \gamma_2 \rhd_\tau \langle \alpha_2, \beta_2\rangle \to \delta_2$ then $\Sigma \vdash (\gamma_1 \wedge \gamma_2) \rhd_\tau \langle \alpha_1 \wedge \alpha_2, \beta_1 \wedge \beta_2\rangle \to \delta_1 \wedge \delta_2$.

($\overline{\mathbf{par}}_\mathbf{7}$) It holds that $\Sigma \vdash \bot \rhd_\tau \langle \alpha, \beta\rangle \to \delta$.

($\overline{\mathbf{par}}_\mathbf{8}$) It holds that $\Sigma \vdash \gamma \rhd_\tau \langle \bot, \beta\rangle \to \delta$.

($\overline{\mathbf{par}}_\mathbf{9}$) It holds that $\Sigma \vdash \gamma \rhd_\tau \langle \alpha, \bot\rangle \to \delta$.

Proposition 12. *Assume that Σ satisfies* (42) *and* (46). *It holds that $\Sigma \vdash \gamma \rhd_\tau \langle \alpha, \beta\rangle \to \delta$ can be derived from the inference rules* ($\overline{\mathbf{par}}_\mathbf{1}$)–($\overline{\mathbf{par}}_\mathbf{9}$) *iff* $\overline{\mathrm{par}_\tau}(set(\alpha), set(\beta), set(\gamma)) \subseteq set(\alpha)$.

Example 18. *Using ($\overline{\mathbf{par}}_\mathbf{1}$), we find*

$$\Sigma \vdash \textit{mid-tan} \rhd_{[0,+\infty[} \langle \textit{low-tan}, \textit{high-tan}\rangle \to (\textit{mid-tan} \vee \textit{high-tan})$$
$$\Sigma \vdash \textit{high-tan} \rhd_{[0,+\infty[} \langle \textit{low-tan}, \textit{high-tan}\rangle \to \textit{high-tan}$$

Combining these two assertions using ($\overline{\mathbf{par}}_\mathbf{5}$) this yields

$$\Sigma \vdash (\textit{mid-tan} \vee \textit{high-tan}) \rhd_{[0,+\infty[} \langle \textit{low-tan}, \textit{high-tan}\rangle \to (\textit{mid-tan} \vee \textit{high-tan}) \tag{60}$$

In entirely the same fashion, we arrive at

$$\Sigma \vdash (\textit{mid-tan} \vee \textit{high-tan}) \rhd_{[0,+\infty[} \langle \textit{mid-tan}, \textit{high-tan}\rangle \to (\textit{mid-tan} \vee \textit{high-tan}) \tag{61}$$

Combining (60) *and* (61) *using* ($\overline{\mathbf{par_3}}$) *gives us*

$$\Sigma \vdash (\textit{mid-tan} \vee \textit{high-tan}) \rhd_{[0,+\infty[} \langle \textit{low-tan} \vee \textit{mid-tan}, \textit{high-tan}\rangle \\ \rightarrow (\textit{mid-tan} \vee \textit{high-tan}) \tag{62}$$

Furthermore, we find using ($\overline{\mathbf{par_1}}$)*:*

$$\Sigma \vdash \textit{med-bod} \rhd_{[0,+\infty[} \langle \textit{med-bod}, \textit{full-bod}\rangle \rightarrow (\textit{med-bod} \vee \textit{full-bod})$$
$$\Sigma \vdash \textit{full-bod} \rhd_{[0,+\infty[} \langle \textit{med-bod}, \textit{full-bod}\rangle \rightarrow \textit{full-bod}$$

which we can combine using ($\overline{\mathbf{par_5}}$) *to obtain:*

$$\Sigma \vdash (\textit{med-bod} \vee \textit{full-bod}) \rhd_{[0,+\infty[} \langle \textit{med-bod}, \textit{full-bod}\rangle \rightarrow (\textit{med-bod} \vee \textit{full-bod}) \tag{63}$$

We can then combine (62) *and* (63) *using* ($\overline{\mathbf{par_6}}$) *to conclude*

$$\Sigma \vdash ((\textit{mt} \vee \textit{ht}) \wedge (\textit{mb} \vee \textit{fb})) \rhd_{[0,+\infty[} \langle ((\textit{lt} \vee \textit{mt}) \wedge \textit{mb}), (\textit{ht} \wedge \textit{fb})\rangle \\ \rightarrow ((\textit{mt} \vee \textit{ht}) \wedge (\textit{mb} \vee \textit{fb})) \tag{64}$$

where we have further abbreviated the labels. Using ($\overline{\mathbf{par_1}}$) *we also find*

$$\Sigma \vdash \textit{dark-red} \rhd_{[0,+\infty[} \langle \textit{light-red}, \textit{medium-red}\rangle \rightarrow (\textit{dark-red} \vee \textit{opaque})$$
$$\Sigma \vdash \textit{opaque} \rhd_{[0,+\infty[} \langle \textit{light-red}, \textit{medium-red}\rangle \rightarrow \textit{opaque}$$
$$\Sigma \vdash \textit{dark-red} \rhd_{[0,+\infty[} \langle \textit{medium-red}, \textit{dark-red}\rangle \rightarrow (\textit{dark-red} \vee \textit{opaque})$$
$$\Sigma \vdash \textit{opaque} \rhd_{[0,+\infty[} \langle \textit{medium-red}, \textit{dark-red}\rangle \rightarrow \vee \textit{opaque}$$
$$\Sigma \vdash \textit{dark-red} \rhd_{[0,+\infty[} \langle \textit{medium-red}, \textit{opaque}\rangle \rightarrow (\textit{dark-red} \vee \textit{opaque})$$
$$\Sigma \vdash \textit{opaque} \rhd_{[0,+\infty[} \langle \textit{medium-red}, \textit{opaque}\rangle \rightarrow \textit{opaque}$$

Applying ($\overline{\mathbf{par_5}}$) *this leads to*

$$\Sigma \vdash (\textit{dark-red} \vee \textit{op}) \rhd_{[0,+\infty[} \langle \textit{light-red}, \textit{med-red}\rangle \rightarrow (\textit{dark-red} \vee \textit{op}) \tag{65}$$
$$\Sigma \vdash (\textit{dark-red} \vee \textit{op}) \rhd_{[0,+\infty[} \langle \textit{med-red}, \textit{dark-red}\rangle \rightarrow (\textit{dark-red} \vee \textit{op}) \tag{66}$$
$$\Sigma \vdash (\textit{dark-red} \vee \textit{op}) \rhd_{[0,+\infty[} \langle \textit{med-red}, \textit{op}\rangle \rightarrow (\textit{dark-red} \vee \textit{op}) \tag{67}$$

and after applying ($\overline{\mathbf{par_4}}$)*, we also find*

$$\Sigma \vdash (\textit{dark-red} \vee \textit{opaque}) \rhd_{[0,+\infty[} \langle \textit{medium-red}, (\textit{dark-red} \vee \textit{opaque})\rangle \\ \rightarrow (\textit{dark-red} \vee \textit{opaque}) \tag{68}$$

We are now ready to show the soundness and completeness of the inference rules proposed in Section 3.1 w.r.t. the semantics introduced in Section 5.

Proposition 13. *Let R, Σ, $\mathcal{B}$ and $\mathcal{C}$ be defined as before, and let $X \subseteq \mathcal{B}$ and $Y \subseteq \mathcal{C}$. It holds that*

$$(R, \Sigma) \vdash \Big(\bigvee_{(x_1,\ldots,x_s)\in X} \bigwedge_i x_i\Big) \rightarrow \Big(\bigvee_{(y_1,\ldots,y_k)\in Y} \bigwedge_i y_i\Big)$$

can be derived from ***(S)***, ***(I')***, ***(E')***, ***(D)***, $(\underline{\mathbf{bet_1}})$–$(\underline{\mathbf{bet_9}})$, $(\overline{\mathbf{bet_1}})$–$(\overline{\mathbf{bet_6}})$, $(\underline{\mathbf{par}_1})$–$(\underline{\mathbf{par}_{12}})$ *and* $(\overline{\mathbf{par}_1})$–$(\overline{\mathbf{par}_9})$ *iff*

$$\widehat{f_R}(X) \subseteq Y$$

Example 19. *We can now formalize the inference about wines from Example 10. Using* ***(S)***, *we find*

$$(R, \Sigma) \vdash \textit{chianti} \rightarrow \textit{low-tan} \wedge \textit{med-body} \tag{69}$$
$$(R, \Sigma) \vdash \textit{merlot} \rightarrow (\textit{low-tan} \vee \textit{mid-tan}) \wedge \textit{med-body} \tag{70}$$
$$(R, \Sigma) \vdash \textit{barolo} \rightarrow \textit{high-tan} \wedge \textit{full-body} \tag{71}$$
$$(R, \Sigma) \vdash \textit{zinfandel} \rightarrow (\textit{mid-tan} \vee \textit{high-tan}) \wedge (\textit{med-body} \vee \textit{full-body}) \tag{72}$$
$$(R, \Sigma) \vdash \textit{low-tan} \wedge \textit{light-body} \rightarrow \textit{light-red} \tag{73}$$
$$(R, \Sigma) \vdash \textit{low-tan} \wedge \textit{med-body} \rightarrow \textit{med-red} \tag{74}$$
$$(R, \Sigma) \vdash \textit{mid-tan} \wedge \textit{med-body} \rightarrow \textit{dark-red} \vee \textit{opaque} \tag{75}$$
$$(R, \Sigma) \vdash \textit{high-tan} \wedge \textit{full-body} \rightarrow \textit{opaque} \tag{76}$$

Combining (47), (54), (69) *and* (70), *we find using* ***(I')***:

$$(R, \Sigma) \vdash \textit{barbera} \rightarrow (\textit{low-tan} \vee \textit{mid-tan}) \wedge \textit{med-body} \tag{77}$$

If we then combine (77) *with* (71), (72), (55) *and* (64), *using* ***(E')*** *gives us*

$$(R, \Sigma) \vdash \textit{bandol} \rightarrow ((\textit{mid-tan} \vee \textit{high-tan}) \wedge (\textit{med-body} \vee \textit{full-body})) \tag{78}$$

Combining (73), (74), (75), (59) *and* (65), *we get using* ***(E')***

$$(R, \Sigma) \vdash \textit{mid-tan} \wedge \textit{full-body} \rightarrow \textit{dark-red} \vee \textit{opaque} \tag{79}$$

Similarly, combining (74), (75), (58) *and* (68)

$$(R, \Sigma) \vdash \textit{high-tan} \wedge \textit{med-body} \rightarrow \textit{dark-red} \vee \textit{opaque} \tag{80}$$

*Repeatedly applying **(D)** to* (75), (76), (79) *and* (81) *yields*

$$(R, \Sigma) \vdash (\textit{mid-tan} \vee \textit{high-tan}) \wedge (\textit{med-body} \vee \textit{full-body}) \rightarrow \textit{dark-red} \vee \textit{opaque} \tag{81}$$

which together with (78)*, gives us*

$$(R, \Sigma) \vdash \textit{bandol} \rightarrow \textit{dark-red} \vee \textit{opaque}$$

*again using **(D)**. Despite that nothing could be concluded about bandol wine using classical deduction, using a combination of interpolative and extrapolative reasoning, we have found that its colour is either dark-red or opaque.*

7. Complexity and implementation

In this section, we first show that interpolative and extrapolative reasoning is PSPACE-hard in general. We then show in Section 7.4 that the complexity crucially depends on the number of attribute domains. In particular, we show that implementations in polynomial time are possible when the number of attribute domains is small enough to be treated as a constant (without placing any bounds on the number of attributes or the number of rules). From this result, it also follows that the inference problem considered in this paper is decidable in EXPTIME. However, whether this problem is also in PSPACE remains currently open.

7.1. Hardness

We prove that interpolative and extrapolative reasoning are PSPACE-hard, by showing a reduction from the dominance problem for CP-nets [53]. We present two such reductions, one which is only based on interpolative reasoning and one which is only based on extrapolative reasoning.

Our terminology and notations are based on the presentation of CP-nets in [54], where a generalization of binary CP-nets is considered. The basic building blocks are *conditional preference rules*, which are expressions of the form $p : x_i > \neg x_i$ (or $p : \neg x_i > x_i$) with p a propositional formula over a set of atoms $V = \{x_1, ..., x_n\}$ and x_i an atom from V. The intuitive meaning is that in situations where p holds, having x_i true is preferred to having x_i false. A (binary) *GCP-net* (generalized ceteris paribus net) over V is a set of such conditional preference rules. An *outcome* is an n-tuple $(I(x_1), ..., I(x_n))$ where I maps each variable to a value from $\{\textit{true}, \textit{false}\}$, i.e. an outcome corresponds to a propositional interpretation. For the ease of presentation, we will identify an outcome with its corresponding mapping I. Let I and J be two outcomes, which differ only in the value of one variable x_i, such that $I(x_i) = \textit{true}$ and $J(x_i) = \textit{false}$. If a GCP-net N contains a rule $p : x_i > \neg x_i$ (resp. $p : \neg x_i > x_i$) such that the propositional formula p is satisfied by both I and J, we say that N sanctions an *improving flip* from J to I (resp. from I to J). Finally, we say that I *dominates* J, given N, if there is a sequence $J = J_0, J_1, ..., J_m = I$ such that N sanctions an improving flip from J_i to J_{i+1} for each i in $\{0, ..., m-1\}$. The problem of deciding, given two outcomes I and J, whether I dominates J is PSPACE-complete [54].

7.2. Reduction to interpolative reasoning

Given a GCP-net N and two outcomes I and J we now construct a rule base R and a set of assertions Σ such that I dominates J iff a particular rule can be derived from R using interpolative reasoning. We consider an attribute domain $Y = \{y_0, y_1, y_2\}$, and one additional attribute domain X_i for each variable $x_i \in V$, which is defined as

$$X_i = \{x_i, \overline{x_i}\} \cup \{x_i^p \mid (p : x_i > \neg x_i) \in N\} \cup \{\overline{x_i}^p \mid (p : \neg x_i > x_i) \in N\}$$

The set of rules R is obtained by adding for each conditional preference rule of the form $p : x_i > \neg x_i$ in N the rule

$$p \wedge x_i^p \rightarrow y_2 \tag{82}$$

and for each conditional preference rule of the form $p : \neg x_i > x_i$, we add to R the rule

$$p \wedge \overline{x_i}^p \rightarrow y_2 \tag{83}$$

Finally, we also add the rule

$$\bigwedge\{x_i \mid J(x_i) = \textit{true}, 1 \leq i \leq n\} \wedge \bigwedge\{\overline{x_i} \mid J(x_i) = \textit{false}, 1 \leq i \leq n\} \rightarrow y_1 \tag{84}$$

The underlying idea is combine rules of the form (84) with rules of the form (82) or (83) using inference rule **(I')** to simulate the idea of an improving flip. The antecedents of each of the newly generated rules will correspond to outcomes (where a conjunct x_i appears if x_i is *true* in the outcome, and a conjunct $\overline{x_i}$ appears otherwise). To this end, for each conditional preference rule of the form $p : x_i > \neg x_i$ we add the assertion $\underline{bet}_{[0,1]}(\overline{x_i}, x_i, x_i^p)$ to Σ and for each conditional preference rule of the form $p : \neg x_i > x_i$ we add the assertion $\underline{bet}_{[0,1]}(x_i, \overline{x_i}, \overline{x_i}^p)$. We also add the assertion $\overline{bet}_{[0,1]}(y_0, y_1, y_2)$.

Proposition 14. *Let I, J, N, R, and Σ be as above. It holds that I dominates J, given N, iff*

$$(R, \Sigma) \vdash \bigwedge\{x_i \mid I(x_i) = \mathit{true}, 1 \leq i \leq n\} \wedge \bigwedge\{\overline{x_i} \mid I(x_i) = \mathit{false}, 1 \leq i \leq n\} \rightarrow y_1 \vee y_2 \tag{85}$$

7.3. Reduction to extrapolative reasoning

We now show that the idea of improving flips can also be simulated using extrapolative reasoning, i.e. using assertions about the relations $\underline{par}$ and $\overline{par}$ in Σ. With each variable x_i we now associate an attribute domain $X_i = \{x_i, \overline{x_i}, x_i', \overline{x_i'}\}$. We consider one additional attribute domain $Y = \{y_0, y_1, y_2, y^-, y^+\}$ whose elements will again appear in the consequent of rules. Intuitively, y_1 is the degree to which outcome J is preferred, y_0 is a lower degree of preference and y_2 is a higher degree of preference. Furthermore, y^- represents a lower degree of preference than y^+, but the relation between y^- and y^+ on the one hand, and y_0, y_1 and y_2 on the other hand will remain unspecified. For each conditional preference rule $p : x_i > \neg x_i$ in N, we add the following two rules to R

$$p \wedge x_i' \rightarrow y^+ \qquad\qquad p \wedge \overline{x_i'} \rightarrow y^-$$

The idea here is to indicate that the direction from having x_i false to having x_i true is towards more preferred outcomes when p is true. By using the attributes x_i' and $\overline{x_i'}$, rather than x_i and $\overline{x_i}$ we can talk about the effect of changing the value of x_i without the need to specify to what degree $p \wedge x_i$ and $p \wedge \overline{x_i}$ are preferred. Similarly, for each preference rule of the form $p : \neg x_i > x_i$, we add

$$p \wedge \overline{x_i'} \rightarrow y^+ \qquad\qquad p \wedge x_i' \rightarrow y^-$$

Finally, we add the rule

$$\bigwedge\{x_i \mid J(x_i) = \mathit{true}, 1 \leq i \leq n\} \wedge \bigwedge\{\overline{x_i} \mid J(x_i) = \mathit{false}, 1 \leq i \leq n\} \rightarrow y_1$$

For each i, we add the assertions $\underline{par_{[0,+\infty[}}(x_i', \overline{x_i'}, x_i; \overline{x_i})$ and $\underline{par_{[0,+\infty[}}(\overline{x_i'}, x_i', \overline{x_i}; x_i)$ to Σ. In addition, we add

$$
\begin{array}{ll}
\overline{par_{[0,+\infty[}}(y_0, y_1, y_1; y_2) & \overline{par_{[0,+\infty[}}(y_2, y_1, y_1; y_0) \\
\overline{par_{[0,+\infty[}}(y^-, y^+, y_0; y_1) & \overline{par_{[0,+\infty[}}(y^-, y^+, y_0; y_2) \\
\overline{par_{[0,+\infty[}}(y^-, y^+, y_1; y_2) & \overline{par_{[0,+\infty[}}(y^+, y^-, y_2; y_1) \\
\overline{par_{[0,+\infty[}}(y^+, y^-, y_2; y_0) & \overline{par_{[0,+\infty[}}(y^+, y^-, y_1; y_0)
\end{array}
$$

Proposition 15. *Let I, J, N, R, and Σ be as above. It holds that I dominates J, given N, iff*

$$(R, \Sigma) \vdash \bigwedge\{x_i \mid I(x_i) = \mathit{true}, 1 \leq i \leq n\} \wedge \bigwedge\{\overline{x_i} \mid I(x_i) = \mathit{false}, 1 \leq i \leq n\} \rightarrow y_1 \vee y_2$$

Corollary 1. *The problem of deciding whether $(R, \Sigma) \vdash \alpha \rightarrow \beta$ is PSPACE-hard, even if either all betweenness information in Σ is trivial, or all information about parallelism in Σ is trivial.*

7.4. Implementation

In propositional logic, the number of possible interpretations is exponential in the number of atoms. In the present setting, on the other hand, the number of interpretations strongly depends on the number of attribute domains. If the number of attribute domains is small compared to the total number of atoms, the number of interpretations is essentially polynomial. For example, if there are only two attribute domains A_1 and A_2, each of which contains $\frac{n}{2}$ atoms, then there are $\left(\frac{n}{2}\right)^2$ interpretations. In such cases, it makes sense to rely on implementation methods that operate at the semantic level, even if that requires an enumeration of all interpretations.

Consider a rule base R, where the antecedents of rules are built from the atoms in $B_1, ..., B_s$ and the consequents are built from $C_1, ..., C_k$. From Proposition 13, we know that we can fully characterize interpolation and extrapolation on R by specifying the value of $\widehat{f_R}$ for each element $(x_1, ..., x_s)$ of $\mathcal{B} = B_1 \times ... \times B_s$. Moreover, from (38), we know that the function $\widehat{f_R}$ can be obtained in an iterative fashion, although the formulation in (38) cannot be used directly, as it involves intersections that range over arbitrary subsets of interpretations (of which there are at least exponentially many) and arbitrary subsets of $[0, 1]$ and $[0, +\infty[$. The following proposition suggests a way to evaluate the right-hand side of (38) in practice.

Proposition 16. *Let* $\mathbf{x} = (x_1, ..., x_s)$*. It holds that*

$$\begin{aligned}\widehat{f_R}^{(i+1)}(\{\mathbf{x}\}) =& \widehat{f_R}^{(i)}(\{\mathbf{x}\}) \\ & \cap \bigcap \{\overline{bet_\sigma}(\widehat{f_R}^{(i)}(set(\alpha)), \widehat{f_R}^{(i)}(set(\gamma))) \mid c_1 \vee c_2\} \\ & \cap \bigcap \{\overline{par_\tau}(\widehat{f_R}^{(i)}(set(\alpha)), \widehat{f_R}^{(i)}(set(\beta)), \widehat{f_R}^{(i)}(set(\gamma))) \mid c_3 \vee c_4\}\end{aligned} \tag{86}$$

where

c_1 *iff* $\alpha = x_{i_1} \wedge ... \wedge x_{i_{j-1}} \wedge a_{i_j} \wedge x_{i_{j+1}} \wedge ... \wedge x_{i_r}$
and $\gamma = x_{i_1} \wedge ... \wedge x_{i_{j-1}} \wedge c_{i_j} \wedge x_{i_{j+1}} \wedge ... \wedge x_{i_r}$
and $\Sigma \vdash x_{i_j} \to a_{i_j} \bowtie_\sigma c_{i_j}$ *with* $B_{i_1}, ..., B_{i_r}$ *all orthogonal*

c_2 *iff* $\alpha = a_{i_1} \wedge ... \wedge a_{i_r}$ *and* $\gamma = c_{i_1} \wedge ... \wedge c_{i_r}$
and $\exists \lambda \in [0, 1] \,.\, \forall j \in \{1, ..., r\} \,.\, \Sigma \vdash x_{i_j} \to a_{i_j} \bowtie_{[\lambda,\lambda]} c_{i_j}$
with $B_{i_1}, ..., B_{i_r}$ *all orthogonal*

c_3 *iff* $\alpha = y_{i_1} \wedge ... \wedge y_{i_{j-1}} \wedge a_{i_j} \wedge y_{i_{j+1}} \wedge ... \wedge y_{i_r}$
and $\beta = y_{i_1} \wedge ... \wedge y_{i_{j-1}} \wedge b_{i_j} \wedge y_{i_{j+1}} \wedge ... \wedge y_{i_r}$
and $\gamma = x_{i_1} \wedge ... \wedge x_{i_{j-1}} \wedge c_{i_j} \wedge x_{i_{j+1}} \wedge ... \wedge x_{i_r}$
and $\Sigma \vdash x_{i_j} \to c_{i_j} \rhd_\tau \langle a_{i_j}, b_{i_j} \rangle$
with $B_{i_1}, ..., B_{i_r}$ *all orthogonal*

c_4 *iff* $\alpha = a_{i_1} \wedge ... \wedge a_{i_r}$ *and* $\beta = b_{i_1} \wedge ... \wedge b_{i_r}$ *and* $\gamma = c_{i_1} \wedge ... \wedge c_{i_r}$
and $\exists \mu \in [0, +\infty[\,.\, \forall j \in \{1, ..., r\} \,.\, \Sigma \vdash x_{i_j} \to c_{i_j} \rhd_{[\mu,\mu]} \langle a_{i_j}, b_{i_j} \rangle$
with $B_{i_1}, ..., B_{i_r}$ *all orthogonal*

Note in particular how Proposition 16 allows us to replace the range of the intersection over arbitrary subsets of $\mathcal{B}$ to a range over conjunctions α, β and γ, of which there are polynomially many if the number of attribute domains is treated as a constant. Moreover, note how for each choice of these conjunctions, there is only one minimal choice for σ or τ, which can easily be found from Σ. This leads to the following procedure to characterize the function $\widehat{f_R}$ at the semantic level: repeat the following until $\widehat{f_R}^{(i)} = \widehat{f_R}^{(i-1)}$, starting with $i = 1$, for each $\mathbf{x} \in \mathcal{B}$:

1. set $S \leftarrow \widehat{f_R}^{(i-1)}(\{\mathbf{x}\})$;
2. for each non-empty subset $\{i_1, ..., i_r\}$ such that $B_{i_1}, ..., B_{i_r}$ are all orthogonal:

(a) for each $j \in \{1, ..., r\}$ and for each a_{i_j} and c_{i_j} in B_{i_j} such that $\underline{bet_\sigma}(a_{i_j}, x_{i_j}, c_{i_j}) \in \Sigma$ for some $\sigma \subseteq [0,1]$:
   i. set $\alpha \leftarrow x_{i_1} \wedge ... \wedge x_{i_{j-1}} \wedge a_{i_j} \wedge x_{i_{j+1}} \wedge ... \wedge x_{i_r}$
   ii. set $\gamma \leftarrow x_{i_1} \wedge ... \wedge x_{i_{j-1}} \wedge c_{i_j} \wedge x_{i_{j+1}} \wedge ... \wedge x_{i_r}$
   iii. set $S \leftarrow S \cap \overline{bet_\sigma}(\widehat{f_R}^{(i-1)}(set(\alpha)), \widehat{f_R}^{(i-1)}(set(\gamma)))$

(b) for each $\lambda \in [0,1]$ for which Σ contains relations of the form $\underline{bet_{[\lambda,\lambda]}}$, and each $a_{i_1}, ..., a_{i_r}, c_{i_1}, ..., c_{i_r}$ such that for each j it holds that $\underline{bet_{[\lambda,\lambda]}}(a_{i_j}, x_{i_j}, c_{i_j}) \in \Sigma$ or $\underline{bet_{[\lambda,\lambda]}}(c_{i_j}, x_{i_j}, a_{i_j}) \in \Sigma$ or $a_{i_j} = x_{i_j} = c_{i_j}$:
   i. set $\alpha \leftarrow a_{i_1} \wedge ... \wedge a_{i_r}$;
   ii. set $\gamma \leftarrow c_{i_1} \wedge ... \wedge c_{i_r}$;
   iii. set $S \leftarrow S \cap \overline{bet_{[\lambda,\lambda]}}(\widehat{f_R}^{(i-1)}(set(\alpha)), \widehat{f_R}^{(i-1)}(set(\gamma)))$;

(c) for each $j \in \{1, ..., r\}$, for each $y_{i_1}, ..., y_{i_{j-1}}, y_{i_{j+1}}, ..., y_{i_r}$ and for each a_{i_j}, b_{i_j} and c_{i_j} in B_{i_j} such that $\underline{par_\tau}(a_{i_j}, b_{i_j}, c_{i_j}; x_{i_j}) \in \Sigma$ for some $\tau \subseteq [0, +\infty[$:
   i. set $\alpha \leftarrow y_{i_1} \wedge ... \wedge y_{i_{j-1}} \wedge a_{i_j} \wedge y_{i_{j+1}} \wedge ... \wedge y_{i_r}$;
   ii. set $\beta \leftarrow y_{i_1} \wedge ... \wedge y_{i_{j-1}} \wedge b_{i_j} \wedge y_{i_{j+1}} \wedge ... \wedge y_{i_r}$;
   iii. set $\gamma \leftarrow x_{i_1} \wedge ... \wedge x_{i_{j-1}} \wedge c_{i_j} \wedge x_{i_{j+1}} \wedge ... \wedge x_{i_r}$;
   iv. set $S \leftarrow S \cap \overline{par_\tau}(\widehat{f_R}^{(i-1)}(set(\alpha)), \widehat{f_R}^{(i-1)}(set(\beta)), \widehat{f_R}^{(i-1)}(set(\gamma)))$;

(d) for each $\mu \in [0, +\infty[$ for which Σ contains relations of the form $\underline{par_{[\mu,\mu]}}$, and each $a_{i_1}, ..., a_{i_r}$, $b_{i_1}, ..., b_{i_r}$, $c_{i_1}, ..., c_{i_r}$ such that for each j it holds that $\underline{par_{[\mu,\mu]}}(a_{i_j}, b_{i_j}, c_{i_j}; x_{i_j})$ or $a_{i_j} = b_{i_j}$ and $c_{i_j} = x_{i_j}$:
   i. set $\alpha \leftarrow a_{i_1} \wedge ... \wedge a_{i_r}$;
   ii. set $\alpha \leftarrow b_{i_1} \wedge ... \wedge b_{i_r}$;
   iii. set $\gamma \leftarrow c_{i_1} \wedge ... \wedge c_{i_r}$;
   iv. set $S \leftarrow S \cap \overline{par_{[\mu,\mu]}}(\widehat{f_R}^{(i-1)}(set(\alpha)), \widehat{f_R}^{(i-1)}(set(\beta)), \widehat{f_R}^{(i-1)}(set(\gamma)))$;

3. set $\widehat{f_R}^{(i)}(\{\mathbf{x}\}) \leftarrow S$.

Note that in step 2(a), it suffices to check whether $\underline{bet_\sigma}(a_{i_j}, x_{i_j}, c_{i_j}) \in \Sigma$ to verify $\Sigma \vdash x_{i_j} \rightarrow a_{i_j} \bowtie_\sigma c_{i_j}$, as all other cases where $\Sigma \vdash x_{i_j} \rightarrow a_{i_j} \bowtie_\sigma c_{i_j}$ holds are covered in one way or another. Indeed, when $\underline{bet_{\sigma'}}(a_{i_j}, x_{i_j}, c_{i_j}) \in \Sigma$, for some $\sigma' \subset \sigma$, we have

$$\overline{bet_{\sigma'}}(\widehat{f_R}^{(i-1)}(set(\alpha)), \widehat{f_R}^{(i-1)}(set(\gamma))) \subseteq \overline{bet_\sigma}(\widehat{f_R}^{(i-1)}(set(\alpha)), \widehat{f_R}^{(i-1)}(set(\gamma)))$$

so we do not need to consider $\underline{bet_\sigma}$ if we already consider $\underline{bet_{\sigma'}}$. The case where $\underline{bet_\sigma}(c_{i_j}, x_{i_j}, a_{i_j}) \in \Sigma$ is handled by swapping the definitions of α and

γ. Finally, the cases where $x_{i_j} = a_{i_j}$ or $x_{i_j} = c_{i_j}$ do not need to be considered because then $\mathbf{x} \in set(\alpha)$ and $0 \in \sigma$, $\mathbf{x} \in set(\gamma)$ and $1 \in \sigma$ or $x_{i_j} = a_{i_j} = c_{i_j}$; in each of these cases we have $\overline{bet_\sigma}(\widehat{f_R}^{(i-1)}(set(\alpha)), \widehat{f_R}^{(i-1)}(set(\gamma))) \subseteq \widehat{f_R}^{(i-1)}(\{\mathbf{x_i}\})$. Similar considerations apply to (b)–(d).

It is clear that each of the steps (a)–(d) is polynomial in the size of $\mathcal{B}$, as is the number of arguments $\mathbf{x}$ for which these steps have to be completed. Finally, after each iteration, there is at least one element $\mathbf{x}$ from $\mathcal{B}$ for which $\widehat{f_R^{(i)}}(\{\mathbf{x}\}) \subset \widehat{f_R^{(i-1)}}(\{\mathbf{x}\})$, unless $\widehat{f_R^{(i)}} = \widehat{f_R}$ after which we can stop. This means that the total number of iterations is upper bounded by $\mathcal{B} \times \mathcal{C}$, and is in particular polynomial in the size of $\mathcal{A}$. This means that the above procedure runs in polynomial time if the number of attribute domains n is small enough to be treated as a constant. In other words, interpolative and extrapolative reasoning is decidable in exponential time, and is polynomial in data complexity.

Corollary 2. *Let R, Σ, $A_1, ..., A_n$, $B_1, ..., B_s$ and $C_1, ..., C_k$ be as before. Let β and γ be propositional formulas such that* $\text{domains}(\beta) \subseteq \{B_1, ..., B_s\}$ *and* $\text{domains}(\gamma) \subseteq \{C_1, ..., C_k\}$. *The problem of deciding whether*

$$(R, \Sigma) \vdash \beta \rightarrow \gamma$$

is in EXPTIME. If the number of attribute domains n is upper bounded by a constant, this problem is in P.

The restriction to have a relatively small number of attribute domains is a natural one in many application contexts. For example, interpolation is often applied to sets of parallel if-then rules, in which case the antecedent of every rule is built from the same attribute domains, and the consequent is a single atom, taken from a fixed attribute domain. In such a case it is not common to have more than a few attribute domains. As another example, consider a rule base $R = R_1 \cup R_2$, such that the antecedents of rules in R_1 are built from the attribute domains $B_1, ..., B_k$, the consequents of rules in R_1 and antecedents of rules in R_2 are built from the attribute domains $C_1, ..., C_s$, and the consequents of rules in R_2 are built from the attribute domains $D_1, ..., D_l$, such that $\{B_1, ..., B_k\} \cap \{D_1, ..., D_l\} = \emptyset$. In many situations, refining R_1 and R_2 separately would be equivalent to refining R as a whole. The following counterexample, however, shows that this is not the case in general.

Example 20. *Let $R = R_1 \cup R_2$ with $R_1 = \{a_1 \to a_2, c_1 \to c_2\}$ and $R_2 = \{a_2 \to a_3, c_2 \to c_3\}$, where the attribute domains are $A_1 = \{a_1, b_1, c_1, d_1\}$, $A_2 = \{a_2, b_2, c_2, d_2\}$ and $A_3 = \{a_3, b_3, c_3, d_3\}$. Assume furthermore that Σ only contains the assertions $\underline{bet}_{[0.5,0.5]}(a_1, b_1, c_1)$, $\overline{bet}_{[0.5,0.5]}(a_2, b_2, c_2)$ and $\overline{bet}_{[0.5,0.5]}(a_3, b_3, c_3)$. Since $R \models a_1 \to c_1$ and $R \models a_3 \to c_3$, using **(S)** and **(I')**, we find $(R, \Sigma) \vdash b_1 \to b_3$. However, using R_1 only trivial information can be derived about b_1, hence refining R_1 and R_2 separately does not allow us to derive anything about b_1.*

8. Discussion

As explained in Section 4.3, interpolative and extrapolative inference should provide sound conclusions as long as the rule base can be seen as the approximation of a mapping m which is linear, or in the case of interpolation, monotonic. In practice, however, this assumption may not be valid, in which case inconsistencies can be introduced by our method. In such cases, interpolative and extrapolative reasoning could still prove useful, although it should be applied more cautiously.

Relaxing the linearity assumption

Consider the following rules, which contain information about the amount of traffic (*light*, *moderate*, *heavy*) at different times during the day:

$$\textit{morning} \to \textit{heavy-traffic} \tag{87}$$
$$\textit{mid-day} \to \textit{moderate-traffic} \tag{88}$$
$$\textit{evening} \to \textit{heavy-traffic} \tag{89}$$

Using **(I)** and the assumption that

$$\textit{mid-day} \to \textit{morning} \bowtie \textit{evening}$$

we then derive the rule

$$\textit{mid-day} \to \textit{heavy-traffic}$$

which is in conflict with (88). This can be explained due to a failure of the monotonicity assumption. In the case of (87)–(89) the underlying mapping is not even deterministic, in the sense that the exact amount of traffic at e.g. 9 am may vary from day to day (even if we assume that the rule base talks about weekdays in a specific city). Nonetheless, even for rules where the linearity assumption fails, interpolation may still be useful. For instance, suppose we introduce the labels *mid-morning* and *mid-afternoon*, which are between *morning* and *mid-day*, and between *mid-day* and *evening* respectively. From (87) and (88) we may derive

$$\textit{mid-morning} \rightarrow \textit{moderate-traffic} \vee \textit{heavy-traffic}$$

Indeed, while the mapping underlying the rule base may, in principle, be arbitrary, it seems natural to assume that more regular mappings would be more likely, i.e. *we could make the assumption that any completion of the knowledge base should not introduce additional irregularities.* In particular, by identifying irregularities with violations of the monotonicity assumption, this leads to the assumption that the conceptual space $\mathcal{C}_1$ corresponding with the antecedent of the rules can be partitioned in a minimal number of segments, such that the mapping is monotonic over these segments. In the traffic example, we would thus assume that the amount of traffic is monotonically decreasing throughout the morning and monotonically increasing throughout the afternoon. While such conclusions would not be valid in general, they are reasonable to make in absence of any other information. Depending on how the rule base (87)–(89) was obtained, we may also argue that the absence of a rule for *mid-morning* suggests that this case is not special, i.e. that those cases which are irregular in some sense would be more likely to be contained in the rule base.

To avoid inconsistencies, the above view suggests that from a rule base R we should try to identify subsets $R_1, ...R_k$ of rules, such that no inconsistencies arise as long as interpolation is applied to two rules from the same set R_i. To be compatible with the above view, we should moreover insist that when $\alpha \rightarrow \alpha_1 \bowtie \alpha_2$, $(\alpha_1 \rightarrow \beta_1) \in R_i$, $(\alpha_2 \rightarrow \beta_2) \in R_i$ and $(\alpha \rightarrow \beta) \in R$, then we should have that $(\alpha \rightarrow \beta) \in R_i$. In other words, the sub-bases R_i should contain all rules that apply to a given (convex) segment of the conceptual space $\mathcal{C}_1$. In this way, we can ensure that when a new rule $\alpha^* \rightarrow \beta^*$ is derived by interpolation from a sub-base R_i, the rules in R_i are indeed the most relevant ones, i.e. that they are the ones whose antecedent is closest to α^* in some sense. In a similar, but slightly less cautious fashion, we may assume that the mapping underlying the rule base R is piecewise linear, and apply extrapolation locally to the sub-bases $R_1, ..., R_k$.

Restricting to the most salient properties

Another reason why inconsistencies may arise is because the information about betweenness or analogical change is not accurate, or, more fundamentally, because it only takes the most salient properties of objects in the account. For example, when we derive betweenness information for wines from wine-food pairings, it will mainly reflect the taste of the wine, and to a much lesser extent properties such as price. As an additional example, we may consider that *coffeehouses* are conceptually between *bars* and *restaurants*, as both *coffeehouses* and *bars* emphasise drinking rather than eating, while *coffeehouses* generally do serve some food (sandwiches, cakes) as well. Nonetheless, we may consider that

$$\mathit{bar} \rightarrow \mathit{serves\text{-}wine} \tag{90}$$

$$\mathit{coffeehouse} \rightarrow \neg \mathit{serves\text{-}wine} \tag{91}$$

$$\mathit{restaurant} \rightarrow \mathit{serves\text{-}wine} \tag{92}$$

Using interpolation and the assumption

$$\mathit{coffeehouse} \rightarrow \mathit{bar} \bowtie \mathit{restaurant}$$

we derive the rule

$$\mathit{coffeehouse} \rightarrow \mathit{serves\text{-}wine}$$

which is in conflict with the rule base. In this case, the inconsistency is mainly due the fact that the property of serving wine was not considered when asserting that coffeehouses are between bars and restaurants. The most natural way to avoid inconsistencies would then be to avoid applying interpolation to derive conclusions from the domain $A = \{\mathit{serves\text{-}wine}, \neg \mathit{serves\text{-}wine}\}$. In absence of any conflicts about attributes from a given domain, we then assume that interpolative and extrapolative conclusions are valid for that domain, an assumption which may need to be revised if additional knowledge became available.

9. Conclusions and future work

The aim of this paper was to study the core principles underlying interpolative and extrapolative reasoning about categorization rules. We have argued that sets of categorization rules can be seen as partial specifications of a linear mapping between conceptual spaces. This view has allowed us to describe interpolation and extrapolation at a purely qualitative level, relying on qualitative spatial relations to encode knowledge about conceptual spaces rather than on degrees of similarity. From a practical point of view, the approach is motivated from the observation that sufficient data to estimate the conceptual relationship between labels from the same attribute domain is often available, e.g. relying on statistical techniques such as multi-dimensional scaling or singular-value decomposition, while knowledge about how different attribute domains are related is often sparse and is usually encoded in a symbolic form. The techniques presented in this paper show how knowledge about conceptual relations between labels of the same attribute domain may be leveraged to refine whatever symbolic knowledge of this kind we have. Although the general inference problem we have considered is PSPACE-hard, we have shown that efficient implementations in polynomial time are possible if the number of attribute domains is sufficiently small.

We may expect that the full generality of our framework would not be needed in many applications. In [34], for instance, an approach is presented to complete rule bases purely based on analogical proportions. The proposal from [34] in fact corresponds to a special case of the approach presented in this paper, where the only non-trivial information in Σ are assertions of the form $\underline{par_{[1,1]}}(a,b,c;d)$ and $\overline{par_{[1,1]}}(a,b,c;d)$. In [34] it is moreover assumed that all labels correspond to intervals in a uni-dimensional space, which implies that the relations $\underline{par_{[1,1]}}$ and $\overline{par_{[1,1]}}$ coincide, and moreover that they exhibit a number of symmetry and transitivity properties that are not generally valid (e.g. $\underline{par_{[1,1]}}(a,b,c;d)$ iff $\underline{par_{[1,1]}}(c,d,a;b)$). These simplifications lead to an inference process which is easier to use in practice, but which is based on assumptions that are not always realistic. By putting the approach from [34] in relation to the approach from this paper, however, it immediately becomes apparent when these simplifications make sense, or how the approach should be adapted when they do not.

In contrast, some applications may require further generalizations of the approach we have presented here. At the level of the conceptual spaces, we have restricted ourselves to Euclidean spaces, whereas arbitrary metric spaces might be considered instead. Moreover, a better understanding is needed of which sets of assertions Σ are actually realizable, either in a Euclidean space or in an arbitrary metric space. Currently, no sound and complete procedures are available to check the consistency of such a set. Although this poses no problems when conceptual relations are obtained from geometric representations, consistency checking procedures may be important when other forms of acquisition are used (e.g. based on natural language processing). In addition to betweenness and parallelism, other types of conceptual relations may also be considered, as the same methodology may be applied to any type of spatial relation that is invariant under linear transformations. At the semantic level (Section 5), we have restricted ourselves to situations where only information about conceptual relation between individual labels is available. If, however, information would be available about the betweenness or parallelism of disjunctions of labels, a refined definition of $\underline{bet}$ and $\underline{par}$ should be used, as e.g. $reg(b)$ may geometrically be between $reg(a_1) \cup ... \cup reg(a_p)$ and $reg(c_1) \cup ... \cup reg(c_q)$ even if $reg(b)$ is not between $reg(a_i)$ and $reg(c_j)$ for any i and j. In such a case, the syntactic characterization of Section 6 should be adapted as well. Given the presented setting, however, such a generalization should be straightforward to formalize.

Appendix A. Proofs

Proof of Proposition 1

Let $X \subseteq \mathcal{B}$ and $Y \subseteq \mathcal{C}$ be such that

$$R \models_{\mathcal{A}} \Big(\bigvee_{(x_1,\ldots,x_s)\in X} \bigwedge_{i=1}^{s} x_i \Big) \rightarrow \Big(\bigvee_{(y_1,\ldots,y_k)\in Y} \bigwedge_{i=1}^{k} y_i \Big) \tag{A.1}$$

By (21) this means that

$$m^*\big(\bigcup_{x\in X} reg(x) \big) \subseteq \bigcup_{y\in Y} reg(y)$$

and given that m^* is a point-wise extension of the mapping m, we also have that

$$\bigcup_{x\in X} m^*\big(reg(x)\big) \subseteq \bigcup_{y\in Y} reg(y)$$

This means in particular that as soon as $m^*(reg(x)) \cap reg(y_0) \neq \emptyset$ for a given $x \in X$ and $y_0 \in \mathcal{C}$, we must have that $y_0 \in Y$. In other words, we have that $\widehat{f}(X) \subseteq Y$, and since this holds for every Y satisfying (A.1), we obtain $\widehat{f}(X) \subseteq f_R(X)$.

Proof of Proposition 2

Suppose that $\overline{bet}(reg(X_1), reg(Y), reg(X_2))$ holds. Then there are $x_1 \in reg(X_1)$, $y \in reg(Y)$ and $x_2 \in reg(X_2)$ such that y is between x_1 and x_2. Due to the fact that the labels of each attribute domain correspond to JEPD properties, there are (unique) elements $\mathbf{a} \in X_1$, $\mathbf{b} \in Y$ and $\mathbf{c} \in X_2$ such that $x_1 \in reg(\mathbf{a})$, $y \in reg(\mathbf{b})$ and $x_2 \in reg(\mathbf{c})$. To complete the proof, it suffices to show that for every i, it holds that $\overline{bet}(a_i, b_i, c_i)$, i.e. that there exist points $x_i \in reg(a_i)$, $y_i \in reg(b_i)$, and $z_i \in reg(c_i)$ such that y_i is between x_i and z_i. As betweenness is preserved under projection, the points x_i, y_i and z_i can be obtained from x, y and z after removing all irrelevant components.

Proof of Proposition 3

Assume that $Y \subseteq \underline{bet}(X_1, X_2)$. We need to show

$$\forall q \in reg(Y) \,.\, \exists p \in reg(X_1), r \in reg(X_2) \,.\, (p = q) \vee (\exists \lambda \in [0,1] \,.\, \overrightarrow{pq} = \lambda \cdot \overrightarrow{pr})$$

If $q \in X_1 \cap X_2$, we can simply take $p = r = q$. If $q \in X_1 \setminus X_2$, then X_2 is realizable and we can take $p = q$ and an arbitrary $r \in reg(X_2) \neq \emptyset$. Similarly, if $q \in X_2 \setminus X_1$ we can take $p \in reg(X_1) \neq \emptyset$ arbitrary and choose $r = q$.

For $q \in reg(\{\mathbf{b} \,|\, \mathbf{a} \in X_1, \mathbf{c} \in X_2, \underline{bet}(\mathbf{a}, \mathbf{b}, \mathbf{c})\})$, note that the fact that $\underline{bet}(\mathbf{a}, \mathbf{b}, \mathbf{c})$ holds entails that the underlying attribute domains are all orthogonal (unless in the entirely trivial case where $\mathbf{a} = \mathbf{b} = \mathbf{c}$). If $\mathbf{a} = \mathbf{b}$, it suffices to take $p = q$ and let r be an arbitrary element from $reg(X_2)$, and similar for the case where $\mathbf{b} = \mathbf{c}$. Otherwise, the vectors $\mathbf{a} = (a_1, ..., a_l)$, $\mathbf{b} = (b_1, ..., b_l)$ and $\mathbf{c} = (c_1, ..., c_l)$ only differ in the i^{th} component for some i, and we have $\underline{bet}(a_i, b_i, c_i)$. Now, given $\underline{bet}(a_i, b_i, c_i)$, we know that for each $q_i \in reg(b_i)$ there are $p_i \in reg(a_i)$ and $r_i \in reg(c_i)$ such that q_i is between p_i and r_i. As the attribute domains are orthogonal, and the vectors $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$ agree on all but the i^{th} component, it is easy to see that we then also have that for each $q \in reg(b)$ there are $p \in reg(a)$ and $r \in reg(c)$ such that q is between p and r; it suffices to extend p_i, q_i and r_i by choosing identical values for the added components.

Finally, for $q \in reg(\{(b_1, ..., b_{i-1}, x, b_{i+1}, ..., b_l) \,|\, \mathbf{b} \in \underline{bet}(X_1^{\downarrow i}, X_2^{\downarrow i}), 1 \leq i \leq l, x \in D_i\})$, we know by induction that for each i and $q' \in reg(Y^{\downarrow i})$ there are $p' \in reg(X_1^{\downarrow i})$ and $r' \in reg(X_2^{\downarrow i})$ such that q' is between p' and r'. By construction, we can extend q' to a point $q \in reg(Y)$ by choosing an arbitrary value for each of the quality dimensions underlying D_i that do not already appear in q'. In the same way, we can then extend p' to $p \in reg(X_1)$ and r' to $r \in reg(X_2)$ by filling in the same values for the added quality dimensions.

Proof of Proposition 4

Assume that $Y \subseteq \underline{par}(X_1, X_2, X_3)$. We need to show

$$\forall s \in \mathit{reg}(Y) \,.\, \exists p \in \mathit{reg}(X_1), q \in \mathit{reg}(X_2), r \in \mathit{reg}(X_3) \,.\, \exists \lambda \geq 0 \,.\, \overrightarrow{rs} = \lambda \cdot \overrightarrow{pq}$$

If $s \in X_1 \cap X_2 \cap X_3$, we can take $p = q = r = s$. If $s \in (X_1 \cap X_2) \setminus X_3$, we know that X_1 is realizable and thus $\mathit{reg}(X_1) \neq \emptyset$. We can then take p arbitrary and choose $q = r = s$. Similarly if $s \in (X_1 \cap X_3) \setminus X_2$, we can choose $p = r = s$ and take $q \in \mathit{reg}(X_2) \neq \emptyset$ arbitrary. If $s \in X_3 \setminus (X_1 \cup X_2)$, we know that X_1 and X_2 are realizable, and we can take $p \in \mathit{reg}(X_1) \neq \emptyset$ and $q \in \mathit{reg}(X_2) \neq \emptyset$ arbitrary, and choose $r = s$. If $s \in X_2 \setminus (X_1 \cup X_3)$, we know that $\mathit{reg}(X_1 \cap X_3) \neq \emptyset$, hence we can take an arbitrary $r \in \mathit{reg}(X_1 \cap X_3)$ and choose $p = r$ and $q = s$.

For $s \in \mathit{reg}(\{\mathbf{d} \,|\, \mathbf{a} \in X_1, \mathbf{b} \in X_2, \mathbf{c} \in X_3, \underline{par}(\mathbf{a}, \mathbf{b}, \mathbf{c}; \mathbf{d})\})$, note that the fact that $\underline{par}(\mathbf{a}, \mathbf{b}, \mathbf{c}; \mathbf{d})$ holds entails that the underlying attribute domains are all orthogonal (with the exception of the trivial case where $\mathbf{a} = \mathbf{b} = \mathbf{c} = \mathbf{d}$). If $\mathbf{a} = \mathbf{c}$ and $\mathbf{b} = \mathbf{d}$, it suffices to take $q = s$ and let $p = r$ be an arbitrary element from $\mathit{reg}(\mathbf{a}) = \mathit{reg}(\mathbf{c})$. If $\mathbf{c} = \mathbf{d}$, we can take $r = s$ and let p and q be arbitrary elements from $\mathit{reg}(\mathbf{a})$ and $\mathit{reg}(\mathbf{b})$ respectively. Otherwise, we have that the vectors $\mathbf{a} = (a_1, ..., a_l)$, $\mathbf{b} = (b_1, ..., b_l)$, $\mathbf{c} = (c_1, ..., c_l)$ and $\mathbf{d} = (d_1, ..., d_l)$ are such that for some j, it holds that $\underline{par}(a_j, b_j, c_j; d_j)$, whereas for all $i \neq j$ we have $a_i = b_i$ and $c_i = d_i$. From $\underline{par}(a_j, b_j, c_j; d_j)$ we know that there are points $p_j \in \mathit{reg}(a_j)$, $q_j \in \mathit{reg}(b_j)$, $r_j \in \mathit{reg}(c_j)$ and $s_j \in \mathit{reg}(d_j)$ such that $\overrightarrow{r_j s_j} = \lambda \cdot \overrightarrow{p_j q_j}$ for some $\lambda \geq 0$. Given the orthogonality of the attribute domains and the fact that $a_i = b_i$ for $i \neq j$, p_j and q_j can be extended to points $p \in \mathit{reg}(\mathbf{a})$ and $q \in \mathit{reg}(\mathbf{b})$ by choosing the same value for all added components. In the same way we can extend r_j and s_j to points $r \in \mathit{reg}(\mathbf{c})$ and $s \in \mathit{reg}(\mathbf{d})$ by choosing the same value for the added components. We then have that $\overrightarrow{rs} = \lambda \cdot \overrightarrow{pq}$, with λ as before.

For $s \in \mathit{reg}(\{(d_1, ..., d_{i-1}, x, d_{i+1}, ..., d_l) \,|\, \mathbf{d} \in \underline{par}(X_1^{\downarrow i}, X_2^{\downarrow i}, X_3^{\downarrow i}), 1 \leq i \leq l, x \in D_i\})$, we know by induction that for each i and $s' \in \mathit{reg}(Y^{\downarrow i})$ there are $p' \in \mathit{reg}(X_1^{\downarrow i})$, $q' \in \mathit{reg}(X_2^{\downarrow i})$ and $r' \in \mathit{reg}(X_3^{\downarrow i})$ such that $\overrightarrow{p'q'} = \lambda \cdot \overrightarrow{r's'}$ for some $\lambda > 0$. By construction, we can extend s' to a point $s \in \mathit{reg}(Y)$ by choosing an arbitrary value for each of the quality dimensions underlying D_i that do not already appear in s'. In the same way, we can then extend p' to $p \in \mathit{reg}(X_1)$, q' to $q \in \mathit{reg}(X_2)$ and r' to $r \in \mathit{reg}(X_3)$ by filling in the same values for the added quality dimensions. The case where $s \in \mathit{reg}(\{(d_1, ..., d_{i-1}, x, d_{i+1}, ..., d_l) \,|\, \mathbf{d} \in \underline{par}(X_1^{\downarrow x}, X_2^{\downarrow y}, X_3^{\downarrow y}), 1 \leq i \leq$

$l, x \in D_i, y \in D_i, D_i$ orthogonal to $D_1, ..., D_{i-1}, D_{i+1}, ..., D_l\}$) is analogous, noting that the orthogonality of D_i means that none of the attribute domains underlying D_i will appear in s', p', q' and r'.

Proof of Proposition 5

Suppose that $\overline{par}(reg(X_1), reg(X_2), reg(X_3); reg(Y))$ holds. Then there are $x_1 \in reg(X_1)$, $x_2 \in reg(X_2)$, $x_3 \in reg(X_3)$ and $y \in reg(Y)$ such that $\overrightarrow{x_3 y} = \lambda \cdot \overrightarrow{x_1 x_2}$ holds. Due to the fact that the labels of each attribute domain correspond to JEPD properties, there are (unique) elements $\mathbf{a} \in X_1$, $\mathbf{b} \in X_2$, $\mathbf{c} \in X_3$ and $\mathbf{d} \in Y$ such that $x_1 \in reg(\mathbf{a})$, $x_2 \in reg(\mathbf{b})$, $x_3 \in reg(\mathbf{c})$ and $y \in reg(\mathbf{d})$. To complete the proof, it suffices to show that for every i, it holds that $\overline{par}(a_i, b_i, c_i; d_i)$, i.e. that there exist points $x_i^1 \in reg(a_i)$, $x_i^2 \in reg(b_i)$, $x_i^3 \in reg(c_i)$ and $y_i \in reg(d_i)$ such that $\overrightarrow{x_i^3 y_i} = \mu \cdot \overrightarrow{x_i^1 x_i^2}$ for some $\mu \geq 0$. As parallelism is preserved under projection, the points x_i^1, x_i^2, x_i^3 and y_i can be obtained from x^1, x^2, x^3 and y after removing all irrelevant components, in which case the latter equality will hold for $\mu = \lambda$.

Proof of Proposition 6

The proof is largely analogous to the proof of Propositions 3 and 4, except for the case of degenerate intervals. For that case, we need to verify that for $\lambda \in [0, 1]$ and $\mu \in [0, +\infty[$, it holds that

$$(\forall i \,.\, \underline{bet_{[\lambda,\lambda]}}(a_i, b_i, c_i)) \Rightarrow \underline{bet_{[\lambda,\lambda]}}(reg(\mathbf{a}), reg(\mathbf{b}), reg(\mathbf{c})) \tag{A.2}$$

$$(\forall i \,.\, \underline{par_{[\mu,\mu]}}(a_i, b_i, c_i; d_i)) \Rightarrow \underline{par_{[\mu,\mu]}}(reg(\mathbf{a}), reg(\mathbf{b}), reg(\mathbf{c}); reg(\mathbf{d})) \tag{A.3}$$

provided that the underlying attribute domains $D_1, ..., D_l$ are all orthogonal. To show (A.2), suppose that $\underline{bet_{[\lambda,\lambda]}}(a_i, b_i, c_i))$ for all i. This means that for every $q_i \in reg(b_i)$, there are $p_i \in reg(a_i)$ and $r_i \in reg(c_i)$ such that $\overrightarrow{p_i q_i} = \lambda \cdot \overrightarrow{p_i r_i}$. Let these points be of the form $p_i = (p_i^1, ..., p_i^{n_i})$, $q_i = (q_i^1, ..., q_i^{n_i})$ and $r_i = (r_i^1, ..., r_i^{n_i})$. Now let q be a point from $reg(\mathbf{b})$. Without lack of generality, given the fact that the attribute domains are orthogonal, we can assume that q is of the form $(q_1^1, ..., q_1^{n_1}, ..., q_l^1, ..., q_l^{n_l})$. We can moreover construct the points $p = (p_1^1, ..., p_1^{n_1}, ..., p_l^1, ..., p_l^{n_l})$ and $r = (r_1^1, ..., r_1^{n_1}, ..., r_l^1, ..., r_l^{n_l})$. By construction, we then have that $p \in reg(\mathbf{a})$, $r \in reg(\mathbf{c})$ and $\overrightarrow{pq} = \lambda \cdot \overrightarrow{pr}$. As we can do this for every $q \in reg(\mathbf{b})$, we have shown (A.2).

The proof of (A.3) is entirely analogous.

Proof of Proposition 7

The proof is entirely analogous to the proof of Propositions 2 and 5.

Proof of Proposition 8

From the definitions of f_R and $\widehat{f_R}$ we immediately find that

$$\widehat{f_R}(X) = \bigcup_{x\in X} \widehat{f_R}(\{x\}) \subseteq \bigcup_{x\in X} f_R(\{x\}) = f_R(X)$$

and thus $\widehat{f_R} \leq f_R$. To show that $\widehat{f} \leq \widehat{f_R}$ we show by induction that $\widehat{f} \leq \widehat{f_R}^{(i)}$ for all $i \in \mathbb{N}$. The case where $i = 0$ was shown in Proposition 1. Assuming that we have already established $\widehat{f} \leq \widehat{f_R}^{(i)}$, from the monotonicity of $\overline{bet_\sigma}$ and $\overline{par_\tau}$ w.r.t. set inclusion we find that

$$\begin{aligned}
\widehat{f_R}^{(i+1)}(\{x\}) &= f_R(\{\mathbf{x}\}) \cap \bigcap_{\sigma\subseteq[0,1]} \{\overline{bet_\sigma}(\widehat{f_R}^{(i)}(Y), \widehat{f_R}^{(i)}(Z)) | \mathbf{x} \in \underline{bet_\sigma}(Y,Z)\} \\
&\qquad \cap \bigcap_{\tau\subseteq[0,+\infty[} \{\overline{par_\tau}(\widehat{f_R}^{(i)}(X), \widehat{f_R}^{(i)}(Y), \widehat{f_R}^{(i)}(Z)) | \mathbf{x} \in \underline{par_\tau}(X,Y,Z)\} \\
&\supseteq f_R(\{\mathbf{x}\}) \cap \bigcap_{\sigma\subseteq[0,1]} \{\overline{bet_\sigma}(\widehat{f}(Y), \widehat{f}(Z)) | \mathbf{x} \in \underline{bet_\sigma}(Y,Z)\} \\
&\qquad \cap \bigcap_{\tau\subseteq[0,+\infty[} \{\overline{par_\tau}(\widehat{f}(X), \widehat{f}(Y), \widehat{f}(Z)) | \mathbf{x} \in \underline{par_\tau}(X,Y,Z)\} \\
&\supseteq \widehat{f}(\{x\})
\end{aligned}$$

where the last step follows from $\widehat{f}(\{x\}) \subseteq f_R(\{x\})$ (since $\widehat{f} \leq f_R$ by Proposition 1), together with (36) and (37). Finally note that $\widehat{f}(X) = \bigcup_{x\in X} \widehat{f}(\{x\})$ by definition of $\widehat{f}$.

Proof of Lemma 1

Assume that α is strongly consistent, and let β be such that $\beta \models \alpha$, β is consistent, and all attribute domains in *domains*(β) are orthogonal. As *set*$(\beta) \subseteq$ *set*(α) we also have $(set(\beta)^{\downarrow i_1})^{\downarrow i_2}...^{\downarrow i_s} \subseteq (set(\alpha)^{\downarrow i_1})^{\downarrow i_2}...^{\downarrow i_s}$ for all $i_1, ..., i_s$. If the positions $i_1, ..., i_s$ refer to all attribute domains that are outside *domains*(β), then clearly $(set(\beta)^{\downarrow i_1})^{\downarrow i_2}...^{\downarrow i_s}$ is non-empty and only contains components referring to orthogonal attribute domains anymore. This means that also $(set(\alpha)^{\downarrow i_1})^{\downarrow i_2 \dots \downarrow i_s}$ is non-empty and refers to orthogonal attribute domains, which means that $set(\alpha)$ is realizable.

Conversely, assume that $set(\alpha)$ is realizable, then there exist positions $i_1, ..., i_s$ such that $(set(\alpha)^{\downarrow i_1})^{\downarrow i_2 \dots \downarrow i_s}$ is non-empty and only refers to orthogonal attribute domains. For $\mathbf{b} \in (set(\alpha)^{\downarrow i_1})^{\downarrow i_2 \dots \downarrow i_s}$ we then find that $conj(\mathbf{b}) \models \alpha$ while all attribute domains in $domains(conj(\mathbf{b}))$ are orthogonal, which means that α is strongly consistent.

Proof of Proposition 9

($\Rightarrow$) Assume that $\Sigma \vdash \alpha \rightarrow \beta \bowtie_\sigma \gamma$ can be derived from the inference rules ($\underline{\mathbf{bet_1}}$)–($\underline{\mathbf{bet_9}}$). We show by induction that $set(\alpha) \subseteq \underline{bet_\sigma}(set(\beta), set(\gamma))$.

If $\Sigma \vdash \alpha \rightarrow \beta \bowtie_\sigma \gamma$ was obtained from ($\underline{\mathbf{bet_1}}$), then α, β and γ are atoms and we have that $\underline{bet_\sigma}(\beta, \alpha, \gamma) \in \Sigma$, from which we find using (39) that $\underline{bet_\sigma}(reg(\beta), reg(\alpha), reg(\gamma))$ holds which is the same as $\underline{bet_\sigma}(set(\beta), set(\alpha), set(\gamma))$ because α, β and γ are atoms. We then easily find that $set(\alpha) \subseteq \underline{bet_\sigma}(set(\beta), set(\gamma))$. Indeed, applying the definition of $\underline{bet_\sigma}$, given by (28), all attribute domains except for the one in which α, β and γ occur are deemed irrelevant, in which case $\underline{bet_\sigma}(set(\beta), set(\gamma))$ reduces to betweenness at the atom level.

If $\Sigma \vdash \alpha \rightarrow \beta \bowtie_\sigma \gamma$ was obtained from ($\underline{\mathbf{bet_2}}$), then there exist α', β', γ' and σ' such that $\alpha \models_{\mathcal{A}} \alpha'$, $\beta' \models_{\mathcal{A}} \beta$ and $\gamma' \models_{\mathcal{A}} \gamma$, $\Sigma \vdash \alpha' \rightarrow \beta' \bowtie_{\sigma'} \gamma'$, and $\sigma' \subseteq \sigma$. By induction, we then have that $set(\alpha') \subseteq \underline{bet_{\sigma'}}(set(\beta'), set(\gamma'))$, while we have $set(\alpha) \subseteq set(\alpha')$, $set(\beta') \subseteq set(\beta)$ and $set(\gamma') \subseteq set(\gamma)$. By the fact that $\underline{bet}$ is clearly monotonic w.r.t. set inclusion, it follows that

$$set(\alpha) \subseteq set(\alpha') \subseteq \underline{bet_{\sigma'}}(set(\beta'), set(\gamma')) \subseteq \underline{bet_\sigma}(set(\beta), set(\gamma))$$

In the case where $\Sigma \vdash \alpha \rightarrow \beta \bowtie_\sigma \gamma$ was obtained from ($\underline{\mathbf{bet_3}}$), we already have $set(\alpha) \subseteq \underline{bet_{1-\sigma}}(set(\gamma), set(\beta))$. From the symmetry of the betweenness relation, we then immediately find that also $set(\alpha) \subseteq \underline{bet_\sigma}(set(\beta), set(\gamma))$.

In the case where $\Sigma \vdash \alpha \rightarrow \beta \bowtie_\sigma \gamma$ was obtained from ($\underline{\mathbf{bet_4}}$), the strong consistency of β ensures that $set(\beta)$ is realizable. By definition, we then have $set(\alpha) \subseteq \underline{bet_\sigma}(set(\alpha), set(\beta))$, given that $0 \in \sigma$.

The case where $\Sigma \vdash \alpha \rightarrow \beta \bowtie_\sigma \gamma$ was obtained from ($\underline{\mathbf{bet_5}}$) directly follows from the fact that we have $set(\alpha) \subseteq \underline{bet_\alpha}(set(\alpha), set(\alpha))$ for any α. Indeed, from (28) is follows that $\underline{bet_\sigma}(X_1, X_2) \supseteq X_1 \cap X_2$.

If $\Sigma \vdash \alpha \rightarrow \beta \bowtie_\sigma \gamma$ was obtained from (**$\underline{\mathbf{bet_6}}$**), then there exist α_1, α_2, β_1, β_2, γ_1 and γ_2 such that $\alpha = \alpha_1 \vee \alpha_2$, $\beta = \beta_1 \vee \beta_2$, $\gamma = \gamma_1 \vee \gamma_2$ and, by induction, $set(\alpha_1) \subseteq \underline{bet_\sigma}(set(\beta_1), set(\gamma_1))$ and $set(\alpha_2) \subseteq \underline{bet_\sigma}(set(\beta_2), set(\gamma_2))$. Now, we have that

$$\begin{aligned}\underline{bet_\sigma}(set(\beta), set(\gamma)) &= \underline{bet_\sigma}(set(\beta_1 \vee \beta_2), set(\gamma_1 \vee \gamma_2)) \\ &= \underline{bet_\sigma}(set(\beta_1) \cup set(\beta_2), set(\gamma_1) \cup set(\gamma_2))\end{aligned}$$

From the monotonicity of $\underline{bet_\sigma}$ w.r.t. set inclusion, we find that the latter expression includes both $\underline{bet_\sigma}(set(\beta_1), set(\gamma_1))$ and $\underline{bet_\sigma}(set(\beta_2), set(\gamma_2))$, and thus also $\underline{bet_\sigma}(set(\beta_1), set(\gamma_1)) \cup \underline{bet_\sigma}(set(\beta_2), set(\gamma_2))$, which in turn includes $set(\alpha_1) \cup set(\alpha_2)$. By definition of *set* the latter expression is equivalent to $set(\alpha_1 \vee \alpha_2)$ and thus to $set(\alpha)$.

If $\Sigma \vdash \alpha \rightarrow \beta \bowtie_\sigma \gamma$ was obtained from (**$\underline{\mathbf{bet_7}}$**), then there exist α', β', γ' and δ such that $\alpha = \alpha' \wedge \delta$, $\beta = \beta' \wedge \delta$, $\gamma = \gamma' \wedge \delta$ and, by induction, $set(\alpha') \subseteq \underline{bet_\sigma}(set(\beta'), set(\gamma'))$. If $set(\alpha) \subseteq set(\beta')$ or $set(\alpha) \subseteq set(\gamma')$, the proof is trivial, so assume that this were not the case. Since δ is strongly consistent, we can assume without lack of generality that the attribute domains in $domains(\delta)$ are orthogonal to each other (since if this were not the case, we could replace δ with a logically equivalent formula for which this is the case). Since the elements of $domains(\delta)$ are moreover orthogonal to the elements of $domains(\beta')$ and $domains(\gamma')$, it is not hard to see that $\underline{bet_\sigma}(set(\beta'), set(\gamma')) \cap set(\delta) \subseteq \underline{bet_\sigma}(set(\beta') \cap set(\delta), set(\gamma') \cap set(\delta))$. Indeed, this is trivial if β' or γ' is not strongly consistent. If both are strongly consistent, then let $\mathbf{b} \in \underline{bet_\sigma}(set(\beta'), set(\gamma')) \cap set(\delta)$ and let $\mathbf{b}'$ be the vector obtained from $\mathbf{b}$ after removing the components of all irrelevant attribute domains. Then there are corresponding subvectors $\mathbf{a}'$ and $\mathbf{c}'$ of elements in $set(\beta')$ and $set(\gamma')$ respectively, such that $\underline{bet_\sigma}(\mathbf{a}', \mathbf{b}', \mathbf{c}')$ holds. Now $\mathbf{a}'$, $\mathbf{b}'$ and $\mathbf{c}'$ can be extended to vectors $\mathbf{a}''$, $\mathbf{b}''$ and $\mathbf{c}''$ by adding the components from $domains(\delta)$, choosing the same values as in $\mathbf{b}$. Then it clearly holds that $\underline{bet_\sigma}(\mathbf{a}'', \mathbf{b}'', \mathbf{c}'')$, while any extension of $\mathbf{a}''$ and $\mathbf{b}''$ to full vectors $\mathbf{a}$ and $\mathbf{b}$ will belong to $set(\delta)$, leading to $\mathbf{b} \in \underline{bet_\sigma}(set(\beta') \cap set(\delta), set(\gamma') \cap set(\delta))$. Hence, we have $set(\alpha') \cap set(\delta) \subseteq \underline{bet_\sigma}(set(\beta') \cap set(\delta), set(\gamma') \cap set(\delta))$, which is equivalent to $set(\alpha' \wedge \delta) \subseteq \underline{bet_\sigma}(set(\beta' \wedge \delta), set(\gamma' \wedge \delta))$ and to $set(\alpha) \subseteq \underline{bet_\sigma}(set(\beta), set(\gamma))$.

If $\Sigma \vdash \alpha \rightarrow \beta \bowtie \gamma$ was obtained from (**$\underline{\mathbf{bet_8}}$**), there are α_1, α_2,

β_1, β_2, γ_1 and γ_2 such that $\alpha = \alpha_1 \wedge \alpha_2$, $\beta = \beta_1 \wedge \beta_2$ and $\gamma = \gamma_1 \wedge \gamma_2$, such that, by induction, $set(\alpha_1) \subseteq \underline{bet}_{[\lambda,\lambda]}(set(\beta_1), set(\gamma_1))$ and $set(\alpha_2) \subseteq \underline{bet}_{[\lambda,\lambda]}(set(\beta_2), set(\gamma_2))$. Given the definition of $\underline{bet}_{[\lambda,\lambda]}$ and given the fact that the attribute domains in $domains(\beta_1) \cup domains(\gamma_1)$ are orthogonal to those in $domains(\beta_2) \cup domains(\gamma_2)$, we can show that $\underline{bet}_{[\lambda,\lambda]}(set(\beta_1), set(\gamma_1)) \cap \underline{bet}_{[\lambda,\lambda]}(set(\beta_2), set(\gamma_2)) \subseteq \underline{bet}_{[\lambda,\lambda]}(set(\beta_1) \cap set(\beta_2), set(\gamma_1) \cap set(\gamma_2))$. If $\lambda = 0$ or $\lambda = 1$ then this is trivial. If $\lambda \in]0,1[$, from $\mathbf{b} \in \underline{bet}_{[\lambda,\lambda]}(set(\beta_1), set(\gamma_1)) \cap \underline{bet}_{[\lambda,\lambda]}(set(\beta_2), set(set(\gamma_2))$, we find that there are two non-overlapping subvectors $\mathbf{b_1}$ and $\mathbf{b_2}$ of $\mathbf{b}$ (containing components for the attribute domains in $domains(\beta_1) \cup domains(\gamma_1)$ and $domains(\beta_2) \cup domains(\gamma_2)$ respectively) and such that $\underline{bet}_{[\lambda,\lambda]}(\mathbf{a_1}, \mathbf{b_1}, \mathbf{c_1})$ and $\underline{bet}_{[\lambda,\lambda]}(\mathbf{a_2}, \mathbf{b_2}, \mathbf{c_2})$ hold, for appropriate subvectors of elements $\mathbf{a}$ and $\mathbf{c}$ from $set(\beta_1 \wedge \beta_2)$ and $set(\gamma_1 \wedge \gamma_2)$. By definition of $\underline{bet}_{[\lambda,\lambda]}$, we then get that $\underline{bet}_{[\lambda,\lambda]}(\mathbf{a_{12}}, \mathbf{b_{12}}, \mathbf{c_{12}})$ holds for the compound vectors, from which we find that $\mathbf{b} \in \underline{bet}_{[\lambda,\lambda]}(set(\beta_1) \cap set(\beta_2), set(\gamma_1) \cap set(\gamma_2))$. Thus we find that

$$set(\alpha_1) \cap set(\alpha_2) \subseteq \underline{bet}_{\sigma}(set(\beta_1) \cap set(\beta_2), set(\gamma_1) \cap set(\gamma_2))$$

which is equivalent to $set(\alpha_1 \wedge \alpha_2) \subseteq \underline{bet}_{\sigma}(set(\beta_1 \wedge \beta_2), set(\gamma_1 \wedge \gamma_2))$ and to $set(\alpha) \subseteq \underline{bet}_{\sigma}(set(\beta), set(\gamma))$.

The case where $\Sigma \vdash \alpha \rightarrow \beta \bowtie \gamma$ was obtained from ($\underline{\mathbf{bet_9}}$) is trivial.

($\Leftarrow$) Now we assume that $set(\alpha) \subseteq \underline{bet}_{\sigma}(set(\beta), set(\gamma))$ and show that $\Sigma \vdash \alpha \rightarrow \beta \bowtie_{\sigma} \gamma$ can be derived from the inference rules ($\underline{\mathbf{bet_1}}$)–($\underline{\mathbf{bet_9}}$). If α is inconsistent, then $\Sigma \vdash \alpha \rightarrow \beta \bowtie_{\sigma} \gamma$ follows from ($\underline{\mathbf{bet_9}}$) and ($\underline{\mathbf{bet_2}}$). Otherwise, given that ($\underline{\mathbf{bet_2}}$) entails syntax-independence, we can assume without lack of generality that α is of the form $\alpha_1 \vee ... \vee \alpha_r$, where each α_i is a conjunction containing exactly one atom from each attribute domain. In other words, $set(\alpha_i) = \{\mathbf{a}\}$ is a singleton. Because of inference rules ($\underline{\mathbf{bet_6}}$) and ($\underline{\mathbf{bet_2}}$), it suffices to show that $\Sigma \vdash \alpha_i \rightarrow \beta \bowtie_{\sigma} \gamma$ can be derived for each i.

The fact that $\mathbf{a} \in \underline{bet}_{\sigma}(set(\beta), set(\gamma))$ either means that (i) $\mathbf{a} \in set(\beta) \cup set(\gamma)$, (ii) there is a $\mathbf{b} \in set(\beta)$ and $\mathbf{c} \in set(\gamma)$ such that $\underline{bet}_{\sigma}(\mathbf{b}, \mathbf{a}, \mathbf{c})$ holds, or (iii) it holds that $\mathbf{a}^{\downarrow j} \in \underline{bet}_{\sigma}(set(\beta)^{\downarrow j}, set(\gamma)^{\downarrow j})$ for some j, where we write $\mathbf{a}^{\downarrow j}$ for the vector $\mathbf{a}$ without the j^{th} component and $set(\beta)^{\downarrow j}$ and $set(\gamma)^{\downarrow j}$ are defined as in (22).

In the first case, if $\mathbf{a} \in \mathit{set}(\beta) \setminus \mathit{set}(\gamma)$ this means that $\mathit{set}(\gamma)$ is realizable and $0 \in \sigma$ (by (28)), hence that γ is strongly consistent (by Lemma 1). This means that we can derive $\Sigma \vdash \alpha_i \rightarrow \alpha_i \bowtie_\sigma \gamma$ using $(\underline{\mathbf{bet_4}})$. Using $(\underline{\mathbf{bet_2}})$, we can then also conclude $\Sigma \vdash \alpha_i \rightarrow \beta \bowtie_\sigma \gamma$ given that $\mathbf{a} \in \mathit{set}(\beta)$. In the same way, if $\mathbf{a} \in \mathit{set}(\gamma) \setminus \mathit{set}(\beta)$, we derive $\Sigma \vdash \alpha_i \rightarrow \gamma \bowtie_{1-\sigma} \beta$, and using $(\underline{\mathbf{bet_3}})$ also $\Sigma \vdash \alpha_i \rightarrow \beta \bowtie_\sigma \gamma$. If $\mathbf{a} \in \mathit{set}(\beta) \cap \mathit{set}(\gamma)$, using $(\underline{\mathbf{bet_5}})$ we derive $\Sigma \vdash \alpha_i \rightarrow \alpha_i \bowtie_\sigma \alpha_i$ from which we derive $\Sigma \vdash \alpha_i \rightarrow \beta \bowtie_\sigma \gamma$ by $(\underline{\mathbf{bet_2}})$.

In the second case, we clearly have that $\mathit{conj}(\mathbf{b}) \models_{\mathcal{A}} \beta$ and $\mathit{conj}(\mathbf{c}) \models_{\mathcal{A}} \gamma$, hence because of inference rule $(\underline{\mathbf{bet_2}})$ it suffices to show that $\Sigma \vdash \mathit{conj}(\mathbf{a}) \rightarrow \mathit{conj}(\mathbf{b}) \bowtie_\sigma \mathit{conj}(\mathbf{c})$ can be derived. By definition, $\underline{\mathit{bet}_\sigma}(\mathbf{b}, \mathbf{a}, \mathbf{c})$ means that all components correspond to orthogonal attribute domains. First assume that σ is a non-degenerate interval. We then have that either (i) $\mathbf{a} = \mathbf{b}$, (ii) $\mathbf{a} = \mathbf{c}$, or (iii) that $\mathbf{a}$ differs only in one component from $\mathbf{b}$ and $\mathbf{c}$, and that component in $\mathbf{a}$ is moreover between the corresponding components in $\mathbf{b}$ and $\mathbf{c}$. In the latter situation, $\Sigma \vdash \mathit{conj}(\mathbf{a}) \rightarrow \mathit{conj}(\mathbf{b}) \bowtie_\sigma \mathit{conj}(\mathbf{c})$ can clearly be derived using inference rules $(\underline{\mathbf{bet_1}})$ (given (43)) and $(\underline{\mathbf{bet_7}})$, whereas when $\mathbf{a} = \mathbf{b}$, it can be derived using $(\underline{\mathbf{bet_4}})$, and when $\mathbf{a} = \mathbf{c}$ it can be derived using $(\underline{\mathbf{bet_4}})$ and $(\underline{\mathbf{bet_3}})$. Now assume that $\sigma = [\lambda, \lambda]$ is a degenerate interval. Then we have $\underline{\mathit{bet}_{[\lambda,\lambda]}}(b_i, a_i, c_i)$ for every i, hence we can derive $\Sigma \vdash \mathit{conj}(\mathbf{a}) \rightarrow \mathit{conj}(\mathbf{b}) \bowtie_\sigma \mathit{conj}(\mathbf{c})$ by repeatedly applying $(\underline{\mathbf{bet_1}})$ and then using $(\underline{\mathbf{bet_8}})$.

In the third case, note that $\mathit{set}(\beta)^{\downarrow j}$ and $\mathit{set}(\gamma)^{\downarrow j}$ correspond to formulas β^* and γ^* such that $\beta^* \models_{\mathcal{A}} \beta$ and $\gamma^* \models_{\mathcal{A}} \gamma$. Moreover, α_i is of the form $\mathit{conj}(\mathbf{a}^{\downarrow j}) \wedge x$ for some atom x, and in particular we have that $\alpha_i \models_{\mathcal{A}} \mathit{conj}(\mathbf{a}^{\downarrow j})$. By induction, we can moreover assume that $\Sigma \vdash \mathit{conj}(\mathbf{a}^{\downarrow j}) \rightarrow \beta^* \bowtie_\sigma \gamma^*$ can be derived. Using $(\underline{\mathbf{bet_2}})$ this means that also $\Sigma \vdash \alpha_i \rightarrow \beta \bowtie_\sigma \gamma$ can be derived.

Proof of Proposition 10

($\Rightarrow$) Assume that $\Sigma \vdash \beta \bowtie_\sigma \gamma \rightarrow \alpha$ can be derived from the inference rules $(\overline{\mathbf{bet_1}})$–$(\overline{\mathbf{bet_6}})$. We show by induction that then $\overline{\mathit{bet}_\sigma}(\mathit{set}(\beta), \mathit{set}(\gamma)) \subseteq \mathit{set}(\alpha)$.

If $\Sigma \vdash \beta \bowtie_\sigma \gamma \rightarrow \alpha$ was obtained from $(\overline{\mathbf{bet_1}})$, then it is not hard to see from (40) that $\overline{\mathit{bet}_\sigma}(\mathit{set}(\beta), \mathit{set}(\gamma)) \subseteq \mathit{set}(\alpha)$.

If $\Sigma \vdash \beta \bowtie_\sigma \gamma \rightarrow \alpha$ was obtained from ($\overline{\mathbf{bet_2}}$), then there exist α', β', γ' and σ' such that $\alpha' \models_\mathcal{A} \alpha$, $\beta \models_\mathcal{A} \beta'$, $\gamma \models_\mathcal{A} \gamma'$, $\sigma \subseteq \sigma'$ and $\Sigma \vdash \beta' \bowtie_{\sigma'} \gamma' \rightarrow \alpha'$. By induction, we then have that $\overline{bet_{\sigma'}}(set(\beta'), set(\gamma')) \subseteq set(\alpha')$, while $set(\alpha') \subseteq set(\alpha)$, $set(\beta) \subseteq set(\beta')$ and $set(\gamma) \subseteq set(\gamma')$. By the fact that $\overline{bet_\sigma}$ is clearly monotonic w.r.t. set inclusion, it follows that $\overline{bet_\sigma}(set(\beta), set(\gamma)) \subseteq set(\alpha)$.

The case where $\Sigma \vdash \beta \bowtie_\sigma \gamma \rightarrow \alpha$ was obtained from ($\overline{\mathbf{bet_3}}$) immediately follows from the definition of $\overline{bet}$.

If $\Sigma \vdash \beta \bowtie_\sigma \gamma \rightarrow \alpha$ was obtained from ($\overline{\mathbf{bet_4}}$), then there exist α_1, α_2, β_1 and β_2 such that $\alpha = \alpha_1 \vee \alpha_2$, $\beta = \beta_1 \vee \beta_2$, and, by induction, $\overline{bet_\sigma}(set(\beta_1), set(\gamma)) \subseteq set(\alpha_1)$ and $\overline{bet_\sigma}(set(\beta_2), set(\gamma)) \subseteq set(\alpha_2)$. We have that $\overline{bet_\sigma}(set(\beta), set(\gamma)) = \overline{bet_\sigma}(set(\beta_1 \vee \beta_2), set(\gamma)) = \overline{bet_\sigma}(set(\beta_1) \cup set(\beta_2), set(\gamma))$. From the definition of $\overline{bet_\sigma}$ it easily follows that the latter expression is equal to $\overline{bet_\sigma}(set(\beta_1), set(\gamma)) \cup \overline{bet_\sigma}(set(\beta_2), set(\gamma))$. By the assumption, this latter expression is known to be included in $set(\alpha_1) \vee set(\alpha_2)$, which is equal to $set(\alpha)$.

If $\Sigma \vdash \beta \bowtie_\sigma \gamma \rightarrow \alpha$ was obtained from ($\overline{\mathbf{bet_5}}$), then there exist α_1, α_2, β_1, β_2, γ_1 and γ_2 such that $\alpha = \alpha_1 \wedge \alpha_2$, $\beta = \beta_1 \wedge \beta_2$, $\gamma = \gamma_1 \wedge \gamma_2$ and, by induction, $\overline{bet_\sigma}(set(\beta_1), set(\gamma_1)) \subseteq set(\alpha_1)$ and $\overline{bet_\sigma}(set(\beta_2), set(\gamma_2)) \subseteq set(\alpha_2)$. Now we have that $\overline{bet_\sigma}(set(\beta_1 \wedge \beta_2), set(\gamma_1 \wedge \gamma_2)) = \overline{bet_\sigma}(set(\beta_1) \cap set(\beta_2), set(\gamma_1) \cap set(\gamma_2)) \subseteq \overline{bet_\sigma}(set(\beta_1), set(\gamma_1)) \subseteq set(\alpha_1)$, using the monotonicity of $\overline{bet_\sigma}$ w.r.t. set inclusion and using induction. For the same reason we also find $\overline{bet_\sigma}(set(\beta_1 \wedge \beta_2), set(\gamma_1 \wedge \gamma_2)) \subseteq set(\alpha_2)$, which allows us to conclude that $\overline{bet_\sigma}(set(\beta_1 \wedge \beta_2), set(\gamma_1 \wedge \gamma_2)) \subseteq set(\alpha_1) \cap set(\alpha_2) = set(\alpha_1 \wedge \alpha_2)$.

The case where $\Sigma \vdash \beta \bowtie_\sigma \gamma \rightarrow \alpha$ was obtained from ($\overline{\mathbf{bet_6}}$) is trivial.

($\Leftarrow$) Now we assume that $\overline{bet_\sigma}(set(\beta), set(\gamma)) \subseteq set(\alpha)$ and show that $\Sigma \vdash \beta \bowtie_\sigma \gamma \rightarrow \alpha$ can be derived from the inference rules ($\overline{\mathbf{bet_1}}$)–($\overline{\mathbf{bet_6}}$).

First assume that β is inconsistent, then $\Sigma \vdash \beta \bowtie_\sigma \gamma \rightarrow \alpha$ can be derived using ($\overline{\mathbf{bet_6}}$). If γ is inconsistent, then $\Sigma \vdash \beta \bowtie_\sigma \gamma \rightarrow \alpha$ can be derived using ($\overline{\mathbf{bet_3}}$) and ($\overline{\mathbf{bet_6}}$). If neither of β and γ is inconsistent (which implies that α cannot be inconsistent either), given that ($\overline{\mathbf{bet_2}}$) entails syntax-independence, we can assume without lack of generality that $\alpha = \alpha_1 \vee ... \vee \alpha_r$, $\beta = \beta_1 \vee ... \vee \beta_s$ and $\gamma = \gamma_1 \vee ... \vee \gamma_t$, where each α_i, β_i and γ_i is a conjunction containing exactly one atom from each attribute domain. In other words, $set(\alpha_i) = \{\mathbf{a_i}\}$, $set(\beta_i) = \{\mathbf{b_i}\}$ and $set(\gamma_i) = \{\mathbf{c_i}\}$ are singletons. By definition of $\overline{bet_\sigma}$, we then have that $\overline{bet_\sigma}(set(\beta), set(\gamma)) = \bigcup_{i,j} \overline{bet_\sigma}(\mathbf{b_i}, \mathbf{c_j})$. In particular, we also find that for every i and j there are $k_1, ..., k_l$ such that $\overline{bet_\sigma}(\mathbf{b_i}, \mathbf{c_j}) \subseteq \{\mathbf{a_{k_1}}, ..., \mathbf{a_{k_l}}\}$. The latter inclusion means that every component of each of $\mathbf{a_{k_1}}, ..., \mathbf{a_{k_l}}$ is between the corresponding components of $\mathbf{b_i}$ and $\mathbf{c_j}$, and furthermore that there are no other such vectors. Using ($\overline{\mathbf{bet_1}}$) (given (44)) and ($\overline{\mathbf{bet_5}}$) we can therefore derive that $\Sigma \vdash \beta_i \bowtie \gamma_j \rightarrow \alpha_{k_1} \vee ... \vee \alpha_{k_l}$, and by ($\overline{\mathbf{bet_2}}$) that $\Sigma \vdash \beta_i \bowtie \gamma_j \rightarrow \alpha$. As we can derive this for every i and j, ($\overline{\mathbf{bet_4}}$) finally allows us to derive $\Sigma \vdash \beta \bowtie \gamma \rightarrow \alpha$.

Proof of Proposition 11

($\Rightarrow$) Assume that $\Sigma \vdash \delta \rightarrow \gamma \rhd_\tau \langle\alpha, \beta\rangle$ can be derived from the inference rules ($\underline{\mathbf{par}}_\mathbf{1}$)–($\underline{\mathbf{par}}_\mathbf{12}$). We show by induction that then $set(\delta) \subseteq \underline{par_\tau}(set(\alpha), set(\beta), set(\gamma))$.

If $\Sigma \vdash \delta \rightarrow \gamma \rhd_\tau \langle\alpha, \beta\rangle$ was obtained from ($\underline{\mathbf{par}}_\mathbf{1}$), then α, β, γ and δ are atoms and we have from (41) that $\underline{par_\tau}(\alpha, \beta, \gamma; \delta)$ holds, from which we find $set(\delta) \subseteq \underline{par_\tau}(set(\alpha), set(\beta), set(\gamma))$.

If $\Sigma \vdash \delta \rightarrow \gamma \rhd_\tau \langle\alpha, \beta\rangle$ was obtained from ($\underline{\mathbf{par}}_\mathbf{2}$), then there exist α', β', γ', δ' and τ' such that $\alpha' \models_\mathcal{A} \alpha$, $\beta' \models_\mathcal{A} \beta$, $\gamma' \models_\mathcal{A} \gamma$, $\delta \models_\mathcal{A} \delta'$, $\tau' \subseteq \tau$ and $\Sigma \vdash \delta' \rightarrow \gamma' \rhd_{\tau'} \langle\alpha', \beta'\rangle$. By induction, we then have that $set(\delta') \subseteq \underline{par_{\tau'}}(set(\alpha'), set(\beta'), set(\gamma'))$, while $set(\alpha') \subseteq set(\alpha)$, $set(\beta') \subseteq set(\beta)$, $set(\gamma') \subseteq set(\gamma)$, and $set(\delta) \subseteq set(\delta')$. By the fact that $\underline{par_\tau}$ is clearly monotonic w.r.t. set inclusion, it follows that $set(\delta) \subseteq \underline{par_\tau}(set(\alpha), set(\beta), set(\gamma))$.

The case where $\Sigma \vdash \delta \rightarrow \gamma \rhd_\tau \langle\alpha, \beta\rangle$ was obtained from ($\underline{\mathbf{par}}_\mathbf{3}$) follows immediately from the definition of $\underline{par_\tau}$, given that the strong consistency of α implies that $set(\alpha)$ is realizable (by Lemma 1).

The case where $\Sigma \vdash \delta \rightarrow \gamma \rhd_\tau \langle\alpha, \beta\rangle$ was obtained from ($\underline{\mathbf{par}}_\mathbf{4}$) follows immediately from the fact that $\underline{par_\tau}(X_1, X_2, X_3) \supseteq X_1 \cap X_2 \cap X_3$ according to (29).

The cases where $\Sigma \vdash \delta \rightarrow \gamma \rhd_\tau \langle\alpha, \beta\rangle$ was obtained from ($\underline{\mathbf{par}}_\mathbf{5}$), ($\underline{\mathbf{par}}_\mathbf{6}$), ($\underline{\mathbf{par}}_\mathbf{7}$) and ($\underline{\mathbf{par}}_\mathbf{8}$) are entirely analogous.

If $\Sigma \vdash \delta \rightarrow \gamma \rhd_\tau \langle\alpha, \beta\rangle$ was obtained from ($\underline{\mathbf{par}}_\mathbf{9}$), then there exist α_1, α_2, β_1, β_2, γ_1, γ_2, δ_1 and δ_2 such that $\alpha = \alpha_1 \vee \alpha_2$, $\beta = \beta_1 \vee \beta_2$, $\gamma = \gamma_1 \vee \gamma_2$,

$\delta = \delta_1 \vee \delta_2$ and, by induction, $set(\delta_1) \subseteq \underline{par}_\tau(set(\alpha_1), set(\beta_1), set(\gamma_1))$ and $set(\delta_2) \subseteq \underline{par}_\tau(set(\alpha_2), set(\beta_2), set(\gamma_2))$. We have

$$\begin{aligned}&\underline{par}_\tau(set(\alpha), set(\beta), set(\gamma))\\&\qquad = \underline{par}_\tau(set(\alpha_1) \cup set(\alpha_2), set(\beta_1) \cup set(\beta_2), set(\gamma_1) \cup set(\gamma_2))\end{aligned}$$

From the monotonicity of $\underline{par}_\tau$ it easily follows that

$$\begin{aligned}&\underline{par}_\tau(set(\alpha_1), set(\beta_1), set(\gamma_1))\\&\qquad \subseteq \underline{par}_\tau(set(\alpha_1) \cup set(\alpha_2), set(\beta_1) \cup set(\beta_2), set(\gamma_1) \cup set(\gamma_2))\\&\underline{par}_\tau(set(\alpha_2), set(\beta_2), set(\gamma_2))\\&\qquad \subseteq \underline{par}_\tau(set(\alpha_1) \cup set(\alpha_2), set(\beta_1) \cup set(\beta_2), set(\gamma_1) \cup set(\gamma_2))\end{aligned}$$

Hence, we obtain $set(\alpha_1) \cup set(\alpha_2) \subseteq \underline{par}_\tau(set(\alpha_1) \cup set(\alpha_2), set(\beta_1) \cup set(\beta_2), set(\gamma_1) \cup set(\gamma_2))$ which is equivalent to what we need to prove.

If $\Sigma \vdash \delta \to \gamma \rhd_\tau \langle\alpha, \beta\rangle$ was obtained from ($\underline{\mathbf{par}}_{\mathbf{10}}$), then there exist α', β' and ϕ such that $\alpha = \alpha' \wedge \phi$ and $\beta = \beta' \wedge \phi$ and, by induction, $set(\delta) \subseteq \underline{par}_\tau(set(\alpha'), set(\beta'), set(\gamma))$. Without lack of generality, we can assume that the elements from $domains(\phi)$ are all orthogonal to each other (because ϕ was assumed to be strongly consistent), and that $\phi = \phi_1 \vee ... \vee \phi_s$ is in disjunctive-normal form. Because of ($\underline{\mathbf{par}}_{\mathbf{9}}$), it is sufficient to show that $set(\delta) \subseteq \underline{par}_\tau(set(\alpha \wedge \delta_i), set(\beta \wedge \delta_i), set(\gamma))$ for each i. Since these elements from $domains(\delta)$ are assumed to be orthogonal to the elements of $domains(\alpha) \cup domains(\beta) \cup domains(\gamma) \cup domains(\delta)$, however, the latter inclusion follows easily from the definition of $\underline{par}_\tau$, given that we already know that $set(\delta) \subseteq \underline{par}_\tau(set(\alpha'), set(\beta'), set(\gamma))$.

The case where ($\underline{\mathbf{par}}_{\mathbf{11}}$) is entirely analogous to the case for ($\underline{\mathbf{par}}_{\mathbf{8}}$) in the proof of Proposition 9.

The case where $\Sigma \vdash \delta \to \gamma \rhd_\tau \langle\alpha, \beta\rangle$ was obtained from ($\underline{\mathbf{par}}_{\mathbf{12}}$) is trivial.

($\Leftarrow$) Now we assume that $set(\delta) \subseteq \underline{par}_\tau(set(\alpha), set(\beta), set(\gamma))$ and show that $\Sigma \vdash \delta \to \gamma \rhd_\tau \langle\alpha, \beta\rangle$ can be derived from the inference rules ($\underline{\mathbf{par}}_{\mathbf{1}}$)–($\underline{\mathbf{par}}_{\mathbf{12}}$). If either of α, β or γ is inconsistent, then we must have that $set(\delta) = \emptyset$, i.e. that δ is inconsistent as well. This means that $\Sigma \vdash \delta \to \gamma \rhd_\tau \langle\alpha, \beta\rangle$ can be derived from ($\underline{\mathbf{par}}_{\mathbf{12}}$) and ($\underline{\mathbf{par}}_{\mathbf{2}}$).

If α, β and γ are consistent, given that ($\underline{\mathbf{par}}_\mathbf{2}$) entails syntax-independence, we can assume without lack of generality that δ is of the form $\delta_1 \vee ... \vee \delta_r$, where each δ_i is a conjunction containing exactly one atom from each attribute domain. In other words, $\mathit{set}(\delta_i) = \{\mathbf{d}\}$ is a singleton. Because of inference rules ($\underline{\mathbf{par}}_\mathbf{9}$) and ($\underline{\mathbf{par}}_\mathbf{2}$), it suffices to show that $\Sigma \vdash \delta_i \rightarrow \gamma \rhd_\tau \langle \alpha, \beta \rangle$ can be derived for each i.

The fact that $\mathbf{d} \in \underline{par_\tau}(\mathit{set}(\alpha), \mathit{set}(\beta), \mathit{set}(\gamma))$ either means that (i) $\mathbf{d} \in \mathit{set}(\beta) \cup \mathit{set}(\gamma)$, (ii) there is an $\mathbf{a} \in \mathit{set}(\alpha)$, $\mathbf{b} \in \mathit{set}(\beta)$ and $\mathbf{c} \in \mathit{set}(\gamma)$ such that $\underline{par_\tau}(\mathbf{a}, \mathbf{b}, \mathbf{c}; \mathbf{d})$ holds, (iii) it holds that $\mathbf{d}^{\downarrow j} \in \underline{par_\tau}(\mathit{set}(\alpha)^{\downarrow j}, \mathit{set}(\beta)^{\downarrow j}, \mathit{set}(\gamma)^{\downarrow j})$ for some j, where we write $\mathbf{d}^{\downarrow j}$ for the vector $\mathbf{d}$ without the j^{th} component and $\mathit{set}(\alpha)^{\downarrow j}$, $\mathit{set}(\beta)^{\downarrow j}$ and $\mathit{set}(\gamma)^{\downarrow j}$ are defined as in (22), or (iv) it holds that $\mathbf{d}^{\downarrow j} \in \underline{par_\tau}(\mathit{set}(\alpha)^{\downarrow j}, \mathit{set}(\beta)^{\downarrow y}, \mathit{set}(\gamma)^{\downarrow y})$ for some j and some y from the corresponding attribute domain.

In the first case, if moreover $\mathbf{d} \in \mathit{set}(\beta) \cap \mathit{set}(\gamma) \cap \mathit{set}(\alpha)$, we can derive $\Sigma \vdash \delta_i \rightarrow \delta_i \rhd_\tau \langle \delta_i, \delta_i \rangle$ using ($\underline{\mathbf{par}}_\mathbf{4}$), and subsequently also that $\Sigma \vdash \delta_i \rightarrow \gamma \rhd_\tau \langle \alpha, \beta \rangle$ using ($\underline{\mathbf{par}}_\mathbf{2}$). If $\mathbf{d} \in (\mathit{set}(\beta) \cap \mathit{set}(\gamma)) \setminus \mathit{set}(\alpha)$, we know that α is strongly consistent and $0 \in \tau$, which means that we can derive $\Sigma \vdash \delta_i \rightarrow \delta_i \rhd_\tau \langle \alpha, \delta_i \rangle$ using ($\underline{\mathbf{par}}_\mathbf{8}$) and thus $\Sigma \vdash \delta_i \rightarrow \gamma \rhd_\tau \langle \alpha, \beta \rangle$ using ($\underline{\mathbf{par}}_\mathbf{2}$). The case where $\mathbf{d} \in (\mathit{set}(\alpha) \cap \mathit{set}(\gamma)) \setminus \mathit{set}(\beta)$ is entirely analogous (using ($\underline{\mathbf{par}}_\mathbf{7}$) instead of ($\underline{\mathbf{par}}_\mathbf{8}$)). If $\mathbf{d} \in \mathit{set}(\beta) \setminus (\mathit{set}(\alpha) \cup \mathit{set}(\gamma))$, we know that $\alpha \wedge \gamma$ is strongly consistent and $1 \in \tau$, hence we can derive $\Sigma \vdash \delta_i \rightarrow \alpha \wedge \gamma \rhd_\tau \langle \alpha \wedge \gamma, \delta_i \rangle$ using ($\underline{\mathbf{par}}_\mathbf{3}$) and again $\Sigma \vdash \delta_i \rightarrow \gamma \rhd_\tau \langle \alpha, \beta \rangle$ using ($\underline{\mathbf{par}}_\mathbf{2}$). If $\mathbf{d} \in \mathit{set}(\gamma) \setminus (\mathit{set}(\alpha) \cup \mathit{set}(\beta))$, we know that either α and β are strongly consistent and $0 \in \tau$ or that $\alpha \wedge \beta$ is strongly consistent, hence we can either derive $\Sigma \vdash \delta_i \rightarrow \delta_i \rhd_\tau \langle \alpha, \beta \rangle$ using ($\underline{\mathbf{par}}_\mathbf{6}$), or we can derive $\Sigma \vdash \delta_i \rightarrow \delta_i \rhd_\tau \langle \alpha \wedge \beta, \alpha \wedge \beta \rangle$ using ($\underline{\mathbf{par}}_\mathbf{5}$). In both cases, we then find $\Sigma \vdash \delta_i \rightarrow \gamma \rhd_\tau \langle \alpha, \beta \rangle$ using ($\underline{\mathbf{par}}_\mathbf{2}$).

In the second case, we clearly have that $\mathit{conj}(\mathbf{a}) \models_\mathcal{A} \alpha$, $\mathit{conj}(\mathbf{b}) \models_\mathcal{A} \beta$ and $\mathit{conj}(\mathbf{c}) \models_\mathcal{A} \gamma$, hence because of inference rule ($\underline{\mathbf{par}}_\mathbf{2}$) it suffices to show that $\Sigma \vdash \mathit{conj}(\mathbf{d}) \rightarrow \mathit{conj}(\mathbf{c}) \rhd_\tau \langle \mathit{conj}(\mathbf{a}), \mathit{conj}(\mathbf{b}) \rangle$ can be derived. From the definition of $\underline{par_\tau}$, we know that all components correspond to orthogonal attribute domains. Moreover, either

- $\mathbf{a} = \mathbf{c}$, $\mathbf{b} = \mathbf{d}$ and $1 \in \tau$,
- $\mathbf{d} = \mathbf{c}$ and $0 \in \tau$,
- $\mathbf{a}$ differs only in one component from $\mathbf{b}$, $\mathbf{c}$ differs from $\mathbf{d}$ only in

that same component, and the respective values a_j, b_j, c_j, d_j of the component satisfy $\underline{par_\tau}(a_j, b_j, c_j; d_j)$,

- $\tau = [\mu, \mu]$ is a degenerate interval, and for each component j, it holds that $\underline{par_\tau}(a_j, b_j, c_j; d_j)$.

In the situation where $\mathbf{a} = \mathbf{c}$ and $\mathbf{b} = \mathbf{d}$, we can derive $\Sigma \vdash \mathit{conj}(\mathbf{d}) \rightarrow \mathit{conj}(\mathbf{c}) \rhd_\tau \langle \mathit{conj}(\mathbf{a}), \mathit{conj}(\mathbf{b})\rangle$ using ($\underline{\mathbf{par_3}}$). In the second situation, we can use ($\underline{\mathbf{par_6}}$). In the third situation, it is clear that $\Sigma \vdash \mathit{conj}(\mathbf{d}) \rightarrow \mathit{conj}(\mathbf{c}) \rhd_\tau \langle \mathit{conj}(\mathbf{a}), \mathit{conj}(\mathbf{b})\rangle$ can be derived by first applying inference rule ($\underline{\mathbf{par_1}}$) (given (45)) and then applying inference rule ($\underline{\mathbf{par_{10}}}$). In the last situation, we can similarly derive $\Sigma \vdash \mathit{conj}(\mathbf{d}) \rightarrow \mathit{conj}(\mathbf{c}) \rhd_\tau \langle \mathit{conj}(\mathbf{a}), \mathit{conj}(\mathbf{b})\rangle$ using ($\underline{\mathbf{par_{11}}}$) instead of ($\underline{\mathbf{par_{10}}}$).

In the third case, note that $\mathit{set}(\alpha)^{\downarrow j}$, $\mathit{set}(\beta)^{\downarrow j}$ and $\mathit{set}(\gamma)^{\downarrow j}$ correspond to formulas α^*, β^* and γ^* such that $\alpha^* \models_\mathcal{A} \alpha$, $\beta^* \models_\mathcal{A} \beta$ and $\gamma^* \models_\mathcal{A} \gamma$. Moreover, δ_i is of the form $\mathit{conj}(\mathbf{d}^{\downarrow j}) \wedge x$ for some atom x, and in particular we have that $\delta_i \models_\mathcal{A} \mathit{conj}(\mathbf{d}^{\downarrow j})$. By induction, we can moreover assume that $\Sigma \vdash \mathit{conj}(\mathbf{d}^{\downarrow j}) \rightarrow \gamma^* \rhd_\tau \langle \alpha^*, \beta^*\rangle$ can be derived. Using ($\underline{\mathbf{par_2}}$) this means that also $\Sigma \vdash \delta_i \rightarrow \gamma \rhd_\tau \langle \alpha, \beta\rangle$ can be derived.

In the fourth case, we have that $\mathit{set}(\alpha)^{\downarrow j}$, $\mathit{set}(\beta)^{\downarrow j}$ and $\mathit{set}(\gamma)^{\downarrow j}$ correspond to formulas α^*, β^* and γ^* such that $\alpha^* \models_\mathcal{A} \alpha$, $\beta^* \wedge y \models_\mathcal{A} \beta$ and $\gamma^* \wedge y \models_\mathcal{A} \gamma$. Using ($\underline{\mathbf{par_{10}}}$) (for $\psi = \top$ and $\phi = y$) and ($\underline{\mathbf{par_2}}$), we again find that $\Sigma \vdash \delta_i \rightarrow \gamma \rhd_\tau \langle \alpha, \beta\rangle$ can be derived.

Proof of Proposition 12

($\Rightarrow$) Assume that $\Sigma \vdash \gamma \rhd_\tau \langle \alpha, \beta\rangle \rightarrow \delta$ can be derived from the inference rules ($\overline{\mathbf{par_1}}$)–($\overline{\mathbf{par_9}}$). We show by induction that then

$$\overline{par_\tau}(\mathit{set}(\alpha), \mathit{set}(\beta), \mathit{set}(\gamma)) \subseteq \mathit{set}(\delta)$$

If $\Sigma \vdash \gamma \rhd_\tau \langle \alpha, \beta\rangle \rightarrow \delta$ was obtained from ($\overline{\mathbf{par_1}}$), then it is not hard to see from (42) that $\mathit{set}(\delta) = \{\mathbf{d} \,|\, \overline{par_\tau}(\mathbf{a}, \mathbf{b}, \mathbf{c}; \mathbf{d}), \mathbf{a} \in \mathit{set}(\alpha), \mathbf{b} \in \mathit{set}(\beta), \mathbf{c} \in \mathit{set}(\gamma)\}$, from which we immediately find

$$\overline{par_\tau}(\mathit{set}(\alpha), \mathit{set}(\beta), \mathit{set}(\gamma)) \subseteq \mathit{set}(\delta)$$

If $\Sigma \vdash \gamma \rhd_\tau \langle \alpha, \beta\rangle \rightarrow \delta$ was obtained from ($\overline{\mathbf{par_2}}$), then there exist α', β', γ', δ' and τ' such that $\alpha \models_\mathcal{A} \alpha'$, $\beta \models_\mathcal{A} \beta'$, $\gamma \models_\mathcal{A} \gamma'$,

$\delta' \models_{\mathcal{A}} \delta$ and $\tau' \subseteq \tau$, and $\Sigma \vdash \gamma' \rhd_{\tau'} \langle \alpha', \beta' \rangle \to \delta'$. By induction, we then have that $\overline{par_{\tau'}}(set(\alpha'), set(\beta'), set(\gamma')) \subseteq set(\delta')$, while $set(\alpha) \subseteq set(\alpha')$, $set(\beta) \subseteq set(\beta')$, $set(\gamma) \subseteq set(\gamma')$ and $set(\delta') \subseteq set(\delta)$. By the fact that $\overline{par_{\tau}}$ is clearly monotonic w.r.t. set inclusion, it follows that $\overline{par_{\tau}}(set(\alpha), set(\beta), set(\gamma)) \subseteq set(\delta)$.

If $\Sigma \vdash \gamma \rhd_{\tau} \langle \alpha, \beta \rangle \to \delta$ was obtained from ($\overline{\mathbf{par}}_{\mathbf{3}}$), then there exist α_1, α_2, δ_1 and δ_2 such that $\alpha = \alpha_1 \vee \alpha_2$, $\delta = \delta_1 \vee \delta_2$ and, by induction, $\overline{par_{\tau}}(set(\alpha_1), set(\beta), set(\gamma)) \subseteq set(\delta_1)$ and $\overline{par_{\tau}}(set(\alpha_2), set(\beta), set(\gamma)) \subseteq set(\delta_2)$. Now, we have that $\overline{par_{\tau}}(set(\alpha), set(\beta), set(\gamma)) = \overline{par_{\tau}}(set(\alpha_1) \cup set(\alpha_2), set(\beta), set(\gamma))$. From the definition of $\overline{par_{\tau}}$ it easily follows that the latter expression is equal to

$$\overline{par_{\tau}}(set(\alpha_1), set(\beta), set(\gamma)) \cup \overline{par_{\tau}}(set(\alpha_2), set(\beta), set(\gamma))$$

By the assumption, this latter expression is known to be included in $set(\delta_1) \vee set(\delta_2)$, which is equal to $set(\delta)$.

The cases where $\Sigma \vdash \gamma \rhd_{\tau} \langle \alpha, \beta \rangle \to \delta$ was obtained from ($\overline{\mathbf{par}}_{\mathbf{4}}$) or from ($\overline{\mathbf{par}}_{\mathbf{5}}$) are entirely analgous.

If $\Sigma \vdash \gamma \rhd_{\tau} \langle \alpha, \beta \rangle \to \delta$ was obtained from ($\overline{\mathbf{par}}_{\mathbf{6}}$), then there exist α_1, α_2, β_1, β_2, γ_1, γ_2, δ_1 and δ_2 such that $\alpha = \alpha_1 \wedge \alpha_2$, $\beta = \beta_1 \wedge \beta_2$, $\gamma = \gamma_1 \wedge \gamma_2$, $\delta = \delta_1 \wedge \delta_2$ and, by induction, $\overline{par_{\tau}}(set(\alpha_1), set(\beta_1), set(\gamma_1)) \subseteq set(\delta_1)$ and $\overline{par_{\tau}}(set(\alpha_2), set(\beta_2), set(\gamma_2)) \subseteq set(\delta_2)$. Given the monotonicity of $\overline{par}$ w.r.t. set inclusion, we have that

$$\begin{aligned}
\overline{par_{\tau}}(set(\alpha_1 \wedge \alpha_2), set(\beta_1 \wedge \beta_2), set(\gamma_1 \wedge \gamma_2)) &\subseteq \overline{par_{\tau}}(set(\alpha_1), set(\beta_1), set(\gamma_1)) \\
&\subseteq set(\delta_1) \\
\overline{par_{\tau}}(set(\alpha_1 \wedge \alpha_2), set(\beta_1 \wedge \beta_2), set(\gamma_1 \wedge \gamma_2)) &\subseteq \overline{par_{\tau}}(set(\alpha_2), set(\beta_2), set(\gamma_2)) \\
&\subseteq set(\delta_2)
\end{aligned}$$

which leads to $\overline{par_{\tau}}(set(\alpha), set(\beta), set(\gamma)) \subseteq set(\delta)$.

The cases where $\Sigma \vdash \gamma \rhd_{\tau} \langle \alpha, \beta \rangle \to \delta$ was obtained from ($\overline{\mathbf{par}}_{\mathbf{7}}$), ($\overline{\mathbf{par}}_{\mathbf{8}}$) or ($\overline{\mathbf{par}}_{\mathbf{9}}$) are trivial.

($\Leftarrow$) Now we assume that $\overline{par_{\tau}}(set(\alpha), set(\beta), set(\gamma)) \subseteq set(\delta)$ and show that $\Sigma \vdash \gamma \rhd_{\tau} \langle \alpha, \beta \rangle \to \delta$ can be derived from the inference rules ($\overline{\mathbf{par}}_{\mathbf{1}}$)–($\overline{\mathbf{par}}_{\mathbf{9}}$). If α, β or γ is inconsistent, $\Sigma \vdash \gamma \rhd_{\tau} \langle \alpha, \beta \rangle \to \delta$ can be inferred immediately using ($\overline{\mathbf{par}}_{\mathbf{7}}$)–($\overline{\mathbf{par}}_{\mathbf{9}}$) and ($\overline{\mathbf{par}}_{\mathbf{2}}$).

If neither of α, β and γ is inconsistent, given that ($\overline{\mathbf{par_2}}$) entails syntax-independence, we can assume without lack of generality that $\alpha = \alpha_1 \vee ... \vee \alpha_r$, $\beta = \beta_1 \vee ... \vee \beta_s$, $\gamma = \gamma_1 \vee ... \vee \gamma_t$, $\delta = \delta_1 \vee ... \vee \delta_u$ where each α_i, β_i, γ_i and δ_i is a conjunction containing exactly one atom from each attribute domain. In other words, $set(\alpha_i) = \{\mathbf{a_i}\}$, $set(\beta_i) = \{\mathbf{b_i}\}$, $set(\gamma_i) = \{\mathbf{c_i}\}$ and $set(\delta_i) = \{\mathbf{d_i}\}$ are singletons. By definition of $\overline{par_\tau}$, we then have that $\overline{par_\tau}(set(\alpha), set(\beta), set(\gamma)) = \bigcup_{i,j,l} \overline{par_\tau}(\mathbf{a_i}, \mathbf{b_j}, \mathbf{c_l})$. In particular, we also find that for every i, j and l there are $k_1, ..., k_r$ such that $\overline{par_\tau}(\mathbf{a_i}, \mathbf{b_j}, \mathbf{c_l}) \subseteq \{\mathbf{d_{k_1}}, ..., \mathbf{d_{k_r}}\}$. The latter inclusion means that the respective components of $\mathbf{a_i}$, $\mathbf{b_j}$, $\mathbf{c_l}$ on the one hand and $\mathbf{d_{k_1}}, ..., \mathbf{d_{k_r}}$ on the other hand all define parallel directions, and in particular that we can derive $\Sigma \vdash \gamma_l \rhd_\tau \langle \alpha_i, \beta_j \rangle \rightarrow \delta_{k_1} \vee ... \vee \delta_{k_r}$ using inference rules ($\overline{\mathbf{par_1}}$) (considering that (46)) and ($\overline{\mathbf{par_6}}$). As we can do this for all i, j and l, ($\overline{\mathbf{par_3}}$), ($\overline{\mathbf{par_4}}$) and ($\overline{\mathbf{par_5}}$) allow us to derive $\Sigma \vdash \gamma \rhd_\tau \langle \alpha, \beta \rangle \rightarrow \delta$.

Proof of Proposition 13

Before we move to the actual proof, we introduce some notations. Recall that for any formula α, $set(\alpha)$ was defined to be a subset of $\mathcal{A}$. However, if $domains(\alpha) \subseteq \{B_1, ..., B_s\}$, then the attribute domains in $\{C_1, ..., C_k\} \setminus \{B_1, ..., B_s\}$ are redundant, as they are not constrained by α. Throughout the proof, we will use the notation $set_{\mathcal{B}}(\alpha)$ to denote the restriction of the vectors in $set(\alpha)$ to the components corresponding to attribute domains from $\mathcal{B}$, and similar for the notation $set_{\mathcal{C}}(\alpha)$. For a vector $\mathbf{x} \in \mathcal{A}$, we also write $\mathbf{x}_{\mathcal{B}}$ and $\mathbf{x}_{\mathcal{C}}$ for its restriction to the attribute domains in $\{B_1, ..., B_s\}$ and in $\{C_1, ..., C_k\}$ respectively. Moreover, it should be clear from the definition of e.g. $\underline{bet_\sigma}$ that when $domains(\beta_1) \subseteq \{B_1, ..., B_s\}$, $domains(\beta_2) \subseteq \{B_1, ..., B_s\}$, and $\underline{bet_\sigma}(set(\beta_1), set(\beta_2))$ then we also have $\underline{bet_\sigma}(set_{\mathcal{B}}(\beta_1), set_{\mathcal{B}}(\beta_2))$, and similar for $\overline{bet_\sigma}$, $\underline{par_\sigma}$ and $\overline{par_\sigma}$, and for situations where the attribute domains are restricted to $\{C_1, ..., C_k\}$.

($\Rightarrow$) Assuming

$$(R, \Sigma) \vdash \beta \rightarrow \gamma \tag{A.4}$$

we show that $\widehat{f_R}(set_{\mathcal{B}}(\beta)) \subseteq set_{\mathcal{C}}(\gamma)$ using structural induction. Note that the soundness part of the proposition follows from this, as

$$set_{\mathcal{B}}(\bigvee_{(x_1,...,x_s)\in X} \bigwedge_i x_i) = X$$
$$set_{\mathcal{C}}(\bigvee_{(y_1,...,y_k)\in Y} \bigwedge_i y_i) = Y$$

1. If the right-hand side of (A.4) has been obtained using **(S)**, we have $R \models_{\mathcal{A}} \beta \rightarrow \gamma$ which means that $f_R(set_{\mathcal{B}}(\beta)) \subseteq set_{\mathcal{C}}(\gamma)$ by definition of f_R. Using Proposition 8, we can then conclude $\widehat{f_R}(set_{\mathcal{B}}(\beta)) \subseteq set_{\mathcal{C}}(\gamma)$.
2. Assume that the last inference rule that was applied to obtain (A.4) was **(I')**. Then there are β_1, β_2, γ_1 and γ_2 such that $(R, \Sigma) \vdash \beta_1 \rightarrow \gamma_1$, $(R, \Sigma) \vdash \beta_2 \rightarrow \gamma_2$, $\Sigma \vdash \beta \rightarrow \beta_1 \bowtie_\sigma \beta_2$ and $\Sigma \vdash \beta_1 \bowtie_\sigma \beta_2 \rightarrow \gamma$. By induction, we know that $\widehat{f_R}(set_{\mathcal{B}}(\beta_1)) \subseteq set_{\mathcal{C}}(\gamma_1)$ and $\widehat{f_R}(set_{\mathcal{B}}(\beta_2)) \subseteq set_{\mathcal{C}}(\gamma_2)$. By construction of $\widehat{f_R}$, this means that $\widehat{f_R}(\underline{bet_\sigma}(set_{\mathcal{B}}(\beta_1), set_{\mathcal{B}}(\beta_2))) \subseteq \overline{bet_\sigma}(set_{\mathcal{C}}(\gamma_1), set_{\mathcal{C}}(\gamma_2))$. Moreover, by Propositions 9 and 10, we know from $\Sigma \vdash \beta \rightarrow \beta_1 \bowtie_\sigma \beta_2$ and $\Sigma \vdash \beta_1 \bowtie_\sigma \beta_2 \rightarrow \gamma$ that $set_{\mathcal{B}}(\beta) \subseteq \underline{bet_\sigma}(set_{\mathcal{B}}(\beta_1), set_{\mathcal{B}}(\beta_2))$ and $\overline{bet_\sigma}(set_{\mathcal{C}}(\gamma_1), set_{\mathcal{C}}(\gamma_2)) \subseteq set_{\mathcal{C}}(\gamma)$. Together with the monotonicity of $\widehat{f_R}$ w.r.t. set inclusion, this leads to $\widehat{f_R}(set_{\mathcal{B}}(\beta)) \subseteq set_{\mathcal{C}}(\gamma)$.
3. The case where **(E')** is the last inference rule that was applied is entirely analogous.
4. Assume that the last inference rule that was applied to obtain (A.4) was **(D)**. Then there are formulas $\beta_1, \beta_2, \gamma_1$, and γ_2 such that

$$(R, \Sigma) \vdash \beta_1 \rightarrow \gamma_1 \tag{A.5}$$
$$(R, \Sigma) \vdash \beta_2 \rightarrow \gamma_2 \tag{A.6}$$
$$\{\beta_1 \rightarrow \gamma_1, \beta_2 \rightarrow \gamma_2\} \models_{\mathcal{A}} \beta \rightarrow \gamma \tag{A.7}$$

Without lack of generality, we can assume that $domains(\beta_1)$ and $domains(\beta_2)$ are subsets of $\{B_1, ..., B_s\}$, while $domains(\gamma_1)$ and $domains(\gamma_2)$ are subsets of $\{C_1, ..., C_k\}$.
Let $\mathbf{x} \in set_{\mathcal{B}}(\beta)$. First suppose that $\mathbf{x} \in set_{\mathcal{B}}(\beta_1)$ and $\mathbf{x} \notin set_{\mathcal{B}}(\beta_2)$. By induction, we know that $\widehat{f_R}(\mathbf{x}) \subseteq set_{\mathcal{C}}(\gamma_1)$. From (A.7), we moreover know that for every $\mathbf{a}$ in $\mathcal{A}$ such that $\mathbf{a}_{\mathcal{B}} = \mathbf{x}$ and $\mathbf{a}_{\mathcal{C}} \in$

$set_{\mathcal{C}}(\gamma_1)$, it holds that $\mathbf{a}_{\mathcal{C}} \in set_{\mathcal{C}}(\gamma)$. This means that $f_R(\mathbf{x}) \cap set_{\mathcal{C}}(\gamma_1) \subseteq set_{\mathcal{C}}(\gamma)$. Together with $\widehat{f_R}(\mathbf{x}) \subseteq set_{\mathcal{C}}(\gamma_1)$, this means in particular that $\widehat{f_R}(\mathbf{x}) \subseteq set_{\mathcal{C}}(\gamma)$.
Entirely analogously, we find $\widehat{f_R}(\mathbf{x}) \subseteq set_{\mathcal{C}}(\gamma)$ when $\mathbf{x} \notin set_{\mathcal{B}}(\beta_1)$ and $\mathbf{x} \in set_{\mathcal{B}}(\beta_2)$.
If $\mathbf{x} \in set_{\mathcal{B}}(\beta_1)$ and $\mathbf{x} \in set_{\mathcal{B}}(\beta_2)$, we find $\widehat{f_R}(\mathbf{x}) \subseteq set_{\mathcal{C}}(\gamma_1) \cap set_{\mathcal{C}}(\gamma_1)$. For every $\mathbf{a}$ in $\mathcal{A}$ such that $\mathbf{a}_{\mathcal{B}} = \mathbf{x}$, $\mathbf{a}_{\mathcal{C}} \in set_{\mathcal{C}}(\gamma_1)$ and $\mathbf{a}_{\mathcal{C}} \in set_{\mathcal{C}}(\gamma_2)$, it then holds that $\mathbf{a}_{\mathcal{C}} \in set_{\mathcal{C}}(\gamma)$ due to (A.7). Again this leads to $\widehat{f_R}(\mathbf{x}) \subseteq set_{\mathcal{C}}(\gamma)$.
Finally, if $\mathbf{x} \notin set_{\mathcal{B}}(\beta_1)$ and $\mathbf{x} \notin set_{\mathcal{B}}(\beta_2)$, we find that any $\mathbf{a} \in \mathcal{A}$ satisfying $\mathbf{a}_{\mathcal{B}} = \mathbf{x}$ is such that $\mathbf{a}_{\mathcal{C}} \in set_{\mathcal{C}}(\gamma)$. Thus again we have $\widehat{f_R}(\mathbf{x}) \subseteq set_{\mathcal{C}}(\gamma)$.

($\Leftarrow$) We show by induction on n that whenever $\widehat{f_R}^{(n)}(X) \subseteq Y$, it holds that

$$(R, \Sigma) \vdash \Big(\bigvee_{(x_1,...,x_s)\in X} \bigwedge_i x_i\Big) \rightarrow \Big(\bigvee_{(y_1,...,y_k)\in Y} \bigwedge_i y_i\Big) \tag{A.8}$$

The base case is straightforward: if $f_R(X) \subseteq Y$, then by construction of f_R, it holds that

$$R \models_{\mathcal{A}} \Big(\bigvee_{(x_1,...,x_s)\in X} \bigwedge_i x_i\Big) \rightarrow \Big(\bigvee_{(y_1,...,y_k)\in Y} \bigwedge_i y_i\Big)$$

and thus (A.8) by inference rule **(S)**.

To show the induction step, because of inference rule **(D)**, it suffices to show that for each $\mathbf{x} = (x_1, ..., x_n)$ in X, it holds that

$$(R, \Sigma) \vdash \bigwedge_i x_i \rightarrow \Big(\bigvee_{(y_1,...,y_k)\in Y} \bigwedge_i y_i\Big) \tag{A.9}$$

From $\widehat{f_R}^{(n+1)}(\{\mathbf{x}\}) \subseteq Y$, by definition of $\widehat{f_R}^{(n+1)}$ we have that there exist $Y_1, ..., Y_m$, $Z_1, ..., Z_m$, $\sigma_1, ..., \sigma_m$, $U_1, ..., U_l$, $V_1, ..., V_l$, $W_1, ..., W_l$ and $\tau_1, ..., \tau_l$ such that

$$\begin{aligned} Y \supseteq\, & f_R(\{x\}) \cap \bigcap_i \overline{bet_{\sigma_i}}(\widehat{f_R}^{(n)}(Y_i), \widehat{f_R}^{(n)}(Z_i)) \\ & \cap \bigcap_i \overline{par_{\tau_i}}(\widehat{f_R}^{(n)}(U_i), \widehat{f_R}^{(n)}(V_i), \widehat{f_R}^{(n)}(W_i)) \end{aligned} \tag{A.10}$$

and moreover $\mathbf{x} \in \underline{bet_{\sigma_i}}(Y_i, Z_i)$ for each $i \in \{1, ..., m\}$ and $\mathbf{x} \in \underline{par_{\tau_i}}(U_i, V_i, W_i)$ for each $i \in \{1, ..., l\}$.

By induction, we know that for each Y_j, it holds that

$$(R, \Sigma) \vdash (\bigvee_{(y_1,...,y_s) \in Y_j} \bigwedge_i y_i) \rightarrow (\bigvee_{(a_1,...,a_k) \in \widehat{f}_R^{(n)}(Y_j)} \bigwedge_i a_i)$$

and similar for Z_j, U_j, V_j and W_j. Furthermore, from $\mathbf{x} \in \underline{bet}(Y_i, Z_i)$ we find using Proposition 9 that

$$\Sigma \vdash \bigwedge_i x_i \rightarrow (\bigvee_{(y_1,...,y_s) \in Y_j} \bigwedge_i y_i) \bowtie_{\sigma_i} (\bigvee_{(z_1,...,z_s) \in Z_j} \bigwedge_i z_i)$$

Similarly, from $\mathbf{x} \in \underline{par}(U_j, V_j, W_j)$, we find using Proposition 11 that

$$\Sigma \vdash \bigwedge_i x_i \rightarrow (\bigvee_{(u_1,...,u_s) \in U_j} \bigwedge_i u_i) \rhd_{\tau_i} \langle (\bigvee_{(v_1,...,v_s) \in V_j} \bigwedge_i v_i), (\bigvee_{(w_1,...,w_s) \in W_j} \bigwedge_i w_i) \rangle$$

Using Proposition 10 and 12, we trivially derive

$$\begin{aligned}
\Sigma \vdash & (\bigvee_{(a_1,...,a_k) \in \widehat{f}_R^{(n)}(Y_j)} \bigwedge a_i) \bowtie_{\sigma_i} (\bigvee_{(b_1,...,b_k) \in \widehat{f}_R^{(n)}(Z_j)} \bigwedge b_i) \\
& \rightarrow (\bigvee_{(a_1,...,a_k) \in \overline{bet_{\sigma_j}}(\widehat{f}_R^{(n)}(Y_j), \widehat{f}_R^{(n)}(Z_j))} \bigwedge_i a_i) \\
\Sigma \vdash & (\bigvee_{(a_1,...,a_k) \in \widehat{f}_R^{(n)}(U_j)} \bigwedge a_i) \rhd_{\tau_j} \langle (\bigvee_{(b_1,...,b_k) \in \widehat{f}_R^{(n)}(V_j)} \bigwedge b_i), (\bigvee_{(c_1,...,c_k) \in \widehat{f}_R^{(n)}(W_j)} \bigwedge c_i) \rangle \\
& \rightarrow (\bigvee_{(a_1,...,a_k) \in \overline{proj_{\tau_j}}(\widehat{f}_R^{(n)}(U_j), \widehat{f}_R^{(n)}(V_j), \widehat{f}_R^{(n)}(W_j))} \bigwedge_i a_i)
\end{aligned}$$

Using inference rules **(I')** and **(E')**, together this allows us to derive

$$(R, \Sigma) \vdash \bigwedge_i x_i \rightarrow (\bigvee_{(a_1,...,a_k) \in \overline{bet_{\sigma_j}}(\widehat{f}_R^{(n)}(Y_j), \widehat{f}_R^{(n)}(Z_j))} \bigwedge_i a_i) \quad \text{(A.11)}$$

$$(R, \Sigma) \vdash \bigwedge_i x_i \rightarrow (\bigvee_{(a_1,...,a_k) \in \overline{par_{\tau_j}}(\widehat{f}_R^{(n)}(U_j), \widehat{f}_R^{(n)}(V_j), \widehat{f}_R^{(n)}(W_j))} \bigwedge_i a_i) \quad \text{(A.12)}$$

Using inference rule **(S)**, we find the trivial rule

$$(R,\Sigma) \vdash \bigwedge_i x_i \rightarrow \bigvee_{(a_1,\ldots,a_k)\in f_R(\{\mathbf{x}\})} \bigwedge_i a_i \tag{A.13}$$

By two repeated applications of **(D)** to combine (A.11), (A.12) and (A.13), we find

$$\begin{aligned}(R,\Sigma) \vdash \bigwedge_i x_i \rightarrow & \Big(\bigvee_{(a_1,\ldots,a_k)\in\overline{bet_{\sigma_j}}(\widehat{f}_R^{(n)}(Y_j),\widehat{f}_R^{(n)}(Z_j))} \bigwedge_i a_i \Big) \\ & \wedge \Big(\bigvee_{(a_1,\ldots,a_k)\in\overline{par_{\tau_j}}(\widehat{f}_R^{(n)}(U_j),\widehat{f}_R^{(n)}(V_j),\widehat{f}_R^{(n)}(W_j))} \bigwedge_i a_i \Big) \\ & \wedge \bigvee_{(a_1,\ldots,a_k)\in f_R(\{\mathbf{x}\})} \bigwedge_i a_i \end{aligned}$$

Finally, from (A.10), it follows that we can apply **(D)** one more time to weaken the latter rule, from which we obtain (A.9).

Proof of Proposition 14

The rule base R initially contains two types of rules. On the one hand, there is (84), whose antecedent corresponds to the outcome J, while on the other hand, there are rules of the form (82) and (83), whose antecedent does not correspond to any outcome at all. Clearly, non-trivial new rules can only be obtained from R by using interpolative reasoning. More specifically, initially, we can only combine the rule (84) with one of the other rules to yield a new rule r_1, whose consequent is of the form $y_1 \vee y_2$ and whose antecedent only differs from the antecedent of (84) in that an atom of the form x_i has been replaced by $\overline{x_i}$, or an atom of the form $\overline{x_i}$ has been replaced by x_i. Moreover, this application of interpolation can only be applied if the antecedents of both (84) and r_1 satisfy the proposition p. Thus it is clear that the antecedent of r_1 corresponds to an outcome J_1 such that there is an improving flip from J to J_1. We can then combine r_1 with some rule of the form (82) or(83) to obtain a new rule r_2 whose antecedent corresponds to an outcome J_2 such that there is an improving flip from J_1 to J_2. It is thus clear that a rule of the form (85) can only be derived if there is a sequence of outcomes $J, J_1, ..., J_m, I$ that corresponds to a sequence of improving flips, i.e. if I dominates J. Conversely, it is also clear that if there is an improving flip from J_i to J_{i+1}, we will be able to derive the corresponding rule r_{i+1} once r_i has been derived.

Proof of Proposition 15

The proof is entirely analogous to the proof of Proposition 14.

Proof of Proposition 16

By construction, it is clear that

$$\begin{aligned}\widehat{f_R}^{(i+1)}(\{\mathbf{x}\}) \subseteq & \widehat{f_R}^{(i)}(\{\mathbf{x}\}) \\ & \cap \bigcap\{\overline{bet_\sigma}(\widehat{f_R}^{(i)}(set(\alpha)), \widehat{f_R}^{(i)}(set(\gamma))) \mid c_1 \vee c_2\} \\ & \cap \bigcap\{\overline{par_\tau}(\widehat{f_R}^{(i)}(set(\alpha)), \widehat{f_R}^{(i)}(set(\beta)), \widehat{f_R}^{(i)}(set(\gamma))) \mid c_3 \vee c_4\}\end{aligned}$$

To show that the inclusion also holds in the other direction, let Y and Z be subsets of $\mathcal{B}$ such that $\mathbf{x} \in \underline{bet_\sigma}(Y, Z)$. We show that there are formulas α and γ such that condition c_1 or c_2 is satisfied and

$$\widehat{f_R}^{(i)}(\{\mathbf{x}\}) \cap \overline{bet_\sigma}(\widehat{f_R}^{(i)}(set(\alpha)), \widehat{f_R}^{(i)}(set(\gamma))) \subseteq \overline{bet_\sigma}(\widehat{f_R}^{(i)}(Y), \widehat{f_R}^{(i)}(Z)) \tag{A.14}$$

First assume that all attribute domains are orthogonal. If $\mathbf{x} \in Y$ and $0 \in \sigma$, then we will have

$$\widehat{f_R}^{(i)}(\{\mathbf{x}\}) \subseteq \overline{bet_\sigma}(\widehat{f_R}^{(i)}(\{\mathbf{x}\}), \widehat{f_R}^{(i)}(Z)) \subseteq \overline{bet_\sigma}(\widehat{f_R}^{(i)}(Y), \widehat{f_R}^{(i)}(Z))$$

and (A.14) is satisfied (regardless of the choice of α and γ). Similarly, we find that (A.14) is satisfied if $\mathbf{x} \in Z$ and $1 \in \sigma$. If $\mathbf{x} \in Y \cap Z$, we find for any (non-empty) σ that

$$\widehat{f_R}^{(i)}(\{\mathbf{x}\}) \subseteq \overline{bet_\sigma}(\widehat{f_R}^{(i)}(\{x\}), \widehat{f_R}^{(i)}(\{x\})) \subseteq \overline{bet_\sigma}(\widehat{f_R}^{(i)}(Y), \widehat{f_R}^{(i)}(Z))$$

and therefore again (A.14). If there are $\mathbf{y} \in Y$ and $\mathbf{z} \in Z$ such that $\underline{bet_\sigma}(\mathbf{y}, \mathbf{x}, \mathbf{z})$, then either we have $\mathbf{y} = \mathbf{x}$ and $0 \in \sigma$ or $\mathbf{y} = \mathbf{z}$ and $1 \in \sigma$ or $\mathbf{x} = \mathbf{y} = \mathbf{z}$, which are already covered, or we have that c_1 or c_2 is satisfied for $\alpha = conj(\mathbf{y})$ and $\gamma = conj(\mathbf{z})$. In the latter two cases we have

$$\overline{bet_\sigma}(\widehat{f_R}^{(i)}(set(\alpha)), \widehat{f_R}^{(i)}(set(\gamma))) \subseteq \overline{bet_\sigma}(\widehat{f_R}^{(i)}(Y), \widehat{f_R}^{(i)}(Z))$$

and thus also (A.14).

Now assume that some attribute domains are not orthogonal, but that $conj(\mathbf{x}) = x_j \wedge conj(\mathbf{x}')$ with $\mathbf{x}' \in \underline{bet_\sigma}(Y^{\downarrow j}, Z^{\downarrow j})$. If $\mathbf{x}' \in Y^{\downarrow j}$, $0 \in \sigma$ and the remaining attribute domains are orthogonal, we have

$$\begin{aligned}\widehat{f_R}^{(i)}(\{\mathbf{x}\}) &\subseteq \widehat{f_R}^{(i)}(set(conj(\mathbf{x}')))\\ &\subseteq \overline{bet_\sigma}(\widehat{f_R}^{(i)}(set(conj(\mathbf{x}'))), \widehat{f_R}^{(i)}(Z))\\ &\subseteq \overline{bet_\sigma}(\widehat{f_R}^{(i)}(Y), \widehat{f_R}^{(i)}(Z))\end{aligned}$$

The case where $\mathbf{x}' \in Z^{\downarrow j}$ and $1 \in \sigma$ is entirely analogous. If the remaining attribute domains are orthogonal and there are $\mathbf{y}' \in Y^{\downarrow j}$ and $\mathbf{z}' \in Z^{\downarrow j}$ such that $\underline{bet_\sigma}(\mathbf{y}', \mathbf{x}', \mathbf{z}')$, we can take $\alpha = conj(\mathbf{y}')$ and $\gamma = conj(\mathbf{z}')$. The case where $conj(\mathbf{x}') = x_l \wedge conj(\mathbf{x}'')$ with $\mathbf{x}'' \in \underline{bet_\sigma}(Y^{\downarrow j \downarrow l}, Z^{\downarrow j \downarrow l})$ is handled by repeating the same argument.

In entirely the same way, we show that for subsets X, Y and Z of $\mathcal{B}$ such that $\mathbf{x} \in \underline{par_\tau}(X, Y, Z)$, there exist conjunctions α, β and γ such that c_3 or c_4 is satisfied and

$$\begin{aligned}&\widehat{f_R}^{(i)}(\{\mathbf{x}\}) \cap \overline{par_\tau}(\widehat{f_R}^{(i)}(set(\alpha)), \widehat{f_R}^{(i)}(set(\beta)), \widehat{f_R}^{(i)}(set(\gamma)))\\ &\quad\subseteq \overline{par_\tau}(\widehat{f_R}^{(i)}(X), \widehat{f_R}^{(i)}(Y), \widehat{f_R}^{(i)}(Z))\end{aligned}$$

which completes the proof.

References

[1] S. Kraus, D. Lehmann, M. Magidor, Nonmonotonic reasoning, preferential models and cumulative logics, Artificial Intelligence 44 (1-2) (1990) 167–207.

[2] A. Collins, R. Michalski, The logic of plausible reasoning: A core theory, Cognitive Science 13 (1) (1989) 1–49.

[3] D. Gentner, K. J. Holyoak, B. N. Kokinov (Eds.), The Analogical Mind: Perspectives from Cognitive Science, MIT Press, 2001.

[4] D. Dubois, H. Prade, What are fuzzy rules and how to use them, Fuzzy Sets and Systems 84 (2) (1996) 169 – 185.

[5] R. Sun, Robust reasoning: integrating rule-based and similarity-based reasoning, Artificial Intelligence 75 (2) (1995) 241–295.

[6] D. Dubois, H. Prade, F. Esteva, P. Garcia, L. Godo, A logical approach to interpolation based on similarity relations, International Journal of Approximate Reasoning 17 (1) (1997) 1 – 36.

[7] J. Medina, M. Ojeda-Aciego, P. Vojtás, Similarity-based unification: a multi-adjoint approach, Fuzzy Sets and Systems 146 (1) (2004) 43 – 62.

[8] A. Tversky, Features of similarity, Psychological review 84 (4) (1977) 327–352.

[9] R. Goldstone, D. Medin, J. Halberstadt, Similarity in context, Memory & Cognition 25 (1997) 237–255.

[10] Y. Shoham, A semantical approach to nonmonotonic logics, in: M. Ginsberg (Ed.), Readings in nonmonotonic reasoning, Morgan Kaufmann Publishers Inc., 1987, pp. 227–250.

[11] R. Reiter, A logic for default reasoning, Artificial Intelligence 13 (1-2) (1980) 81 – 132.

[12] M. Gelfond, V. Lifschitz, The stable model semantics for logic programming, in: Proceedings of the Fifth International Conference and Symposium on Logic Programming, 1988, pp. 1081–1086.

[13] P. Gärdenfors, Conceptual Spaces: The Geometry of Thought, MIT Press, 2000.

[14] R. Speer, C. Havasi, H. Lieberman, Analogyspace: reducing the dimensionality of common sense knowledge, in: Proceedings of the 23rd AAAI Conference on Artificial intelligence, 2008, pp. 548–553.

[15] S. Schockaert, H. Prade, Qualitative reasoning about incomplete categorization rules based on interpolation and extrapolation in conceptual spaces, in: Proceedings of the Fifth International Conference on Scalable Uncertainty Management, 2011, pp. 303–316.

[16] L. Wittgenstein, Some remarks on logical form, Proceedings of the Aristotelian Society, Supplementary Volumes 9 (1929) 162–171.

[17] R. Carnap, A basic system of inductive logic, part 1, in: R. Carnap, R. Jeffrey (Eds.), Studies in Inductive Logic and Probability, Vol. 1, University of California Press, 1971, pp. 33–165.

[18] W. V. Quine, Natural kinds, in: Ontological Relativity and Other Essays, Columbia University Press, 1969, pp. 114–138.

[19] F. Lehmann, A. G. Cohn, The EGG/YOLK reliability hierarchy: semantic data integration using sorts with prototypes, in: International Conference on Information and Knowledge Management, 1994, pp. 272–279.

[20] S. Schockaert, H. Prade, Solving conflicts in information merging by a flexible interpretation of atomic propositions, Artificial Intelligence 175 (11) (2011) 1815 – 1855.

[21] E. Clementini, S. Skiadopoulos, R. Billen, F. Tarquini, A reasoning system of ternary projective relations, IEEE Transactions on Knowledge and Data Engineering 22 (2) (2010) 161 –178.

[22] T. Hahmann, M. Gruüninger, Multidimensional mereotopology with betweenness, in: Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, 2011, pp. 906–9011.

[23] E. Ruspini, On the semantics of fuzzy logic, International Journal of Approximate Reasoning 5 (1991) 45–88.

[24] T. Vetterlein, A logic of the similarity with prototypes and its relationship to fuzzy logic, in: Proceedings of the EUSFLAT-LFA Conference, 2011, pp. 196–202.

[25] D. Dubois, H. Prade, Similarity versus preference in fuzzy set-based logics, in: E. Orlowska (Ed.), Modelling Incomplete Information: Rough Set Analysis, Physica Verlag, Heidelberg, 1998, pp. 441–461.

[26] L. Zadeh, A theory of approximate reasoning, in: J. Hayes, D. Michie, L. Mikulich (Eds.), Machine Intelligence, Vol. 9, Elsevier, 1979.

[27] D. G. Schwartz, A system for reasoning with imprecise linguistic information, International Journal of Approximate Reasoning 5 (5) (1991) 463 – 488.

[28] I. Perfilieva, D. Dubois, H. Prade, F. Esteva, L. Godo, P. Hodáková, Interpolation of fuzzy data: Analytical approach and overview, Fuzzy Sets and Systems 192 (2012) 134–158.

[29] M. Sheremet, D. Tishkovsky, F. Wolter, M. Zakharyaschev, Comparative similarity, tree automata, and diophantine equations, in: Proceedings of the 12th International Conference on Logic for Programming, Artificial Intelligence, and Reasoning, Vol. 3835 of Lecture Notes in Computer Science, 2005, pp. 651–665.

[30] R. Alenda, N. Olivetti, C. Schwind, Preferential semantics for the logic of comparative concepts similarity, in: Proceedings of the International Conference on Topology, Algebra and Category in Logic, 2010.

[31] F. Dupin de Saint-Cyr, J. Lang, Belief extrapolation (or how to reason about observations and unpredicted change), Artificial Intelligence 175 (2) (2011) 760 – 790.

[32] L. Miclet, H. Prade, Handling analogical proportions in classical logic and fuzzy logics settings, in: C. Sossai, G. Chemello (Eds.), European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty, 2009, pp. 638–650.

[33] H. Prade, G. Richard, B. Yao, Classification by means of fuzzy analogy-related proportions – a preliminary report, in: Proceedings of the International Conference on Soft Computing and Pattern Recognition, 2010, pp. 297–302.

[34] H. Prade, S. Schockaert, Completing rule bases in symbolic domains by analogy making, in: Proceedings of the EUSFLAT-LFA Conference, 2011, pp. 928–934.

[35] D. Zhou, B. Schölkopf, A regularization framework for learning from graph data, in: ICML Workshop on Statistical Relational Learning, 2004, pp. 132–137.

[36] R. Gérard, S. Kaci, H. Prade, Ranking alternatives on the basis of generic constraints and examples: a possibilistic approach, in: International Joint Conference on Artifical intelligence, 2007, pp. 393–398.

[37] M. Abraham, D. Gabbay, U. Schild, Analysis of the talmudic argumentum a fortiori inference rule (kal vachomer) using matrix abduction, Studia Logica 92 (2009) 281–364.

[38] K. D. Forbus, Qualitative process theory, Artificial Intelligence 24 (1-3) (1984) 85 – 168.

[39] B. Kuipers, Qualitative simulation, Artificial Intelligence 29 (3) (1986) 289 – 338.

[40] H. Prade, G. Richard, Analogical proportions: another logical view, in: M. Bramer, R. Ellis, M. Petridis (Eds.), Research and Development in Intelligent Systems XXVI, Proc. 29th Ann. Inter. Conf. on AI (SGAI'09), Cambridge, UK, December 2009, Springer, 2010, pp. 121–134.

[41] R. Laymon, Using Scott domains to explicate the notions of approximate and idealized data, Philosophy of Science 54 (2) (1987) 194–221.

[42] P. Gärdenfors, M. Williams, Reasoning about categories in conceptual spaces, in: International Joint Conference on Artificial Intelligence, 2001, pp. 385–392.

[43] J. Aisbett, G. Gibbon, A general formulation of conceptual spaces as a meso level representation, Artificial Intelligence 133 (1-2) (2001) 189 – 232.

[44] M. Raubal, Formalizing conceptual spaces, Formal Ontology in Information Systems 114 (2004) 153–164.

[45] R. Billen, E. Clementini, Semantics of collinearity among regions, in: On the Move to Meaningful Internet Systems 2005: OTM Workshops, Vol. 3762 of Lecture Notes in Computer Science, 2005, pp. 1066–1076.

[46] Z. Pawlak, Rough Sets: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, 1992.

[47] P. D. Turney, Measuring semantic similarity by latent relational analysis, in: Proceedings of the 19th International Joint Conference on Artificial Intelligence, 2005, pp. 1136–1141.

[48] M. Bayoudh, H. Prade, G. Richard, Evaluation of analogical proportions through Kolmogorov complexity, Knowledge-Based Systems 29 (2012) 20–30.

[49] R. McDonald, F. Pereira, S. Kulick, S. Winters, Y. Jin, P. White, Simple algorithms for complex relation extraction with applications to biomedical ie, in: Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics, 2005, pp. 491–498.

[50] S. Schockaert, H. Prade, Interpolation and extrapolation in conceptual spaces: A case study in the music domain, in: Proceedings of the 5th International Conference on Web Reasoning and Rule Systems, 2011, pp. 217–231.

[51] T. R. Davies, S. J. Russell, A logical approach to reasoning by analogy, in: International Joint Conference on Artificial Intelligence, 1987, pp. 264–270.

[52] A. Tarski, A lattice theoretical fixpoint theorem and its application, Pacific Journal of Mathematics 5 (1955) 285–309.

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

[54] J. Goldsmith, J. Lang, N. Wilson, The computational complexity of dominance and consistency in CP-nets, Journal of Artificial Intelligence Research 33 (2008) 403–432.