A Qualitative Approach to Vague Spatio-Thematic Query Processing

Rolf Grütter and Thomas Scharrenbach

Swiss Federal Research Institute WSL, an Institute of the ETH Domain, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland
{Rolf.Gruetter, Thomas.Scharrenbach}@wsl.ch

Abstract. In order to support the processing of spatial queries, spatial knowledge must be represented in a way that machines can make use of it. In ontology-based geographic information systems, a challenge thus is to enhance thematic knowledge with spatial knowledge. A way to achieve this is to combine existing approaches to spatial knowledge representation with ontologies. In this paper an implementation of the Region Connection Calculus (RCC) in the Web Ontology Language (OWL) augmented by DL-safe rules is used in order to represent spatio-thematic knowledge. It is shown how the represented knowledge supports the processing of queries using (possibly vague) spatial concepts. Thereby, the division of land into administrative regions, rather than, for instance, a metric system, is taken as a frame of reference for evaluating closeness. Hence, closeness is evaluated based on purely qualitative criteria. Since colloquial descriptions typically involve qualitative concepts, the presented approach is expected to align better with the way human beings deal with closeness than a quantitative approach. The paper is discussed w.r.t. related work and an overview of possible future extensions is provided.

1 Introduction

Fueled by a joint initiative of research institutes and industrial organizations towards the Semantic Web, knowledge representation has regained considerable attention through the last decade [1]. Technically, the initiative was committed to advance Description Logics (DLs), f.k.a. terminological systems, as a means for capturing the terminological and assertional knowledge of a domain and for inferring new knowledge from existing. This kind of knowledge has been (and continues to be) made available by so-called ontologies [2].

Ontologies are increasingly integrated into applications in order to support semantic interoperability and to provide a homogenous view of heterogeneous data [3, 4]. In geographic information systems, where mereological considerations are all-important, for instance, in order to process spatial queries, a challenge is to enhance ontologies, usually representing purely thematic knowledge, with spatial knowledge [5, 6]. A way to achieve this is to combine existing logic-based approaches to spatial knowledge representation with description logics [7, 8, 9, 10].

In this paper, an implementation of the Region Connection Calculus (RCC) in the Web Ontology Language (OWL) augmented by DL-safe rules is used in order to represent spatio-thematic knowledge. We show how such a representation can be applied for answering queries involving (possibly vague) spatial concepts. The primary goal is to demonstrate how state-of-the-art technology can be used for this purpose. While pursuing this goal, some additions to the theory of spatial knowledge representation are made. The basic idea underlying our approach is to take the division of land into administrative regions, rather than, for instance, a metric system, as a frame of reference for evaluating closeness. Accordingly, closeness is evaluated based on purely qualitative criteria. Since colloquial descriptions typically involve qualitative concepts, our approach is expected to align better with the way human beings deal with closeness than a quantitative approach.

The paper is organized as follows: Section 2 provides an overview of recent work on vague spatial concepts. Insights gained from this overview are used later in the paper when describing and implementing the notion of spatial closeness. Section 3 provides a short introduction into DL-safe rules which are used – together with OWL DL and RCC – for the representation of qualitative spatio-thematic knowledge. In Section 4, the vague notion of spatial closeness is introduced into RCC and its implementation in DL augmented by DL-safe rules is outlined. In Section 5, the approach is applied to spatio-thematic query answering in the Web. Section 6 discusses the approach and Section 7 concludes with an overview of future work.

2 Related Work

There is a bulk of work about vague spatial concepts in both philosophy and geographic information science. In this paper, we limit our discussion to the most recent works. For a comprehensive survey, particularly of approaches using fuzzy logic or contextual information, the interested reader is referred to [11].

In [12] an experiment with human subjects concerning the vague spatial relation "near" between places in environmental space is presented.[1] An environmental space is referred to as the space of buildings, neighbourhoods, and cities, without consideration of symbolic representations such as maps. In order to better understand how humans conceptualize nearness and to test the fit of formal theories to human concepts, the author seeks to apply appropriate theories to the data resulting from the experiment. Amongst other insights into the conceptualization of "nearness", the experiment shows the importance of scale factors introduced by the context of the reference place. This supports our claim made in Section 4 that scales, more precisely, the categories in which terms human subjects think, are an important contingency of the notion of closeness.[2]

In [13] a qualitative representation for spatial proximity that accounts for absolute binary nearness relations is introduced. The formalism is based on the notion of perceived points, called sites, in a point based universe. Proximity concepts are determined by the parameters of distance between two sites and weight of each of those

[1] In the literature the term "near" is often used to denote the same concept as "close to".

[2] cf. "In the discussion of proximity perception and cognition, the term 'scale' refers to the size of the frame of reference in a perceiver's mental map, i.e. the spatial extent of the area that is considered when a distance is assessed." [11, p. 162]

sites. These parameters are drawn from the concept of Generalized Voronoi Diagrams, i.e., Power Diagrams.

The approach introduced in [13] and that presented here have in common that the qualitative description of nearness is based on a *qualitative* representation of distance: in their case Voronoi diagrams transform (quantitative) distances into a network of (qualitative) topological relations. This is different from all other approaches discussed in this paper, where a mapping mechanism between qualitative and *metric* distance measures is established (or implied). While the authors link their concept of nearness to the topological relations equality, external connectedness and inclusion – which can also be expressed by RCC – this link is established by the *areas of influence* of perceived points in a point based universe. Polygons and spatial relations between polygons are not considered. While the approach is appealing in its formal strictness and the cognitively useful models and interpretations provided, it does not address the issue of grounding. In particular, it is not clear how the weights $w(p)$ associated with every site p are retained from the abstracted "real world" entity.

In order to enable metric systems (such as GIS) to translate between linguistic proximity measures and metric distance measures, a statistical approach to context-contingent proximity modelling is presented in [11]. Relevant context factors are chosen that influence, according to empirical studies, the way human beings reason about proximity. The presented translation mechanism works in one direction only: Given the corresponding metric distance measures and context information, linguistic proximity measures are "predicted". This direction does not support the obvious translation of local prepositions, such as "near" or "far", used in natural language queries into distance measures processed in metric systems. In [14] the authors note that the information required to bind the numerous context variables may not be available in a practical application, and hence it is difficult to see how they would be implemented on a large scale.

According to [15], the answer to whether something is near or not depends on the context in which the question is asked and the nature of the objects being compared. In [14] ontologies are used to make explicit the vague spatial relation "near" for database querying. In order to keep it implementable in a practical system, the algorithm used to calculate the relation "near" is relatively simple. It only uses two contextual parameters, namely Euclidean distance from a reference point and density of the feature class. Despite its simplicity it achieves perfect precision and recall when applied to a (small) number of test sets obtained by asking people which objects were near to each of the reference points. Different from the approach presented here, the authors base their algorithm on quantitative parameters, namely Euclidean distance and gravity (which is a measure of how objects are distributed), and not on qualitative relations. Furthermore, they reduce the objects considered to centroids and the calculations to point calculations whereas we consider spatial regions (i.e., polygons).

In [16] an approach to geographic information retrieval integrating topological, geographical and conceptual matching is presented. For topological matching topological relations are extracted from overlaying data layers; for geographical matching constraints are obtained from dictionaries; for conceptual matching a geographic ontology is used. A constraint, provided as an example, defines two geographic objects (points or polygons) as near provided they are connected by a third object (an arc, e.g., a road), the length of which is less than 1 kilometre. Different from the approach presented here, a metric distance measure thus is a necessary condition for nearness, although not a sufficient. However, the framework seems general enough to be aligned with that presented here.

3 Preliminaries

The presented framework uses a number of spatial relations from different RCC sublanguages, particularly RCC-8, and a composition rule. It also uses the subsumption hierarchy of RCC relations and a sum function as introduced in [17]. These spatial notions are implemented in OWL DL augmented by DL-safe rules. We thus assume that the reader is familiar with RCC [17, 18] and description logics [19], particularly OWL DL [20].[3]

As mentioned, the composition rule of the framework is implemented as a DL-safe rule. DL-safe rules are function-free Horn rules with the restriction that each variable in the rule occurs in a non-DL-atom in the rule body [22]. This is achieved by adding special non-DL-literals such as $\mathcal{O}(x)$ to the rule body, and by adding a fact $\mathcal{O}(a)$ for each individual a to the knowledge base. For instance, the RCC-5 composition rule $\forall x \forall y \forall z\ [\mathsf{PP}(x, y) \wedge \mathsf{EQ}(z, y) \rightarrow \mathsf{PPi}(z, x)]$ is implemented as the DL-safe rule $\mathsf{properPartOf}(x, y) \wedge \mathsf{equalTo}(z, y) \wedge \mathcal{O}(x) \wedge \mathcal{O}(y) \wedge \mathcal{O}(z) \rightarrow \mathsf{inverseProperPartOf}(z, x)$ where $\mathcal{O}(x)$, $\mathcal{O}(y)$ and $\mathcal{O}(z)$ are non-DL-literals.[4] While in theory DL-safe rules support complex, i.e., disjunctive, heads (respectively negation in the rule body) [23, 24], there is currently no implementation that supports this feature. However, since RCC relations describe a closed world [17], it is always possible to replace a negative atom, for instance $\neg\mathsf{disconnectedFrom}(z, y)$, by a, possibly auxiliary (cf. Section 4.1), positive atom, for instance $\mathsf{connectsWith}(z, y)$.

4 Representing Spatio-Thematic Knowledge

4.1 Defining Closeness in RCC

A basic assumption underlying our approach is that administrative regions are social artifacts and their organization is largely, if not entirely, motivated by the property of spatial closeness. To be more precise, administrative regions are assumed to mirror how a collective perceives spatial closeness on increasing scales of social organization. We will come back to this when applying the approach to an example query.

Since administrative regions are typically organized in partitions it is necessary to introduce the notion of a partition and to reformulate it in a way that is compliant with a model-theoretic interpretation of RCC, the formalism used for expressing closeness.

---

[3] An approach similar to RCC which is based on the description of topological relations between two spatial regions was introduced as the 9-intersection model in [21].

[4] In DL-safe rules, all variables are universally quantified.

Compliance with this kind of interpretation is a requirement for the implementation of RCC in a DL knowledge base and rule base in Section 4.2. Further, for asserting closeness between individual regions, we must slightly extend RCC and introduce closeness by an additional relation. The idea is that, given the conceptualization of a user in terms of a query and a partially ordered and typed system of partitions, closeness can be evaluated by a composition rule.

Definition 1a (Partition). A partition is defined as a (possibly improper) subset of the power set of a set Y, denoted by $(Y_i)_{i \in I} \subseteq \mathcal{P}(Y)$, for which holds

- $Y = \bigcup_{i \in I} Y_i$ where I is a finite index set;
- $Y_i \cap Y_j = \emptyset$ for $i \neq j$;
- $Y_i \neq \emptyset$ for all $i \in I$.

In this definition, Y_i and Y refer to sets of points in a point-based universe. As mentioned, for reasons of compatibility with the model-theoretic semantics of DL, we use a non-standard interpretation where regions are interpreted as *individuals*, and not as sets, in an abstract domain. We thus reformulate the definition using the Boolean RCC function SUM and the RCC relation DR (i.e., "discrete from"). As is customary, we use lower case letters for variables denoting individuals.

Definition 1b (Partition in RCC). A family of regions $(x_i)_{i \in I}$ is a partition of a region y if the following holds:

- $y = \mathsf{SUM}_{i \in I}\, x_i$ where I is a finite index set;[5] [6]
- $\forall x_i \forall x_j\, \mathsf{DR}(x_i, x_j)$ for $i \neq j$;
- regions $(x_i)_{i \in I}$ are named for all $i \in I$.

We only consider partitions where the elements are typed by kinds of administrative regions, for instance, $\mathsf{Commune}(x_i)$ says that x_i is of type Commune. For regions we do not allow multiple typing, that is, the concepts used for typing are mutually disjoint. Similarly, a given type is used for a single partition only. This allows distinguishing the partitions by their types.

In order to account for the different scales of social organization we define a *partial order* on the system of partitions in RCC by comparing partitions w.r.t. their granularity.

Definition 2 (Partial Order on Typed Partitions in RCC). Let $C(x_i)_{i \in I}$ and $D(y_j)_{j \in J}$ be partitions of the same region of types C and D, respectively. We say that $C(x_i)_{i \in I}$ is *more fine-grained* than $D(y_j)_{j \in J}$, denoted by $C(x_i)_{i \in I} \preceq D(y_j)_{j \in J}$, if each element of $C(x_i)_{i \in I}$ is a (possibly improper) subset of an element of $D(y_j)_{j \in J}$. A partial order on typed partitions is reflexive, transitive and antisymmetric.

This means that each element of $D(y_j)_{j \in J}$ is partitioned by elements of $C(x_i)_{i \in I}$. For instance, $\mathsf{Commune}(x_i)_{i \in I}$ and $\mathsf{District}(y_j)_{j \in J}$ are typed partitions of a canton and each element of $\mathsf{District}(y_j)_{j \in J}$ is partitioned by elements of $\mathsf{Commune}(x_i)_{i \in I}$.

---

[5] $\mathsf{SUM}_{i \in I}\, x_i$ is defined as $\forall z\, [\mathsf{C}(z, y) \leftrightarrow \bigvee_{i \in I} \mathsf{C}(z, x_i)]$ [17]. C stands for "connects with".

[6] This implies $\forall x_i\, \mathsf{P}(x_i, y)$. P stands for "part of".

Definition 3 (Minimal Partial Order on Typed Partitions in RCC). We say that a partial order on typed partitions is *minimal* w.r.t. a given conceptualization, denoted by $C(x_i)_{i \in I} \preceq_{\min} D(y_j)_{j \in J}$, if the conceptualization does not provide a type for any $(w_k)_{k \in K}$ such that $C(x_i)_{i \in I} \preceq (w_k)_{k \in K} \preceq D(y_j)_{j \in J}$. A minimal partial order on typed partitions is *intransitive*.

For instance, if a given conceptualization provides the administrative types District and Commune, any partial order comprising a non-typed partition of intermediate granularity is not minimal.

For asserting closeness between individual regions we slightly extend RCC and introduce the relation $\mathsf{CL}(x, y)$ which is read as "x is close to y". In accordance with empirical evidence [12], closeness is introduced as a *weakly asymmetrical* relation. This means that the relation is symmetrical, if x and y are members of the same partition, but asymmetrical, if y is a member of a more fine-grained partition than x or else, if x is a non-administrative region.[7]

Definition 4 (Closeness in RCC). Given a region x_i of a partition used as a referent in a query, a type C of a conceptualization for x_i and a minimal partial order on typed partitions $C(x_i)_{i \in I} \preceq_{\min} D(y_j)_{j \in J}$, closeness in RCC can be inferred by the composition rule $\forall x_i \forall y_j \forall z\ [\mathsf{P}(x_i, y_j) \wedge \mathsf{XC}(z, y_j) \rightarrow \mathsf{CL}(z, x_i)]$.

In this definition, $\mathsf{XC}(z, y_j)$, read as "exclusively connects with", is an auxiliary relation. Its main purpose is to prevent the transitive property of $\mathsf{P}(x_i, y_j)$, which has been overridden by the definition of a minimal partial order on typed partitions, from being reintroduced through the backdoor of the composition rule. Note that since the relation is directed from z to y_j, transitivity is excluded by removing $\mathsf{Pi}(z, y_j)$ (i.e., "inverse part of") and its subrelations from $\mathsf{C}(z, y_j)$.

Definition 4 shows that closeness depends on the type of region used as a referent in a query, hence on the way a user conceptualizes a domain. This includes the scale on which spatial relations are to be evaluated. It also depends on partitions into administrative regions reflecting how a collective perceives spatial closeness on increasing scales of social organization. As a result of this dependency, $\mathsf{CL}(z, x_i)$ is undefined unless it is related to a minimal partial order on typed partitions. A comprehensive example is provided in Section 5.

4.2 A DL Knowledge Base and Rule Base for RCC

The knowledge required for answering (possibly vague) spatio-thematic queries can be represented by a DL knowledge base $\mathcal{KB}$ consisting of a TBox $\mathcal{T}$ and ABox $\mathcal{A}$, $\mathcal{KB} = \{\mathcal{T}, \mathcal{A}\}$, and a rule base $\mathcal{RB}$ for DL-safe rules.

Among other things $\mathcal{T}$ contains a number of concept inclusion axioms that introduce kinds of regions. It is worth recalling the definition of an ontology as an explicit, formal specification of a shared conceptualization [2]. Accordingly, the introduced categories are not arbitrary. They are social artifacts and reflect how a collective

[7] In [12] it is argued that for nearness the subject-referent dichotomy plays a dominant role in that the referent creates the scale in which the relation has context.

thinks that the world (or a piece thereof) is structured. In the long run, a collectively shared conceptualization is furthermore not invariant but evolves together with the development of a society and a country.

In order to implement RCC in DL, the subsumption hierarchy of RCC relations [17] is represented as a hierarchy of binary role inclusion axioms in the TBox. The RCC relation $\mathsf{P}(x_i, y_j)$ and its subrelations are implemented as functional roles, thereby ensuring that an individual x_i is only part of a single region y_j. This overrides the transitivity of the RCC relation $\mathsf{P}(x, y)$, which prevents, for instance, communes to be related to cantons (or to countries or continents if these were represented). Partitions are represented in $\mathcal{T}$ by (anonymous) concepts that are made up of individual names, also called *nominals*, $\{x_1, \ldots, x_n\}$. Nominals are linked to types by concept inclusion axioms of the form $C \sqsubseteq \{x_1, \ldots, x_n\}$ stating that the set of individuals in the interpretation of C is a (possibly improper) subset of the individuals in the interpretation of $\{x_1, \ldots, x_n\}$. In order to disallow multiple typing the concepts used for typing are defined as mutually disjoint, $C \sqsubseteq \neg D$.

In order to populate the ABox, known RCC relations between individual regions are asserted as role assertions. Particularly, partitions are asserted as of $\mathsf{partOf}(x_i, y_j)$, or any of its subrelations, for all applicable $x_i \in \{x_1, \ldots, x_n\}$ and $y_j \in \{y_1, \ldots, y_m\}$. In so doing, $\mathcal{A}$ is closed w.r.t. nominals denoting administrative regions.[8] A minimal partial order on typed partitions is implemented by asserting $\mathsf{partOf}(x_i, y_j)$, or any of its subrelations, exclusively for those pairs of individuals (x_i, y_j) for which hold $C(x_i)_{i \in I} \preceq_{\min} D(y_j)_{j \in J}$. $\mathcal{A}$ also contains facts about individual regions in terms of concept assertions.

The composition rule $\forall x \forall y \forall z\ [\mathsf{P}(x, y) \wedge \mathsf{XC}(z, y) \rightarrow \mathsf{CL}(z, x)]$ is implemented in $\mathcal{RB}$ by the DL-safe rule $\mathsf{partOf}(x, y) \wedge \mathsf{exclusivelyConnectsWith}(z, y) \wedge \mathcal{O}(x) \wedge \mathcal{O}(y) \wedge \mathcal{O}(z) \rightarrow \mathsf{closeTo}(z, x)$ where $\mathcal{O}(x)$, $\mathcal{O}(y)$ and $\mathcal{O}(z)$ are non-DL-literals. In order to make the rule DL-safe, a fact $\mathcal{O}(a)$ is asserted for each individual a in the ABox. The rule is read as "A region z is close to a region x if x is part of a region y and z exclusively connects with y *where the identity of all regions is known*."

4.3 Processing Vague Spatio-Thematic Queries

The concepts implicitly and explicitly used in a query reveal how a user conceptualizes a domain. Thereby the user is assumed to be a member of the social collective in question. Query concepts can be used to determine the scale on which closeness is to be evaluated. They translate what is often referred to as the *context* of a vague concept, such as closeness, from contingencies in the real world into linguistic constraints. We assume a query of the form $\forall z\ [Q(z) \wedge \mathsf{CL}(z, a)]$ which is expected to return the set of those individuals of type Q that are close to a given individual a of a partition. In this query, the type of individual a, for instance $C(a)$, sets the scale for the evaluation of closeness.

[8] Note that we use $\mathsf{partOf}(x_i, y_j)$ and its subrelations only for asserting partitions into administrative regions.

Algorithm 1. Function **CLOSETO** computes $(Q \sqcap \exists \mathsf{closeTo}.\{a\})(z)$ from $\mathcal{KB}$ and $\mathcal{RB}$.

FUNCTION CLOSETO

INPUT: Knowledge Base $\mathcal{KB} = \{\mathcal{T}, \mathcal{A}\}$, Rule Base $\mathcal{RB}$, Concept Q, Individual a

OUTPUT: Set<Individual>

1. $\{b\} \leftarrow \{b \mid \mathcal{A} \vDash \mathsf{partOf}(a, b)\}$
2. $V \leftarrow \{v_{i \in I} \mid \mathcal{A} \vDash \mathsf{exclusivelyConnectsWith}(v_i, b)\}$
3. $W \leftarrow \{w_{j \in J} \mid \mathcal{A} \vDash Q(w_j)\}$
4. $Z \leftarrow V \cap W$
5. **OUTPUT** Z

The query $\forall z\, [Q(z) \wedge \mathsf{CL}(z, a)]$ is implemented in DL by the concept description $Q \sqcap \exists \mathsf{closeTo}.\{a\}$. Given an ABox $\mathcal{A}$ and a concept description $Q \sqcap \exists \mathsf{closeTo}.\{a\}$, the retrieval problem is thus to find all individuals z in $\mathcal{A}$ such that $\mathcal{A} \vDash (Q \sqcap \exists \mathsf{closeTo}.\{a\})(z)$. Algorithm 1 shows the steps (1–5) to take when processing a query. Note that in order to process steps 1 and 2 the composition rule is required.

5 Applying the Approach to an Example Query

As stated in Section 4, we assume that administrative regions mirror how a collective perceives spatial closeness on increasing scales of social organization. In order to verify our assumption, it is necessary to recall the design principles underlying the organization into administrative units. For example, the following official statement clarifies the purpose of districts as administrative units:

"The administrative districts ... perform decentralized administrative tasks of the cantons, particularly in the areas of health (district hospitals, public health), to some extent education (district schools), judiciary (district courts) and general administration (taxation, business failures, etc.). In several cantons the administrative districts, furthermore, correspond to the electoral wards." [25]

The statement implies that the organization of districts is largely, if not entirely, motivated by the spatial closeness of the communes. District hospitals, for instance, are decentralized entities of the health care delivery system. They might have been established with the intention to keep the distance between the patients (living in the communes) and the care providers small. This obviously reflects the experience of human subjects who perceive this distance as small. Similar arguments apply to district schools, district courts and electoral wards. It is, therefore, reasonable to claim that the organization of districts is motivated by the property of spatial closeness. The organization of other administrative regions can be motivated in a similar way.

Search engines on the Web are weak in supporting spatial queries. This can be demonstrated with the following example. Suppose we are looking for an answer to the question "Which landscapes and natural monuments of national importance are close to Aesch ZH?" where Aesch ZH is a commune in the canton of Zurich, Switzerland. Searching with the strings `<Landschaften "in der Nähe von"`

`Aesch ZH>` (i.e., landscapes close to Aesch ZH) returns 1,350 matches.[9] None of the top 30 deal with a landscape.

This poor answer could be a result of there being no landscapes close to Aesch ZH or there being no resources related to landscapes close to Aesch ZH on the Web. In an attempt to exclude the second, we search for `<Albiskette-Reppischtal>` which is the name of a landscape of national importance potentially close to Aesch ZH. Searching for Albiskette-Reppischtal returns 230 matches. None of these matches appears among the top 30 of the initial search. Worse, none of the scanned results from the initial search indicates that the search engine really understood the query.

The knowledge required for answering the example query is represented in a consistent DL knowledge base $\mathcal{KB} = \{\mathcal{T}, \mathcal{A}\}$ and DL-safe rule base $\mathcal{RB}$ as published on http://www.wsl.ch/personal_homepages/scharren/download/terra_cognita_2009.owl.[10] The description language used for the $\mathcal{KB}$ is OWL DL [20]; the DL expressivity is $\mathcal{ALCHOIF}$. It is clear that the representation of relevant knowledge is not sufficient. In order to answer the example query a search engine must be able to make use of it. The search engine used above does not have this ability. We used Pellet 2.0 in order to test the presented approach. Pellet 2.0 is a DL reasoner that integrates a SPARQL query-engine and a rule engine for the processing of DL-safe rules.[11] It is able to handle all aspects of the introduced framework. Note that a detailed discussion of successful search algorithms is outside the scope of this paper.

The example query can be formalized in DL as $(\mathsf{Landscape} \sqcap \exists \mathsf{closeTo}.\{\mathsf{Aesch}\})(z)$ where Aesch is an element of a partition. The query is processed as indicated in Section 4.3 and the result set $\{\mathsf{Albiskette\text{-}Reppischtal}\}$ contains a landscape (of national importance) which is close to Aesch. Accordingly, a logically enabled search engine operating on the $\mathcal{KB}$ and $\mathcal{RB}$ is expected to return the 230 matches for Albiskette-Reppischtal when queried with the strings `<Landschaften "in der Nähe von" Aesch ZH>`, which it does not at present.[12]

6 Discussion

Valid evaluations of the example query include cases where Aesch is a (possibly inverse) part of, partially overlaps or is externally connected to a landscape. The current release of the reasoner used for the example does not allow excluding these cases in a single query (cf. Section 3). They could still be excluded by querying all landscapes that are not in any of the above relations to Aesch and by intersecting the result set with the result set from the example query.

Closeness is a *vague* concept, in the sense that there exist borderline cases for which it is difficult to decide whether they are covered by the concept or not [12].

---

[9] http://www.google.ch

[10] Please contact the authors if you wish to access the Web page and it is no longer maintained.

[11] http://clarkparsia.com/pellet

[12] Further queries that are successfully processed include communes, districts or landscapes adjacent to a given commune or district; communes close to a given commune; districts close to a given commune or district; landscapes close to a given district.

While our approach takes a Boolean decision on closeness it still accounts for borderline cases by using a qualitative formalism. Whether a region is close to another region or not depends on the size and shape of the administrative units serving as a frame of reference. When comparing the evaluation of closeness, even on the same scale of social organization, a given metric distance may in one case be interpreted as close and in another case as not close. Our claim is that the size and shape of administrative regions are not arbitrary but reflect how a collective perceives spatial closeness on increasing scales of social organization. Whether our claim is empirically well founded or not remains to be shown.

The concept of closeness evolves over time. What is perceived as close by the members of a social collective (at least in the industrialized countries) has been subject to change for decades. Similarly, at the institutional level, the concept of closeness evolves. In recent years, several cantons in Switzerland, for instance, have revised their administrative structures or established a legal basis for future revisions. A result of these revisions is a reduction in the number of districts. It is important to note that the societal change precedes the institutional. If this was not the case, proposals for structural revisions would not obtain a majority of popular votes.[13] Since our approach evaluates closeness within the frame of administrative structures and the institutional change lags behind the societal change, it tends to underestimate closeness.[14] According to our approach, if one region is identified as being close to another, it has been perceived as such possibly for a long time. While it takes into account evolution of closeness, it does so at the slow pace of the institutions and not at the fast pace of the social collective.

7 Conclusion and Outlook

In this paper an implementation of RCC in OWL DL augmented by DL-safe rules is used in order to represent spatio-thematic knowledge. It is demonstrated how such a representation can be used for answering queries involving (possibly vague) spatial concepts. Thereby, the division of land into administrative regions rather than, for instance, a metric system is taken as a frame of reference for evaluating closeness. Accordingly, closeness is evaluated based on purely qualitative criteria. Since colloquial descriptions typically involve qualitative concepts, the presented approach is expected to align better with the way human beings deal with closeness than a quantitative approach.

So far the presented approach supports the evaluation of closeness of regions *w.r.t. an administrative region*. Evaluation of closeness between arbitrary regions would be desirable. Exploring whether and how the frame of reference can be leveraged to support evaluation of closeness between arbitrary regions is left to future work. Likewise, the scalability of the implementation and, possibly, alternative implementation strategies remain to be explored.

---

[13] This might be a peculiarity of the Swiss political system.

[14] In a practical application, if the data base is not updated at the time when the revisions take effect the lag is even longer.

The paper only considers the concept of closeness. There are a number of additional vague spatial concepts such as “near”, “next to”, “a little distance outside”, “a long way off”, and “far away from”. It would be interesting to formalize these concepts in a similar way as demonstrated for “close to”. Such a formalization might result in a theory of vague spatial concepts in RCC which could be implemented, for instance, in OWL DL augmented by DL-safe rules.

Acknowledgments. The authors sincerely thank Jürg Schenker (FOEN), Martin Hägeli (WSL) and Bettina Waldvogel (WSL) for the fruitful discussions and leadership that made this research possible. This research has been funded and conducted in cooperation with the Swiss Federal Office for the Environment (FOEN).

References

1. Berners-Lee, T.: What a semantic web is not. World Wide Web Consortium, Design Issues, http://www.w3.org/DesignIssues/RDFnot.html (1998)
2. Gruber T.R.: A translation approach to portable ontology specifications. Knowledge Acquisition 5(2), 199–220 (1993)
3. Wache, H., Vögele, T., Visser, U., Stuckenschmidt, H., Schuster, G., Neumann, H., Hübner, S.: Ontology-based Integration of Information—A Survey of Existing Approaches. In: Gómez-Pérez, A., Gruninger, M., Stuckenschmidt, H., Uschold M. (eds.) IJCAI-01 Workshop: Ontologies and Information Sharing. CEUR-WS, vol. 47, pp. 108–118. CEUR-WS, RWTH Aachen (2001)
4. Fonseca, F., Egenhofer, M., Agouris, P., Câmara, C.: Using Ontologies for Integrated Information Systems. Transactions in GIS 3(6), 231-257 (2002)
5. Jones, C.B., Purves, R., Ruas, A., Sanderson, M., Sester, M., van Kreveld, M., Weibel, R.: Spatial Information Retrieval and Geographical Ontologies: An Overview of the SPIRIT Project. In: Proceedings of the 25th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 387-388. ACM Press (2002)
6. Bishr, Y.: Geospatial Semantic Web. In: Rana, S., Sharma, J. (eds.) Frontiers of Geographic Information Technology, pp. 139–154. Springer, Berlin Heidelberg (2006)
7. Haarslev V., Lutz C., Möller R.: Foundations of Spatioterminological Reasoning with Description Logics. In: Cohn A.G., Schubert L.K., Shapiro S.C. (eds.) Principles of Knowledge Representation and Reasoning, pp. 112–123. Morgan-Kaufmann, San Mateo, CA (1998)
8. Lutz, C., Miličić, M.: A Tableau Algorithm for Description Logics with Concrete Domains and General TBoxes. Journal of Automated Reasoning 38(1-3), 227-259 (2007)
9. Grütter, R., Bauer-Messmer, B.: Towards Spatial Reasoning in the Semantic Web: A Hybrid Knowledge Representation System Architecture. In: Fabrikant, S., Wachowicz M. (eds.) The European Information Society: Leading the Way with Geo-information. LNGC, pp. 349–364. Springer, Berlin Heidelberg (2007)
10. Möller, R., Wessel, M.: Flexible Software Architectures for Ontology-Based Information Systems. Journal of Applied Logic, Special Issue on Emperically Sucessful Systems (2007)
11. Yao, X., Thill, J.-C.: How Far Is Too Far? – A Statistical Approach to Context-contingent Proximity Modeling. Transactions in GIS 9(2), pp. 157–178 (2005)
12. Worboys, M.F.: Nearness Relations in Environmental Space. International Journal of Geographical Information Science 15(7), pp. 633–651 (2001)
13. Brennan, J., Martin, E.: Foundations for a Formalism of Nearness. In: AI 2002: Advances in Artificial Intelligence. LNCS, vol. 2557, pp. 71–82. Springer, Heidelberg (2002)

14. Dolbear, C., Hart, G., Goodwin, J.: From Theory to Query: Using Ontologies to Make Explicit Imprecise Spatial Relationships for Database Querying. In: COSIT 2007: International Conference on Spatial Information Theory. LNCS, vol. 4736. Springer, Heidelberg (2007)
15. Hart, G., Dolbear, C.: What's so Special about Spatial? In: Scharl, A., Tochtermann, K. (eds.) The Geospatial Web: How Geobrowsers, Social Software and the Web 2.0 are Shaping the Network Society, pp. 39–44. Springer-Verlag, London Limited (2006)
16. Mata, F.: Geographic Information Retrieval by Topological, Geographical, and Conceptual Matching. In: Fonseca, F., Rodríguez, M.A., Levashkin, S. (eds.) GeoS 2007, LNCS 4853, pp. 98–113. Springer-Verlag, Berlin, Heidelberg (2007)
17. Randell, D.A., Cui, Z., Cohn, A.G.: A Spatial Logic based on Regions and Connections. In: Nebel, B., Rich, C., Swartout W. (eds.) Principles of Knowledge Representation and Reasoning, pp. 165–176. Morgan Kaufmann, San Mateo, CA (1992)
18. Bennett, B.: Logics for Topological Reasoning. In: 12th European Summer School in Logic, Language and Information (ESSLLI-MM), August 6–18, 2000, Birmingham, UK (2000)
19. Baader, F., Nutt, W.: Basic Description Logics. In: Baader, F., Calvanese, D., McGuinness, D.L., Nardi, D., Patel-Schneider, P.F. (eds.) The Description Logic Handbook, pp. 47–100. Cambridge University Press (2003)
20. Patel-Schneider, P.F., Hayes, P., Horrocks, I.: OWL Web Ontology Language: Semantics and Abstract Syntax. W3C Recommendation 10 February 2004. World Wide Web Consortium (2004)
21. Egenhofer, M., Franzosa, R.: Point-Set Topological Spatial Relations. International Journal of Geographical Information Systems 5(2), 161–174 (1991)
22. Motik, B., Sattler, U., Studer, R.: Query Answering for OWL-DL with Rules. Journal of Web Semantics, 549–563 (2004)
23. Motik, B., Horrocks, I., Rosati, R., Sattler, U.: Can OWL and Logic Programming Live Together Happily Ever After? In: Cruz, I.F., Decker, S., Allemang, D., Preist, C., Schwabe, D., Mika, P., Uschold, M., Aroyo, L. (eds.) Proceedings of the 5th International Semantic Web Conference (ISWC 2006). LNCS, vol. 4273, pp. 501–514. Springer, Heidelberg (2006)
24. Motik, B., Rosati, R.: A Faithful Integration of Description Logics with Logic Programming. In: Veloso, M.M. (ed.) Proceedings of the 20th International Joint Conference on Artificial Intelligence (IJCAI 2007), pp. 477–482. Morgan Kaufmann Publishers (2007)
25. Bundesamt für Statistik: Institutionelle Gliederungen der Schweiz: Schweiz – die Bezirke. Bundesamt für Statistik, Neuchâtel, http://www.bfs.admin.ch/bfs/portal/de/index/regionen/11/geo/institutionelle_gliederungen/01a.html (2008)