Spatial Knowledge Representation on the Semantic Web

Frederik Hogenboom, Bram Borgman, Flavius Frasincar, and Uzay Kaymak
Erasmus University Rotterdam
PO Box 1738, NL-3000 DR
Rotterdam, the Netherlands
Email: {fhogenboom, borgman, frasincar, kaymak}@ese.eur.nl

***Abstract*—The Region Connection Calculus based on 8 relations (RCC8) is one of several extensively researched methods to use for qualitative spatial representation and reasoning. We discuss several issues arising when representing RCC8 in OWL DL, a decidable fragment of OWL. There is no direct encoding of such a calculus in OWL DL, as the language lacks required features such as role reflexivity, role Boolean operators, and role inclusion axioms. Some of these features are to be included in the new version of the OWL standard, OWL 2, but this language still lacks the expressive power to support role negations, conjunctions, disjunctions, and complex role inclusion axioms. Recently, advances in description logics languages as SROIQBs have made possible expressing some of the above constructs, while maintaining the decidability of the language. In this paper, we exploit these new opportunities by providing qualitative spatial knowledge representation on the Semantic Web.**

***Keywords*-RCC8; OWL; spatial knowledge representation; SROIQBs**

I. Introduction

In this era of fast growing information needs, it is important to be able to represent data in a proper, meaningful way. The Semantic Web opens the possibility to find, share, and combine information more easily. It is already possible to represent and reason with many types of information, but there is no solution to do so with spatial knowledge, which is knowledge acquired by someone or something in a spatial environment. Although we can represent some spatial knowledge on the Web by storing spatial features of subjects in ontologies (e.g., geo-ontologies), these ontologies cannot capture the semantics of spatial relations in a computer "reasonable" manner. This spatial qualitative knowledge can be generated from geographical information systems by using translation schemes, implemented using for example Prolog [1].

The Region Connection Calculus (RCC) [2] is one of several extensively researched methods to use for qualitative spatial or topological representation and reasoning, and it might be suitable to use on the Semantic Web too. Unfortunately, this is not yet possible. In this paper, we discuss several issues arising when trying to enable topological spatial reasoning on the Semantic Web using RCC, in particular in conjunction with Web Ontology Language (OWL). We focus on RCC instead of alternative calculi like 9-Intersection [3], because of its popularity and simplicity. Also, RCC is based on logical theory, which makes it a more suitable candidate to be used for the Semantic Web, which is based on Description Logics (DL), than 9-Intersection, which is based on elementary geometry.

There are different versions of OWL we can consider in this paper. A commonly used OWL version is OWL DL, which is a decidable fragment of OWL and is based on the description logics language $\mathcal{SHOIN}$. Also, OWL 2 (previously known as as OWL 1.1) is often mentioned in literature. OWL 2 is an extension of OWL DL and is based on the $\mathcal{SROIQ}$ description logics language and its specifications have already reached a recommendation status [4].

Being able to model spatial information on the Web using RCC relations is important in order to support spatial reasoning, as a lot of spatial information is already available on the Web in an unstructured form. Spatial relations can help one to make decisions related to a wide variety of topics, for instance the identification of geographical neighbors. If one is starting a new business in a country, it is preferred that the latter country is not in an armed conflict with any of its neighboring countries. Another example can be found in cadaster applications, e.g., laying a road in such a way that it does not overlap with any of the existing buildings.A space-based representation of weather information and the spreading of epidemics are other possible applications.

Much work has already been done on the subject of representing RCC relations in OWL [5]–[7]. However, none of the suggested representations in OWL investigated so far provide for a smooth and full integration of RCC into OWL. In [8] we touched upon several of these issues. In this paper, we discuss these issues in more detail and provide solutions to some of the problems encountered, without breaking the design philosophy of the OWL language.

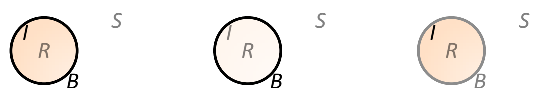
The paper is organized as follows. First, the main relations of the Region Connection Calculus are introduced in Sect. II. Then, we elaborate on state-of-the-art approaches to express RCC relations in specific versions of OWL and their underlying description logics languages (i.e., $\mathcal{SHOIN}$ and $\mathcal{SROIQ}$) in Sect. III. Section IV proposes a way to represent RCC relations by using $\mathcal{SROIQB}_s$ [9], which is a description logics language that has some additional features over $\mathcal{SROIQ}$ that are currently not supported by OWL. Finally, we wrap up in Sect. V by drawing conclusions and by giving directions for future research.

II. RCC

This section discusses the main relations of the Region Connection Calculus (RCC). Also, we distinguish between different versions of RCC, amongst which is RCC8. During the rest of the paper, we will focus on the relations of RCC8. This section concludes with a composition table of the RCC8 relations.

A. Introduction

The Region Connection Calculus serves for qualitative spatial or topological representation and reasoning and was introduced in 1992 by Randell et al [2]. RCC abstractly describes regions (in Euclidean or topological space) by their possible relations to each other.



(a) Closure (b) Boundary (c) Interior

Figure 1. Region basics

Figure 1 illustrates the basics of regions. Regions are defined as regular non-empty subsets of a topological set. The closure of region (or subset) R in the topological space S can be defined as the smallest closed subset of S which contains R. A region also has a boundary B, which is the closure of the region minus its interior I. A region always contains more than one element (interior point), and is possibly composed of multiple disconnected pieces that may also contain holes. Figure 1(a) shows that the closure of the region R in the topological space S is equal to boundary B plus interior I. The RCC spatial relations operate on regions, i.e., subsets that contain the closure of their interior.

B. RCC Relations

We now continue to elaborate on the basic relations of the Region Connection Calculus. Table I shows all RCC relations. Note that inverse relations are omitted in this section, because these relations are trivially defined based on the corresponding direct relations. The relation between regions R_1 and R_2 is equal to the inverse relation between regions R_2 and R_1, e.g., TPP(R_1, R_2) is equal to TPPi(R_2, R_1).

Table I
RCC RELATIONS

Description	Relation
Connected	C
Part of	P
Overlaps	O
Externally connected to	EC
Disconnected from	DC
Discrete from	DR
Proper part	PP
Tangential proper part of	TPP
Non-tangential proper part of	NTPP
Equal to	EQ
Partially overlaps	PO
Overlaps not equal	ONE
Spatially related	SR

1) Connected: RCC is based on a single primitive relation between spatial regions: the connected relation. We will refer to this connected relation as C. Two regions are connected if and only if their topological closures share a common point [10]. The relation C has to be reflexive and symmetric. A reflexive relation implies that the relation should not only be applicable to two different spatial regions, but that it should also be possible for a region to be connected to itself. When region R_1 is connected with region R_2 and region R_2 is also connected with region R_1, the relation between both regions is symmetric.

2) Part of: The part of relation indicates that every region R_3 which connects to R_1, also connects to R_2, thus making R_1 part of R_2. The relation P translated into First Order Logic (FOL) is:

$$\mathrm{P}(R_1, R_2) \equiv_{def} \forall R_3[\mathrm{C}(R_3, R_1) \rightarrow \mathrm{C}(R_3, R_2)]$$

Figures 4, 5, and 6 illustrate the part of relation. Region R_1 is inside region R_2 and thus is part of region R_2. Please note that the relation P can also be a PP(i), TPP(i), NTPP(i), or EQ relation, where "i" stands for inverse of the relation.

3) Overlaps: If two regions have more than one point in common (i.e., another region), the regions are overlapping (O). Regions are considered to be overlapping when the relation between the regions can be defined as a TPP(i) (Fig. 4), NTPP(i) (Fig. 5), EQ (Fig. 6), or PO (Fig. 7) relation. In FOL:

$$\mathrm{O}(R_1, R_2) \equiv_{def} \exists R_3[\mathrm{P}(R_3, R_1) \wedge \mathrm{P}(R_3, R_2)]$$

4) Externally connected to: Two regions are said to be externally connected (EC) if they only share borders and thus both regions are connected but are not overlapping. Figure 2 displays a typical EC relation, as the borders of regions R_1 and R_2 touch each other. The relation EC can be described in FOL as:

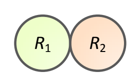


Figure 2. EC

$$\mathrm{EC}(R_1, R_2) \equiv_{def} \mathrm{C}(R_1, R_2) \wedge \neg\mathrm{O}(R_1, R_2)$$

5) Disconnected from: If regions R_1 and R_2 are disconnected (DC), no points exist that are in regions R_1 and R_2 at the same time and thus the intersection of R_1 and R_2 is empty. Figure 3 represents this relation in a graphical way. Both regions are, according to the definition of the relation DC, not connected to each other. The relation DC is in FOL:

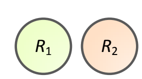


Figure 3. DC

$$\mathrm{DC}(R_1, R_2) \equiv_{def} \neg\mathrm{C}(R_1, R_2)$$

6) Discrete from: Region R_1 can be discrete from (DR) R_2 if both regions are not overlapping. Both in the relation EC and DC (Figs. 2 and 3, respectively) regions are not overlapping, since they are only touching with bordering points or not touching at all. The relation DR is in FOL:

$$\mathrm{DR}(R_1, R_2) \equiv_{def} \neg\mathrm{O}(R_1, R_2)$$

7) Proper part: Region R_1 is said to be a proper part of region R_2 if R_1 is fully inside R_2, but is not equal to R_2. The regions shown in Figs. 4 and 5 illustrate the proper part of (PP) relation: R_1 is inside of R_2 and is smaller than R_2. In FOL, the PP relation can be defined as:

$$\mathrm{PP}(R_1, R_2) \equiv_{def} \mathrm{P}(R_1, R_2) \wedge \neg\mathrm{P}(R_2, R_1)$$

8) Tangential proper part: The tangential proper part relation indicates that one region is a subset of another region and that they share some points on the borders, so the tangential proper part relation is basically a proper part relation where bordering points are shared. For the regions from our previous examples, this relation can be denoted as:

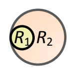


Figure 4. TPP

$$\mathrm{TPP}(R_1, R_2) \equiv_{def} \mathrm{PP}(R_1, R_2) \wedge \exists R_3[\mathrm{EC}(R_3, R_1) \wedge \mathrm{EC}(R_3, R_2)]$$

The relation is illustrated in Fig. 4, region R_1 is inside region R_2 and R_1 touches the border of R_2.

9) Non-tangential proper part: The NTPP (non-tangential proper part) relation is almost similar to the TPP relation, differing in the fact that no bordering points are shared. Figure 5 shows an example of an NTPP relation, where region R_1 is inside region R_2, without sharing any bordering points. In FOL notation NTPP can be defined as:

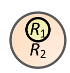


Figure 5. NTPP

$$\mathrm{NTPP}(R_1, R_2) \equiv_{def} \mathrm{PP}(R_1, R_2) \wedge \neg\exists R_3[\mathrm{EC}(R_3, R_1) \wedge \mathrm{EC}(R_3, R_2)]$$

10) Equal to: Two regions, say R_1 and R_2, are equal if and only if they are exactly the same or, in other words, identical. Region R_1 is part of R_2 and vice versa (Fig. 6). In FOL, we can define this relation as:

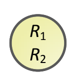


Figure 6. EQ

$$\mathrm{EQ}(R_1, R_2) \equiv_{def} \mathrm{P}(R_1, R_2) \wedge \mathrm{P}(R_2, R_1)$$

11) Partially overlaps: It is possible for regions to share interior points. When more points than just border points, but not all points of the regions are shared, one can say the regions are partially overlapping. Figure 7 illustrates two partially overlapping regions R_1 and R_2. We can denote this relation as:

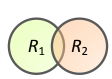


Figure 7. PO

$$\mathrm{PO}(R_1, R_2) \equiv_{def} \mathrm{O}(R_1, R_2) \wedge \neg\mathrm{P}(R_1, R_2) \wedge \neg\mathrm{P}(R_2, R_1)$$

12) Overlaps not equal: Furthermore, it is possible for two regions to overlap, but not to be equal, i.e., the relation ONE holds. Region R_1 can have the relation ONE with region R_2 when either R_1 is partially overlapping R_2 or R_1 is a proper part of R_2 (or R_2 is a proper part of R_1). These relations are illustrated in Figs. 4, 5, and 7. ONE can be denoted as:

$$\mathrm{ONE}(R_1, R_2) \equiv_{def} \mathrm{O}(R_1, R_2) \wedge \neg\mathrm{EQ}(R_1, R_2)$$

13) Spatially related: In RCC, two regions always have the SR (spatially related) relation, since SR contains every relation of RCC: it is the universal relation. Each of the previously introduced illustrations of relations applies to this universal relation. In FOL, SR can be denoted as:

$$\mathrm{SR}(R_1, R_2) \equiv_{def} \top(R_1, R_2)$$

C. RCC Versions

A few versions of the region connection calculus are circulating, for instance RCC1, RCC2, RCC3, RCC5, RCC8, RCC15, and RCC23. The numbers indicate the amount of relations defined in the connection calculi. In this paper, we focus on RCC8, which is widely used and researched. RCC8 has several nice properties, e.g., reasoning over its relations is decidable [11], [12]. Also, RCC8 is one of the smallest sets of topological base relations which makes topological distinctions rather than just mereological ones and RCC8 can be regarded as the spatial counterpart of the temporal (one-dimensional) interval algebra [12]. In addition, reasoning over the general RCC calculus is known to be undecidable [13].

Using the relation C, eight Jointly Exhaustive and Pairwise Disjoint (JEPD) relations (i.e., there is no relation between the domain objects that can not be described by a predicate of the JEPD set, and only one relation can hold at a time) can be defined for RCC8: EC, DC, TPP, TPPi, NTPP, NTPPi, EQ, and PO. The RCC5 relation set contains five relations: PP, PPi, PO, EQ and DR. The RCC3 relations are: ONE, EQ and DR. RCC2 contains the O and DR relations and the RCC1 relation set only contains one relation, i.e., SR. RCC15 and RCC23 can be partially described using the previously introduced relations [14]. All these relations (except for the inverses) have been discussed in Subsect. II-B.

D. Composition Table of RCC8

In order to check the consistency of a knowledge base holding spatial relations, so-called composition tables are used [7], [15]. The compositions of all RCC8 relations can be found in Table II.

We explain the principle of composition and the correctness of the composition table by using an example: the composition of EC (second row) and NTPP (sixth column). For convenience, the corresponding cell is highlighted in

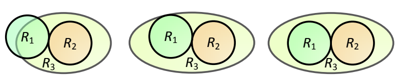
Table II
Composition table for RCC8 relations

∘	EC	DC	TPP	TPPi	NTPP	NTPPi	EQ	PO
EC	EC, DC, TPP, TPPi, EQ, PO	EC, DC, TPPi, NTPPi, PO	EC, TPP, NTPP, PO	EC, DC	TPP, NTPP, PO	DC	EC	EC, DC, TPP, NTPP, PO
DC	EC, DC, TPP, NTPP, PO	EC, DC, TPP, TPPi, NTPP, NTPPi, EQ, PO	EC, DC, TPP, NTPP, PO	DC	EC, DC, TPP, NTPP, PO	DC	DC	EC, DC, TPP, NTPP, PO
TPP	EC, DC	DC	TPP, NTPP	EC, DC, TPP, TPPi, EQ, PO	NTPP	EC, DC, TPPi, NTPPi, PO	TPP	EC, DC, TPP, NTPP, PO
TPPi	EC, TPPi, NTPPi, PO	EC, DC, TPPi, NTPPi, PO	TPP, TPPi, EQ, PO	TPPi, NTPPi	TPP, NTPP, PO	NTPPi	TPPi	TPPi, NTPPi, PO
NTPP	DC	DC	NTPP	EC, DC, TPP, NTPP, PO	NTPP	EC, DC, TPP, TPPi, NTPP, NTPPi, EQ, PO	NTPP	EC, DC, TPP, NTPP, PO
NTPPi	TPPi, NTPPi, PO	EC, DC, TPPi, NTPPi, PO	TPPi, NTPPi, PO	NTPPi	TPP, TPPi, NTPP, NTPPi, EQ, PO	NTPPi	NTPPi	TPPi, NTPPi, PO
EQ	EC	DC	TPP	TPPi	NTPP	NTPPi	EQ	PO
PO	EC, DC, TPPi, NTPPi, PO	EC, DC, TPPi, NTPPi, PO	TPP, NTPP, PO	EC, DC, TPP, TPPi, EQ, PO	TPP, NTPP, PO	EC, DC, TPPi, NTPPi, PO	PO	EC, DC, TPP, TPPi, NTPP, NTPPi, EQ, PO

the table. Each table entry is a theorem on relations ρ and regions R of the form:

$$\forall R_1 \forall R_2 \forall R_3 [(\rho_1(R_1, R_2) \wedge \rho_2(R_2, R_3)) \rightarrow \rho_3(R_1, R_3)]$$

As stated in [16], the proofs which underly the entries in the composition table are both tedious to do and in some cases difficult to secure. Often a difficult proof is only obtained via lemmas. Therefore, we do not give formal proofs for (every entry of) the composition table, but we explain our example graphically.



(a) PO (b) TPP (c) NTPP

Figure 8. The composition of EC and NTPP

When we compose EC and NTPP graphically, we obtain the relations shown in Fig. 8. In any case, the theorem of the composition holds. It is possible for R_1 to partially overlap R_3, while R_1 is externally connected to R_2 and R_2 is a non-tangential proper part of R_3 (Fig. 8(a)). Also, R_1 can have a (N)TPP relation with R_3, while R_1 is externally connected to R_2 and R_2 is a non-tangential proper part of R_3 (Figs. 8(b) and 8(c)). The situation can be denoted in First Order Logic as follows:

$$\forall R_1 \forall R_2 \forall R_3 [(\text{EC}(R_1, R_2) \wedge \text{NTPP}(R_2, R_3)) \rightarrow$$

$$\text{PO}(R_1, R_3) \vee \text{TPP}(R_1, R_3) \vee \text{NTPP}(R_1, R_3)]$$

In this section, we have shown that one entry of the composition table of the RCC8 relations (Table II) is correct by evaluating it graphically. The same approach can be used for every table entry.

III. Translating RCC8 Relations into OWL

We now continue with elaborating on ways to translate RCC8 into OWL. In this section, different versions of OWL are discussed in relation to RCC8 representation, i.e., OWL DL (which is based on the logic $\mathcal{SHOIN}$) in Sect. III-A and OWL 2 (based on $\mathcal{SROIQ}$) in Sect. III-B. Section III-C discusses a concrete domains extension to $\mathcal{ALC}$, a subset of OWL DL.

A. Translating RCC8 into OWL DL

Katz and Grau [5] present a way of representing RCC8 in OWL DL. It is shown that while OWL DL already has much of the features needed to represent RCC8 relations, it lacks a necessary ability to represent reflexive roles, i.e., roles that hold always when the subject and object are equal. This is because it is known from previous research that RCC8 can be translated into *S4*, a modal logic [17]. *S4* is basically an extension of the description logic *S*, which is supported by OWL. The difference between *S* and *S4* is the possibility to represent reflexive relations. It is claimed that the reflexive property could be added to OWL in a future version without much trouble. Indeed, OWL 2 supports reflexivity, as it is based on the logic $\mathcal{SROIQ}$ [18] which includes reflexivity.

Katz et al. then proceed to give a translation of the RCC8 relations into OWL DL. Table III shows the DL-variants with R_1 and R_2 as regular regions between which the relation should hold, and $R_3, \ldots, R_8$ as unique concepts, which do not appear anywhere else in this knowledge base. Note that $R_3, \ldots, R_8$ may not be empty, so ABox assertions requiring them to contain at least one individual are needed.

B. A Hybrid Knowledge Representation System Architecture

Grütter and Bauer-Messmer [6], [7] present an approach for combining RCC8 with OWL 2. They identify some problems with the previously presented approach when translating RCC8 into OWL 2. First of all, the regions are

Table III
TRANSLATIONS OF RCC8 RELATIONS INTO OWL DL

Relation	OWL DL
EC(R_1, R_2)	$\forall C.R_1 \sqsubseteq \exists C.\neg R_2; R_3 \equiv R_1 \sqcap R_2$
DC(R_1, R_2)	$R_1 \sqsubseteq \neg R_2$
TPP(R_1, R_2)	$R_1 \sqsubseteq R_2; R_4 \equiv R_1 \sqcap \exists C.\neg R_2$
TPPi(R_1, R_2)	$R_2 \sqsubseteq R_1; R_5 \equiv R_2 \sqcap \exists C.\neg R_1$
NTPP(R_1, R_2)	$R_1 \sqsubseteq \forall C.R_2$
NTPPi(R_1, R_2)	$R_2 \sqsubseteq \forall C.R_1$
EQ(R_1, R_2)	$R_1 \equiv R_2$
PO(R_1, R_2)	$R_6 \equiv \forall C.R_1 \sqcap \forall C.R_2; R_7 \equiv R_1 \sqcap \neg R_2;$ $R_8 \equiv \neg R_1 \sqcap R_2$

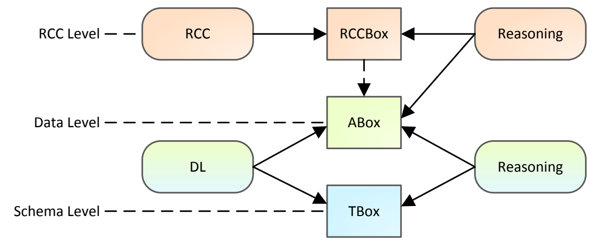


Figure 9. The hybrid system architecture of Grütter and Bauer-Messmer

sets in the abstract object domain (OWL TBox, or schema level) and not in a concrete domain (OWL ABox, or data level). Current versions of OWL do not allow classes to also be individuals (or properties) at the same time. The punning mechanism of OWL 2 is not useful here as this allows the same name to be used for two different entities: class and individual [19]. This prevents RCC8 from being used with domain ontologies, as these require the regions to be represented as individuals [5], [6]. Another problem with OWL 2 is that it does not support the kind of role inclusion axioms required for the RCC composition tables of the form $S \circ T \sqsubseteq R_1 \sqcup \ldots \sqcup R_n$ as given in Table II.

In addition to this, we find the approach of Katz et al. is difficult to interpret. Also, there is an explosion of TBox axioms, as you have to introduce for each RCC8 role association one TBox axiom, instead of defining the relations of an RCC8 role only once and apply it for different instances. Also, for some of the RCC8 axioms additional TBox axioms specifying the non-emptiness of some regions need to be introduced. Therefore, the translation of RCC8 into OWL lacks practicability, which is also emphasized by Stocker and Sirin [20].

For these reasons, instead of translating RCC into $\mathcal{SROIQ}$ or OWL, the authors propose to combine RCC with OWL at the level of the knowledge representation system architecture and not at the level of the formalisms. Therefore, the architecture of OWL is extended with RCC-specific components. These are implemented in what the authors call an *RCCBox*, which is similar to the RBox (role box) in $\mathcal{SROIQ}$ [18], in which the RCC relations and composition tables (for RCC1, 2, 3, 5, and 8) are specified. The combination of this RCCBox with OWL is the hybrid aspect of this approach. Figure 9 shows the relations between the RCCBox and the ABox and TBox.

Also, it was found that this approach also does not work well with OWL DL. One major problem is the inability of OWL DL to check whether two regions are connected or not – unless all connected relations have been explicitly defined – because role negation is not natively supported at ABox level. However, to explicitly define all the combinations of two disconnected regions would be impossible with a large number of regions. OWL 2 does support role negation at ABox level, so it is then possible to check whether any of the other 7 RCC8 relations hold, and if not, to then automatically state that the disconnected relation takes place between the two regions. This is just an approximation, which might turn out to be wrong when new facts are added to the knowledge base, but it is a consequence of the difficulties of working with the open world assumption [21].

C. Concrete Domains and GCIs in DL

Kutz et al. introduce an approach for combining several abstract description systems (domains), such that the combination is decidable if its components are decidable [22]. This means the combination method can be considered to show robust computational behavior. The authors show that their approach, the $\mathcal{E}$-connection method, has a wide range of applications, amongst which are spatial logics.

However, according to Lutz and Miličić [23], combining concrete domains (constraint systems) with General Concept Inclusions (GCI) usually leads to undecidability. A particular class of systems is based on predicates that are interpreted as jointly exhaustive and pairwise disjoint. This so-called JEPD property is also applicable to RCC8, as stated in Sect. II-C, and leads to ω-admissibility. It is stated that ω-admissibility is sufficient for proving decidability of description logics equipped with concrete domains and GCIs. Examples of ω-admissible constraint systems are Allen relations (temporal), which were implemented in $\mathcal{SHIN}$, a subset of OWL, in [24], and RCC8 relations (spatial).

The authors describe $\mathcal{ALC}(\mathcal{C})$, an extension of $\mathcal{ALC}$ which is able to handle ω-admissible constraint systems and GCIs. Two concepts are added to the $\mathcal{ALC}$ DL: $\exists U_1, U_2.r$ and $\forall U_1, U_2.r$, where r is a binary concrete domain predicate and U_1 and U_2 represent feature paths. In $\mathcal{ALC}(\mathcal{C})$, a general concept inclusion axiom is an expression of the form $C \sqsubseteq D$, where C and D are concepts without limitations (e.g., cycles are allowed). Lutz and Miličić present a tableau algorithm for $\mathcal{ALC}(\mathcal{C})$ that proves the decidability of the language. This has been generalized to $\mathcal{SHIN}$ [25].

IV. TRANSLATING RCC8 INTO DL

In the previous section, we have encountered a few major issues when representing RCC8 relations in currently

Table IV
TRANSLATIONS OF RCC ELEMENTS INTO DL (INVERSE RELATIONS ARE OMITTED)

Relation	FOL	DL
$\mathrm{C}(R_1, R_2)$	$\mathrm{C}(R_1, R_2)$	C
$\mathrm{P}(R_1, R_2)$	$\forall R_3[\mathrm{C}(R_3, R_1) \rightarrow \mathrm{C}(R_3, R_2)]$	$\neg(\mathrm{C} \circ \neg\mathrm{C})$
$\mathrm{O}(R_1, R_2)$	$\exists R_3[\mathrm{P}(R_3, R_1) \wedge \mathrm{P}(R_3, R_2)]$	$\mathrm{P}^{-} \circ \mathrm{P}$
$\mathrm{EC}(R_1, R_2)$	$\mathrm{C}(R_1, R_2) \wedge \neg\mathrm{O}(R_1, R_2)$	$\mathrm{C} \sqcap \neg\mathrm{O}$
$\mathrm{DC}(R_1, R_2)$	$\neg\mathrm{C}(R_1, R_2)$	$\neg\mathrm{C}$
$\mathrm{DR}(R_1, R_2)$	$\neg\mathrm{O}(R_1, R_2)$	$\neg\mathrm{O}$
$\mathrm{PP}(R_1, R_2)$	$\mathrm{P}(R_1, R_2) \wedge \neg\mathrm{P}(R_2, R_1)$	$\mathrm{P} \sqcap \neg(\mathrm{P}^{-})$
$\mathrm{TPP}(R_1, R_2)$	$\mathrm{PP}(R_1, R_2) \wedge \exists R_3[\mathrm{EC}(R_3, R_1) \wedge \mathrm{EC}(R_3, R_2)]$	$\mathrm{PP} \sqcap (\mathrm{EC} \circ \mathrm{EC})$
$\mathrm{NTPP}(R_1, R_2)$	$\mathrm{PP}(R_1, R_2) \wedge \neg\exists R_3[\mathrm{EC}(R_3, R_1) \wedge \mathrm{EC}(R_3, R_2)]$	$\mathrm{PP} \sqcap \neg(\mathrm{EC} \circ \mathrm{EC})$
$\mathrm{EQ}(R_1, R_2)$	$\mathrm{P}(R_1, R_2) \wedge \mathrm{P}(R_2, R_1)$	$\mathrm{P} \sqcap \mathrm{P}^{-}$
$\mathrm{PO}(R_1, R_2)$	$\mathrm{O}(R_1, R_2) \wedge \neg\mathrm{P}(R_1, R_2) \wedge \neg\mathrm{P}(R_2, R_1)$	$\mathrm{O} \sqcap \neg\mathrm{P} \sqcap \neg(\mathrm{P}^{-})$
$\mathrm{ONE}(R_1, R_2)$	$\mathrm{O}(R_1, R_2) \wedge \neg\mathrm{EQ}(R_1, R_2)$	$\mathrm{O} \sqcap \neg\mathrm{EQ}$
$\mathrm{SR}(R_1, R_2)$	$\top(R_1, R_2)$	$\top$

available OWL variants, i.e., OWL DL and OWL 2. First of all, RCC requires reflexivity (for the C relation), but in $\mathcal{SHOIN}$, the description logic on which OWL DL is based, this property is not included. Also, the required role negation at ABox level for the definition of the various RCC8 relations and complex role inclusion axioms for the composition of RCC8 relations are not natively supported by OWL DL. However, $\mathcal{SROIQ}$ – and thus OWL 2 – does support role negation at ABox level, reflexivity, and role inclusion axioms while maintaining decidability. However, $\mathcal{SROIQ}$ does not fully meet our needs, as RCC8 also requires role negation, conjunction, and disjunction at an abstract level, and we also need complex role inclusion axioms. The recently proposed extension for $\mathcal{SROIQ}$, $\mathcal{SROIQB}_s$ [9], as explained below, fulfills some of our needs.

The expressive description logic $\mathcal{SROIQB}_s$ can be obtained from $\mathcal{SROIQ}$ by allowing arbitrary Boolean constructors on simple roles, of which the addition is denoted by $\mathcal{B}_s$. Up until now, Boolean constructors (i.e., negation, conjunction, and disjunction) on roles have been used sporadically in description logics. Recently, in [9] the authors proposed to incorporate these Boolean constructors into DL languages, enabling one to enhance the expressivity of these languages without significantly increasing the reasoning complexity. It is stated in [9] that satisfiability checking, instance retrieval, and computing class subsumptions for $\mathcal{SROIQB}_s$ knowledge bases is `N2ExpTime`-complete, while for OWL DL it is `NExpTime`-complete.

One problem still remains in $\mathcal{SROIQB}_s$, i.e., composition is only allowed to exist in the left hand side of a role inclusion axiom. However, when translating RCC8 relations into a description logics language, we need composition in the right hand side as well (e.g., the equivalents given by the composition tables or RCC8 relation definitions are in fact double inclusions between the left hand side and the right hand side). Furthermore, $\mathcal{SROIQB}_s$ allows only conjunction of simple role expressions. However, for expressing RCC8 relations, we also need role conjunctions that contain role composition.

Wessel shows extensions to various $\mathcal{ALC}$ variants in order to implement the composition-based role inclusion axioms of the form we need for representing RCC8 [26]. It is concluded that implementing such axioms often leads to undecidability. However, Horrocks and Sattler show that $\mathcal{SHIQ}$ can be extended with complex role inclusion axioms, without endangering the decidability [27]. This is true when restricting the set of role inclusion axioms to be acyclic.

In the rest of this section we will prove based on the FOL definitions of RCC8 relations given in Sect. II-B that their DL counterparts presented in Table IV are sound.

Proposition 1: $\mathrm{C} \equiv \mathrm{C}$

Proof: In FOL we define the relation C as $\mathrm{C}(R_1, R_2)$. This relation is a primitive relation and thus it cannot be rewritten. In DL the relation C is defined as C.

■

Proposition 2: $\mathrm{P} \equiv \neg(\mathrm{C} \circ \neg\mathrm{C})$

Proof: The relation P is defined in FOL as $\mathrm{P}(R_1, R_2) \equiv \forall R_3[\mathrm{C}(R_3, R_1) \rightarrow \mathrm{C}(R_3, R_2)]$. Rewriting yields $\mathrm{P}(R_1, R_2) \equiv \neg\mathrm{C}(R_3, R_1) \vee \mathrm{C}(R_3, R_2) \equiv \neg(\mathrm{C}(R_3, R_1) \wedge \neg\mathrm{C}(R_3, R_2))$ (De Morgan's law). Note that the universal quantifier has been removed by skolemization. Continuing, we arrange the arguments of C for composition and obtain $\mathrm{P}(R_1, R_2) \equiv \neg(\mathrm{C}^{-}(R_1, R_3) \wedge \neg\mathrm{C}(R_3, R_2))$. Then, we can translate the relation into DL, and thus we obtain $\mathrm{P} \equiv \neg(\mathrm{C}^{-} \circ \neg\mathrm{C})$. Omitting the inverse (because C is symmetric) yields $\mathrm{P} \equiv \neg(\mathrm{C} \circ \neg\mathrm{C})$.

■

Proposition 3: $\mathrm{O} \equiv \mathrm{P}^{-} \circ \mathrm{P}$

Proof: The FOL definition of the relation O is $\mathrm{O}(R_1, R_2) \equiv \exists R_3[\mathrm{P}(R_3, R_1) \wedge \mathrm{P}(R_3, R_2)]$. When rewriting this FOL expression to a DL expression, we can apply composition. Therefore, the quantifier disappears and after arranging the arguments of P for composition, we obtain $\mathrm{O}(R_1, R_2) \equiv \mathrm{P}^{-}(R_1, R_3) \wedge \mathrm{P}(R_3, R_2)$ and thus translating the definition into DL, $\mathrm{O} \equiv \mathrm{P}^{-} \circ \mathrm{P}$.

■

Proposition 4: EC $\equiv$ C $\sqcap \neg$O

Proof: In FOL the relation EC between externally connected regions R_1 and R_2 is defined as EC$(R_1, R_2) \equiv$ C$(R_1, R_2) \wedge \neg$O(R_1, R_2). Rewriting into DL gives EC $\equiv$ C $\sqcap \neg$O.

■

Proposition 5: DC $\equiv \neg$C

Proof: The DC relation is defined as DC$(R_1, R_2) \equiv$ $\neg$C(R_1, R_2) in FOL. This relation is a primitive relation and thus it cannot be rewritten. Therefore, in DL the relation DC is defined as DC $\equiv \neg$C.

■

Proposition 6: DR $\equiv \neg$O

Proof: We define the DR relation between discrete regions R_1 and R_2 in FOL as DR$(R_1, R_2) \equiv \neg$O(R_1, R_2). Rewriting into DL gives DR $\equiv \neg$O.

■

Proposition 7: PP $\equiv$ P $\sqcap \neg$P$^-$

Proof: In FOL, PP is defined as PP$(R_1, R_2) \equiv$ P$(R_1, R_2) \wedge \neg$P(R_2, R_1). Here, P(R_2, R_1) indicates an inverse relation between the two regions, and thus in DL this translates to P$^-$. Rewriting PP to a DL expression yields PP $\equiv$ P $\sqcap \neg$P$^-$.

■

Proposition 8: TPP $\equiv$ PP $\sqcap$ (EC $\circ$ EC)

Proof: The FOL definition of the relation TPP is TPP$(R_1, R_2) \equiv$ PP$(R_1, R_2) \wedge \exists R_3$[EC$(R_3, R_1) \wedge$ EC(R_3, R_2)]. When rewriting this FOL expression to a DL expression, we can apply composition. Therefore, the existential quantifier disappears and after rearranging the arguments of EC for composition, we obtain TPP$(R_1, R_2) \equiv$ PP$(R_1, R_2) \wedge$ (EC$^-(R_1, R_3) \wedge$ EC(R_3, R_2)). By translating the previous formula into DL, we get TPP$(R_1, R_2) \equiv$ PP $\sqcap$ (EC$^-$ $\circ$ EC). Because EC is a symmetric relation, we can omit the inverse, and thus TPP$(R_1, R_2) \equiv$ PP $\sqcap$ (EC $\circ$ EC).

■

Proposition 9: NTPP $\equiv$ PP $\sqcap \neg$(EC $\circ$ EC)

Proof: The NTPP relation is defined in FOL as NTPP$(R_1, R_2) \equiv$ PP$(R_1, R_2) \wedge \neg\exists R_3$[EC$(R_3, R_1) \wedge$ EC(R_3, R_2)]. By skolemization, the quantifier disappears and after arranging the arguments of EC for composition, we obtain NTPP$(R_1, R_2) \equiv$ PP$(R_1, R_2) \wedge \neg$(EC$^-(R_1, R_3) \wedge$ EC(R_3, R_2)) and thus translating the definition into DL, NTPP $\equiv$ PP $\sqcap \neg$(EC$^-$ $\circ$ EC). Because EC is symmetric, NTPP $\equiv$ PP $\sqcap \neg$(EC $\circ$ EC).

■

Proposition 10: EQ $\equiv$ P $\sqcap$ P$^-$

Proof: In FOL, EQ is defined as EQ$(R_1, R_2) \equiv$ P(R_1, R_2) $\wedge$ P(R_2, R_1). Here, P(R_2, R_1) indicates an inverse relation between the two regions, and thus in DL this translates to P$^-$. Rewriting EQ to a DL expression yields EQ $\equiv$ P $\sqcap$ P$^-$.

■

Proposition 11: PO $\equiv$ O $\sqcap \neg$P $\sqcap \neg$(P$^-$)

Proof: The FOL definition of the relation PO is PO$(R_1, R_2) \equiv$ O$(R_1, R_2) \wedge \neg$P$(R_1, R_2) \wedge \neg$P(R_2, R_1). When rewriting this FOL expression to a DL expression, we obtain PO $\equiv$ O $\sqcap \neg$P $\sqcap \neg$(P$^-$).

■

Proposition 12: ONE $\equiv$ O $\sqcap \neg$EQ

Proof: In FOL the relation ONE is defined as ONE$(R_1, R_2) \equiv$ O$(R_1, R_2) \wedge \neg$EQ(R_1, R_2). Rewriting into DL gives ONE $\equiv$ O $\sqcap \neg$EQ.

■

Proposition 13: SR $\equiv \top$

Proof: In DL, SR is expressed as the top relation, denoted as $\top$, and therefore SR $\equiv \top$.

■

V. Conclusions

In this paper, we have identified several issues encountered when representing RCC8 in OWL. We have reviewed different approaches to implement RCC8 relations in OWL. OWL DL, which is based on the description logics language $\mathcal{SHOIN}$ is not suitable for implementing RCC8 relations, as the language lacks required features such as role reflexivity, role Boolean operators, and role inclusion axioms. Some of these features are to be included in the forthcoming OWL 2 (based on $\mathcal{SROIQ}$), but this language still lacks the expressive power to support role negations, conjunctions, and disjunctions at an abstract level, as well as complex role inclusion axioms. However, a recently developed extension for $\mathcal{SROIQ}$, $\mathcal{SROIQB}_s$, seems to support some of our needs, as the language adds role Boolean operators to $\mathcal{SROIQ}$ while maintaining the decidability of the language. Therefore, we propose to extend OWL 2 with role Boolean operators. However, if we want to express RCC8 relations in a description logics language, complex role inclusion axioms need to be supported on the right hand side, a feature that is not included in $\mathcal{SROIQB}_s$. Also, Boolean role operators on complex roles are needed to be added as well (e.g., for defining P, PP, TPP, etc.).

Therefore, as future work we would like to investigate how to extend $\mathcal{SROIQ}$ with more complex role inclusion axioms and Boolean operators on complex roles, while maintaining decidability. Also, it is worth investigating how one can represent more powerful spatial knowledge representation formalisms like RCC15 and 9-Intersection on the Semantic Web. One of the advantages of 9-Intersection is that it allows to capture the spatial relations between not only regions, but also between regions and points or between regions and lines. Furthermore, we would like to implement the tableau algorithm given in [9] and experiment with RCC8-based reasoning in different domains (business, cadaster, weather, health, etc.). Finally, the decidability of the proposed OWL 2 extensions is the subject of future work.

ACKNOWLEDGEMENT

The authors would like to thank Chuck Ng and Marc Kuo for their contribution to this paper.

REFERENCES

[1] W. van Hage, J. Wielemaker, and G. Schreiber, "The Space Package: Tight Integration Between Space and Semantics," in *Terra Cognita Workshop (Terra 2009)*. CEUR-WS.org, 2009.

[2] D. A. Randell, Z. Cui, and A. G. Cohn, "A Spatial Logic Based on Regions and Connection," in *3rd Int. Conf. on Knowledge Representation and Reasoning (KR 1992)*. Morgan Kaufmann, 1992, pp. 165–176.

[3] M. J. Egenhofer, J. Sharma, and D. M. Mark, "A Critical Comparison of the 4-Intersection and 9-Intersection Models for Spatial Relations: Formal Analysis," in *11th Int. Symposium on Computer-Assisted Cartography (AUTO-CARTO 1993)*. American Society for Photogrammetry and Remote Sensing, 1993, pp. 1–11.

[4] B. Motik, P. F. Patel-Schneider, and B. Parsia, "OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax - W3C Recommendation 27 October 2009," 2009.

[5] Y. Katz and B. C. Grau, "Representing Qualitative Spatial Information in OWL-DL," in *International Workshop on OWL: Experiences and Directions (OWLED 2005)*, ser. CEUR Workshop Proceedings, vol. 188. CEUR-WS.org, 2006.

[6] R. Grütter and B. Bauer-Messmer, "Combining OWL with RCC for Spatioterminological Reasoning on Environmental Data," in *3rd Int. Workshop on OWL: Experiences and Directions (OWLED 2007)*. CEUR-WS.org, 2007.

[7] ——, "Towards Spatial Reasoning in the Semantic Web: A Hybrid Knowledge Representation System Architecture," in *10th AGILE Int. Conf. on Geographic Information Science (AGILE 2007)*. Springer, 2007, pp. 349–364.

[8] F. Hogenboom, F. Frasincar, and U. Kaymak, "A Review of Approaches for Representing RCC8 in OWL," in *25th Symposium On Applied Computing (SAC 2010)*, S. Y. Shin, S. Ossowski, M. Schumacher, M. J. Palakal, and C.-C. Hung, Eds. ACM, 2010, pp. 1444–1445.

[9] S. Rudolph, M. Krötzsch, and P. Hitzler, "Cheap Boolean Role Constructors for Description Logic," in *11th European Conf. on Logics in Artificial Intelligence (JELIA 2008)*. Springer, 2008, pp. 362–374.

[10] J. Renz and B. Nebel, "Qualitative Spatial Reasoning using Constraint Calculi," in *Handbook of Spatial Logics*, M. Aiello, I. E. Pratt-Hartmann, and J. F. van Benthem, Eds. Springer, 2007, ch. 4, pp. 161–215.

[11] Brandon Bennett, "Spatial Reasoning with Propositional Logics," in *4th Int. Conf. of Principles of Knowledge Representation and Reasoning (KR 1994)*. Morgan Kaufmann, 1994, pp. 51–62.

[12] J. Renz, *Qualitative Spatial Reasoning with Topological Information*, ser. Lecture Notes in Computer Science, G. Goos, J. Hartmanis, and J. van Leeuwen, Eds. Springer, 2002, vol. 2293.

[13] A. Grzegorczyk, "Undecidability of some Topological Theories," *Fundamenta Mathematicae*, vol. 38, no. 2, pp. 137–152, 1951.

[14] O. Stock, *Spatial and Temporal Reasoning*. Springer, 1998.

[15] A. G. Cohn, J. Gooday, and B. Bennett, "A Comparison of Structures in Spatial and Temporal Logics," in *16th Int. Wittgenstein Symposium (IWS 1994)*, R. Casati, B. Smith, and G. White, Eds. Holder-Pichler-Tempsky, 1994, pp. 1–14.

[16] D. A. Randell, A. G. Cohn, and Z. Cui, "Computing Transitivity Tables: A Challenge for Automated Theorem Provers," in *11th Int. Conf. on Automated Deduction (CADE 1992)*, D. Kapur, Ed. Springer, 1992, pp. 786–790.

[17] C. I. Lewis and C. H. Langford, *Symbolic Logic*. The Century Co, 1932.

[18] I. Horrocks, O. Kutz, and U. Sattler, "The Even More Irresistible $\mathcal{SROIQ}$," in *10th Int. Conf. on Principles of Knowledge Representation and Reasoning (KR 2006)*. AAAI Press, 2006, pp. 57–67.

[19] C. Golbreich and E. K. Wallace, "OWL 2 Web Ontology Language New Features and Rationale - W3C Recommendation 27 October 2009," 2009.

[20] M. Stocker and E. Sirin, "PelletSpatial: A Hybrid RCC-8 and RDF/OWL Reasoning and Query Engine," in *6th Int. Workshop on OWL: Experiences and Directions (OWLED 2009)*. CEUR-WS.org, 2009.

[21] L. Baresi, E. Di Nitto, and C. Ghezzi, "Toward Open-World Software: Issue and Challenges," *Computer*, vol. 39, no. 10, pp. 36–43, 2006.

[22] O. Kutz, C. Lutz, F. Wolter, and M. Zakharyaschev, "E-connections of Abstract Description Systems," *Artificial Intelligence*, vol. 156, no. 1, pp. 1–73, 2004.

[23] C. Lutz and M. Miličić, "A Tableau Algorithm for Description Logics with Concrete Domains and General TBoxes," *J. Autom. Reason.*, vol. 38, no. 1–3, pp. 227–259, 2007.

[24] F. Frasincar, V. Milea, and U. Kaymak, *Semantic Web Information Management: A Model-Based Perspective*. Springer, 2010, ch. tOWL: Integrating Time in OWL, pp. 225–246.

[25] C. Lutz, "Adding numbers to the SHIQ description logic: First results," in *8th International Conference on Principles of Knowledge Representation and Reasoning (KR 2002)*, 2002, pp. 191–202.

[26] M. Wessel, "Obstacles on the Way to Qualitative Spatial Reasoning with Description Logics: Some Undecidability Results," in *Int. Description Logics Workshop (DL 2001)*. CEUR-WS.org, 2001.

[27] I. Horrocks and U. Sattler, "Decidability of SHIQ with Complex Role Inclusion Axioms," *Artificial Intelligence*, vol. 160, no. 1–2, pp. 79–104, 2003.