Formal Qualitative Spatial Augmentation of the Simple Feature Access Model

Shirly Stephen
School of Computing and Information Science, University of Maine, Orono, ME 04469, USA
shirly.stephen@maine.edu

Torsten Hahmann
School of Computing and Information Science, University of Maine, Orono, ME 04469, USA
torsten.hahmann@maine.edu

Abstract

The need to share and integrate heterogeneous geospatial data has resulted in the development of geospatial data standards such as the OGC/ISO standard Simple Feature Access (SFA), that standardize operations and simple topological and mereotopological relations over various geometric features such as points, line segments, polylines, polygons, and polyhedral surfaces. While SFA's supplied relations enable qualitative querying over the geometric features, the relations' semantics are not formalized. This lack of formalization prevents further automated reasoning – apart from simple querying – with the geometric data, either in isolation or in conjunction with external purely qualitative information as one might extract from textual sources, such as social media. To enable joint qualitative reasoning over geometric and qualitative spatial information, this work formalizes the semantics of SFA's geometric features and mereotopological relations by defining or restricting them in terms of the spatial entity types and relations provided by CODIB, a first-order logical theory from an existing logical formalization of multidimensional qualitative space.

2012 ACM Subject Classification Computing methodologies → Ontology engineering; Computing methodologies → Spatial and physical reasoning; Information systems → Geographic information systems

Keywords and phrases space, geometry, geospatial semantics, qualitative spatial representation (QSR), simple feature access, topological relations, formal ontology

Digital Object Identifier 10.4230/LIPIcs.COSIT.2019.15

Funding This material is based upon work supported by the National Science Foundation under Grant Number III-1565811.

Acknowledgements The authors are grateful for the four anonymous reviewer's thoughtful comments that helped improve the final version of the paper.

1 Introduction

The need to share and integrate the large amounts of heterogeneous geospatial data has resulted in the development of geospatial data standards, such as the OGC's GeoSPARQL standard [27] and the shared OGC/ISO standards Geography Markup Language (GML) [23] and Simple Feature Access [22]. All of these standards include some types of simple and complex geometric features – often simply referred to as *geometries* – for representing geographic objects. The most commonly used features include points, line segments and aggregations into polylines, and polygons and aggregations into polyhedral surfaces. Primarily concerned with interoperability across spatial databases and geographic information systems, these standards also prescribe a number of common spatial operators, e.g., for calculating intersections, differences, buffers, or distances between features.

Many of these standards have further incorporated a number of simple mereotopological relations (with Boolean values), such as intersects, contains, overlaps, meets, or crosses.

14th International Conference on Spatial Information Theory (COSIT 2019).
Editors: Sabine Timpf, Christoph Schlieder, Markus Kattenbeck, Bernd Ludwig, and Kathleen Stewart; Article No. 15; pp. 15:1–15:18
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

These are based on results from the Region Connection Calculus (RCC) [28] and the almost equivalent topological relations defined by the 9-intersection method [8, 9] and its dimensionally extended refinements (DE-9I) [5, 6] and further extensions [26, 29].

The Simple Feature Access (SFA) model [22], an OGC and ISO standard for vector-based encoding of 2D geometric data, is one of the most widely implemented standards for facilitating geospatial data interoperability. It is at least partially implemented by a wide range of geographic information systems and spatial database systems, including ArcSDE (the spatial database system that ArcGIS uses), PostGIS, and the spatial extensions of MySQL, Oracle, and IBM Db2. Other geospatial standards, like GeoJSON[1] and GeoSPARQL [27], also build on SFA.

However, the mereotopological relations provided by SFA and similar standards use them as query operators only[2]. This enables more natural access to geometric data but without formalizing the relationships between geometric representations and the mereotopological or other qualitative relations, these approaches cannot support qualitative reasoning over the queried information. Moreover, storing "native" topological information – for example as provided from textual sources where precise locations or spatial extents are unknown or unknowable – is currently not possible without having to *invent geometric objects*. For example, the spatial content of the two statements "Lot A is for sale and abuts Broadway." and "Lot B that does not border Broadway is not for sale." cannot be represented in GIS without assigning geometries to the named objects.

Frameworks for qualitative spatial representation and reasoning (see, e.g., the overview in [7]) such as the RCC support direct reasoning about topological and other kinds of qualitative spatial information (e.g., direction), but cannot easily mix geometric data sources (e.g., the precise location of "Broadway") and qualitative information (the fact that "Lot A" and Broadway are connected) to infer which lots on a property map may be for sale. Similar interpretation of qualitative spatial information on a geometric dataset is needed during natural disasters, when interpreting human reports (e.g., from social media or news reports) on road networks, elevation data, and hydrological data, to help answer simple queries, such as "is any part of the historic center flooded?".

The presented work is a step in this direction by developing a first-order logical theory[3] that treats geometric features (e.g., polylines, polygons) and relations between them as specializations of more general types of features (e.g., any kind of 2D regions or 1D features) and mereotopological relations between them. Key to this endeavour is the use of a *multidimensional* theory of space wherein, unlike traditional logical theories of mereotopology (including the RCC), spatial entities of different dimensions can co-exist and be related. We choose the theory CODIB (based on CODI [17, 16] with an extension by boundary/interior distinctions [15]) as suitable multidimensional theory of qualitative space and test to what extent geometric features from SFA [22] can be treated as specializations of CODIB's more general non-geometric spatial feature types from CODIB. For example, SFA's line segments or polylines should specialize the general one-dimensional spatial features, called "curves", from CODIB. Specifically, we want to leverage the detailed formal semantics encoded in CODIB to capture the semantics of SFA's various geometric feature types and mereotopological relations in greater detail. Currently, much of these semantics are described in natural language and mathematical notation in the standard, but are not accessible to automated

[1] http://geojson.org/

[2] Most GIS support the RCC or DE-9I relations, with recent progress on storing the computed relations more efficiently [24]. There has also been a call to extend this to a larger set of qualitative relations [11].

[3] The term "theory" refers throughout the paper to a logical theory. The terms "theory", "ontology" and "axiomatization" are used synonymously.

reasoning. Wherever possible, we logically define SFA's geometric features in terms of CODIB's spatial concepts and, where that is not possible, treat them as specializations with suitable constraints.

Our specific contributions are: (1) develop a first-order logic axiomatization, called SF-FOL, of SFA; (2) in the process, show that all of the geometric feature types from SFA specialize or map to types of spatial entities definable in CODIB; (3) fully define SFA's mereotopological relations in CODIB and thus provide computer-interpretable semantics of these qualitative relations; and (4) verify the consistency of SF-FOL. This makes both SFA's and CODIB's mereotopological relations applicable to geometric and qualitative data alike and allows using automated first-order logic theorem provers (ATPs) for integrated mereotopological reasoning over combinations of qualitative and geometric data from any sources that adhere to the SFA standard.

2 Background and Related Work

Mereotopological relations are among the most common qualitative spatial relations [25], and have been incorporated into virtually all upper ontologies [14]. They include purely topological relations such as contact/connection or disconnection, and purely mereological relations such as parthood, containment, or inside, as well as relations that describe the interaction of topology and mereology such as overlap (i.e., contact via sharing a part). Simple mereotopological relations have also been included in popular geospatial data standards thanks to seminal work on the 9-intersection method [8, 9], its dimension-extended refinement (DE-9I) [5] and extensions thereof [6, 26, 29]. However, the 9-intersection method determines these relations from geometric data by computing a matrix of values that indicate the pairwise intersections of two object's *interior* ($\circ$), *boundary* (∂), and *complement* ($'$). Each of the nine pairs have either Boolean values – empty nor non-empty intersection – as in the original 9-intersection framework, or have dimensional values – either -1 (empty intersection), 0, 1, or 2 – as in the dimension-extended method. This way of determining the qualitative relations requires an underlying geometric representation, with associated operations for determining their boundary and interior, for all involved objects. Moreover, the semantics of the mereotopological relations, especially their interaction (e.g. parthood specializes overlap or a whole is in contact with everything any of its parts is on contact with), are never explicitly captured and thus not available for qualitative reasoning with the underlying data. Moreover, the relations cannot be used for reasoning where geometric data models are not the only source of qualitative information.

This is in sharp contrast with axiomatic treatments of mereotopology, which constrain the interpretations of one or two primitive relations, such as contact and/or parthood, and define other relations, such as overlap or external contact, in terms of the primitive ones [3]. By explicitly formalizing relationships between the relations, axiomatic frameworks permit reasoning with qualitative information even in the absence of geometric information. The most well-known axiomatic theory is the RCC [28] that defines eight mereotopological relations similar to the ones from the basic 9-intersection model. The variety of existing axiomatic theories are more thoroughly reviewed in [20][4]. However, axiomatic theories of mereotopology

[4] Qualitative spatial calculi (see, e.g., the overview in [7]) are yet another approach to qualitative spatial reasoning, but they can only incorporate qualitative information and cannot make use of geometric information without first translating it to qualitative information. A hybrid reasoning system utilizing a constraint network reasoning approach for reasoning with both geometric and qualitative information has been presented in [10]. This work here goes a step further by explicitly formalizing the semantic

have, in the philosophical tradition of Whitehead, been often married to strict region-based conceptualizations of space wherein extended spatial entities – typically called regions – are the only first-class entities of the domain, while points and other lower-dimensional entities are not entities in the domain. This prevents full integration with geometric data standards, such as SFA, that permit entities of different dimensions. The idea of *multidimensional mereotopology* [12, 13, 17, 30] aims to overcome this restriction by axiomatically formalizing mereotopological relations not just between entities of equal dimensions but also between entities of different dimensions. This work utilizes the multidimensional mereotopology CODIB [17, 16, 15], which has been specifically developed to qualitatively generalize geometric data models, as basis for formalizing SFA's semantics. CODIB is based on the three primitive relations of **CO**ntainment, relative **DI**mension, and **B**oundary containment [15], which give the theory its name. CODIB builds on and extends the theory CODI (without any notion of boundaries) [17, 16] by the additional relation of boundary containment. Unlike other multidimensional theories [12, 30], CODIB allows entities of lower dimensions to exist independent of entities of higher dimension, similar to how such entities (e.g., polylines or points) are used in geometric data standards. [12, 30] require each line or curve to be part of the boundary of some 2D region and each point to be the endpoint of some curve in a model. The INCH calculus [13], on the other hand, does not model boundaries at all. Another alternative formalization of multidimensional mereotopology is provided by the space ontology (GFO space) [1] that is part of the General Formal Ontology (GFO). However, GFO space is primarily concerned with *physical*, *phenomenal space* (i.e., the space of material objects), which is different from the kind of *abstract*, *extensional space* that geometric data models describe[5][15, 1].

3 Preliminaries

We now review and formalize the relevant aspects of the SFA standard, namely its classes of geometric features and its qualitative relations. In particular, Section 3.1 formalizes the intrinsic semantics of the UML subclass hierarchy from the standards document in first-order logic as starting point for its semantic enhancement. Subsequently, Section 3.2 reviews key relations and concepts from the CODI and CODIB ontologies and provides definitions of novel concepts that are necessary to draw some of the distinctions that SFA makes. These concepts and relations will be used as basis for elaborating the SFA semantics and making its geometric features available for integration with purely qualitative information and for general qualitative reasoning.

All logical sentences throughout our exposition are assumed to be universally quantified. They are labeled in the format '[theory]-[type][number]' (e.g. SFC-T1) where the first letter(s) indicate the theory (e.g. SFC=simple features concept, SFR=simple features relation, PO=partial overlap, D=dimension), while the type distinguishes axioms (A), definitions (D: defining a concept or relation), theorems (T: a property provable from the axioms and definitions), and mappings (M: an axiom that establishes some relationship between SFA and CODIB). All theories are available in modularized form in the Common Logic syntax from the COLORE repository[6].

---

relationships between the two types of information for reuse with any logic-based reasoner.

[5] For example, in phenomenal space, any road would be a 3D object, whereas in abstract space it is typically modeled as a 1D spatial feature.

[6] In `https://colore.oor.net/`. Note that all of axioms are specified using only the classical first-order logic syntax of Common Logic and without use of any of Common Logic's specialized features such as restricted module import or use of sequence markers. This allows easy translation to pure first-order

3.1 Semantics of Simple Feature Concepts and Spatial Relations

SFA [21] is an OGC and ISO standard for vector-based encoding of 0-2D geometric data that aims to facilitate interoperability across GIS and spatial databases. SFA is at least partially implemented by ArcGIS, PostGIS, and the spatial extensions of MySQL, Oracle, and IBM Db2. Other standards, like GeoSPARQL [27] and GeoJSON, build on it.

3.1.1 Semantics of Concepts (Classes) from Simple Features

At the core of SFA lies a set of simple geometries such as individual points (*sf_point*), polylines (*sf_line_string*: a sequence of straight line segments), and polyhedral surfaces (*sf_polyhedral_surface*: a connected, possibly non-planar 2D area obtained by stitching polygons together). *Sf_line_string* and *sf_polyhedral_surface* specialize the general classes *sf_curve*, which may include non-straight segments, and *sf_surface*, which may include 2D areas with non-straight boundary segments, respectively (SFC-A1,A2). These two classes capture all kinds of 1D and 2D spatial objects. Note that at this point, we only formalize the relationships between the classes as we cannot capture their detailed semantics. Only later on, with the help of CODIB concepts and relations, can we formalize the classes in more detail.

In addition to the three classes of simple features, collections of simple features can be modeled using the *sf_geometry_collection* class. The four specializations of the abstract class *sf_geometry* are mutually disjoint (SFC-A3–A6) and jointly exhaustive (SFC-D1).

(SFC-D1) $sf\_geometry(x) \leftrightarrow$
$sf\_point(x) \vee sf\_curve(x) \vee sf\_surface(x) \vee sf\_geometry\_collection(x)$

(SFC-A1) $sf\_line\_string(x) \rightarrow sf\_curve(x)$

(SFC-A2) $sf\_polyhedral\_surface(x) \rightarrow sf\_surface(x)$

(SFC-A3) $sf\_point(x) \rightarrow \neg sf\_curve(x) \wedge \neg sf\_surface(x) \wedge \neg sf\_geometry\_collection(x)$

(SFC-A4) $sf\_curve(x) \rightarrow \neg sf\_point(x) \wedge \neg sf\_surface(x) \wedge \neg sf\_geometry\_collection(x)$

(SFC-A5) $sf\_surface(x) \rightarrow \neg sf\_point(x) \wedge \neg sf\_curve(x) \wedge \neg sf\_geometry\_collection(x)$

(SFC-A6) $sf\_geometry\_collection(x) \rightarrow \neg sf\_point(x) \wedge \neg sf\_curve(x) \wedge \neg sf\_surface(x)$

Sf_line_string is further specialized into *sf_line* (SFC-A7), which represents a single straight line segment, and *sf_linear_ring* (SFC-A8), a linear feature that is closed, that is, its start and end points coincide and thus its boundary is empty. The intended semantics of *sf_line* and *sf_linear_ring* will be more fully formalized in Section 4.1 by establishing mappings to CODIB concepts that are more densely axiomatized. For example, SFC-M3, M4, M8, and M9 together with CODIB's formalization (including the definitions of AtomicS-D, SimpleS-D, BranchedS-D, ConS-D, and the formalization of the predicate *ICon* from [15]) entail that any *sf_line* is a connected curve with two distinct end points. Likewise, *sf_polygon* is a specialization of *sf_polyhedral_surface* (SFC-A9), capturing a planar 2D area with a single closed polyline as exterior boundary[7]. Another specialization of *sf_polyhedral_surface* is *sf_tin* (SFC-A10), a triangulated irregular network (TIN), which consists of triangles. A single triangle, described by *sf_triangle*, is a polygon and the simplest kind of a TIN (SFC-D2). It is bounded by a closed polyline (i.e., a *sf_linear_ring*) that consists of exactly three line segments (i.e., *sf_line*) – this will be formalized by SFC-M13 in Section 4.1.

logic representations such as the TPTP format [31] supported by many theorem provers and model finders.

[7] SFA models *sf_polygon* and *sf_polyhedral_surface* as separate specializations of *sf_surface*, but permits polyhedral surfaces to consist of a single polygon, in which case it is spatially a polygon.

(SFC-A7) $sf\_line(x) \rightarrow sf\_line\_string(x)$

(SFC-A8) $sf\_linear\_ring(x) \rightarrow sf\_line\_string(x)$

(SFC-A9) $sf\_polygon(x) \rightarrow sf\_polyhedral\_surface(x)$

(SFC-A10) $sf\_tin(x) \rightarrow sf\_polyhedral\_surface(x)$

(SFC-D2) $sf\_triangle(x) \leftrightarrow sf\_polygon(x) \wedge sf\_tin(x)$

Sf_multi_point, *sf_multi_curve* and *sf_multi_surface* specialize *sf_geometry_ collection* (SFC-A11); they are aggregations of only *sf_points*, *sf_curves*, or *sf_surfaces*, respectively. *Sf_multi_curve* and *sf_multi_surface* are again abstract classes in SFA, with only the specializations *sf_multi_line_string* (SFC-A12) and *sf_multi_polygon* (SFC-A13) being instantiable. The latter two consist only of *sf_line_strings* and *sf_polygons*, respectively, as axiomatically captured in Section 4.2.

(SFC-A11) $sf\_geometry\_collection(x) \leftrightarrow$
$sf\_multi\_point(x) \vee sf\_multi\_curve(x) \vee sf\_multi\_surface(x)$

(SFC-A12) $sf\_multi\_line\_string(x) \rightarrow sf\_multi\_curve(x)$

(SFC-A13) $sf\_multi\_polygon(x) \rightarrow sf\_multi\_surface(x)$

The axioms SFC-A1 to SFC-A13 together with SFC-D1,D2 form the ontology SFC-Core[8] that serves as basis for our semantic elaboration of SFA in Section 4.

3.1.2 Spatial Relations in Simple Features

In addition to various geometric spatial operations (e.g., buffer, intersection, convexHull), which are only well-defined on geometric features (e.g., on polygons rather than general surfaces), SFA includes eight named qualitative spatial relations based on the dimensionally extended 9-intersection method [5] that can equally be applied to generalizations of geometric features such as arbitrary curves and surfaces. SFA's relations include the five primitive relations *disjoint*, *touches*, *within*, *overlaps*, and *crosses*, with three additional relations *contains* (inverse of *within*), *intersects* (negation of *disjoint*), and *equals* (conjunction of *within* and *contains*) being defined[9]. SFA expresses the semantics of these relations using the interior, boundary, and exterior of the related objects [22], but does not formally relate the relations to one another as we will do in Section 4.3. Three dimensional constraints are explicitly mentioned in SFA: touches does not apply to points (or *sf_multi_points*), overlaps requires the involved entities to be of equal dimension, and crosses is not applicable to two surfaces (or *sf_multi_surfaces*). These constraints will become provable as theorems of our CODIB-based formalization of these relations.

3.2 Dimensional Features and Qualitative Spatial Relations in CODIB

This subsection reviews CODIB by first reviewing its core CODI and then additional relation of boundary containment. A computer-readable encoding of the axioms are provided in the Common Logic syntax in the COLORE repository to facilitate automated verification and reasoning.

---

[8] Available from `https://colore.oor.net/simple_features`.
[9] See the definitions provided in SFR-M6–M8. We have only decided to map *contains* to CODIB's *Cont* relation and then define *within* as its inverse.

3.2.1 CODI

Core to the multidimensional mereotopology CODIB is the theory CODI[10] of containment and dimension that axiomatizes mereotopological relations in a dimension-independent way using two primitive relations: (1) the mereological notion of containment, $Cont(x, y)$[11], and a relation $\leq_{\text{dim}} (x, y)$, read as "x has the same or a lower dimension than y", to compare the dimension of two entities [16, 17]. In addition, the primitive unary predicate $S(x)$ is used to denote spatial regions, which capture mathematical regions of space whose existence is independent of whether an actual physical object occupies a spatial region or not. *Cont* is reflexive, symmetric, and transitive (Cont-A1–A3) and allows defining the zero (i.e., null) region denoted by the unary predicate *ZEX* (ZEX-D). Containment requires the contained entity to be of the same or a lower dimension than the entity it is contained in (CD-A1).

The relative dimension $\leq_{\text{dim}} (x, y)$ alone can define additional relations of equal dimension $=_{\text{dim}} (x, y)$, lesser dimension $<_{\text{dim}} (x, y)$, minimal dimension $MinDim(x)$ (i.e., the dimension of a point; D-D6), and next-lower dimension $\prec_{\text{dim}} (x, y)$ (D-D7). The relation $\leq_{\text{dim}} (x, y)$ is axiomatized to form a discrete (i.e., there is a next-lower dimension for every non-minimal entity) and bounded (i.e., a lowest and highest dimension exists) pre-order over all spatial regions. That also implies that every spatial region must be of uniform dimension, i.e., all components (i.e. parts) thereof are of the same dimension, precluding objects such as a region consisting of a 2D region and a separate, isolated point or linear feature. Spatial regions can still *contain* lower-dimensional entities (e.g., a 2D region containing 1D features and points). Using the relative dimension of the involved entities, we can specialize containment to parthood (i.e., equidimensional containment; EP-D) and proper parthood (EPP-D). Minimal spatial entities have no proper parts (ME-D2), that is, they are indivisible. There can be minimal entities within each dimension.

(Cont-A1) $S(x) \land \neg ZEX(x) \leftrightarrow Cont(x, x)$
(containment is reflexive for all nonzero spatial regions)

(Cont-A2) $Cont(x, y) \land Cont(y, x) \rightarrow x = y$ ***(containment is antisymmetric)***

(Cont-A3) $Cont(x, y) \land Cont(y, z) \rightarrow Cont(x, z)$ ***(containment is transitive)***

(ZEX-D) $ZEX(x) \leftrightarrow S(x) \land \forall y[\neg Cont(x, y) \land \neg Cont(y, x)]$ ***(zero region)***

(CD-A1) $Cont(x, y) \rightarrow x \leq_{\text{dim}} y$ ***(interaction between** Cont **and** $\leq_{\text{dim}}$**)***

(D-D6) $MinDim(x) \leftrightarrow \neg ZEX(x) \land \forall y\,[\neg ZEX(y) \rightarrow x \leq_{\text{dim}} y]$ ***(minimal-dimensional entities)***

(D-D7) $x \prec_{\text{dim}} y \leftrightarrow (\leq_{\text{dim}} y \land \neg(y \leq_{\text{dim}} x) \land \forall z\,[z \leq_{\text{dim}} x \lor y \leq_{\text{dim}} z]$ ***(next-lower dimension)***

(EP-D) $P(x, y) \leftrightarrow Cont(x, y) \land x =_{dim} y$ ***(parthood: equidimensional containment)***

(EPP-D) $PP(x, y) \leftrightarrow P(x, y) \land x \neq y$ ***(proper parthood)***

(ME-D2) $Min(x) \leftrightarrow \neg ZEX(x) \land \forall y\,[\neg PP(y, x)]$ ***(minimal entities within a dimension)***

Contact, $C(x, y)$, as the most general topological relation is definable as x and y sharing some contained object (C-D) and is provably reflexive and symmetric. Specialized types of contact can be distinguished based on the relative dimension: partial overlap $PO(x, y)$ holds only between entity of equal dimension and requires them to share a part (PO-D); incidence $Inc(x, y)$ holds between entities of different dimension and requires a part of the lower-dimensional entity to be shared with the higher-dimensional entity (Inc-D); and superficial contact $SC(x, y)$ requires the shared entity to be of a lower dimension than both of the entities in contact (SC-D).

[10] colore.oor.net/multidim_mereotopology_codi/codi.clif
[11] The relation *Cont* is the qualitative generalization of *SFA*'s *contains* relation.

(C-D) $C(x,y) \leftrightarrow \exists z[Cont(z,x) \wedge Cont(z,y)]$ ***(contact)***

(PO-D) $PO(x,y) \leftrightarrow \exists z[P(z,x) \wedge P(z,y)]$ ***(overlap in a part)***

(Inc-D) $Inc(x,y) \leftrightarrow \exists z[(Cont(z,x) \wedge P(z,y) \wedge z <_{dim} x) \vee (P(z,x) \wedge Cont(z,y) \wedge z \prec_{\text{dim}} y)]$ ***(incidence)***

(SC-D) $SC(x,y) \leftrightarrow \exists z[Cont(z,x) \wedge Cont(z,y)] \wedge \forall z[Cont(z,x) \wedge Cont(z,y) \rightarrow z \prec_{\text{dim}} x \wedge z \prec_{\text{dim}} y]$ ***(superficial contact)***

While CODI does not distinguish different primitive types of entities, they can be defined: *PointRegions* (which encompass individual points and sets of points) are of minimal dimension, *Curves* are of next higher dimension, and so forth [19]. All of these primitive classes specialize the class S of abstract spatial regions.

(PR-D) $PointRegion(x) \leftrightarrow S(x) \wedge MinDim(x) \wedge \neg ZEX(x)$ ***(point sets)***

(Point-D) $Point(x) \leftrightarrow PointRegion(x) \wedge Min(x)$ ***(individual points)***

(Curve-D) $Curve(x) \leftrightarrow S(x) \wedge \forall y[PointRegion(y) \rightarrow y \prec_{\text{dim}} x]$ ***(curves as 1D entities)***

(AR-D) $ArealRegion(x) \leftrightarrow S(x) \wedge \forall y[Curve(y) \rightarrow y \prec_{\text{dim}} x]$ ***(areal regions as 2D entities)***

3.2.2 CODIB

CODIB[12] is a logical extension of CODI that introduces an additional primitive relation of *boundary containment*, $BCont(x,y)$. *BCont* specializes containment by requiring the contained entity to be of a lower dimension than the containing entity (BC-A1), though the contained entity does not need to be of the next-lower dimension. For example, an areal (i.e., 2D) region can contain both curves and points in its boundary. Additional axioms (BC-A2–A5) that constrain the interaction of *BCont* with other relations, including incidence, parthood, partial overlap and containment are not shown here, they are documented in [15]. *BCont* is primitive because it cannot be defined in CODI, that is, in some models of CODI it cannot be determined whether a contained entity is actually contained in the boundary or interior of some containee.

(BC-A1) $BCont(x,y) \rightarrow Cont(x,y) \wedge x \prec_{\text{dim}} y$

3.2.3 Refined Spatial Region Concepts in CODIB

In order to express the SFA concepts in detail, we further refine the basic dimensionally defined classes from CODIB based on whether and how their parts are connected, resulting in the subclass hierarchy shown in Figure 1. A connected region is one that is internally connected (ConS-D), while a region that is not internally connected is called a multipart region (MS-D). The property of *Internal connectedness* (ICon-D) from CODI requires each proper part y of an entity x to be connected to its complement $x - y$ such that the shared entity (denoted by the intersection of y and $x - y$) is of exactly one dimension lower than x[13]. For example, two polygons that share a line segment as boundary are internally connected, but if they only share a point, they are not.

A connected region that contains at least three non-overlapping proper parts that share an entity of lower dimension is called a branched region (BranchedS-D). A simple region is one that is connected and not branched (Simple-D). An atomic region is a simple region without any proper parts (Atomic-D).

[12] `colore.oor.net/multidim_mereotopology_codib/codib.clif`
[13] See [16] for the full axiomatization of the intersection and complement operations in CODI.

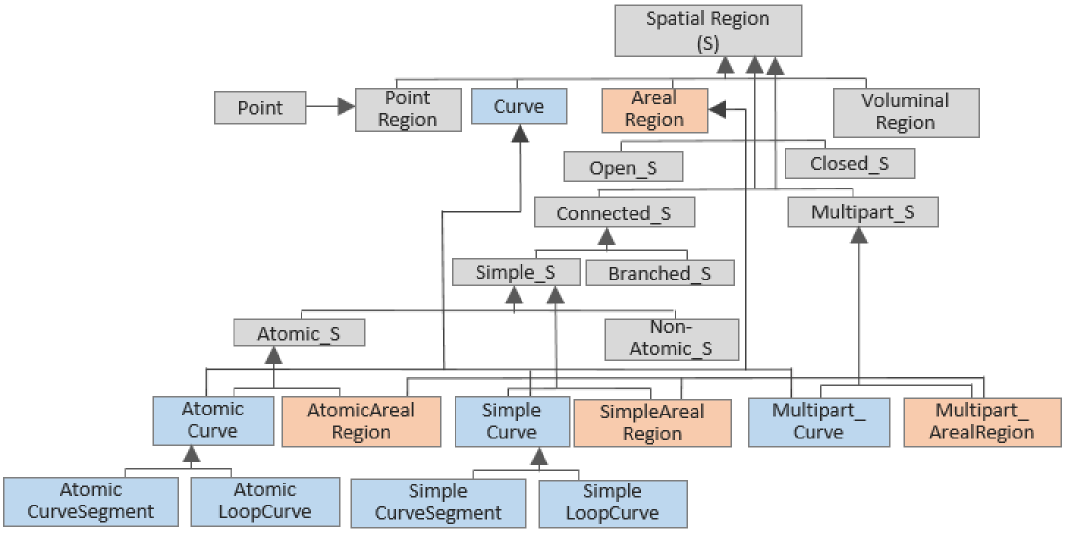

Figure 1 Taxonomy of refined CODIB spatial region concepts classified based on presence/absence of boundaries, connectedness, branching and parts.

(ICon-D) $ICon(x) \leftrightarrow \forall y[PP(y,x) \rightarrow C(y, x-y) \wedge y \cdot (x-y) \prec_{\text{dim}} x]$ ***(internally connected)***

(ConS-D) $Connected\_S(x) \leftrightarrow S(x) \wedge ICon(x)$ ***(connected spatial region)***

(MS-D) $Multipart\_S(x) \leftrightarrow S(x) \wedge \neg Connected\_S(x)$ ***(multipart spatial region)***

(BranchedS-D) $Branched\_S(x) \leftrightarrow Connected\_S(x) \wedge \exists p,q,r,s[PP(p,x) \wedge PP(q,x) \wedge PP(r,x) \wedge \neg PO(p,q) \wedge \neg PO(p,r) \wedge \neg PO(q,r) \wedge s \prec_{\text{dim}} p \wedge s \prec_{\text{dim}} q \wedge s \prec_{\text{dim}} r \wedge Cont(s,p) \wedge Cont(s,q) \wedge Cont(s,r)]$ ***(A branched spatial region is a connected region that has three distinct non-overlapping parts p, q, r that all share a common lower-dimensional entity s. For example, a branched curve has three non-overlapping segments that all share a point.)***

(SimpleS-D) $Simple\_S(x) \leftrightarrow Connected\_S(x) \wedge \neg Branched\_S(x)$ ***(simple spatial region)***

(AtomicS-D) $Atomic\_S(x) \leftrightarrow Simple\_S(x) \wedge Min(x)$ ***(an atomic spatial region is a simple spatial region that is minimal, i.e., has no proper parts)***

These properties are now used to define specialized classes of curves and areal regions.

(SCS-D) $SimpleCurveSegment(x) \leftrightarrow Curve(x) \wedge Simple\_S(x) \wedge \exists p,q[BCont(p,x) \wedge BCont(q,x) \wedge p \neq q]$ ***(Simple curve segment has two distinct end points)***

(SLC-D) $SimpleLoopCurve(x) \leftrightarrow Curve(x) \wedge Simple\_S(x) \wedge \forall y[Point(y) \rightarrow \neg BCont(y,x)]$

(Simple loop curve is closed: it does not contain any point in its boundary)

(ACS-D) $AtomicCurveSegment(x) \leftrightarrow SimpleCurveSegment(x) \wedge Atomic\_S(x)$

(ALC-D) $AtomicLoopCurve(x) \leftrightarrow SimpleLoopCurve(x) \wedge Atomic\_S(x)$

(SAR-D) $SimpleArealRegion(x) \leftrightarrow ArealRegion(x) \wedge Simple\_S(x)$

(MC-D) $Multipart\_Curve(x) \leftrightarrow Curve(x) \wedge Multipart\_S(x)$

(MAR-D) $Multipart\_ArealRegion(x) \leftrightarrow ArealRegion(x) \wedge Multipart\_S(x)$

4 Axiomatization of Simple Feature as Extension of CODIB

In this section we present the core of our formalization that elaborates the semantics of the concepts in the skeleton axiomatization of SFA from Section 3.1 using qualitative concepts and relations from $CODI(B)$. This results in two new ontologies that logically extend SFC-Core and CODIB: SFC-FOL, which includes the more detailed axiomatization of SFA's concepts, and SFR-FOL, which axiomatizes SFA's mereotopological relations. Figure 2 summarizes the taxonomic relationships between the SFA and $CODI(B)$ concepts, but the real contribution are the detailed axiomatic mappings.

4.1 Axiomatization of Simple Feature's Simple Geometric Features

SFA's most general spatial entity is the class *sf_geometry*, which can be mapped to CODI's (and CODIB's) most general class of a *spatial region S* (SFC-M1). *Sf_point* and *sf_surface* map one-to-one to CODI's *Point* and *ArealRegion* (SFC-M2,M6), respectively. CODI's *Curve* captures any kind of one-dimensional features, that may be bounded segments (e.g., a *CurveSegment*), closed (e.g., a *LoopCurve*), infinite (e.g., a ray or a line in the mathematical sense), or branching with more than three endpoints. *Sf_curve* is much more restricted in scope in that it explicitly requires a start and an end point, though the points may coincide as in a closed curve. SFA's definition of *sf_curve* rules out infinite or branching curves. Thus, *sf_curve* maps to the union of *SimpleCurveSegment* and *SimpleLoopCurve* (SFC-M3). SFC-M4 and SFC-M5 elaborate the two cases in more detail. A *sf_curve* that is a *SimpleCurveSegment* has distinct start and end points (SFC-M4), while one that is a *SimpleLoopCurve* has identical[14] start and end points (SFC-M5) and does not contain any points in its boundary (SFC-T1). The axioms SFC-M1 to M6 tie in SFA's simple features with the qualitative spatial ontologies CODIB and allows using CODIB's mereotopological relations in conjunction with SFA features.

(SFC-M1) $sf\_geometry(x) \leftrightarrow S(x)$

(***sf_geometry is equivalent to CODIB's Spatial Region class***)

(SFC-M2) $sf\_point(x) \leftrightarrow Point(x)$ (***sf_point is equivalent to CODIB's Point***)

(SFC-M3) $sf\_curve(x) \leftrightarrow SimpleCurveSegment(x) \vee SimpleLoopCurve(x)$

(***an sf_curve is either a SimpleCurveSegment or SimpleLoopCurve in CODIB***)

(SFC-M4) $sf\_curve(x) \wedge SimpleCurveSegment(x) \rightarrow \exists p1, p2[sf\_point(p1) \wedge sf\_point(p2) \wedge sf\_start\_point(p1, x) \wedge sf\_end\_point(p2, x) \wedge BCont(p1, x) \wedge BCont(p2, x) \wedge p1 \neq p2]$

(***an sf_curve that is a simple curve segment has distinct start and end points that are boundary contained***)

(SFC-M5) $sf\_curve(x) \wedge SimpleLoopCurve(x) \rightarrow \big[\exists p1, p2[sf\_point(p1) \wedge sf\_point(p2) \wedge sf\_start\_point(y, x) \wedge sf\_end\_point(z, x)]\big] \rightarrow y = z$

(***an sf_curve that is a simple loop curve has the same start and end point***)

(SFC-T1) $sf\_curve(x) \wedge SimpleLoopCurve(x) \rightarrow \neg\exists y[sf\_point(y) \wedge BCont(y, x)]$

(***an sf_curve that is a loop curve does not contain any point in its boundary***)

(SFC-T2) $sf\_curve(x) \rightarrow \forall y[PP(y, x) \wedge Min(y) \rightarrow AtomicCurveSegment(y)]$

(***any sf_curve has AtomicCurveSegments as only minimal parts***)

(SFC-M6) $sf\_surface(x) \leftrightarrow ArealRegion(x)$

(***sf_surface is equivalent to CODIB's ArealRegion***)

[14] Note that in CODIB, two points are identical if they are co-located.

The SFA concepts at the next, more refined level of the hierarchy in Figure 2 use CODIB's distinctions between (1) open and closed, (2) atomic, simple (atomic or not), and branched. For example, the SFA concept *sf_line_string* refines the union of CODIB's *SimpleCurveSegment* and *SimpleLoopCurve* and *sf_line* refine *AtomicCurveSegment*, respectively (SFC-T3,M7). The only added constraints are that each segment is a linear approximation between two points – a fact that cannot be expressed within a qualitative representation of space. *Sf_linear_ring* is a *sf_line_string* that is closed and thus a *SimpleLoopCurve* (SFC-M8).

(SFC-T3) $sf\_line\_string(x) \rightarrow SimpleCurveSegment(x) \lor SimpleLoopCurve(x)$

(***from SFC-A1, SFC-M3***)

(SFC-M7) $sf\_line(x) \rightarrow AtomicCurveSegment(x)$

(***sf_line specializes CODIB's AtomicCurveSegment***)

(SFC-M8) $sf\_linear\_ring(x) \rightarrow SimpleLoopCurve(x)$

(***sf_linear_ring specializes CODIB's SimpleLoopCurve***)

Sf_polygons are simple areal regions with a single exterior boundary and with each boundary piece being a *sf_linear_ring* (SFC-M9). A *sf_polyhedral_surface* is a simple areal region formed by "stitching" together *sf_polygons* along their common boundaries (SFC-M10). Such surfaces in a 3-dimensional space may not be planar as a whole. An *sf_triangle* is a *sf_polygon* (SFC-M11) with exactly three non-overlapping lines forming their boundary. The exterior boundary defines the "top" of the surface which is the side of the surface from which the exterior boundary appears to traverse the boundary in a counter clockwise direction. The interior boundary will have the opposite orientation, and appear as clockwise when viewed from the "top". *Sf_tin* is a *sf_polyhedral_surface* whose minimal parts are *sf_triangles* (SFC-M12).

(SFC-M9) $sf\_polygon(x) \rightarrow SimpleArealRegion(x) \land \exists y,z[sf\_linear\_ring(y) \land BCont(y,x) \land boundary(z) = y \land P(x,z)] \land \forall v\big[BCont(v,x) \rightarrow \exists w[P(v,w) \land BCont(w,x) \land sf\_linear\_ring(w)]\big]$

(***sf_polygon specializes CODIB's SimpleArealRegion such that some linear ring in its boundary bounds a region z of which x is part. This accommodates polygons with and without holes. For polygons with holes, some linear ring describes the polygons "outer boundary", whereas for polygons without holes $z = x$ can be chosen such that z is the entire boundary of x. The second condition expresses that every entity v in the boundary of x must be part of some linear ring that that describes a continuous piece of internal or external boundary of x's entire boundary.***)

(SFC-M10) $sf\_polyhedral\_surface(x) \leftrightarrow SimpleArealRegion(x) \land ICon(x) \land \forall y[P(y,x) \land Min(y) \rightarrow sf\_polygon(y)]$ (***sf_polyhedral_surface is equivalent to CODIBs SimpleArealRegion that is internally-connected and is an aggregation of sf_polygons***)

(SFC-M11) $sf\_triangle(x) \leftrightarrow sf\_polygon(x) \land \exists p,q,r[\neg PO(p,q) \land \neg PO(p,r) \land \neg PO(q,r) \land sf\_line(p) \land sf\_line(q) \land sf\_line(r) \land BCont(p,x) \land BCont(q,x) \land BCont(r,x) \land \forall s(sf\_line(s) \land BCont(s,x) \rightarrow s = p \lor s = q \lor s = r)]$

(***sf_triangle is a sf_polygon with exactly three non-overlapping lines bounding it***)

(SFC-M12) $sf\_tin(x) \leftrightarrow sf\_polyhedral\_surface(x) \land \forall y[Min(y) \land PP(y,x) \rightarrow sf\_triangle(y)]$

(***sf_tin is a polyhedral surface consisting only of sf_triangles as minimal parts***)

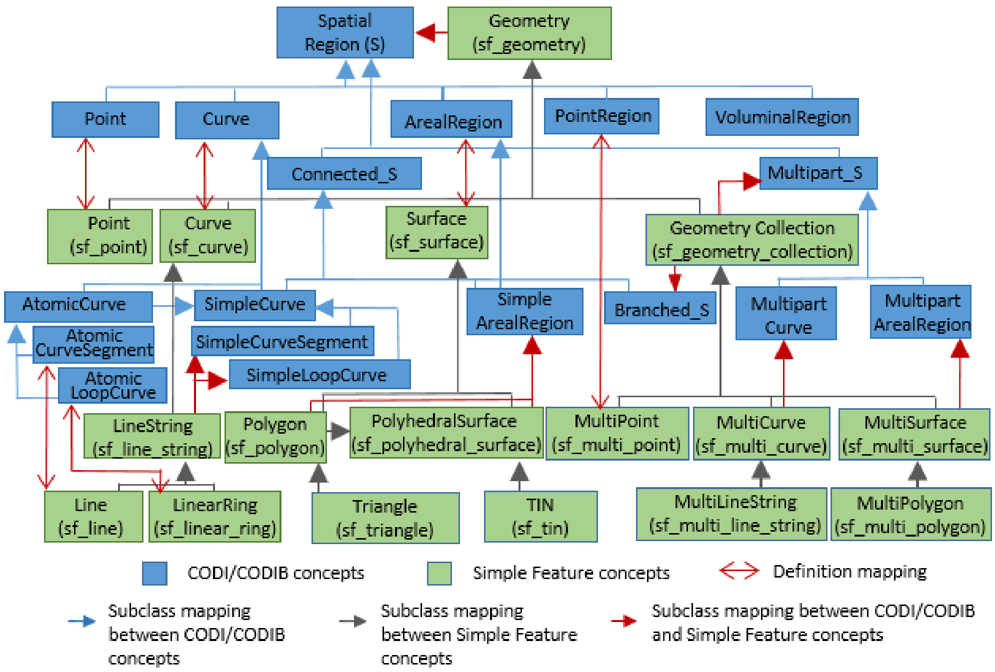

Figure 2 Hierarchy of SF-FOL indicating subclass relationships among SFA concepts, among CODI and CODIB concepts and between SFA and CODI(B) concepts.

4.2 Axiomatization of Simple Feature's Simple Feature Collections

Sf_geometry_collection includes all multipart or branched spatial regions (SFC-M13). Its subclasses map to CODIB's *PointRegion* (SFC-M14) or refine its *Multipart_Curve* (SFC-M15) or *Multipart_ArealRegion* (SFC-M16), respectively, which exhaustively classify *sf_geometry_collection* (SFC-T4). These mappings are not one-to-one because unlike the corresponding CODIB concepts, the SFA concepts restrict how the components can be spatially configured. For example, SFA does not include "branching", non-planar constructions consisting of multiple 2D regions (e.g., three 2D regions meeting in a single line segment) or non-planar arrangements of points. *Sf_multi_line_string* and *sf_multi_polygon* refine *sf_multi_curve* and *sf_multi_surface* (SFC-M17,M18) in that they are constituted only from line strings (i.e., linearly approximated curves) and polygons (i.e., surfaces with linear approximated boundaries).

(SFC-M13) $sf\_geometry\_collection(x) \rightarrow Multipart\_S(x) \vee Branched\_S(x)$

(***sf_geometry_collection specializes CODIB's multipart or branched spatial region***)

(SFC-M14) $sf\_multi\_point(x) \rightarrow PointRegion(x)$

(SFC-M15) $sf\_multi\_curve(x) \rightarrow Multipart\_Curve(x)$

(SFC-M16) $sf\_multi\_surface(x) \rightarrow Multipart\_ArealRegion(x)$

(SFC-T4) $sf\_geometry\_collection(x) \rightarrow PointRegion(x) \vee Multipart\_Curve(x) \vee Multipart\_ArealRegion(x)$ (***SFA's geometry collection is either a PointRegion, Multipart_Curve or Multipart_ArealRegion***)

(SFC-M17) $sf\_multi\_line\_string(x) \leftrightarrow sf\_multi\_curve(x) \wedge \forall y[P(y,x) \wedge Min(y) \rightarrow sf\_line\_string(y)]$

(***sf_multilinestring is a sf_multicurve with minimal parts that are sf_line_strings***)

(SFC-M18) $sf\_multi\_polygon(x) \leftrightarrow sf\_multi\_surface(x) \wedge \forall y[P(y,x) \wedge Min(y) \rightarrow sf\_polygon(y)]$

(***sf_multipolygon is a sf_multisurface with minimal parts that are sf_polygons***)

Table 1 SFA's mereotopological relations, their equivalent *Egenhofer* relations, and the developed mappings to CODIB's relations. The relations in the bottom part are all defined in terms of the top five relations.

SFA	9IM	Definition in terms of CODIB relations and additional theorems
disjoint	disjoint	**(SFR-M1)** $sf\_disjoint(x,y) \leftrightarrow S(x) \wedge S(y) \wedge \neg C(x,y)$
touches	meet	**(SFR-M2)** $sf\_touches(x,y) \leftrightarrow S(x) \wedge S(y) \wedge \forall z[Cont(z,x) \wedge Cont(z,y) \rightarrow BCont(z,x) \wedge BCont(z,y)]$ **(SFR-T1)** $sf\_touches(x,y) \rightarrow SC(x,y)$ **(SFR-T2)** $sf\_touches(x,y) \rightarrow sf\_point(x) \wedge \neg sf\_point(y)$
crosses	-	**(SFR-M3)** $sf\_crosses(x,y) \leftrightarrow S(x) \wedge S(y) \wedge \big[[Inc(x,y) \wedge \neg Cont(x,y) \wedge \neg Cont(y,x)] \vee \forall z[Cont(z,x) \wedge Cont(z,y) \rightarrow Curve(x) \wedge Curve(y) \wedge (z <_{\dim} x \wedge z <_{\dim} y \wedge \neg BCont(z,x) \wedge \neg BCont(z,y))]\big]$ **(SFR-T3)** $x <_{\dim} y \wedge sf\_crosses(x,y) \rightarrow Inc(x,y) \wedge \neg Cont(x,y)$ **(SFR-T4)** $x =_{\dim} y \wedge sf\_crosses(x,y) \rightarrow SC(x,y)$ **(SFR-T5)** $sf\_crosses(x,y) \wedge sf\_curve(x) \wedge sf\_curve(y) \rightarrow SC(x,y)$
overlaps	overlap	**(SFR-M4)** $sf\_overlaps(x,y) \leftrightarrow S(x) \wedge S(y) \wedge PO(x,y) \wedge \neg P(x,y) \wedge \neg P(y,x)$
contains	contains ∨ covers	**(SFR-M5)** $sf\_contains(x,y) \leftrightarrow S(x) \wedge S(y) \wedge Cont(x,y)$
within	inside ∨ coveredBy	**(SFR-M6)** $sf\_within(x,y) \leftrightarrow sf\_contains(y,x)$
equals	equal	**(SFR-M7)** $sf\_equals(x,y) \leftrightarrow sf\_contains(x,y) \wedge sf\_within(x,y)$
intersects	¬ disjoint	**(SFR-M8)** $sf\_intersects(x,y) \leftrightarrow \neg sf\_disjoint(x,y)$ **(SFR-T6)** $sf\_intersects(x,y) \leftrightarrow sf\_touches(x,y) \vee sf\_crosses(x,y) \vee sf\_overlaps(x,y) \vee sf\_contains(x,y) \vee sf\_within(x,y)$ **(SFR-T7)** $sf\_intersects(x,y) \leftrightarrow S(x) \wedge S(y) \wedge C(x,y)$
relate (any)	-	**(SFR-M9)** $sf\_relate(x,y) \rightarrow sf\_intersects(x,y) \vee sf\_disjoint(x,y))$ **(SFR-T8)** $sf\_intersects(x,y) \leftrightarrow S(x) \wedge S(y)$

The axioms of SFC-Core together with the mappings SFC-M1 to SFC-M18 form the ontology SFC-FOL[15]. The theorems SFC-T1 to SFC-T4 can be proved from SFC-FOL.

4.3 Axiomatization of Simple Feature's Qualitative Spatial Relations

So far we have focused on elaborating the semantics of SFA's feature types using CODIB. But SFA's mereotopological relation can, likewise, be expressed using CODIB's relations as summarized in Table 1, similar to the mapping between the DE-I9 relations and *CODIS* [18]. All SFA relations, except for *sf_disjoint*, are specializations of contact (C). *Sf_disjoint* is the negation of contact (SFR-M1), which places no dimensional restriction on the involved entities. The relation *sf_touches* relates two connected features who share parts of their boundaries (i.e., $\partial x \cap \partial y \neq \emptyset$) but no parts of their interiors ($x^\circ \cap y^\circ = \emptyset$). This specializes CODIB's superficial contact relation SC that holds for objects that are in contact but do not share a part of either object. But SC is not sufficient as it allows the lower-dimensional entity to share part of its interior with the higher-dimensional entity (e.g., a curve segment tangential

[15] Available from `https://colore.oor.net/simple_features`.

to a region). Instead, *sf_touches* needs to express that any shared entities are boundary contained in both of the participating entities (SFR-M2). Then, *SC* becomes provable from it (SFR-T1). From the definition of *SC* it can further be inferred that *sf_touches* applies to entities of any dimension except between two points (SFR-T2).

Sf_crosses is a specialization of one of two of CODIB's relation: (1) incidence *Inc* for two entities of different dimension, where a part of the lower-dimensional entity is contained in the higher-dimensional one (e.g., a curve being incident with a polygon by a segment of the curve being contained in the polygon), or (2) superficial contact *SC* for two entities of equal dimension that share only a lower-dimensional entity (e.g., two curves intersecting in a point) (SFR-M3).

Sf_overlaps is a stronger contact relation that only applies to two equidimensional entities and is equivalent to CODIB's partial overlap *PO* when neither entities is a part of the other (SFR-M4). Full containment of an entity inside another entity of the same spatial dimension is represented in CODI by its primitive containment relation, which maps to *sf_contains* (SFA-M5) and to *sf_within* for its inverse (SFR-M6). The special case of spatial equality is captured by *sf_equals* (SFR-M7). *sf_intersects* is the negation of *sf_disjoint* (SFR-M8), which means it generalizes *sf_touches*, *sf_crosses*, *sf_overlaps*, *sf_contains*, *sf_within*, and, indirectly, *sf_equals* (SFR-T6) and is logically equivalent to CODIB's contact relation (SFR-T7). *sf_relate* describes any of SFA's mereotopological relations (SFR-M9), which maps to any pair of spatial entities in CODIB no matter how they are spatially related (SFR-T8).

The axioms of SFC-Core together with the mappings SFR-M1 to SFR-M9 form the ontology SFR-FOL[16]. The theorems SFR-T1 to SFR-T8 can be proved from SFR-FOL.

4.4 Logical Verification

Our primary tool for evaluating the developed first-order ontology SF-FOL are different variants of consistency checking summarized in Table 2. In its simplest form, consistency checking verifies that an ontology is free of internal contradiction. This typically involves constructing some small finite model using a finite model finder. A known problem with this approach is that it aims to construct the smallest models, which are often trivial in the sense that the extension of many classes and relations therein are empty or universal. For example, one trivial model for CODIB consists of a set of isolated points, but without any curves or areal regions. Moreover, most of the CODIB relations, such as *BCont*, *SC*, or *Inc*, may not be used at all in a trivial model whereas other relations, such as *Cont* or *P*, may relate objects only to themselves. Such a model does not prove that all classes may indeed be instantiated (i.e., some curve, areal region, or more specialized defined subclasses such as a branched curve) and all relation may apply to pairs of distinct entities. One can force the creation of non-trivial models by adding existential axioms of the form $\exists x P(x)$ and $\exists x, y[R(x, y) \wedge x \neq y]$ to the theory. This approach has been implemented in the Macleod suite of tools[17] and previously been utilized to prove CODI's and CODIB's nontrivial consistency with the help of the finite model finder Paradox3 [4]. Here, the same approach is used to prove SF-FOL's nontrivial consistency.

An additional way to verify an ontology is to prove its consistency with some sample datasets. Rather than constructing an arbitrary model that satisfies certain constraints, this external verification ensures that the ontology is actually consistent with the kind of

[16] Available from `https://colore.oor.net/simple_features`.
[17] `https://github.com/thahmann/macleod`

Table 2 Overview of the employed consistency checking methods for verification of the developed first-order logic ontology.

Type	Task	Description
Internal verification	Consistency checking	Ascertains the ontology is free of internal contradictions
	Non-trivial consistency checking	Ascertains that a model exists that instantiates each class and each relation positively and negatively by pairs of distinct objects
External verification	Consistency checking with data	Ascertains that the ontology is consistent with a set of assertions describing a dataset

model encountered in the domain. This has not been done previously for CODI or CODIB as real-world purely qualitative information is hard to come by. However, by mapping SFA concepts to CODIB as qualitative generalization thereof, we can now exploit the abundance of geometric data already stored in GIS or geospatial databases.

In this work SF-FOL is verified internally, nontrivially and externally with Paradox3. Proving nontrivial consistency of SF-FOL ensures that instantiation of all the axiomatically defined or restricted Simple Feature types and SFA's mereotopological relations is possible and the new mappings and axioms do not contain any contradictions. In addition, we employed small subsets of data, consisting of samples of 20 to 40 geometric features, to externally verify SF-FOL. The data is extracted from publicly hosted shapefiles[18] that includes polygon representations of counties and subdivisions, polyline representations of major roads, and point representations of schools and other civic buildings within the state of Maine. Only the type of geometry and the SFA relations to other, nearby geometries are stored as assertions. The extracted assertions (i.e., the ABox) were added to SF-FOL (i.e. the TBox) and handed to the model finder to construct a model. As an additional step, we encoded sample queries, such as '*What are the areal regions within Penobscot county that intersect I-95?*', which can be expressed logically in CODIB as $ArealRegion(s) \wedge sf\_within(s,' PenobscotCounty') \wedge sf\_intersects(s,' I95')$. This allows retrieving possible instantiations of s, which were manually inspected to identify any unintended models, such as schools being returned as possible solutions, that helped refine the axiomatization.

Generally, the utilized ontology verification techniques are somewhat similar to software testing techniques: they can help identify problematic models of an ontology that require changing or adding axioms but do not prove that the ontology is fully correct. This would require a full representation theorem describing the structure of *all the models* of SF-FOL, which is beyond the scope of this paper. The completeness of SF-FOL is not verified as this would require alternative characterization of all models.

5 Conclusion and Future Work

A core component of many geospatial data models and standards used to store and analyze conventional GIS data are taxonomic classifications of geometric feature types and basic mereotopological relations to support qualitative querying of the geometric data. However, the semantics of the mereotopological relations are not explicitly formalized and thus not

[18] `https://www.maine.gov/megis/catalog/`

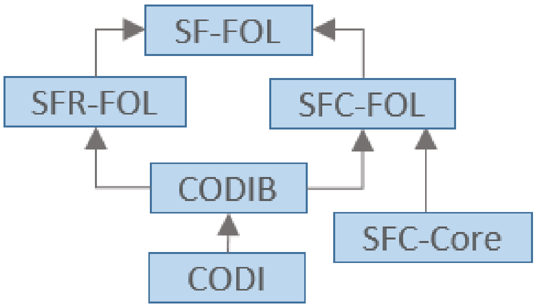

Figure 3 The relationships between the developed and reused axiomatic theories.

accessible for further automated reasoning. Because of this limitation, purely qualitative spatial information, i.e. spatial information that relates objects for which no geometric information is available in the data store, cannot be easily reasoned over in conjunction with existing geometric data. To address this challenge, this paper presents a semantically augmented formalization, SF-FOL, of the basic geometric feature types (axiomatized in SFC-Core) and qualitative spatial relations (axiomatized in SFR-Core) of the Simple Features Access (SFA) standard. This augmented formalization is provided as an extension of the CODIB theory, a qualitative axiomatization of mereotopological space in first-order logic. The relationships between the developed theories is illustrated in Figure 3.

It is shown that all of SFA's geometric features specialize the more general, only dimensionally-constrained, classes of spatial entities from CODIB and its subtheory CODI. The distinctions between "straight line segments" and "curve segments" and, analogously, between "fully bounded regions" and "polygons" are the only ones that are not fully definable in CODIB because they are inherently geometric[19]. But because these distinctions are irrelevant to mereotopological relations, all of CODIB's spatial relations can be evaluated over geometric features in SF-FOL. Likewise, all of *SFA*'s mereotopological relations are fully defined in the SFR-FOL module of SF-FOL and thus can be employed for querying over both geometric and qualitative data.

Future Work: While the mereotopological approach of describing geometric concepts and spatial relations enhances spatial reasoning capabilities, formalization in a language such as first-order logic and relying on general-purpose automated theorem provers and model finders for reasoning comes with the cost of intractability of reasoning. The number of first-order logic (FOL) assertions explodes even when reasoning with a very small spatial dataset. Preliminary experiments with Paradox, one of the best performing FOL model finders, show that reasoning with data against a fairly complex ontology such as CODIB often terminates without success except for the tiniest datasets. In ongoing work, we systematically test how to improve model finding performance by explicitly using the qualitative abstractions and "throwing away" geometric information and by converting data into logically equivalent formats that are less taxing on a model finder.

[19] One cannot distinguish a straight line from a curve without a metric in the space that defines the shortest segment between two points, see the discussion of such issues in [2, 20]

References

1 Ringo Baumann, Frank Loebe, and Heinrich Herre. Towards an Ontology of Space for GFO. In *Conf. on Formal Ontology in Inf. Systems (FOIS-16)*, pages 53–66, 2016.

2 Stefano Borgo and Claudio Masolo. Full mereogeometries. *Rev. Symb. Logic*, 3(4):521–567, 2010.

3 Roberto Casati and Achille C. Varzi. *Parts and Places.* MIT Press, 1999.

4 Koen Claessen and Niklas Sörensson. New techniques that improve MACE-style finite model building. In *Workshop on Model Computation at CADE 2003*, 2003.

5 Eliseo Clementini and Paolino Di Felice. A comparison of methods for representing topological relationships. *Inf. Sci.*, 3(3):149–178, 1995.

6 Eliseo Clementini and Paolino Di Felice. A model for representing topological relationships between complex geometric features in spatial databases. *Inf. Sci.*, 90(1):121–136, 1996.

7 Anthony G. Cohn and Jochen Renz. Qualitative Spatial Representation and Reasoning. In F. van Harmelen, V. Lifschitz, and B. Porter, editors, *Handbook of Knowledge Representation.* Elsevier, 2008.

8 Max J. Egenhofer. Reasoning about binary topological relations. In *Symp. on Large Spatial Databases (SSD'91)*, LNCS 525, pages 141–160. Springer, 1991.

9 Max J. Egenhofer and John Herring. Categorizing binary topological relations between regions, lines, and points in geographic databases. Technical report, Department of Surveying Engineering, Univ. of Maine, 1991.

10 Giorgio De Felice, Paolo Fogliaroni, and Jan Oliver Wallgrün. A Hybrid Geometric-Qualitative Spatial Reasoning System and Its Application in GIS. In *Conf. on Spatial Inf. Theory (COSIT-09)*, 2009.

11 Paolo Fogliaroni, Paul Weiser, and Heidelinde Hobel. Qualitative Spatial Configuration Search. *Spatial Cognition & Computation*, 16(4):272–300, 2016. doi:10.1080/13875868.2016.1203327.

12 Anthony Galton. Taking dimension seriously in qualitative spatial reasoning. In *Europ.Conf. on Artif. Intell. (ECAI-96)*, pages 501–505, 1996.

13 Nicholas M. Gotts. Formalizing commonsense topology: the INCH calculus. In *Int. Symp. on Artif. Intell. and Math.*, pages 72–75, 1996.

14 Michael Gruninger, Carmen Chui, and Megan Katsumi. Upper Ontologies in COLORE. In *Proc. of the Joint Ontology Workshops (JOWO 2017)*, 2017.

15 Torsten Hahmann. *A Reconciliation of Logical Representations of Space: from Multidimensional Mereotopology to Geometry.* PhD thesis, Univ. of Toronto, 2013.

16 Torsten Hahmann. On Decomposition Operations in a Theory of Multidimensional Qualitative Space. In *Int. Conf. on Formal Ontology in Inf. Syst. (FOIS 2018)*, pages 173–186, 2018.

17 Torsten Hahmann and Michael Grüninger. A naïve theory of dimension for qualitative spatial relations. In *Symp. on Logical Formalizations of Commonsense Reasoning (CommonSense 2011).* AAAI Press, 2011.

18 Torsten Hahmann and Michael Grüninger. A Theory of Multidimensional Qualitative Space: Semantic Integration of Spatial Theories that Distinguish Interior from Boundary Contact (Extended Abstract). In *Proc. of the Int. Conference on Spatial Information Theory (COSIT 2011), Belfast, Maine, September 12-16, 2011*, 2011.

19 Torsten Hahmann and Michael Grüninger. Multidimensional mereotopology with betweenness. In *Int. Joint Conf. on Artif. Intell. (IJCAI-11)*, pages 906–911, 2011.

20 Torsten Hahmann and Michael Grüninger. Region-based Theories of Space: Mereotopology and Beyond. In Shyamanta M. Hazarika, editor, *Qualitative Spatio-Temporal Representation and Reasoning: Trends and Future Directions*, pages 1–62. IGI, 2012.

21 J Herring. OpenGIS Implementation Standard for Geographic information - Simple feature access-Part 1: Common architecture, 2011.

22 International Electrotechnical Commission (ISO/IEC). ISO 19125:2004 geographic information – simple feature access, 2004.

23 International Electrotechnical Commission (ISO/IEC). ISO 19136:2007 geographic information – Geography Markup Language (GML), 2007.

24 Zhiguo Long, Matt Duckham, Sanjiang Li, and Steven Schockaert. Indexing large geographic datasets with compact qualitative representation. *International Journal of Geographical Information Science*, 30(6):1072–1094, 2016. doi:10.1080/13658816.2015.1104535.

25 David M. Mark and Max J. Egenhofer. Modeling spatial relations between lines and regions: Combining formal mathematical models and human subjects testing. *Cartogr. Geogr. Inf. Sci.*, 21(3):195–212, 1994.

26 Mark McKenney, Alejandro Pauly, Reasey Praing, and Markus Schneider. Dimension-refined topological predicates. In *Conf. on Advances in Geographic Information Systems (GIS-05)*, pages 240–249. ACM, 2005.

27 Matthew Perry and John Herring. OGC GeoSPARQL – a geographic query language for RDF data, 2012.

28 David A. Randell, Zhan Cui, and Anthony G. Cohn. A spatial logic based on regions and connection. In *KR'92: Principles of Knowledge Representation and Reasoning*, pages 165–176, 1992.

29 Markus Schneider and Thomas Behr. Topological relationships between complex spatial objects. *ACM Trans. Database Systems*, 31(1):39–81, 2006.

30 Barry Smith. Mereotopology: a theory of parts and boundaries. *Data Knowl. Eng.*, 20(3):287–303, 1996.

31 G Sutcliffe. The TPTP Problem Library and Associated Infrastructure. From CNF to TH0, TPTP v6.4.0. *Journal of Automated Reasoning*, 59(4):483–502, 2017.