Classification, Individuation and Demarcation of Forests: Formalising the Multi-Faceted Semantics of Geographic Terms

Lucía Gómez Álvarez[1] **and Brandon Bennett**[2]

1 **School of Computing, University of Leeds, Leeds, UK**
`sc14lga@leeds.ac.uk`
2 **School of Computing, University of Leeds, Leeds, UK**
`B.Bennett@leeds.ac.uk`

Abstract

Many papers have considered the problem of how to define *forest*. However, as we shall illustrate, while most definitions capture some important aspects of what it means to be a forest, they almost invariably omit or are very vague regarding other aspects. In the current paper we address this issue, firstly by providing a definitional framework based on spatial and physical properties, within which one can make explicit the implicit variability of the natural language forest concept in terms of explicit parameters. Our framework explicitly differentiates between the functions of *classification*, *individuation* and *demarcation* that comprise the interpretation of predicative terms. Whereas ontologies have traditionally concentrated predominantly on classification, we argue that in many cases (especially in the case of geographic concepts) criteria for individuation (i.e. establishing how many distinct individual objects of a given type exist) and demarcation (establishing the boundary of an object) require separate attention, involve their own particular definitional issues and are affected by vagueness in different ways. We also describe a prototype Prolog system that illustrates how our framework can be implemented.

1998 ACM Subject Classification I.2.4 Knowledge Representation Formalisms and Methods

Keywords and phrases Forest, Definition, Vagueness, Ontology, GIS

Digital Object Identifier 10.4230/LIPIcs.COSIT.2017.8

1 Introduction

There are many definitions of forest in the literature [29, 9]. While most of these capture some important aspects of what it means to be a forest, they almost invariably omit or only very loosely specify other important requirements. In particular, constraints on the overall size, shape and topology of a forest are often omitted. In the current paper we address this issue, firstly by providing a framework within which one can give formal definitions which combine both spatial and physical properties; and secondly, by making explicit the implicit variability of the natural language forest concept in terms of explicit parameters; and thirdly, by differentiating between the functions of *classification*, *individuation* and *demarcation* that comprise our interpretation of predicative terms. Whereas formal definitions have traditionally concentrated predominantly on classification, we argue that in many cases criteria for individuation (i.e. establishing how many distinct individual objects of a given type exist) and demarcation (establishing the boundary of an object) require separate attention and involve their own particular definitional issues, particularly in the domain of geographical objects. In the current paper we propose the use of a *supervaluation semantics*, which we

13th International Conference on Spatial Information Theory (COSIT 2017).
Editors: Eliseo Clementini, Maureen Donnelly, May Yuan, Christian Kray, Paolo Fogliaroni, and Andrea Ballatore;
Article No. 8; pp. 8:1–8:15
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

suggest provides a natural and flexible framework within which to represent the wide variety of different interpretations of the term ‘forest’.

2 Defining ‘forest’

How much forest is there in the world? This is a surprisingly difficult question to answer [1]. A broad range of forest concepts and definitions (more than 600 were reported in [29]) have been specified for different purposes, thereby leading to very different estimates [19]. Thus, it is not surprising that confusion arises over both global forest extent and its spatial distribution [17, 18].

Traditionally, two main categories of forest definitions have been discussed in the literature: *land cover* and *land use* [28]. While the former defines forest in terms of the ecological layer and the physical characteristics of the land, the latter does it with regard to the purpose to which the land is put to use by humans [32]. Definitions favouring one or another (or both) approaches, together with other relevant features, are linked to different perspectives and management objectives, the most relevant ones being: timber management, conservation of ecosystems, increasing carbon stocks and landscape restoration [9]. For example, definitions used for the analysis of carbon stocks generally focus on land cover, but ignore aspects like connectedness or distinctions between natural or planted forests because they are not relevant to describe the carbon potential. The opposite happens when defining forest for landscape restoration purposes, where together with land use information, they become crucial aspects to understand the effects on ecosystem services and forest-based livelihoods. These differences are linked to scale and disciplinary compartmentalization and, besides responding to the purpose for which the definitions were created [9], they pose limitations on the construction of global knowledge [20] and data interoperability.

Moreover, beyond the semantic ambiguity of the concept itself, reflected in a wide variety of specifications, most definitions of forest (and other geographic features), both in the academic literature and in administrative regulations, are not precise [26, 34] which questions consistency even within a particular research community or monitoring project. This limits understanding of data and may impair management decisions or distort research findings.

Awareness of these issues exists and there is an extensive literature reporting it, both in academia [7, 6, 29] and in policy [25, 14]. However, global agreement on the meaning of such words has not been reached. Part of the community has focused on precisely defining forest together with other natural resource terms. Other research has focused on examining the reasons for definitional problems. Many papers have advocated accommodating a variety of definitions [9] and pointed to consequent challenges, particularly for data integration and multidisciplinary work [20]. The current investigation follows these lines and proposes a framework that enables the coexistence, analysis and comparison of different precise definitions of forest.

3 Forest definitions from a logical and ontological point of view

Concepts and definitions of forests and other geographical features have received wide attention in the domain of Knowledge Representation, Geographical Information Science and Philosophy, and the pervasive vagueness that affects them has been highlighted [13, 15]. In this section we first briefly discuss different aspects of vagueness. We then introduce supervaluation semantics and discuss its ability to express the variety of possible meanings of concepts, which will provide us with a flexible framework for implementing forest definitions.

Finally, we analyse how the research on ontology and, more specifically, on geographical ontology, can contribute to the discussion on appropriate forest definitions, providing insight into aspects of geographical features which are often overlooked on most characterisations.

3.1 Vagueness and logics of vagueness for the study

The topic of vagueness has a long philosophical history, dating back nearly two and a half thousand years [38]. Although not often explicitly mentioned in the literature, it is important for our approach to make the distinction proposed in [5] between *sorites* vagueness and *conceptual* vagueness:

Sorites[1] *vagueness* occurs when the applicability of a predicate depends on specific measurable parameters but their thresholds are undetermined, thus creating borderline cases. For example, if we define forest as a large expanse densely populated by trees, there is no exact specification of exactly how many trees it must contain. In fact, giving a precise definition seems to be contrary to the way that the term 'forest' is used. It would seem very odd to claim that a particular number of trees is insufficient to form a forest, whereas if there were one more we would have a forest, as described by *Fisher* in [15].

Another kind of vagueness arises when there is a lack of clarity on which attributes or conditions are essential to the meaning of a given term, so that it is controversial how it should be defined. Thus, there is indeterminism regarding to which property or logical combination of properties is relevant to determining whether a concept is applicable. We call this *conceptual vagueness.* It is akin to ambiguity in that it occurs where a term has multiple meanings that are qualitatively distinct, except that with ordinary ambiguity the different meanings apply to completely different objects or situations, whereas with conceptual vagueness there is considerable overlap between the sets of cases to which the different meanings apply. This kind of vagueness underlies the controversy about whether to define forest in terms of land cover or land use, and this case is a good illustration of the overlapping of the applicability of the two interpretations, since the use to which land can be put depends to a large extend on the material and ecological properties of its land cover.

In order to provide formalisations of vague predicates and vague concepts we follow the lines of [3] by taking a supervaluationistic approach. Supervaluation semantics is based on the idea that a vague language can be interpreted in many different precise ways, each of which can be logically conceptualised in a precisification, which determines precise truth conditions for each predicate of the language. By incorporating a set of many possible precisifications into the semantics, a supervaluation model can accommodate major differences between perspectives of multiple disciplines and management objectives as well as more minor variations in the use of vague terms. Although the truth of propositions describing properties of a situation will typically vary from precisification to precisification, there will be some propositions that are true within the whole set of what are considered to be reasonable (a.k.a. *admissible*) propositions. Such propositions may be called *supertrue* and constitute a common consensus of accepted facts.

3.2 From an Ontological point of view

The value of ontologies in developing advanced information systems is now well established [21]. A formal ontology can provide both a solid conceptual layer to clarify the forest

---

[1] Referred to a puzzle known as The Heap: Would you describe a single grain of wheat as a heap? No. Two grains? No. ... You must admit the presence of a heap sooner or later, so where do you draw the line? (Stanford Dictionary of Philosophy).

definitions and computational support for the integration of information coming from remote sensing and traditional survey data. More importantly, it gives us a framework within which we can systematically compare formal forest definitions and their spatial projections by making implicit assumptions as explicit as possible.

One of the main advantages of ontologies is that they improve the interoperability of the information systems that use them, acting to enforce a consensus view reached by a community regarding a certain domain. By formalising the semantics of terminology, ontologies provide a well defined framework for the interchange of information between different systems. Moreover, logic-based ontologies support automated reasoning (the ability to infer logical consequences) over the formal definitions and axioms of the domain. These advantages have been reported for ontologies in information systems [31, 39] in general and in GIS [16, 27] in particular.

In this paper we suggest that, by enriching a Forest Ontology with supervaluation semantics, we provide the basis for a flexible, adaptive and unbiased representation of vague terms, such as *Forest* or *Tree*. This brings the opportunity of moving from the need for a (more or less) consensual and precise but artificial definition of a vague term like *Forest* to a more realistic formalisation capturing a whole set of views coming from multiple disciplines and stakeholders. This is, overall, a first step towards a framework explicitly supporting the coexistence of different forest definitions within information systems. Moreover, the use of supervaluation semantics, which largely preserves the inferential capabilities of first order logic, should facilitate reasoning both within particular precise interpretations and with regard to comparisons between different interpretations.

Moreover, in order to provide a sound characterisation of forests we can learn from the literature on formal ontology in general and in the geographic domain in particular. Given the complexity of the geographic space, its special characteristics [12] and the variety of 'things' it can include [33, 36], much discussion has been raised when trying to answer concisely the question of what is a geographic concept [35]. Some of the most interesting characteristics pointed by ontologists and geographic information scientists include location, topology, boundaries and mereology. It is surprising thus that when defining forest, little attention has been paid to those aspects.

Among the most challenging issues affecting forest definitions are the following:

- *Existence of vague objects.* There is an ongoing philosophical discussion on whether the vagueness exhibited by geographic names and descriptions is merely linguistic or ultimately ontological, that is, its terms are vague because they refer to vague objects, objects with fuzzy boundaries [37]. Beyond the philosophical discussion, fuzzy and crisp representations of forest have different uses both in academia and research.
- *The dichotomy of the object-field representations.* In order to represent geographic phenomena, ontologies have to encapsulate not only the meanings linked to specific concepts but also the way these meanings are handled and represented [2]. Thus, a precisification of a concept such as forest must embed information about its mode of specification, typically either in terms of an object model or a field.
- *Scale and granularity.* A conceptualization of geographic space may have several levels of granularity, each of which will be appropriate for problem solving at different levels of detail [13].
- *Endurants and perdurants.* Endurants are entities that persist through time and are regarded as wholly present at each moment of their existence (matter and objects), while perdurants are never fully present at any one given moment in time, but instead 'unfold' themselves in successive phases or temporal parts (processes and events) [2]. Although

geographic entities are usually described as endurants, they may also be considered as processess. This is especially the case for forests, which undergo continual change, as trees grow and die.

- *Individuation criteria.* How are entities such as mountains, rivers and forests individuated within a landscape? Although the possession of a boundary is one mark of individuality, in the geographical domain boundaries give rise to a number of ontological conundrums and may themselves be difficult to individuate [8].
- *Topology, Mereology and Location.* In the geographic space, topology is considered to be first-class information, whereas metric properties, such as distances and shapes, are used as refinements that are frequently less exactly captured [13]. A general theory of spatial location is necessary to relate an entity with the spatial region that it occupies and, finally, topology is crucial as mereology alone cannot account for some very basic spatial relations, such as the relationship of continuity between two adjacent objects or the relation of one thing being entirely inside or surrounding some other thing [8].
- *Identity over time.* A simple model of composite objects as mathematical sets (where sets are identical if and only if they have exactly the same elements) does not account for more complex unity and identity criteria allowing one to accept the continued existence of geographic objects even after the loss of or gain certain parts [8].

4 A framework for the formalisation of forest definitions

In order to provide a precise account of what a forest is, many aspects of its nature must be specified. We propose a framework that compiles a collection of relevant features to characterise a forest, which tries to unify different perspectives. The framework, shown in Table 1, is by no means exhaustive.

Moreover, we divide these aspects into three main groups, using the notions of classification, individuation and demarcation. We suggest that there are a large number of natural language predicative words and phrases for which a reasonably clear distinction can be made between some or all of these different aspects, and each has different ways of being made precise. Thus, we review the general notion of individuality as it has been discussed in Philosophy, complemented by the notions of identity and unity [22, 23] and by certain existential conditions. We discuss the difference between *individuation* and *classification*, and how this affects information modelling for in GIS and Forest monitoring. We propose a framework that clearly differentiates these aspects and also distinguishes the aspect of *demarcation*, which is particularly relevant for GIS. We suggest that the use of these notions as guiding tools for building ontology-driven GIS can provide clarity on the ontological commitments assumed in research and forest monitoring.

We finally show that in considering precisification of a vague predicate, these aspects are often seperable, in that it may be possible to make one or two of the aspects precise without committing to a precisification of the other aspects(s). For example we might individuate and classify a forest entity without demarcating a precise boundary; or we might demarcate a region of vegetation land cover without committing to whether it should be classified as a forest. We argue that such distinctions are necessary to clarify the ontological commitment made when using vague concepts in research.

5 Main purposes of forest definitions

The features included in the framework in Table 1 are grouped in three main categories, namely classification, individuation and demarcation. These categories refer to three main

Table 1 Compilation of forest definitions from different sources.

Aspects of forest concept definitions	
1. Classification	
1.1. Qualitative characteristics (global to the whole object)	A typical example would be the land use
1.2 Presence (or absence) of features	E.g. roads, trees of more than 5m, shrubs, ...
1.3 Density, uniformity and scale of features	Measures like the canopy cover should be measured not only in terms of the density but also the uniformity and or scale, given that the predicate can be applied to regions with different extensions.
1.4 Location restrictions	Some definitions are contextualised in one area, like tropical forest
2. Individuation	
2.1 Morphological restrictions.	Such as shape or minimal area
2.2 Metrical restrictions	We may want to evaluate the proximity of constituents
2.3 Topological restrictions	Is the forest necessarily self-connected? Does the forest have holes?
2.4 Mereological restrictions	Is the forest the same forest (whole) if it loses one part?
2.5 Rough location	Part of the identity of the object is linked to its geographical rough position
3. Demarcation	
3.1 Fine grained threshold	Determines the precise boundary of the forest
3.2 Fuzzy threshold?	We may allow for fuzzy boundaries

purposes of definition identified as particularly relevant for the field of naive geography. Moreover, our understanding of *classification* and *individuation* match with the general notions used to evaluate and validate ontologies suggested in [23], and the additional category of *demarcation* is specific to the geographical domain of the current research.

5.1 Questions from a naive geographer

We consider three simple questions using the term forest and the different aspects of its meaning to which they relate. These questions are used as guidelines to link cognitive conceptualizations of the geographical space with relevant notions in the domain of philosophy and ontology, as well as with actual research questions around the topic of global forest monitoring.

(a) How much forest is there?
(b) How many forests are there in this region?
(c) What area is occupied by *this* forest?

Question (a) should be interpreted as *How much forestland is there?*, where the mass noun ‘forestland’ is typically interpreted in terms of a field conceptualisation of the geographical space. A possible answer to that question can be found in [24], where the global land-area is organised as a grid of land-pieces which are systematically classified in terms of the canopy cover (it must be stressed that the classified entities in this case are not forests but land parcels). A relevant precisification of ‘forest’ or ‘forestland’ for this study is exclusively concerned with fixing the characteristics that a piece of land needs to display in order to satisfy this classification.

Question (b), however, requires the *individuation* of forests in order to be able to count them, thus taking an object model approach. Characterising individuation criteria of objects is hard and, as seen in Table 1, relies strongly on spatial (and also temporal) factors. As we can see in [9], identity criteria are necessary in order to characterise forest for many management objectives. For example, for landscape restoration purposes and conservation of natural ecosystems, it is important to understand the dynamics of individual forests, and to track whether they merge, split, appear or disappear, even if the global amount of forestland is preserved.

Question (c) asks for the demarcation (or extension) of *a forest.* To give a precise answer, appropriate thresholds and footprint algorithms need to be selected. Demarcation criteria is particularly relevant for assessing forest loss and gain among others. It is often presumed that forests need to be demarcated in order to answer question (b). However, as will be shown in section 5.4, this need not necessarily be the case: we may be able to differentiate (and hence count) forests, without fully establishing their exact boundaries.

The ability of an intelligent agent to interpret these questions and select the relevant aspects of forest definitions, relies on a capacity to understand that geographic terms may present different aspects of a complex multi-faceted semantics in different contexts. In other words, it requires certain understanding of a naive geography.

5.2 Classification

Almost all predicates, and certainly all the notions of forest, incorporate some kind of classification. In this paper we consider that an object x is classified under a predicate ϕ if it satisfies the necessary and sufficient conditions that govern ϕ's applicability. This can be formalised as

$$\forall x[\phi(x) \rightarrow \Phi(x)] \text{ and } \forall x[\Psi(x) \rightarrow \phi(x)] (\text{where } \phi \text{ does not occur in either } \Phi \text{ or } \Psi)\,.$$

Here we understand that the predicate ϕ does not carry identity or unity criteria on its own. Rather, any object denoted by x to which ϕ may be applied is a member of a domain of individuals that has already been fixed (either by direct stipulation of the domain or by axioms involving other predicates). Following the nomenclature in [23] we express it as ϕ**-O-U**. Moreover, classificational predicates are either semi-rigid -R or anti-rigid ~R[2].

Common examples of classification tasks in the geographic domain include both the assignation of a category to an already individuated geographical object, such as classifying a particular forest *forest#23* into a forest type *tropicalForest*(*forest#23*) or a tree *tree#5* into its species *oak*(*tree#5*), and the assignation of a category to a portion of a mass term, typically a certain region or land area *landpiece*#8, for example into *forestland*(*landpiece#8*). The latter, which focuses on the properties that characterise whether the concept *foresthood* is applicable to a given land parcel, is certainly the one that has received more attention within the Forestry literature. This kind of characterisation does not incorporate any specification of individuality; and it seems this is not required to answer the naive question (a). It assumes that an appropriate division of land into parcels has already been made (e.g. as raster cells) and characterises 'forest' or 'forestland' as a mass term. Thus, the predicate is not concerned with forest *objects* as such and does not provide a means to answer the question of how many

---

[2] The notion of rigidity is introduced in [23]: 'A property is rigid if it is essential to all its possible instances. There are also properties that are essential to some entities and not essential to others (semi-rigid), and properties that are not essential to all their instances (anti-rigid)'

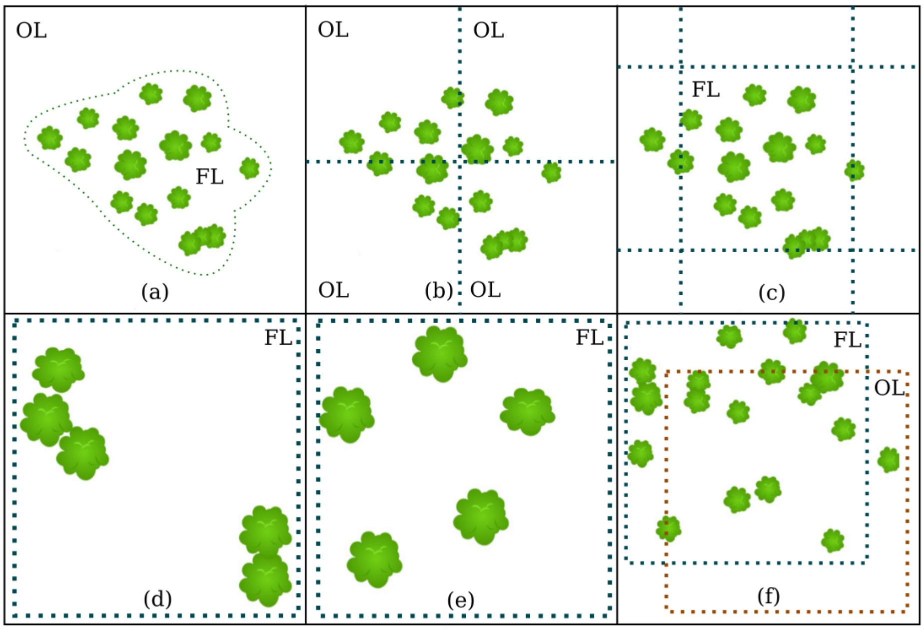


Figure 1 Different land parcels (dashed lines) into forestland (FL) or otherland (OL) following FAO's requirement of a 10% of canopy cover.

forests there are. Moreover, although ϕ can be used to determine the total area of forestland over the entire domain under consideration, it does not determine the extension of individual forests, and, in some cases, the total area of forestland will be different from the total area of the forests contained, for example when some parcels of forestland are isolated from any major forest.

5.2.1 Classificational characterisation of forests

In order to provide a background for any specification of the classification criteria for a precisification of *forestland*, some basic characteristics are discussed below, especially those relating to the points listed in Table 1.

Consider the task of classification applied over portions of the mass concept of land. In the framework proposed in Table 1 we consider certain aspects of the classification that tend to be overlooked even in attempts to provide concise definitions, such as [14]. Following the framework, a precisification of forestland combines, first, (1.1) qualitative attributes, such as the legal land use of the area. Then, (1.2) the presence of certain features, such as the trees and the absence of other elements such as roads or buildings (these classificatory features may of course make reference to other kinds of object or land cover defined in the ontology, which in turn may be also subject to issues of vagueness and of finding an appropriate individuation). Following, (1.3) the density, uniformity and scale of features. As illustrated in Figure 1 and described below, differences in scale and uniformity may result in substantial variations in the meaning of the categories. Finally, (1.4) some location restrictions (e.g. the area must be within the tropics for *tropical-forestland*(x)) may be added in order to improve contextual adaptation.

Figure 1 shows some classifications of pieces of land in terms of their canopy cover, set to the 10% as required in the FAO definition of forest [14]. Figures (a), (b) and (c) display the same piece of land. (a) shows a common sense demarcation and, while (b) and (c) classify that same land using a grid of the same size, both result in different information: (b) is fully covered by 'otherland' (OL) while (c) is mostly covered by 'forestland' (FL). Similarly, two pieces that greatly overlap can be classified as different land types, as in (f). A change in scale such as from (c) to (e) implies a gain on precision but also a variation on the meaning of the 'forestland'. In an extreme scenario, a single tree could constitute 'forestland', which could be misleading. Finally, uniformity issues can arise when the density is concentrated in clusters (d) instead of being evenly distributed across the piece of land (e). This could lead to the confusion of open forests with small and isolated clusters of trees in grasslands.

While any of the previous observations challenges the application of these techniques, they highlight the level of uncertainty of the produced information and inform of the need for the specification of the admissible scale and uniformity requirements for the classification to be meaningful under a certain precisification. A more detailed account of the ontological issues involved on the mapping of land cover is provided in [10].

5.3 Individuation

The notion of individuality is fundamental in the study of ontology and essential when adopting an object model. However, formally characterising the full criteria for the individuation of particular classes of object tends to be extremely hard [23] and is not addressed in the majority of actual ontologies. Studies in Cognitive Science show that humans identify and individuate objects using at least three sources of information: spatiotemporal information, property (featural) information, and sortal information [40]. Moreover, among them, spatial features such as shape are typically more salient than other properties [40].

Within the philosophical literature it is considered that individuation requires both identity and unity, where the former is related to the problem of distinguishing a specific instance of a certain class from other instances (by means of a characteristic property unique to that object) and the latter is related to the problem of distinguishing the parts or constituents of an instance from the rest of the world, which are bound together creating a whole. We now propose a particular form of existence condition to axiomatically specify the domain of individuals that are instances of a given concept.

Existence conditions differ from classification in that the latter express the necessary and sufficient conditions for an object to be an instance of a class while the former explicitly specify the necessary and sufficient conditions to infer the existence of an object. Below is a constructive existential axiom that specifies that whenever a set of conditions $\Phi(x_1, \dots, x_n)$ are satisfied for some original concepts $x_1, \dots, x_n$, then an object of kind K exists and a relation holds between the original group and the existent object $\Psi(x_1, ..., x_n, y)$.

$$\forall x_1 ... \forall x_n [\Phi(x_1, \dots, x_n) \rightarrow \exists y [K(y) \land \Psi(x_1, \dots, x_n, y)]],$$
$$\forall y [K(y) \rightarrow \exists x_1 ... \exists x_n [\Phi(x_1, \dots, x_n) \land \Psi(x_1, \dots, x_n, y)]]\,.$$

The identity criteria I of a concept determine the conditions under which it can be established that two references refer to the same object, that is, the characteristics that are unique to a single specific instance [22].

$$\forall x \forall y [(K(x) \land K(y)) \rightarrow (I_k(x, y) \leftrightarrow (x = y))\,.$$

Finally, the notion of unity refers to the problem of describing the parts of objects and the specific conditions (UC) under which the object constitutes a whole. A general axiomatic characterisation of this, in terms of a unifying relation among the parts of a whole is given in

[23]. In modelling a particular domain or type of object, it is likely that more specific unity criteria will be required. For instance, a forest may be regarded as a spatially connected region of forested land, which is of maximal extent (i.e. is not part of a larger spatially connected forested region). Thus, assuming a predicate Forested has been defined and applied to all parcels of forestland, then the following axiom (where P is the parthood relation and SCON is the property of being spatially self connected) captures a unity condition for a possible precisification of forest:

$$\mathsf{Forest}(x) \rightarrow \mathsf{Forested}(x) \wedge \mathsf{SCON}(x) \wedge \neg\exists y[\mathsf{P}(x,y) \wedge \neg(x=y) \wedge \mathsf{Forested}(y) \wedge \mathsf{SCON}(y)]\,.$$

Another approach to the unity of forests becomes available if we have forest detailed information at the level of the location of individual trees. We can then define a forest as a maximal collection of proximal trees, parameterised by some threshold of proximity. This is the approach taken in our prototype software implementation, which will be described in Section 6.

A variety of situations can hinder the specification of a criteria for identity and unity. Some of them are drastic evolutions of objects through time, situations in which objects merge or split and objects whose boundaries are ill defined or affected by sorites vagueness thus creating confusion about self-connectedness and parthood.

Underlying axiomatisations of the space and mereotopology are key to provide appropriate notions of parthood. In some cases, a set theoretical view where two sets are the same if and only if they have exactly the same elements is appropriated to model the space. However, for most objects a looser identity criteria that allows one to accept the continued existence of an object even after the loss of certain sorts of parts is necessary.

5.3.1 Individuation of geographical objects

The consideration of the individuation of forests entangles in the extensive bibliography about the ontology of geographical features, their characterisation and their boundaries. Difficulties tend to arise both regarding the unification of geographical features (e.g. deciding whether something is part or not of a forest) and their indentity (e.g. deciding whether a forest now is the same forest as one that existed 100 years ago) particularly if there have been substantial changes in vegetation or location [4]. Moreover, while most of the objects in the physical world have a *bona fide* boundary that acts as one of the main marks of their individuality, geographical boundaries are often fuzzy or otherwise indeterminate [8], which makes the individuation even more challenging and the demarcation of most geographical objects non trivial.

It may seem that a demarcation is required in order to individuate a feature such as a *forest*. However, this is not necessarily the case. We show how individuation is possible without committing to a specific boundary, particularly in cases in which we don't encounter borderline cases. We first describe the example of *mountain*, widely used in the literature, to show how individuation and demarcation are not necessarily related. We then analyse the case of *forest* and discuss how in certain cases it may be appropriate to use different thresholds for the same parameters in the different modes of predication. Although *mountain* and *forest* involve very different attributes, the two cases are analogous in that individuation can be achieved prior to demarcation, and, if needed, a demarcation can subsequently be obtained by further precisification.

In a prototypical case of *mountain* individuation, the main characterisation for both the existence and the individuation can be done in terms of the peak and a minimum

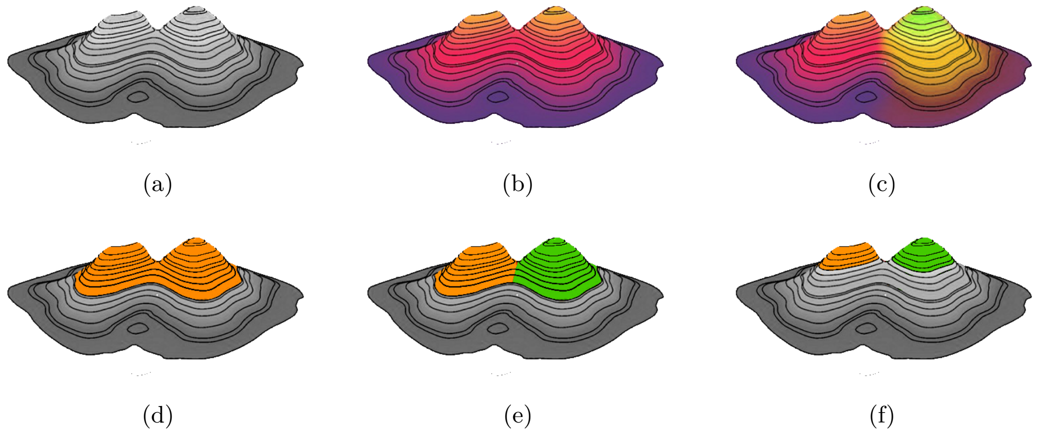


Figure 2 Mountain individuation and demarcation. Each individual is shown with a gradient of different colours, where the colour on the peak (orange and green) are the ones characterising the object and used in the demarcated images (d), (e) and (f).

prominence[3]. Thus, in Figure 2 we see (a) a landscape in which we can identify, depending on the precisification, (b) one or (c) two mountains. In (b) the minimum prominence required is not satisfied and thus the whole ensemble is assumed to be a single mountain, and the contrary occurs in (c). It is not until (d), (e) and (f) that we specify a demarcation strategy. (d) and (e) apply the same demarcation strategy to the two different individuations (b) and (c), and an alternative demarcation for (c) is shown in (f). The identity criteria for the mountain is to share the highest point, that is, the peak. In terms of unity, mountain can be characterised as a self-connected whole extending from the peak to whichever precise boundary.

The identification of *forest* without relaying on the demarcation is harder because the most salient features are its shape, size and parts. However, as we can see in Figure 3 different looser metrics can be established, in that case nearness between the members, to detect rough shape, location, size and fragments. Different nearness thresholds and minimal size criteria determine the individuation of two forests in Figure 3(a) and three forests in Figure 3(b). In addition, different precisifications of tree are used in Figure 3(a) and Figure 3(b), implying that in the former some elements are not considered members of the collective. The actual strategy for the demarcation (green shade connecting the trees) is not done until a later stage and depends on the selection of a suitable footprint strategy. In the figure, different thresholds for the demarcation are shown in different tones of green. Finally, it must be noted that it is not necessary to analyse forest in terms of a collective of trees in order to be able to individuate it. A similar approach to the one in mountains could be done for forests by interpolating a field of canopy cover in which similar measures to the ones used in Figure 2 can be used.

Although identity is likely to be characterised mainly through the rough location of the object, both identity and unity criteria are less intuitive for *forest* than for *mountain*, and are expected to vary between precisifications according to the particular management objectives.

---

[3] characterizes the height of a mountain or hill's summit by the vertical distance between it and the lowest contour line encircling it but containing no higher summit within it. It is a measure of the independence of a summit.

Thus, whether forest must be self-connected or not and whether it must comply with either morphological or metrical restrictions must be made explicit for all the precisifications that attempt to refer to *forests* as individuals.

5.4 Demarcation

Finally, by *demarcation* we mean the act of determining the spatial extension of an object, or equivalently of establishing its boundary. Once this extension/boundary is established, it may be referred to as 'the demarcation' of the object; or, in cases where the boundary is unclear or debatable, it may be regarded as one of many possible demarcations of the object.

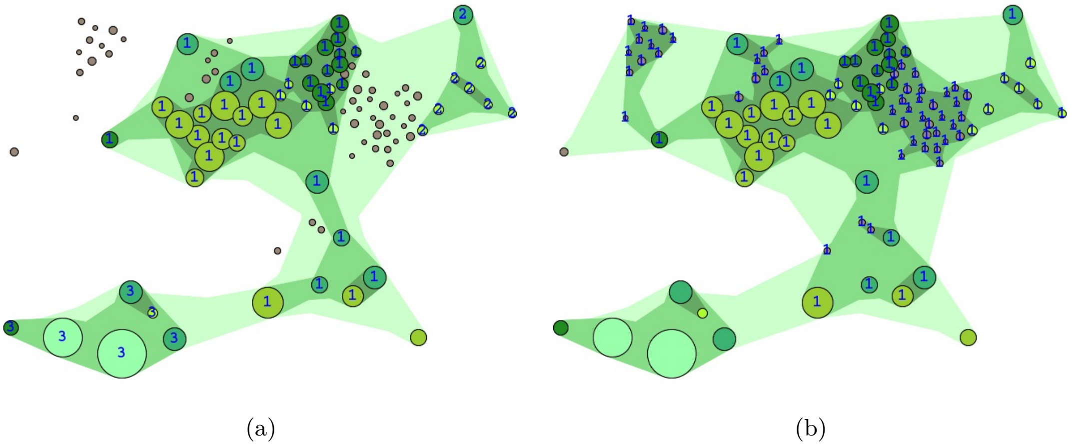
Establishing an object's demarcation may be straightforward or extremely problematic. Moreover, although this may be because of the characteristics of the particular object under consideration, it is usually strongly related to the ontological category of object. For example, physical artifacts (e.g. cups, tables) are typically easily demarcated because they consist of solid matter forming an integral whole that is not physically connected to any other matter. Even with artifacts, the demarcation may not be completely clear. For example, some tables have a glass top that just rests on the wood below. It is debatable whether the glass is actually part of the table. For aggregate objects, such as a school of fish, or indeed a forest, demarcation is very often problematic, both because it may not be clear which entities should be counted as constituents and because there is no unique way to determine the spatial extension of something that is made up of many disconnected constituents (e.g., in demarcating a forest, one may want to include the space that lies between the trees within the demarcation of the forest, as can be seen in Figure 3). Distinguished regions within field like objects, which are again common in geography (e.g. soil type regions) also give rise to significant demarcation problems [30].

Although the study of suitable algorithms for the demarcation go beyond the scope of this paper, it must be noted that, in most of the studies, the specification of these strategies are often reduced to a set of thresholds [14, 24]. In some cases, particularly in human drawn maps, it has even been guided by intuitive and aesthetic judgements. We consider that a more careful analysis and specification of the algorithms used can be important for an appropriate characterisation of forests. For instance, a survey on the strategies that can be used to demarcate the forests identified in Figure 3 can be found in [11].

5.5 Interactions among classification, individuation and demarcation

We do not see classification, individuation and demarcation as cleanly separable aspects of a predicate's semantics. As we have seen, a reasonable classification often presupposes an appropriate individuation. Individuation of an object typically goes some way towards determining its demarcation; and conversely by demarcating objects we usually also individuate them. Consequently, these different aspects may coexist within a definition and be entangled among different terms within an ontology. This may result in a complex albeit comprehensive ontology. But if one does not pay some attention to ensuring that all aspects are accounted for, they may be omitted. As we have seen, it is typically classification that receives the focus of attention, whereas individuation is taken for granted.

From a purely logical point of view, the difference between individuation and classification is also not entirely clear cut. This is because we often have choices as to what we take to be the universe of quantification. What is considered to be an individuating criterion in one choice of universe could be considered as a classification criterion applied to a larger universe of entities from which we want to select a significant subset.



(a) (b)

Figure 3 Forest individuation and demarcation.

6 Implementation of a Supervaluationistic Geographic Query System

We have implemented in Prolog a prototype system for individuating, demarcating and classifying forest regions. The system interprets a set of spatially located plant objects in terms of a precisification specification of the form $\langle \Theta(\mathsf{tree}),\ \Gamma(\nu),\ \sigma \rangle$, where $\Theta(\mathsf{tree})$ is a classification of tree objects in terms of more basic properties (e.g. species, height), $\Gamma(\nu)$ is a tree grouping algorithm parametrised by a nearness threshold ν, and σ gives the minimum number of trees for a tree group to be considered as a forest region.

The images in Figure 3 were generated by this system and illustrate how altering the threshold for 'nearness' used in an aggregation algorithm and the classification criteria of constituents affect individuation and demarcation of aggregates. The mid green region indicates the grouping obtained using the nearness threshold, ν, that has been applied in grouping and determines the number of tree groups counted. (We may regard the tree groups as 'forests', although to keep the computation simple and the images clear, we are individuating much smaller groups of trees than would normally be considered to be forests.) The light green area is computed with a nearness threshold 2ν, which incorporates all trees into one group, and the dark area with $\nu/3$ shrinks the tree group demarcations to include only the more dense areas. In visualisation (a) the small brown circles (thorn bushes) are not counted as forest constituents according to the chosen version of $\Theta(\mathsf{tree})$. Hence, we get a split between forest region 1 and forest region 2. In visualisation (b), using an alternative theory $\Theta'(\mathsf{tree})$, thorn bushes are treated as forest constituents, so regions 1 and 2 become merged. Also, the minimum number of trees required to count as a forest region, σ, has been increased in the (b) precisification. Because of this, the group counted as 3 in (a) is no longer considered to be a forest region in (b).

7 Conclusions

We have discussed several challenging problems that obstruct the task of giving precise definitions of geographic terms, such as 'forest'. To address these challenges we have provided a framework within which one can specify a range of possible interpretations of the 'forest', and which makes explicit how the semantic aspects of *classification*, *individuation* and *demarcation* interact and combine within possible definitions. We have indicated how this variability can be modelled within an ontological theory augmented with supervaluation semantics incorporating explicit specification of precisifications. The proposed framework has been implemented within a prototype Prolog-based GIS. In future work we intend to give a fully formalised theory, which will form an ontology module within our system or could equally be used within a different (non-Prolog) implementation based on the same general principles. Because of its generality and flexibility, this framework could be applied to characterising a wide range of geographic and other spatial objects, even where significant vagueness and ambiguity is present and where complex individuation and demarcation criteria may be required.

Since spatial properties and relationships often play an essential role in specifying individuation and demarcation criteria, we believe that theoretical study of these aspects of predicate semantics will play a key role in establishing more comprehensive and robust ontologies of geographic and other spatially related terminology. Moreover, the development of foundational spatial information theory, which unites both geometrical and cognitive aspects of space, will play a key role in addressing this challenge.

References

1 F. Achard. Forest definition and extent. In *Vital Forest Graphics*, pages 6–9. UNEP/Earthprint, 2009.

2 P. Agarwal. Ontological Considerations in GIScience. *International Journal of Geographical Information Science*, 19(5):501–536, 2005.

3 B. Bennett. What is a Forest? On the vagueness of certain geographic concepts. *Topoi*, 20(2):189–201, 2001.

4 B. Bennett. Physical objects, identity and vagueness. In *Principles of Knowledge Representation and Reasoning: Proceedings of KR2002*, pages 395–406, San Francisco, CA, 2002. Morgan Kaufmann.

5 B. Bennett. Modes of concept definition and varieties of vagueness. *Applied Ontology*, 1(1):17–26, 2005.

6 D. Bruce and L. C. Wensel. Modelling forest growth: approaches, definitions, and problems. *USDA Forest Service general technical report NC*, 1988.

7 D. C. Buckland, D. R. Redmond, and R. Pomerleau. Definitions of Terms in Forest and Shade Tree Diseases. *Canadian Journal of Botany*, 35(5):675–679, 9 1957.

8 R. Casati, B. Smith, and A. Varzi. Ontological Tools for Geographic Representation. *Formal Ontology in Information Systems*, pages 77–85, 1998.

9 R. Chazdon, P. Brancalion, and L. Laestadius. When is a forest a forest? Forest concepts and definitions in the era of forest and landscape restoration, 2016.

10 A. Comber, P. Fisher, and R. Wadsworth. What is land cover? *Environment and Planning B: Planning and Design*, 32(2):199–209, 2005.

11 M. Dupenois and A. Galton. Assigning Footprints to Dot Sets: An Analytical Survey. In *COSIT 2009*, pages 227–244. Springer, Berlin, Heidelberg, 2009.

12 M. Egenhofer. What's special about spatial? *ACM SIGMOD Rd*, 22(2):398–402, 6 1993.

13 M. Egenhofer and D. Mark. Naive Geography. *International COSIT*, pages 1–15, 1995.

14 FAO. *Global Forest Resources Assessment.* 2015.

15 P. Fisher. Sorites paradox and vague geographies. *Fuzzy Sets and Systems*, 113:7–18, 2000.

16 F. Fonseca and M. Egenhofer. Ontology-driven geographic information systems. *Proceedings of 7th ACM ISA in GIS*, 35(4):14–19, 1999.

17 D. Fuller. Tropical forest monitoring and remote sensing: A new era of transparency in forest governance? *Singapore Journal of Tropical Geography*, 27(1):15–29, 2006.

18 C. Giri, Z. Zhu, and B. Reed. A comparative analysis of the Global Land Cover 2000 and MODIS land cover data sets. *Remote Sensing of Environment*, 94(1):123–132, 2005.

19 A. Grainger. Difficulties in tracking the long-term global trend in tropical forest area. *Proceedings of the NAS of the USA*, 105(2):818–823, 2008.

20 A. Grainger. Uncertainty in the construction of global knowledge of tropical forests. *Progress in Physical Geography*, 34(6):811–844, 2010.

21 N. Guarino. Formal Ontology and Information Systems. *Proceedings of FOIS*, 98:6–8, 1998.

22 N. Guarino and C. Welty. Identity, Unity, and Individuality: Towards a Formal Toolkit for Ontological Analysis. *Ecai 2000*, pages 219–223, 2000.

23 N. Guarino and C. Welty. An Overview of OntoClean. In *Handbook on Ontologies*, pages 201–220. Springer Berlin Heidelberg, 2009.

24 M. C. Hansen et al. High-Resolution Global Maps of 21st-Century Forest Cover Change. *Science*, 850(27):123–134, 2013.

25 P. Holmgren and R. Persson. Evolution and prospects of global forest assessments. *UNASYLVA-FAO*, 2003.

26 Z. Hu, G. Tang, and G. Lu. A new geographical language: A perspective of GIS. *Journal of Geographical Sciences*, 24(3):560–576, 2014.

27 M. Kavouras and M. Kokla. A method for the formalization and integration of geographical categorizations. *International Journal of GIS*, 16(5):439–453, 2002.

28 H. Gyde Lund. When Is a Forest not a forest? *Journal of Forestry*, 100(8):21–27, 2002.

29 H. Gyde Lund. Definitions of Forest, Deforestation, Afforestation and Reforestation. *Forest Information Services*, pages 1–159, 2007.

30 D. M. Mark and F. Csillag. The nature of boundaries on 'area-class' maps. *Cartographica*, 26:65–78, 1989.

31 E. Mena and V. Kashyap. Domain Specific Ontologies for Semantic Information Brokering on the Global Information Infrastructure. In *Proceedings of FOIS'98*, pages 269–83. 1998.

32 W. Meyer and B. Turner. *Changes in land use and land cover: a global perspective*. Cambridge University Press, 1994.

33 B. Smith and D. M. Mark. Ontology and geographic kinds. In *Proceedings of the 8th International Symposium on Spatial Data Handling (SDH'98)*, 1998.

34 D. Tagliapietra et al. A review of terms and definitions to categorise estuaries, lagoons and associated environments. *Marine and Freshwater Research*, 60(6):497–509, 2009.

35 E. Tomai and M. Kavouras. From Onto-GeoNoesis to Onto-Genesis: The Design of Geographic Ontologies. *GeoInformatica*, 83(281), 2004.

36 A. Varzi. Philosophical Issues in Geography – An Introduction. *Topoi*, 2:119–130, 2001.

37 A. Varzi. Vagueness in geography. *Philosophy & Geography*, 4(1):49–65, 2001.

38 A. Varzi. Vagueness, Logic, and Ontology. *Yearbooks for Philosophical Hermeneutics*, 1:135–154, 2001.

39 G. Wiederhold. Interoperation, Mediation, and Ontologies. In *FGCS 1994 Workshop on Heterogeneous Cooperative Knowledge Base*, volume 3, pages 33–48. ICOT, 1994.

40 Fei Xu. Sortal concepts, object individuation, and language. *Trends in Cognitive Sciences*, 11(9):400–406, 2007.