Anchoring: A New Approach to Handling Indeterminate Location in GIS

Antony Galton and James Hood

University of Exeter, UK, {A.P.Galton, J.M.Hood}@exeter.ac.uk

Abstract. We describe a new approach to representing vague or uncertain information concerning spatial location. Locational information about objects in information space is expressed through various *anchoring relations* which enable us to state exactly what is known regarding the spatial location of an object without forcing us to identify that location with either a precise region in the embedding space or any precise mathematical construct from such regions, such as rough sets or fuzzy sets. We describe the motivation for introducing Anchoring, propose the beginnings of a formal theory of the anchoring relations, and illustrate some of the ideas with examples typical of the real-life use of GIS.

1 Introduction

In this paper we introduce Anchoring, a new approach to handling vague and uncertain location in GIS. We use Anchoring to frame and approach one particular problem: Within the spatial data model of our GIS, how is it possible to reference spatial information whose location cannot be assigned to precise coordinates?

A key motivation behind Anchoring was to develop a framework or model that preserves any imprecision or vagueness, and does not eliminate it through any forced approximation to precise regions. Many existing approaches are based on the faulty assumption that vague or imprecise objects can, indeed must, be associated with regions which represent their spatial extent. No matter whether these regions are 'fuzzy', or the intersection of a set of 'precisifications', we feel that any precise representation of a vague or imprecise spatial extent, though useful at a stage of analysis, fails to capture a key aspect of the region (namely, its vagueness or imprecision), and should therefore be omitted from any spatial data model.

With Anchoring we try to keep the vagueness or uncertainty of an object's location 'alive' by saying only what we can say about it precisely and leaving everything else for subsequent analysis. What can be said precisely about a region whose boundary is imprecise or vague sometimes includes its topological relationships to regions whose boundary is precise. Anchoring uses these relationships to give vague or imprecise regions reference within a spatial data model.

2 The problem of indeterminate location

Frank [3] articulates the problem thus:

> 'Space in today's GIS is always seen as Euclidean space, represented with Cartesian coordinates . . . Data that are not related to a point with given position cannot be entered in a GIS. This limits the use of GIS for combining verbal reports about events, such as accidents.' [3]

To illustrate, suppose we wish to record in our GIS information about a car accident, but all we know about the location of the accident is that it is 'somewhere' on the stretch of motorway which lies between two given junctions. The data model of our GIS, quite standardly, uses the coordinatised Euclidean plane; accidents would be represented as geometric points, and stretches of motorway as connected sets of geometric line segments. Our problem is that without assigning the accident to a particular point, specified by a pair of Cartesian coordinates, we cannot represent the accident in the spatial data model. Of course, in reality we expect the emergency services to arrive at the scene of such an accident quickly, and they could then give us a precise description of the accident's location. But what if the data we are entering into our GIS is from historical sources, and the GIS is being used for analysis of this data? Or if the accident was not on a motorway but in a remote area like, say, a national park, and the information is provided through possibly impressionistic verbal reports lacking precise coordinates? In such situations our information might never be precise enough to represent the accident as a geometric point.

This example involves uncertainty, but related problems arise in connection with vagueness—although vagueness and uncertainty are quite distinct concepts, they have enough commonality in the kinds of representational problems they raise to justify our treating them together, at least in some respects. In both cases we see a kind of *indeterminacy*, and it is this aspect which is primarily addressed by our Anchoring theory.

Vagueness arises, amongst other places, in connection with descriptive terms such as 'hill', 'valley', or 'town centre', whose referents, while certainly located in space, cannot be assigned precisely delineated locations, short of an arbitrary and contestable decision. In Figure 1(a), for example, the name 'Greendale' clearly refers to the valley whose presence is evident to the map user from the contours shown; the map rightly makes no attempt to assign a region with precise boundaries as the referent of 'Greendale', and yet it cannot truly be claimed that the map is on *that* account ambiguous or incomplete: after all, we all know what a valley is, and can see from the particular configuration of contours that there is a unique valley associated with the name 'Greendale'. This does not mean that we can always give definite answers to questions such as whether or not Green House is in Greendale: it is part of our common-sense understanding of the notion of a valley that answers to such questions will sometimes be indeterminate. To be entirely certain *which* valley we are talking about, we can always refer to it as the valley of the River Green; and while this fixes the valley beyond doubt it still does nothing to make more precise its exact extent.

Likewise, in Figure 1(b), Green Hill is the hill which has a summit at elevation 220m at the exact coordinates where this is marked, and from our general understanding of

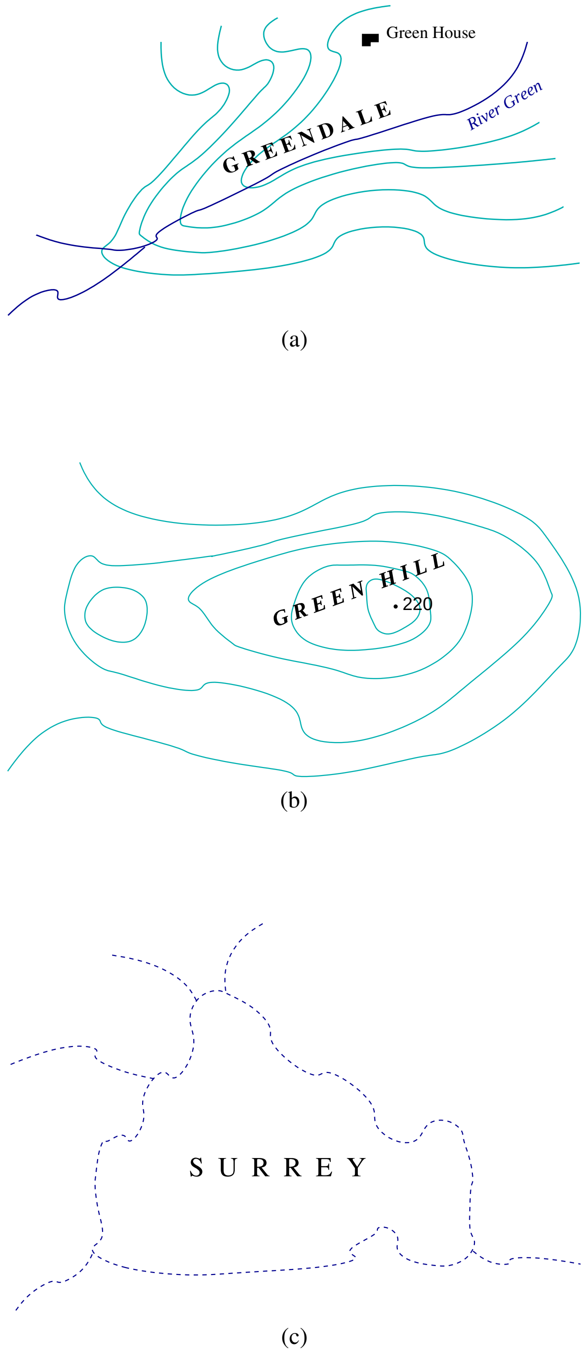


Fig. 1. Placement of names on maps. Unlike ‘Surrey’, neither ‘Greendale’ nor ‘Green Hill’ is associated with a region having a precisely delineated boundary. None the less, all three maps are both clear and informative.

what hills are this suffices to establish which hill bears that name, without in any way importing the spurious precision that would come from delineating some precise outer boundary and thereby assigning to the hill some exact spatial extent.

Contrast these cases with Figure 1(c), where the word ‘Surrey’ is printed to indicate the English county of that name, and from our understanding that counties do have precise boundaries we can identify the relevant boundary on the map and understand the spatial extent of Surrey to be precisely what is contained within that boundary.

3 Outline of some existing approaches

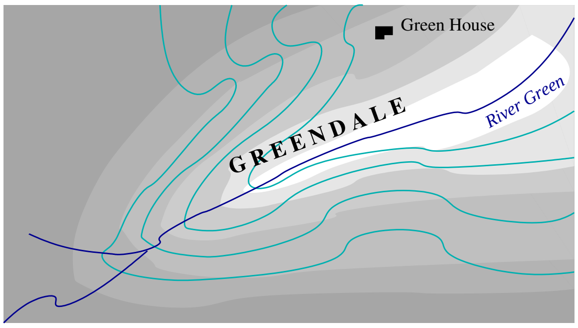
A number of methods have been proposed for handling locational vagueness and uncertainty in the context of information systems in which spatial location can only be specified precisely. We shall review a few of these here, and in particular shall examine how they fare in relation to the examples presented above.

Fuzzy sets

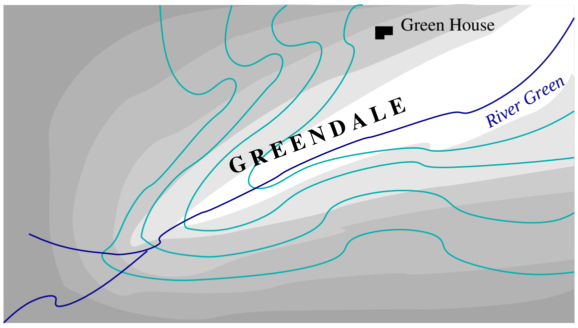
If we think of our space as a set S of points, each endowed with precise numerical coordinates, then a precise location can be understood as some subset L of that space. For each point $p \in S$, we have either $p \in L$ or $p \notin L$. This is what is meant by saying that L is a *crisp* set, and standard set theory admits no other kind. A crisp set L is completely specified once it is determined, for each $p \in S$, which of $p \in L$ and $p \notin L$ holds. A fuzzy set [8], on the other hand, is a more complex beast: to specify it, we must state, for each $p \in S$, the *degree* to which p may be regarded as belonging to L. This is a number in the range $[0, 1]$, and in principle all real numbers in this range are possible fuzzy membership degrees. Crisp sets arise as the special case in which the number assigned is always either 0 (corresponding to $p \notin L$) or 1 (corresponding to $p \in L$), whereas in a general fuzzy set, the points assigned value 0 are separated from those assigned 1 by a zone consisting of points assigned intermediate values. It is normally assumed, moreover, that the values are assigned in accordance with some continuous function, so that close together points will be assigned close together values.

Referring back to our examples, the location of the valley can be represented as a fuzzy set, as shown in Figure 2(a), where the fuzzy membership values for the valley’s location are presented by means of shading, with lighter shades for higher values. A big problem here is that there does not seem to be any principled basis on which to decide exactly how the fuzzy values should be assigned; this is illustrated by Figure 2(b), where a completely different, but equally ‘plausible’ assignment is used.

In the case of the motorway accident, if all we know about the location of the accident is that it is somewhere between, say, junctions 26 and 27 on the M5, then there does not seem to be *any* set, either crisp or fuzzy, which can represent the location of the accident. Since we know that (at the granularity we are working with) this location is a point, and are merely uncertain as to the exact location of that point, a better way of expressing this knowledge might be in the form of a probability distribution—in this case a flat distribution on the relevant stretch of motorway, and zero elsewhere.



(a) A fuzzy membership function for ‘Greendale’.



(b) Another fuzzy membership function for ‘Greendale’.

Fig. 2. Arbitrariness of fuzzy membership functions: How do we choose between the functions shown here, or any of the infinitely many others we could have drawn?

Rough sets, egg-yolk theory, and supervaluation

Under this heading we include all those approaches which assign to a vague or uncertain entity not one precise region, but *two*: an ‘inner’ region containing all points which are uncontestably within the location of the entity of interest, and an ‘outer’ region containing all points which are *not* uncontestably *outside* that location. In egg-yolk theory [1], the former is likened to an egg-yolk, the latter to the whole egg; in rough set theory [4], they are the inner and outer approximations respectively.

Supervaluation semantics [2], when applied to spatial location (as opposed to its more usual application area of vague concepts), ends up with a similar picture. If the location of an object is vague, then there may be many different ways of making it precise—these are called ‘precisifications’ of the location. Supervaluation theory works with propositions, so we must consider the truth or falsity of propositions of the form ‘point p is in region R’ (i.e., $p \in R$) under different precisifications of R. Propositions which come out true under every precisification are called ‘supertrue’—this corresponds to our earlier use of the phrase ‘uncontestably true’, and therefore the points p for which the proposition ‘$p \in R$’ is supertrue correspond to the egg-yolk; likewise, the whole egg consists of those points for which ‘$p \notin R$’ is *not* supertrue.

Applying this to our hill example, we see that there are many different ways in which we could place the inner and outer boundaries. We are confronted with a similar problem to the fuzzy set case: the theory forces us to make some arbitrary decisions which embody more precise information than we are entitled to assume.

We can, however, salvage something of value from this situation by interpreting the two boundaries somewhat differently. The yolk must certainly only contain points for which $p \in R$ is uncontestably true, but we should not insist that it contain *all* such points, since to determine precisely which those points are is by nature an impossible task (put another way, the predicate ‘uncontestably true’ is itself vague). Likewise, the outer boundary must be placed somewhere so that it only *excludes* points for which ‘$p \in R$’ is uncontestably *false*, but we should not insist that it excludes all such points. This gives considerable lee-way in the placement of the boundaries, and indeed points the way to the freer approach embodied in our Anchoring theory. This understanding of vague boundaries was well expressed by Wittgenstein:

> The boundaries . . . are still only like the walls of the *forecourts*. They are drawn arbitrarily at a point where we can still draw something firm. — Just as if we were to border off a swamp with a wall, where the wall is not *the* boundary of the swamp, it only stands around it on firm ground. It is a sign which shows there is a swamp inside it, but not, that the swamp is exactly the same size as that of the surface bounded by it. [6, §211]

Information space and precise space

By way of introduction to Anchoring, it will be helpful to adopt the following perspective on spatial location. We are talking about geographical features that are located in space. We have a descriptive language for talking about these features, and in particular this allows us to say things about where the features are located in relation to other features, e.g., ‘My house is north of the town centre’. If we consider all the things we can say about these features, short of assigning exact absolute locations, we can think of this as constituting an *information space*—it is a space in the sense that it includes information about the locations of features relative to each other, but also in the more abstract sense of being an assignment of properties and relations (spatial or otherwise) to a collection of entities. It is the sort of information that could be captured in a relational database. However a GIS is not just a relational database; the entities it handles are embedded in a space—an 'embedding space' [7]—whose structure constrains the properties and relations that it is possible to assign to those entities. To use a GIS to record information about the objects in the information space, we have to bring those objects into relation with the locations we can define in this embedding space.

We define a *precise space* as any embedding space in which phenomena can be assigned exact positions relative to some numerical coordinate system. The coordinatised Euclidean plane is an example of a precise space. So, too, is the surface of a sphere under the latitude and longitude coordinate system. And aside from mathematical objects, any map, drawing or other representation of the earth's surface can function as a precise space, so long as each point on the representation corresponds to a fixed geographic position according to a fixed scale or projection.

In what we might call the 'classical' GIS data model, the embedding space is always a precise space, and there is only *one* kind of relation between information space and that precise space, namely *exact location*. That is, each element of information space is assigned to a spatial location (point, line, polygon, . . .) defined in the precise space. In effect the object is *identified* with the location it is assigned to (see Figure 3).

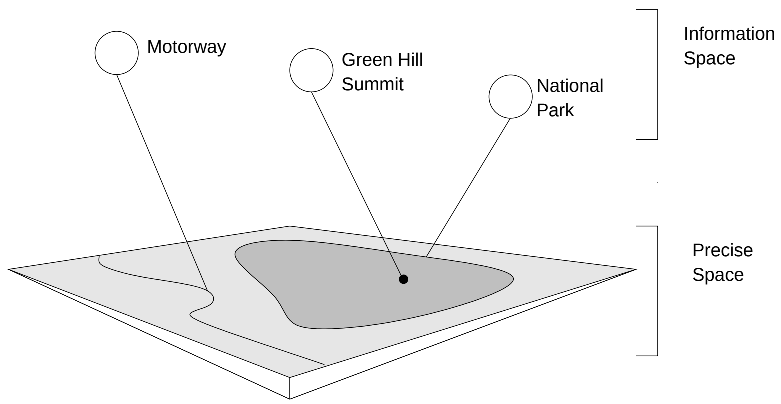


Fig. 3. The 'classical' view of a GIS data model with just one kind of relation (exact location) between objects in information space and locations in precise space.

In fuzzy set approaches, it is recognised that for many objects in information space it is not possible to assign an exact location in precise space without falsifying reality. They therefore replace exact location with *fuzzy location*, which assigns to an object O in information space not one exact location but rather, in effect, an infinite set of exact locations, each indexed by a degree of fuzzy membership. On the other hand, rough set and egg yolk approaches assign a *rough location* consisting of just *two* exact locations, the inner and outer approximations.

By presenting things in this way, we intend to highlight the fact that the problem of vague location is *to establish the nature of the relationship between information space and precise space*. Our guiding insight is that rather than constraining this relationship to some particular very precise form (be it exact location, fuzzy location, or rough location), we should take a freer and more relaxed view of the possible relationships between the two spaces, in particular allowing for *more than one* type of relationship between an entity in information space and an exact spatial location.

4 Anchoring Relations

The underlying ontology of Anchoring consists of two levels, representing information space and precise space. The elements of precise space are points and various kinds of sets of points which can function as precise spatial locations; the elements of information space are all the various geographical objects and phenomena of interest. The key idea is that an object in information space may be *anchored* to locations in precise space in various different ways. We leave it open at this stage just how many different ways we will need, but here are some we have found useful so far:

- Object O is *anchored in* region R; that is, whatever may be the nature of O's spatial location, we can certainly say that O is located within R.
- Object O is *anchored over* region R; that is, we can certainly say that R falls within the location of O (in many cases, R will in fact be a point).
- Object O is *anchored outside* region R, i.e., no part of O is located within R.
- Object O is *anchored alongside* region R, meaning that O is so situated that it abuts R (in many cases, R will be a line).

Some of these relations are illustrated, using examples from §5, in Figure 4.

Superficially, it might be thought that all we have done here is to replicate standard relations between regions such as the RCC relations $\mathsf{P}(O, R)$, $\mathsf{PI}(O, R)$, $\mathsf{DC}(O, R)$, $\mathsf{EC}(O, R)$. However, this would be to radically misunderstand our intentions: the O term here is *not* a precise spatial location—that is, it is not that we are merely ignorant of its location, but rather that in the case that O is irreducibly vague, it simply does not have a precise location. This prevents us from writing, e.g., $\mathsf{P}(O, R)$, since 'P' in RCC denotes a relation between precise locations; but we can still say, in our theory, that O (an object) is anchored in R (a precise region).

Even though objects in information space may be vaguely located, they may bear definite spatial relations to each other. For example, even if the terms 'Central London' and 'Greater London' are not precisely defined, we can certainly say that Central London is part of Greater London, and this places constraints on the possible anchorings: for example, Greater London must be anchored over any precise location that Central London is anchored over, and Central London must be anchored in any precise location that Greater London is anchored in. Such observations suggest the possibility of developing a formal calculus of the anchoring relations. In the next section we sketch the beginning of such a calculus; its further development is work for the future.

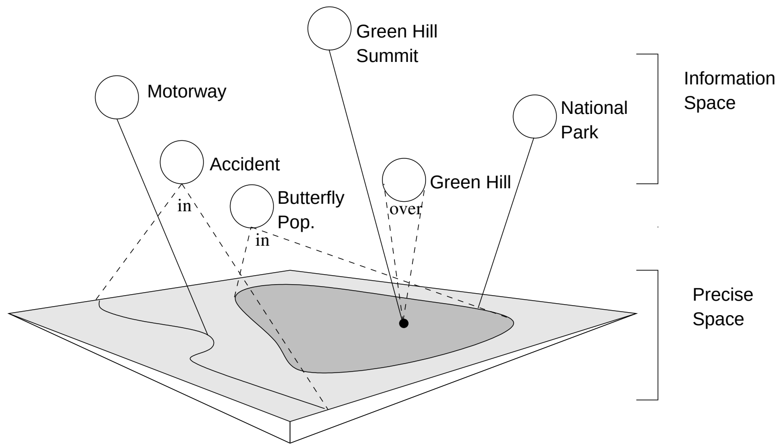


Fig. 4. Broken lines indicate anchoring relations, solid lines indicate exact location.

An object O that is exactly located at L is anchored *both in and over* L; this provides the Anchoring definition of exact ('classical') location. O still belongs to information space, but now we can reasonably *identify* it with L in the manner of classical GIS. For example, assuming the boundaries of the county of Surrey are precisely delineated, we can use 'Surrey' as the name, not only of a county (i.e., a particular type of administrative unit) but also of an exact spatial location, a subset of precise space. We can go on to say things like 'Box Hill is anchored in Surrey' as a shorthand for 'Box Hill is anchored in the region R such that Surrey is anchored both in and over R'. Bearing this convention in mind, here are some examples of anchoring:

- Brighton is anchored alongside the Sussex coastline.
- The North Atlantic Ocean is anchored outside the Southern Hemisphere.
- Eastern Europe is anchored over both Latvia and Ukraine.
- The Black Forest is part of Southern Germany, and Southern Germany is anchored in Germany, so the Black Forest is anchored in Germany.
- In Figure 1, Greendale is anchored over a stretch of the River Green, and Green Hill is anchored over the summit point at 220m.

4.1 Towards an axiomatic treatment of anchoring relations

We list here a plausible set of axioms for the part of Anchoring concerned with anchoring in and anchoring over, together with some of their consequences. Our notational conventions are as follows. We write

- bold lower-case letters (e.g., $\mathbf{v}$) to denote elements of information space (these are elements whose locations are described using anchoring relations);
- italic upper-case letters (e.g., P) to denote precise regions in the precise space;
- $\mathbf{v} \rhd P$ to mean that $\mathbf{v}$ is anchored over P.
- $\mathbf{v} \lhd P$ to mean that $\mathbf{v}$ is anchored in P.

- P and O for the RCC relations 'is part of' and 'overlaps'; these are used exclusively between elements of information space, since in the precise space, whose elements are sets of points, we can use standard set-theoretic notations;
- $\mathbf{v} + \mathbf{w}$ and $\mathbf{v} \cdot \mathbf{w}$ to denote the mereological sum and product of $\mathbf{v}$ and $\mathbf{w}$. Again, these are used exclusively over the information space; in precise space we can use the set-theoretical notations $\cup$ and $\cap$.

Using this notation, our proposed axioms are as follows:

1. If any part of $\mathbf{w}$ is anchored over P then $\mathbf{w}$ is anchored over any part of P:
$$\mathbf{v} \rhd P \wedge \mathsf{P}(\mathbf{v}, \mathbf{w}) \wedge Q \subseteq P \rightarrow \mathbf{w} \rhd Q$$
2. If $\mathbf{v}$ is anchored in part of P then any part of $\mathbf{v}$ is anchored in P:
$$\mathbf{v} \lhd Q \wedge \mathsf{P}(\mathbf{w}, \mathbf{v}) \wedge Q \subseteq P \rightarrow \mathbf{w} \lhd P$$
3. If two objects are anchored over the same region, then they overlap, and their common part is also anchored over that region:
$$\mathbf{v} \rhd P \wedge \mathbf{w} \rhd P \rightarrow \mathsf{O}(\mathbf{v}, \mathbf{w}) \wedge \mathbf{v} \cdot \mathbf{w} \rhd P$$
4. The sum of two objects anchored in the same region is also anchored in that region:
$$\mathbf{v} \lhd P \wedge \mathbf{w} \lhd P \rightarrow \mathbf{v} + \mathbf{w} \lhd P$$
5. An object anchored over two regions is anchored over their union:
$$\mathbf{v} \rhd P \wedge \mathbf{v} \rhd Q \rightarrow \mathbf{v} \rhd P \cup Q$$
6. An object anchored in two regions is anchored in their intersection:
$$\mathbf{v} \lhd P \wedge \mathbf{v} \lhd Q \rightarrow \mathbf{v} \lhd P \cap Q$$
7. An object anchored in a region lies within any object anchored over that region:
$$\mathbf{v} \lhd P \wedge \mathbf{w} \rhd P \rightarrow \mathsf{P}(\mathbf{v}, \mathbf{w})$$
8. Any region an object is anchored over is part of any region it is anchored in:
$$\mathbf{v} \lhd P \wedge \mathbf{v} \rhd Q \rightarrow Q \subseteq P$$

Some straightforward consequences of these axioms (together with standard set theory and mereology) are:

9. $\mathbf{v} \rhd P \cup Q \leftrightarrow \mathbf{v} \rhd P \wedge \mathbf{v} \rhd Q$ (From 1 and 5).
10. $\mathbf{v} \lhd P \cap Q \leftrightarrow \mathbf{v} \lhd P \wedge \mathbf{v} \lhd Q$ (From 2 and 6).
11. $\mathbf{v} \rhd P \wedge \mathbf{w} \rhd Q \rightarrow \mathbf{v} + \mathbf{w} \rhd P \cup Q$ (From 1 and 5)
12. $\mathbf{v} \lhd P \wedge \mathbf{w} \lhd Q \wedge \mathsf{O}(v, w) \rightarrow \mathbf{v} \cdot \mathbf{w} \lhd P \cap Q$ (From 2 and 6)

Anchoring is more general than theories based on the 'egg-yolk' idea or rough sets. In such theories, a vague region $\mathbf{v}$ is represented as a pair of precise regions $(P_{\mathbf{v}}^-, P_{\mathbf{v}}^+)$, where $P_{\mathbf{v}}^- \subset P_{\mathbf{v}}^+$. In anchor notation we would write $\mathbf{v} \rhd P_{\mathbf{v}}^- \wedge \mathbf{v} \lhd P_{\mathbf{v}}^+$. But Anchoring does not impose any requirement that there should be, for each vague region $\mathbf{v}$, exactly two precise regions P and Q such that $\mathbf{v} \rhd P \wedge \mathbf{v} \lhd Q$, nor, even if this is the case for a particular $\mathbf{v}$, would it in any way equate $\mathbf{v}$ with the pair (P, Q), instead leaving it open that at a later time, more precise anchoring information concerning $\mathbf{v}$ may be acquired without the identity of $\mathbf{v}$ as a single vague region being disrupted.

5 Illustrative studies

Although the axioms above may seem straightforward—even trivial—any implementation of Anchoring should respect them, for they will play a role in almost any inference from anchoring information. For example, suppose we know that a certain hut is located on Green Hill, and that the hill is anchored over the hill summit, whose exact location is the region A. If we now learn that the precise location of the hut is a region B, then axiom 5 assures us that the hill is anchored over the union $A \cup B$.[1]

5.1 Uncertain Location: Motorway Accidents

Suppose we want to use an anchoring-enabled GIS to record road accidents. In our example, all we know about a certain accident is that it occurred on the M5 between junctions 26 and 27. To enter this into our system, we activate the Location dialogue, which gives us the option of either entering Exact Location or Anchoring Information. We choose the latter, and select the 'Anchored In' option. (Since the location of the accident is effectively a point, anchoring-in implies uncertainty rather than vagueness here.) We specify the stretch of the M5 between junctions 26 and 27. Similarly, we can enter a second accident known to have taken place between junctions 28 and 30.

Later, we want to query our system to find all motorway accidents in Devon during the period of interest. The boundary between Devon and Somerset crosses the M5 between junctions 26 (which is in Somerset) and 27 (in Devon). The system searches the database and identifies our second accident, between junctions 28 and 30, as definitely located in Devon; but it also finds the first one and returns it as 'possible' (pending further information), since it is known to have occurred on a stretch of road part of which is in Devon. Suppose we now add to the database the information that this accident took place west of the point where the A38 crosses over the M5. We now know that this object is anchored in the stretch of motorway between junctions 26 and 27, and also in the stretch of motorway west of the A38 crossing. By axiom 6, the system can infer that the accident is anchored in the stretch of motorway between the A38 crossing and junction 27, and since this stretch is entirely within Devon, the query would this time return that accident as definitely in Devon.

Suppose now that our query is for accidents in the Exeter area (Exeter is situated close to junction 30 of the M5). To define the Exeter area, we again select Anchoring Information rather than Exact Position. Here is a good case for 'egg-yolk' style inner and outer bounds (interpreted in the more flexible Wittgensteinian manner). We say that the Exeter area is anchored over the city of Exeter itself (some suitable exactly-located administrative boundary is presumably available for this in our database), and anchored in a circle of radius 10 miles, say, centred on the centre of Exeter. Then the system will give us a list of definites and possibles, and the possibles can be annotated with the source of uncertainty: some will be possible because their exactly-known location falls within the outer region but not the inner, some will be possible because their own

[1] The further, non-monotonic, inference that the hill is anchored over some larger connected region incorporating both A and B (e.g., their convex hull) is not licensed by Anchoring. In future work we propose to explore such non-monotonic extensions to the theory.

locations are uncertain but they are anchored in a region which overlaps the definite part of the Exeter area, and some will suffer from both types of uncertainty.

It might be objected that we have no business to be handling data as ill-defined as ‘an accident on the M5 somewhere between junctions 28 and 30’; but of course we may have no choice. This will often be the case with historical data, where greater precision is not available. But even with more up-to-date information such as recent motorway accidents, we may be at the mercy of information supplied by members of the public, who may be less than precise in their recall of the relevant details. If your GIS insists on only having precise information then you either have to make some of the details up (and no doubt this happens quite often) or you must ignore any information that lacks the required precision (this doubtless happens too). Anchoring provides a way of *not being embarrassed* by the uncertainty or vagueness that inevitably attends much of our information.

5.2 Vague location: Butterfly population

Assume that our GIS is used to record information about rare species in a national park. The boundary of the national park is precisely defined, so it can be assigned a precise location. Suppose that a population of Heath Fritillary (*Melitaea athalia*), a rare species of butterfly, lives within the boundary of the national park.[2] By nature, the Heath Fritillary population cannot be assigned a precise location, consisting as it does of many scattered individuals which are constantly moving around. On the other hand, there *are* some definite facts we can record; for example (1) the population is confined within the boundaries of the national park, (2) there are certain point locations within the park where the butterflies have been observed, and (3) there are areas of unsuitable habitat within the park where the butterflies have never been, and are never likely to be observed (see Figure 5). Thus the object P, representing the Heath Fritillary population, is (1) anchored in the precisely delineated region which constitutes the geographical extent of the national park, (2) anchored over each of the individual points where the butterflies have been observed, and (3) anchored outside the regions where the habitat is unsuitable. In the third case, of course, it may be that ‘areas of unsuitable habitat’ are themselves vague and have to be anchored to some more precisely specifiable regions, in which case it will be these regions which P must be anchored outside.

Given all this information, one could, if necessary, construct a region which amounts to an informed guess as to the location of the population, using the GIS to derive this region if desired. But it would be a mistake to enter this region as the main information concerning the location of the population: it is better to maintain in the database only those pieces of information which are definitely known—in this case the three types of anchor information—and use these to make plausible inferences which may well have to be rescinded if more detailed anchoring information becomes available.

[2] In Exmoor, Devon, populations of Heath Fritillary are protected by the Devon Wildlife Trust (http://www.devonwildlifetrust.org), and GIS is used in this work.

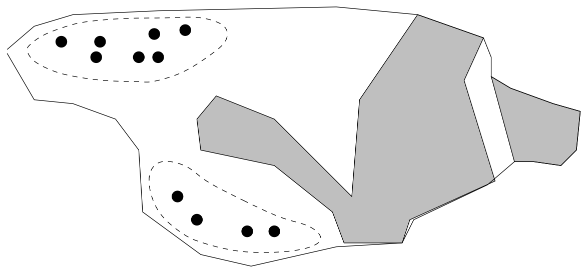


Fig. 5. Butterfly population in a National Park. The black dots represent locations where the butterflies have been observed; the shaded areas represent unsuitable habitat. The broken lines represent a guess as to the region occupied by the butterfly population.

6 Conclusions and Future Work

In this paper we have described a new approach to representing vague or uncertain information concerning spatial location. Locational information about objects in information space is expressed through various anchoring relations which enable us to state exactly what is known regarding the spatial location of an object without forcing us to identify that location with either a region in precise space or any mathematical construct from such regions, such as rough sets or fuzzy sets.

We have described the motivation for Anchoring, begun work on a formal theory of anchoring relations, and illustrated these ideas with examples typical of real-life use of GIS. Much remains to be done. First, the formal theory must be extended to cover anchor relations such as anchoring outside and anchoring alongside, and linked more firmly with both the mereotopological structure of information space and the set-theoretic structure of precise space. Second, systematic methods must be developed for applying Anchoring to a wide range of examples of interest to GIS users. Third, we must explore ways of incorporating the theory into existing spatial data models, allowing the kinds of dialogues we have envisaged to be handled correctly and efficiently.

The problem of how to represent vague location has its roots in a much wider problem, namely the integration of field-based and object-based geographical ontologies or data models. Smith and Mark [5] note that

> [t]he completion of such an ontology is a challenging task, involving serious research challenges .... But its realization would bring significant benefits, because many serious professional users of geographic information, such as pilots, soldiers, and scientists, hikers, wildfire fighters, and naturalists, must communicate about particular parts of the landscapes as if they were objects, while at the same time drawing on the resources of field-based topographic databases.

The relationship between a hill or a valley and the underlying configuration of contours—as in our Greendale and Green Hill examples—is a good example of this. That valleys, hills, and other geomorphological features do not have precise boundaries reflects their nature as features of the underlying, continuous field of land elevation. Just as people identify hills and valleys on a contour map, all we can ask is that our data models ‘point’ to areas of precise space in some way associated with these features. This is achieved in our theory by the use of anchoring relations rather than exact location.

References

1. A. G. Cohn and N. M. Gotts. The ‘egg-yolk’ representation of regions with indeterminate boundaries. In Peter A. Burrough and Andrew U. Frank, editors, *Geographic Objects with Indeterminate Boundaries*, GISDATA 2, pages 171–187. Taylor & Francis, 1996.
2. Kit Fine. Vagueness, truth, and logic. *Synthese*, 30:263–300, 1975.
3. Andrew U. Frank. Spatial ontology: A geographical information point of view. In Oliviero Stock, editor, *Spatial and Temporal Reasoning*. Kluver Academic Publishers, 1997.
4. Z. Pawlak. Rough sets. *International Journal of Computational Information Science*, 11:341–356, 1982.
5. Barry Smith and David M. Mark. Do mountains exist? Towards an ontology of landforms. *Environment and Planning B: Planning and Design*, 30(3):411–237, May 2003.
6. Ludwig Wittgenstein. *Philosophical Remarks*. Basil Blackwell, Oxford, 1975. Edited from his posthumous writings by Rush Rhees and translated into English by Raymond Hargreaves and Roger White.
7. Michael Worboys and Matt Duckham. *GIS: A computing perspective*. CRC Press, Boca Raton, Florida, 2004. Second edition.
8. L. A. Zadeh. Fuzzy sets. *Information and Control*, 8:338–353, 1965.