Presented at 'the Manchester meeting' 2001

Communicating geographic information in context Extended abstract

Michael F. Worboys
Department of Computer Science, Keele University, Staffordshire, UK ST5 5BG
email: michael@cs.keele.ac.uk

Abstract

This paper examines extensions by Dretske, Barwise, Seligman and others to the Shannon-Weaver model of information communication, with special reference to geographic information. Classifications of entities into types are taken as the basis for information repositories, while structure-preserving mappins called infomorphism provide the linkages. A key notion is that of a channel, which can be used to model the notions of contextuality and spatial representations. Context is a key issue in communication, and the approach here is to allow the communication channel to mediate between the contexts of the communicators. In the case of geographic information, particular features of channel, such as accuracy and level of detail, will be examined, and the paper will discuss how vagueness and multi-cnextuality can be handled by the theory.

1 Introduction

This paper is concerned with the transmission of geographic information, from person to person or from system to person. Consider the following dialogue, taking place on Keele campus:

Person A: Where is the bookshop?

Person B: The bookshop is 100m from the chapel.

If we assume some common context between A and B, such as the metric system of length, and knowledge of the position of the chapel (a significant landmark on the Keele campus.), then it is likely that useful information will flow from person A to person B. the context provides the basis for the flow.

Suppose now that person A had met person C rather than B, and the following dialogue ensued.

Person A: Where is the bookshop?

Person C: The bookshop is near the chapel.

In this case, assuming the same common knowledge of the chapel, less useful information about the location of the bookshop is conveyed, as person A must guess person C's notion of what constitutes 'near'. To exaggerate the point, person C might be an airline pilot, and anything within an hour's flight of Manchester Airport is near the Keele chapel. (In fact, there is evidence [9] that people vary their notion of nearness to take account of the context implied by the question, in this case the Keele campus). The context of the dialogue is weaker than in the first conversation between A and B, and correspondingly different information is conveyed.

The point of this preamble is to suggest that context is a key facilitator of information flow, and without context there is no flow. The purpose of this paper is to provide a formal foundation for these ideas, with emphasis on modelling information source and destination as well as the channel that provides the context in which the information can flow.

Underlying the paper is the thesis that any foundational work on geographic information science, with the emphasis on information, should draw on general theories of information science. In particular, discussions of communication of geographic information need to be founded in the general work of information scientists on this topic. The classic work was done by Shannon and Weaver [5, 6]. Several very important concepts for our own work were established at that time, in particular the notion of channel, which has close connections to our own discussion of context. However, the emphasis of this early work was on quantity of information, rather than the more interesting issue regarding its content. Later work by Dretske [4], Barwise and Seligman [1], Devlin [3], and others, gives us an approach to information content. Given the mutable nature of information, it is only a discussion of content that can help us answer such questions as whether the information in two different data sets is the same or different.

This paper explores work on information flow and content, with particular reference to geographic information. We begin by explaining the principle constructs of classification, infomorphism and channel, and show how the idea of a representation can be approached using this apparatus. We draw examples from familiar representations of the geographic world, such as maps, but show how the framework developed here is quite general and may be extended to less familiar spatial representations (e.g. audio cues for in-car navigation systems). The paper concludes by showing how contextual properties of representations, such as accuracy, precision and vagueness may be formulated, and conclude with some further avenues to explore.

2 Shannon's model of information communication

Figure 1 shows the primary elements and relationships of the Shannon-Weaver theory of information communication. In this theory, a channel provides the dependency between source and destination, and this dependency is expressed as a probability that represents the 'surprisal value' (*entropy*) of the information transmitted. Entropy is about the amount of uncertainty that can on average be eliminated by a piece of information. Quantities that relate to the flow of information between source and destination, such as capacity, noise, and equivocation, are measured by probabilities, usually averaging over all possible eventualities. Thus the theory is quantitative, concerned with amounts of information, rather than information content.

Shannon-Weaver theory has been used as a basis for the calculation amounts of information conveyed by a cartographic data source. For example, Clarke [2] defines the 'coordinate digital density function' as a metric based on entropy that allows feature-to-feature and map-to-map comparisons of spatial information quantity.

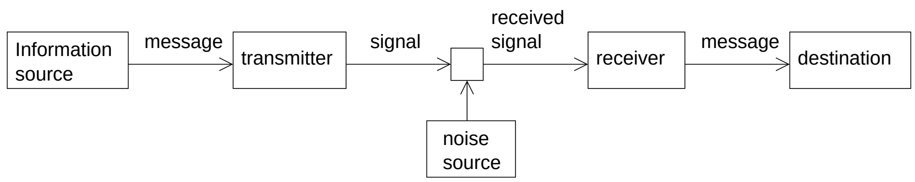

Figure 1: Shannon-Weaver model of information communication

3 Extension to Shannon-Weaver theory

The rationale for the extended theory of information flow was provided by Dretske [4] as the need to discuss the content of information communication, rather than just its quantity. Dretske also provides a critique of the concept of entropy as an appropriate measure of information quantity. His theory is summed up in the following four principles.

1. Information flow results from regularities in a distributed system.
2. Information flow crucially involves types and their particulars.
3. It is by virtue of regularities among connections that information about some components of a distributed system carry information about other components.
4. The regularities of a given distributed system are relative to its analysis in terms of information channels.

Principle 2 suggests classification of entities into types as the basis of a theory of information flow, and it is this notion that is now discussed.

3.1 Classification

A *classification* $\mathbf{A}$ is a triple $\langle A, \Sigma_A, \models_A \rangle$, where A is a set of objects of $\mathbf{A}$ to be classified, called *tokens* of $\mathbf{A}$, Σ_A are the *types* of $\mathbf{A}$ used to classify the tokens, and $\models_A$ the typing relationship. A classification may be shown diagrammatically in figure 2.

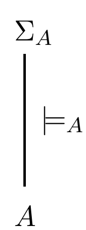

Figure 2: Classification $\mathbf{A}$ of tokens into types

A relevant, simple example of a classification is that of part of the real world into a set of features. In this case the classification $\mathbf{W}$ consists of tokens that are real world objects and types that are features (e.g. churches, roads and mountains) that classify the objects.

Another classification of the same set of tokens is the classification $\mathbf{L}$ of real world objects based upon their location, either using some quantitative classification such as a coordinate system or a qualitative classification based for example on linguistic descriptions of spatial relationships to other objects, such as 'near the centre of town' or 'in England'. Notice that such classifications might well be vague, inaccurate, imprecise and context-dependent. One of the principal reasons for formalizing classification systems in this way is to understand such contextual properties.

Our third example demonstrates that a classification can represent spatial relationships as well as one-place spatial properties. Let $\mathbf{W}^*$ consist not only of tokens that are single real world objects, but also ordered pairs, triples, and so on. Then we can extend out classification to include binary and higher-order spatial relationships, so a spatial relationship such as 'Keele is in Staffordshire', can be formalised as $(\mathsf{Keele}, \mathsf{Staffordshire}) \models_{W^*} \mathsf{inregion}$, where $\mathsf{inregion}$ is the type indicating the spatial relationship.

3.2 Infomorphism

A classification provides a model of an information system, whether human or machine, source, destination or mediator. The next part of the theory provides a mechanism for moving relating information systems having similar informational structure. Figure 3 shows the formal way, known as an *infomorphism*, of structurally relating two classifications. Notice that the functions f^+ from types to types and f^- from tokens to tokens work together contravariantly. In order that classification structure is preserved, the pair of functions f^+ and f^- satisfy the following condition.

$$\text{For each token } b \in B \text{ and } \alpha \in \Sigma_B.\ f^-(b) \models_A \alpha \text{ iff } b \models_B f^+(\alpha) \quad (1)$$

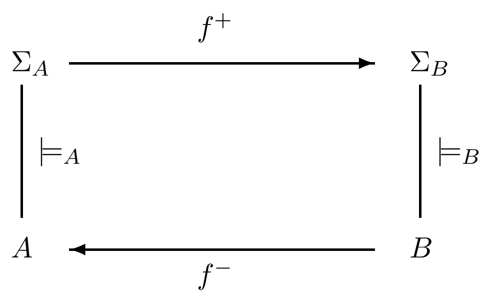

Figure 3: Infomorphism **f** from **A** to **B**

A good example of an infomorphism is provided by the relationship between a map and the part of the real world that it represents. In this case, let A be a set of map elements and $Sigma_A$ be the set of map symbol types to which the map elements may belong. Let B be a set of real world entities and $Sigma_B$ be the set of feature types that the real world entities may belong. Condition 1 gives the infomorphism constraint that ensures that the map is set up correctly in relation to the world. For example is b is an entity in the real world and $f^-(b)$ is a particular red line on a map, then condition 1 gives $f^-(b) \models_A \mathsf{redline}$ iff $b \models_B \mathsf{road}$. So, this infomorphism will be satisfied if map elements corresponding to roads are shown as red lines.

Of course, maps and cartography is more complex than the example indicates. For example, the correspondence between real world and cartographic artifact may be imperfect in several different ways. To models this more complex scenario, and to explore the role of context (in this case the cartographic context of map production), we need the concept of ‘channel’, to be developed next.

3.3 Channel

The concept of channel will be used to model the context in which transmission of information takes place. Figure 4 shows the general configuration in the binary case. The *channel core* C acts as a medium for the transmission of information between A and B. The tokens of C are called *connections*. A connection c *connects* f(c) and g(c).

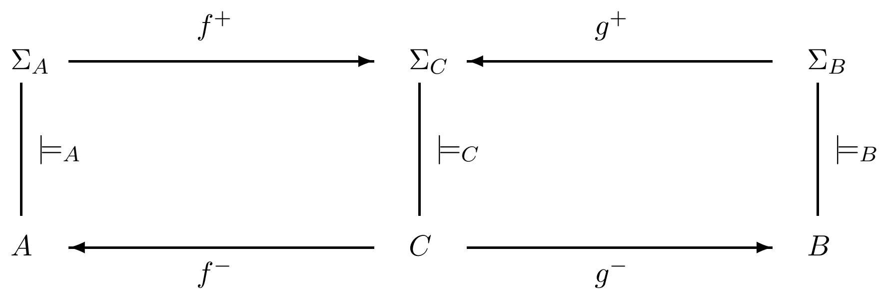

Figure 4: Channel

Channels allow information flow, and in the general theory may be n-ary. However, only binary channels are needed for this discussion. We will return to channels in section 4, when we will see how they are used to make precise the notion of a representation, and in section 6 where we see channels used to model multi-contextual situations.

3.4 Local logics

Each classification has a logic associated with its typing system, called its *local logic*. As we will see later, it is the local logic that holds the informational structure of the classification system, and movements between local logics indicate the information flow. Local logics allow inferences in the type system, and will be represented using sequents, in the style of Gentzen. For types α, β, the sequent $\alpha \vdash \beta$ indicates that the inference from α to β is valid. For example, suppose that the types were real world features, then the sequent house $\vdash$ building indicates that houses are buildings. The sequents in the local logic are called *constraints*.

Consider the classification $\mathbf{A} = \langle A, \Sigma_A, \models_A \rangle$. The standard case is that, for types $\alpha, \beta \in \Sigma_A$, the sequent $\alpha \vdash \beta$ is valid iff $\forall a \in A.\ a \models_A \alpha$ implies $b \models_A \beta$. So, for our example, house $\vdash$ building would be valid if all real world objects in the system are buildings. A token is called *normal* if it satisfies all constraints in the logic.

In general, not all tokens may be normal. In our example, suppose that there exists a real world object that is a house but not a building. These non-normal exceptions may be used to model kinds of imperfection in the informational structure and flows, such as inaccuracy and incompleteness. They are considered in section 5

4 Representation

The reader who has been patient enough to follow us to this point will we hope now begin to see some rewards for the pain of all this formal machinery, in that it allows us to capture with considerable richness the notion of *representation*. The notion of representation is central to geographic information science, with the map as its paradigm.

Formally, a *representation system* consists of a binary channel (see subsection 3.3), on the core of which is imposed a local logic. Consider the binary channel shown in figure 4. Suppose that the core **C** has a local logic associated with its types. Call the classification **A** the *source* and **B** the *target* of the representation system.

The *representations* are the tokens of A. We say that a is a is a representation of b if there is a connection from a to b. A type α in the source *indicates* type β in the target if $f(S) \vdash g(b)$ in the local logic of the core.

As an example, let **A** be the classification of map elements and **B** be a set of real world objects and features. Then, the representations are the map objects, and map objects are representing real world objects. Let the core **C** be the cartographic process of surveying real world objects and associating them with map elements. The cartographic process might contain rules such as associating a red line on the map with a road on the ground. So a constraint in the local logic might be $\mathsf{draw\ red\ line} \vdash \mathsf{surveyed\ road}$. In that case, using our definition of 'indicates', *red lines indicate roads*.

The treatment is necessarily brief, but shows the possibilities of formally modelling flows between geographic representations, e.g. between graphical and textual representations of geographic phenomena.

5 Properties of representations

5.1 Accuracy

A representation is accurate if there is an appropriate correspondence between the representation and the target domain being represented. The theory of representations developed in the preceding section gives us a succinct way of characterizing accuracy. We will use the representation system provided by the channel in figure 4, with a local logic imposed upon its core. The token $a \in A$ is an accurate representation of token $b \in B$ if firstly a is a representation of b and secondly the token that connects a and b is normal.

Returning to our 'maps of the real world' example, with the cartographic local logic containing the single constraint $\mathsf{draw\ red\ line} \vdash \mathsf{surveyed\ road}$. Consider a representation of a road as a red line on a map. Then, the token c that connects this representation to the real world object is normal, as it satisfies all constraints in the local logic. So, as expected, the representation is accurate.

On the other hand, consider a representation of a representation of a lake as a red line on a map. Then, the token c that connects this representation to the real world object is not normal, as it does not satisfy the constraint $\mathsf{draw\ red\ line} \vdash \mathsf{surveyed\ road}$. So, as expected, the representation is not accurate.

While this example is simple, it points the way to analysis of the accuracy of complex representations (maybe non-graphical), in which local logics are more interesting.

5.2 Precision

Precision refers to the level of detail that a representation captures from the target domain. It is independent of accuracy. Objects that are discernibly distinct in the target domain may or may not be differentiated in the representation domain, resulting in more or less precision. In terms of the theory developed here, differentiation of tokens occurs by means of their types, so imprecision will result if two tokens with distinct type sets in the target domain have the same type sets in the representation domain. There are two cases to consider here:

identity indiscernibility The representation fuses the identities of objects in the target domain.

typological indiscernibility The representation fuses the types of objects in the target domain.

In our mapping example, a case of identity indiscernibility occurs when a collection of distinct buildings (church, pub and school) in the world are merged into a single map object. Typological indiscernibility might occur when although the objects are represented as distinct map elements, they are assigned the same type, say, building.

Using the representation system provided by the channel in figure 4, with a local logic imposed upon its core, we can express formally these two kinds of imprecision. Let tokens $b, b' \in B$ be distinct elements of the target domain. Then, b and b' are *identity indiscernible* in the representation system if all representations of b are representations of b', and conversely. Also, b and b' are *typologically indiscernible* if all representations a of b and a' of b' have the same type set, so that $\forall \alpha.\ a \models \alpha$ iff $a' \models \alpha$.

A more refined analysis of imprecision would include such subcases as indiscernibility due to survey and indiscernibility due to cartographic rules, and this analysis is the subject of ongoing work.

6 Vagueness and contextuality

Vague predicates admit borderline cases for which it is not clear whether the predicate is true of false . Vague entities have parts for which it is unclear whether or not the part is part of the entity [8, 7]. There are many examples of vagueness in geographic phenomena. For example, a vague predicate is 'nearness' and vague entities are a mountain or 'the North of England'. The sorites paradox shows that vagueness is not easy to handle using classical reasoning approaches, and many more or less successful attempts have been made to provide a principled account of the semantics of vagueness, and reasoning with vague notions. Vague notions often have associations with context dependence and subjectivity. What may count as the North of England for me may be different for you (*subjectivity*). What may count as near home may depend on whether I am thinking of walking of taking the space shuttle *context dependence*.

Even though classical approaches to reasoning have difficulties with vagueness, that does not mean that it is impossible to have information flow with vague ideas. For example, if you phone me to tell me that you are currently in the North of England, and I know that I am a few miles to the North of you and in England, then I can reply to you that I am also in the North of England, *even if our notions of the North of England are quite different*

We will consider how vagueness may be discussed using the theoretical ideas so far developed, and briefly show how the theory can be used to model information exchange in the cases where the context involves vague notions. We will use as a running example the vague spatial relation of nearness on Keele campus [9] introduced at the outset of the paper. The point that we are emphasizing here is that the theory of information flow allows us to reason with context dependent ideas, of which nearness is a good example.

Suppose we have two people, A and B, with their own notions of nearness. We will set up classification systems connected by a channel, as in figure 4, to model the nearness notions of A and B, whose contexts are mediated by the channel **C**.

The classification systems **A** and **B**, modelling the nearness notions of A and B, respectively, are set up as follows. Let the tokens of A be locations on Keele campus. Let Σ_A consist of the type: near. The idea of the **A** classification system is that:

$$\forall a = (l, m) \in A.\ \ a \models \mathsf{near} \text{ iff person A thinks } l \text{ is near the chapel} \tag{2}$$

We have a similar set up for the classification system **B**: Let the tokens of B be locations on Keele campus and Σ_B consist of the type near. Then:

$$\forall b = (l, m) \in B.\ \ b \models \mathsf{near} \text{ iff person B thinks } l \text{ is near the chapel} \tag{3}$$

Of course, because of the subjectivity and context dependence of nearness, these two classification systems will in general be different. What they have in common is expressed by the mediating channel **C**. To construct channel **C**, we need some preliminary definitions.

Definition
For points $x, y \in \Re^2$. $\mathsf{line}(x, y) = \{z \in \Re^2 \mid z \text{ is on the closed straight line segment } xy\}$

Definition
For a point $x \in \Re^2$, a *neighbourhood of* x is a region S of $\Re^2$ with the following properties:

1. $x \in S$
2. If $y \in S$, then $\mathsf{line}(x, y) \subseteq S$

The idea is that neighbourhoods in $\Re^2$ model the idea of a region of locations that are near to a particular location. The second property of neighbourhoods is an attempt to model the idea of 'downward closure' with respect to nearness. We can note that every neighbourhood is connected (but not every connected region is a neighbourhood). We are now ready to define the mediating classification system **C**.

Let the tokens of **C** be points $x \in \Re^2$. Let Σ_C consist all the neighbourhoods of the origin $\Re^2$.

If we assume that person A has a 'consistent' notion of nearness, then it will be possible to set up an infomorphism from **A** to **C**. Similarly, we should be able to set the infomorphism from **B** to **C** (Note that in general it is *not* possible to set up an infomorphism directly from **A** to **B**). We now have a configuration in which limited information can flow between person A and person B, using the channel **C** that expresses some common geometrical understanding about nearness.

Once again, this construction is merely a sketch, indicating the possibilities of these ideas. It has at least two shortcomings, firstly missing out the full details of the construction through lack of space, but more importantly lacking capacity to model more complex nearness ideas (e.g. location l is nearer the chapel than location l'). A fuller development will be given in the final paper.

7 Conclusions and future directions

This extended abstract has sketched out the application of recent work on information flow to geographic information. It has argued for an approach that would handle information content rather than quantity. The theory of information flow, due to Dretske, Barwise, Seligman and others, has been developed with reference to examples from geographic information. We have shown how it can be used to examine a range of issues involved in representation of spatial phenomena, including accuracy and precision. We have also used the theory to discuss the key issue of context, and shown how multi-contextual situations can be approached.

While this is only a beginning, it seems clear to us that the theory developed here can be used to address several key questions in geographic information science, including the following.

- How can we determine the informational content of a spatial dataset?
- Can we determine whether two spatial datasets contain the same information?
- How can we treat in detail the data quality issues arising in spatial data representations?
- How can we ensure that the informational content of two different representations (say, graphic and linguistic) of geographic phenomena are equivalent in information content?
- How can we represent and reason with information arising from vague representations?

The reader will realize that some of these questions are ill-formulated. For example, as we have seen the information content of an information source depends on the channel and destination. So, one of the key issues that arises is the need for further work on context, how it is specified, and how it affects information flow.

References

[1] J. Barwise and J. Seligman. *Information Flow.* Cambridge University Press, Cambridge, England, 1997.

[2] K.C. Clarke. The coordinate digit density function as a tool for analytical cartography. In *Proceedings of GIScience2000, Savannah, GA*, 2000.

[3] K. Devlin. *Logic and Information.* Cambridge University Press, Cambridge, England, 1991.

[4] F. Dretske. *Knowledge and the Flow of Information.* Bradford Brooks, MIT Press, Cambridge, MA, 1981.

[5] C.E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423, 623–656, 1948.

[6] C.E. Shannon and W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, 1949.

[7] A.C. Varzi. Vagueness in geography. *Philosophy and Geography*, 2001. Forthcoming.

[8] T. Williamson. *Vagueness*. Routledge, London, 1994.

[9] M.F. Worboys. Nearness relations in environmental space. *International Journal of Geographical Information Science*, 2001. In print.