Applying semantic concepts to analysing media and modalities

Keith Stenning Robert Inder Irene Neilson*

Abstract

Our long term goal is an understanding of human communication in terms which would provide the basis for rational design. The kernel would be a theory of the cognitive consequences of allocating the same information to different media and modalities, based on the user's information processing characterised in computational terms. Our theory of the cognitive consequences of media/modality allocation starts from an analysis of differences in logical expressiveness of graphical and linguistic representations (Stenning & Oberlander (1994, in press)).

This semantic approach requires conceptualisations of *medium* and *modality* that can be related to representation systems. We propose that media are the physical/perceptual aspects of representations; modalities are classes of interpretation function which map media onto meanings. These interpretations of the terms contrast with existing HCI usage. The curtailments of abstraction observed in graphics arise from interactions between medium and modality.

Graphical modalities are distinguished from sentential languages by the nature of their interpretation functions. A hierarchy of expressiveness of interpretations of graphics is defined, and compared with interpretations of sentential languages. Using the expressiveness of representations to predict their cognitive properties also requires reference to the availability of constraints of their interpretation to users. Further contrasts between graphics and language emerge in the availability of their constraints.

Three example domains of graphical representations are analysed from this perspective—matrix graphics; logic diagrams; and semantic networks. Some empirical evidence of the usability of these notations is reviewed as evidence that the proposed conceptualisation offers powerful generalisations.

*The support of the Economic and Social Research Council UK (ESRC) is gratefully acknowledged. The work was partly supported by the ESRC funded Human Communication Research Centre. We also wish to acknowledge partial support from Grants # 9018050 from the JCI in Cognitive Science and HCI and # GR/J586129 from SERC/DTI

1 Introduction

Interest in multi-media and multi-modal communication systems is currently being driven by the advent of new technologies for human computer interaction. New ways are being developed for us to communicate with machines, and with each other. These innovations present both problems and opportunities for the design of information technology, since they allow designers to use poorer, as well as better, combinations of media in their designs. Designers trying to navigate among the peaks and troughs of this new terrain need principled grounds for evaluating different combinations of media and modalities for expressing the same information. We seek to understand human communication in terms that will provide those grounds, and thus form a basis for rational IT design. But solutions to this practical problem are also a central part of the theoretical enterprise of cognitive science. The ability to construct and interpret new representations for new domains is a distinctively human achievement. Even mundane human reasoning consists largely of selecting appropriate representations which make reasoning tractable for each task that presents itself. This work on the cognitive consequences of allocating the same information to different media and modalities develops work such as Larkin & Simon 1987, and Palmer 1978 in a more semantically oriented direction (Stenning & Oberlander (in press)).

Our work is based on the application of techniques developed for the analysis of languages to the analysis of multi-media communication. In turn, we believe that the study of multi-media communication will produce insights into the fundamentals of linguistic communication and give rise to a richer theoretical framework encompassing both sets of phenomena.

A semantic approach to understanding multi-media communication relies on the logical concept of a *representation system*—an explicit definition of what forms representations within a system can take, coupled with an *interpretation function* which maps elements of a representation onto the entities and relationships they represent. This idea leads us to give very specific meanings to the terms “medium” and “modality”: *media* are categories of representing entities and *modalities* are categories of interpretation function. These interpretations conflict with some of the many (and sometimes inconsistent) uses of these terms in the HCI and multi-media communications literature. One of our first goals, therefore, is to clarify these conflicts and explain our chosen interpretation of the terms.

Our analysis is based on an account of the semantics of graphical representations—of the way their parts combine to convey meaning—and we seek to base empirical predictions of the usability of representations on this analysis of their expressive power, as defined in logical/computational theory. Our comparison of the expressive power of graphical and linguistic representations

(Stenning & Oberlander (1992)). suggests that graphical representations cannot express arbitrary abstractions. A form of graphic that cannot express an abstraction necessary for some task will prove pathological, and will be better replaced by a more expressive language. But where a graphic can express the necessary abstractions, we predict that its weak expressiveness will make it a more efficient representation for performing the necessary inferences. This is so just because weakly expressive systems are computationally more tractable.

Suggesting that forms of representation should be compared on the basis of their logical expressiveness raises an immediate question for a theory of modality allocation. If a certain sentential system is found to have the same logical expressive power as a particular graphical system, such a theory appears to suggest that they should have the same cognitive properties. So why should graphical systems be preferable to weak sentential ones? The answer lies in the *cognitive availability* of the limits of the expressive power. Availability determines whether the weakness of the representation can be exploited. We will offer examples of how graphical notations 'wear their weakness on their sleeves': the nature of these notations makes their weakness available for exploitation. Weak sentential systems, in contrast, lack this property—nothing in the arrangement of symbols on a page constrains the expressive power of the systems. Their weakness cannot be exploited without considerable learning, although where such learning does occur, we would predict it would greatly reduce the discrepancy between equivalent sentential and graphical systems.

Although we discuss only the expressive power of information presentations, we acknowledge that many other factors combine to determine their cognitive characteristics—for example, low-level perceptual properties, aesthetic appeal, ... We do not claim that the expressive properties of representations are the whole of the story, but merely that they must form the core of an account of a cognitive theory, and that they have been given less explicit attention than they deserve.

The next section of the paper sets out our theoretical perspective in detail. Our proposed interpretation of *medium* and *modality* and their relation to representation systems is elaborated and related to previous work on the effects of assigning information to different media/modalities. The need for conceptual clarification of the terms 'medium' and 'modality' and their relations to knowledge representation systems is illustrated by reference to inconsistencies and conflicting influences on the usage of these terms in the literature. Particular attention is paid to differences between the system design and cognitive perspectives on these concepts.

Section 3 analyses the basis for the contrast between two exemplary modalities—

graphics and language. We survey several properties of graphical interpretations which define an ordering on the expressiveness of systems. We focus especially on the enforcement of information presentation by some styles of interpretation. Information enforcement emerges as a critical characteristic of graphical modalities, whereas the nature of linguistic interpretation means that linguistic modalities always can avoid information enforcement. Information enforcement can be related to weak expressiveness and inferential tractability.

In order to link expressiveness to usability it is essential to consider what constraints on expression are *available* to which users with what background information. The observation that the constraints on expression of graphical representations are generally more available to the relatively naive user can be explained in terms of our characterisation of graphical modalities. Thus the account explains why graphical representations are often easier to interpret for novices than sentential equivalents, but without resorting to the naive theory that graphical semantics is transparent, and 'realistic'.

Our analysis of graphical semantics ranges examples along a dimension of expressiveness ranging from primitive graphics to powerfull languages. In Section 4, we illustrate our conceptualisation by presenting analyses of representations which are drawn from widely different parts of the dimension: matrix graphics, elementary logic diagrams, and semantic networks. In each case we illustrate how these systems can be analysed in terms of their constraints to reveal their expressiveness, as defined in logical/computational terms. Some existing empirical evidence on the usability of these different types of representation is discussed, and gaps in current empirical investigation pointed out.

Our concept of modality divides up the space of representations somewhat differently from pre-theoretical intuition. In particular, the more logically expressive interpretations of some kinds of abstract graphics such as semantic networks have the characteristics of linguistic modalities. We believe that this divergence from pre-theoretical intuition may have a role in explaining the mixed results in observation of the usability of these notations. This is an example of the sort of regimentation of of psychological data that one might expect to issue from this theoretical approach.

2 Semantic analysis, media and modalities

Because HCI research and computer system design brings together workers from several disciplines, the terms "medium" and "modality" are used in several different ways that each appear reasonable against their own background. Clarification of the way we use these terms is essential, but we do so in the belief that no definition of these terms can be faithful to all usages in this multi-disciplinary field. For example, the psychological terminology of *sensory modalities* is close to the computer scientists use of *media*. Having set out our use of these terms, we relate it to some of the other ways they are used in the literature.

2.1 Media and modalities meet in representation systems

We start by presenting some intuitions as landmarks against which we believe definitions of "medium" and "modality" should encompass. For each term we present two intuitions, one which underlies distinctions between cases, and one which underlies abstraction over cases.

1. The very same physical representation can constitute more than one medium when perceived by different sensory systems. Thus embossed diagrams can be 'read' by eye or by finger, and constitute different media in the two cases. [basis of distinction for media]
2. At some useful level of abstraction, different representations that are visually perceived in two dimensions can be regarded as the same medium. [basis of abstraction for media]
3. Diagrams and texts are different modalities even though they may be represented in overlapping ranges of media. [basis of individuation for modalities]
4. Some presentations of language in acoustic and in printed form are examples of information in the *same* modality presented in *different* media [basis of abstraction for modalities]

The first intuition emphasises the perceptual basis of our concept of medium. Media are about physical energy impinging on users, but a rational theory must distinguish them by how physical energy is perceived. The second makes the same point for abstractions. We can regard computer screens, paper, vapour trails in the sky, and other physical expressions of 'writing/drawing' as a single abstract medium, because they share a perceptual basis. Media can of course be more finely divided, and our emphasis on the abstraction inherent in our conceptualisation does not deny the cognitive significance of finer divisions within these general categories.

Our third intuition which underlies distinctions between modalities is that there are clearly different ways of expressing information in the same medium, and these differences do not depend on the medium itself. For example, both diagrams and sentences can be presented in visual and tactile media, though

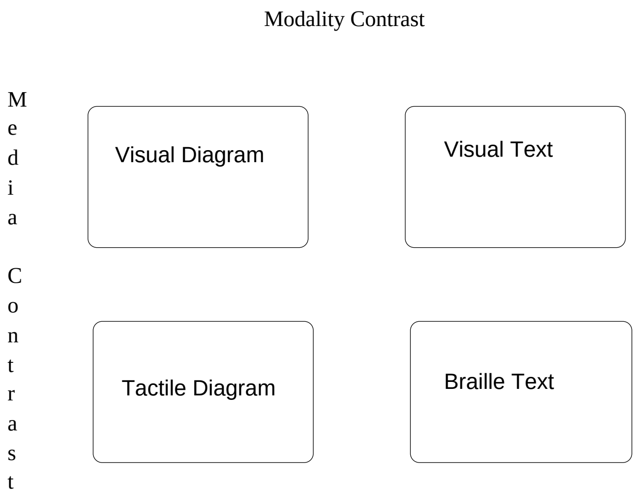

Figure 1: The concepts of *medium* and *modality* applied to tactile and visual diagrams and texts

their perceptual basis is then quite different. But the *difference* between diagrams and language in each medium is invariant, and it is the basis of this invariant contrast which constitutes modality. Finally, our fourth intuition provides an example of what the basis of this invariance across media might be—language is a modality of information presentation which is invariant across different the visual and acoustic media.

As we shall see below, language is a modality peculiarly free to find expression in different media. Most other modalities are more circumscribed in their media appearances, and an account of the styles of semantics which constitute each modality must explain why this is so. Figure 1 summarises by example the distinctions between a representation's "medium", which is based on how it is perceived, and its "modality", which is based on how it is interpreted.

Cognitive theories of information presentation must explain differences between information displays in both *media-variance* and *modality-variance*. Our focus on the expressive power of representations emphasises the latter, which has been relatively neglected. By defining modalities as categories of interpretation we can conceptualise cognitive differences between representations which are distinguished from perceptual properties of media.

2.2 Interpretation functions

A representation system consists of a set of physically/perceptually defined representations mapped on to a set of meanings by an interpretation function. In the best understood (and most regimented) examples of representation systems—logics and computer languages—a finite alphabet is used to define a vocabulary which is concatenated into a set of sentences according to syntactic rules. The interpretation function for such a language gives the vocabulary items meaning by mapping them onto (sets of) members of a specified domain. Finally, the statements are given meaning by semantic rules that parallel the syntactic rules for forming them. Thus the physical/perceptual and the interpretational aspects of representation are brought together in interpretation functions.

But languages are a special type of representation system because they are almost independent of their physical form. The only constraint on an alphabet is that the *functional* distinctions and relationships between members are preserved. Thus the shapes of the letters or the acoustic properties of the sounds do not matter, provided one can tell which letter is which, and which follows which. This insight, from Saussure (1916), explains why a language is a modality which is independent of any medium of expression. This independence of physical medium is not universal. Graphical representation systems are typically interpreted by direct, uniform mappings between relations in the representation and relations in the represented domain. This style of interpretation limits the scope for substituting one physical form of representation for another, because the new form has to provide relations which, within the capabilities of the human perceptual system, precisely mirror those in the old form. This is, for example, why there are visual and tactile diagrams but there is no acoustic equivalent.

Modalities are to be defined in terms of kinds of interpretation functions whereas media are to be understood as physical/perceptual implementations of the representations of those systems. It is the concept of representational system that brings defines relations between media and modalities. Although choice of media may constrain choice of modality, there is still wide choice of kind of interpretation with a medium, and this choice plays a powerful role in determining usability.

2.3 Influences on current usage of terms in HCI, or, Multimedia is what we have—Multi-modality is what we want

Our use of the "medium" and "modality" conflicts with common usages in HCI. The practice of describing systems which can handle information in a variety of forms—typically combining text with video and audio—as *multi-media* is close to our usage. But few of these systems can relate information in different media: there is no *fusion* of data. Thus multi-modality has become the banner of designers who want to create systems which can interrelate information received though different media on the basis of meaning (see e.g. Coutaz 1992) . But this means that any multi-modal system must, by this definition, be multi-media.

We believe that modality and medium are best defined independently. While the above HCI definition would class a system that can integrate speech and text as multi-modal, we would not. We see them as examples of a single modality—language—expressed in different media. On the other hand, since we regard text and diagrams as different modalities, we would ascribe multi-modality to a system which could interrelate facts in a text with information extracted from a diagram, even though that system might involve only a single medium. This is conspicuously something that current systems generally cannot do: the issue of machine access to meaning is just as difficult within a single medium as between media.

There are obvious reasons for those engaged in constructing software systems to treat speech and written language as separate. Speech is very hard to process by machine, and the procedures for translating between speech and text are necessarily complex, and possibly incomplete. This is because it is, in general, impossible to establish mappings between the two at the level of "alphabetic" correspondences. Interpreting either properly requires taking account of features of extended portions of the representation, such as intonation and stress in speech, and punctuation and layout in text. For this reason, lexically equivalent samples of speech and text may not be equivalent in meaning. Nevertheless, we contend that when speech and writing samples are deemed equivalent they then differ in medium, not modality.

The computer science approach in which media are *physically* defined tends to underplay the abstraction that is involved in the concept of medium. A screen and a sheet of paper are very different physically but we need to abstract over these differences to capture what is in common between 2-D visual displays. Exactly what abstraction is appropriate will depend on what range of possibilities are actually being exploited in using the displays. In one context, the fact that the computer screen is capable of dynamic and unobtrusive change, whereas the paper is a static display, may be of vital importance. In another context where the dynamic characteristics of the screen are only used for 'scrolling' through a document, the difference may be unimportant. We should not generalise from the legitimate interests of system builders and assume that *medium* is a simple physical concept that does not have an ineluctably perceptual and therefore cognitive component.

The psychological approach to conceptualising media and modalities also runs counter to our usage. Sensory *modalities* provide the physical/perceptual access of user to representation. Psychology does not have a systematic use of the term media, nor a systematic term for what we have called modalities of information presentation. Near reversal of terminology relative to computer science appears to result from the difference between experimental analytic as opposed to engineering synthetic methodologies in the two disciplines. The result in psychology has been neglect of systematic study of the nature of styles of interpretation of different representation systems.

Tensions between ways the terms "medium" and "modality" are used make it unlikely that any definitions of medium, multi-media, modality and multimodality will gain universal acceptance in our multi-disciplinary field. For our present purposes, what is critical is to relate both medium and modality to the concepts of *representational system* and *interpretation function*. When we do this (below) it emerges that the concepts we need for our theory lead to unexpected classifications for many currently available systems. Given the pragmatic influences on the current uses of the terms, this seems to be a price worth paying for a perspicuous theory.

3 Graphical and Linguistic Modalities

Modalities are kinds of interpretation function relating representing things to represented things. That different kinds of interpretation function have different consequences for the tractability of computations based on them is a commonplace of computer science. In this section, our aim is to develop a theory for predicting the cognitive characteristics of representation systems based on the nature of their interpretation functions, and on users' access to them.

The contrast between graphical and linguistic modalities is perhaps the touchstone modality distinction and has received much attention from philosophers throughout the history of semantics. We here assume that at least the outlines of an adequate account of this contrast will transfer to treatments of other modalities.

We take the basic structure of the interpretation functions for sentential languages, sketched briefly in the previous section, to be relatively well understood. We therefore focus initially on accounts of the semantics of graphical representations and then return to compare what we find with the situation with sentential languages.

The traditional view of graphical representations is that they represent by virtue of their resemblance to what they represent. As Goodman (1968) says "Vestiges of this view, with assorted refinements, persist in most writings on representation" (page 3). As Goodman elegantly illustrates, this view suffers from basic problems as an account of even any subspecies of representation. Closeness of resemblance is neither necessary nor sufficient for establishing *denotation*, the core of representation. Resemblance is a reflexive relation whereas representation is not. Things are most closely resembled by themselves but are rarely represented by themselves. Tokens of types may resemble each other as closely as we wish but rarely represent each other—each of the identical cars rolling off a production line resembles all the others but does not, in general, represent them.

Global properties such as resemblance are unsatisfactory for theory. They prove hard to apply in detail to actual cases because they fail to distinguish between:

1. 'lexical' and 'propositional' semantics (roughly between unstructured and structured semantic relations);
2. properties of *systems* of representation and *token* representations;
3. what various different classes of user may or may not know about interpetations;
4. different grades of correspondence between representation and represented; and
5. different interpreting relations operating on different aspects of the same complex representation system.

We take each of these issues in turn:

1) The contrast between icons and words is one of the points at which modality differences arise in HCI discussions most frequently. Paradoxically, the semantics of most computer icons is essentially lexical semantics. The relevant issues are about how a visual pattern denotes a class of objects, rather than about how a configuration of parts of a pattern expresses a relation between elements in a domain. There are differences between how icons do this and how words do this, differences that have an impact on memorability, but semantically there is little difference because the lexical relation of denotation has relatively little structure. In what follows we will mainly be concerned with propositional semantics.

2) The contrast between the level of representation *system* and the level of *individual* representations can be illustrated by Figure 2. Figure 2 i) is intended to be interpreted as a token representation from a system of representations in which placement of square icons in a field represents the size of the elements they denote—say, larger to the left. Other individual representations from this system would have different numbers of icons with different labels in different arrangements. There is an isomorphism between every individual representation of the system, and the sizes of any set of truthfully represented objects. For such an inexpressive system, there is

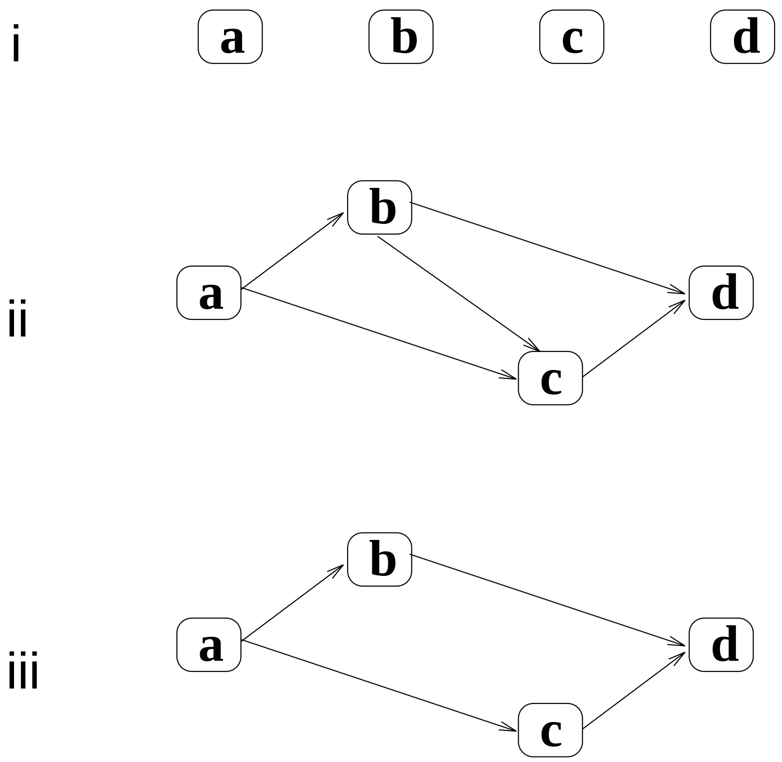

Figure 2: Three diagrammatic systems with potentials for interpretations of contrasting expressiveness

also a system-level isomorphism between the left-of representing relation and the bigger-than represented relation.

However, for the system of representation from which Figure 2 ii) and iii) are drawn from, the system-level and the token-level relations diverge. The intended interpretation of these diagrams is that the representing relation (arrow-from-x-to-y) represents that the thing denoted by x is bigger than the thing denoted by y[1]. For the individual representation in Figure 2 ii) and any of its models, there is still a complete correspondence between the arrow relation and the bigger than relation—if there is an arrow from 'a' to 'b' then a is bigger than b, *and vica versa*. But this correspondence breaks down at the system-level as witnessed in iii) where the diagram does not represent whether a is bigger than, equal to, or smaller than b. Individual representations *may* correspond completely to their models, but other representations of the same system do not.

3) This distinction between token-level and system-level relations is critical for applications of semantic theory to the cognitive properties of representations. Users often have to infer the scope of systems of representation and the properties of their interpretation from examples. In doing so they can extrapolate from their experience in different ways. If we only knew of Figure 2 ii), we might suppose that it was drawn from a representation system in which all individual representations had sets of icons which were completely and consistently connected, thus excluding iii), and restoring a system-level correspondence between 'arrow' and 'greater than'. Since knowledge of system will prove a critical determinant of cognitive properties of modalities, we require a treatment which clearly distinguishes system- and token-level correspondences.

Approaching graphical modalities in terms of the *enforcement of information* by representation systems has just this effect of distinguishing system and token levels of analysis. Enforcement only makes sense at a system level. As far as token representations go, if it is possible to represent a relation between some elements, it is *possible* to represent it between all elements. The issue is whether representation is *avoidable*. Are there other tokens of the same system which fail to enforce the expression of the information?

4) Distinguishing token- and system-levels of analysis reveals the need to distinguish different grades of correspondence between representations and what they represent. How does the enforcement of information arise in systems of representation like that in Figure 2 i) but not in the system from which ii) and iii) are drawn? Why can the second system express abstractions (assert only partial orderings of elements) while the first must

[1] More precisely a correspondence between 'x is connected to y by a directed arrow-path, and 'x is greater than y.

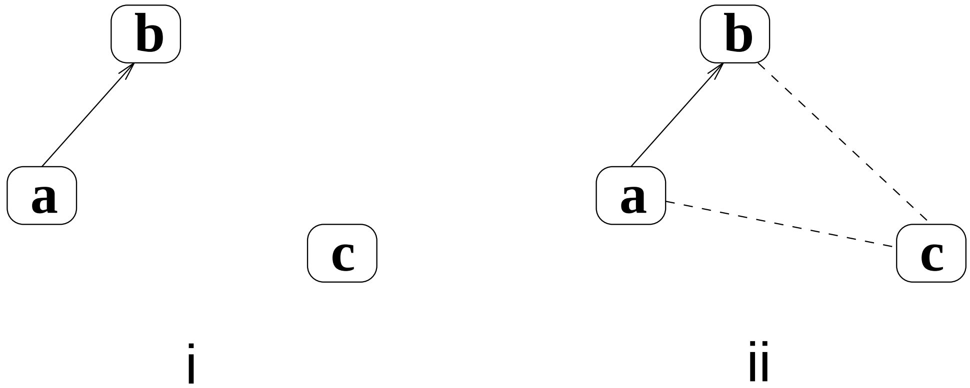

Figure 3: The diagram in i is inconsistent under an exhaustive interpretation of arrows. The lack of an arrow implies equivalence, which means that **c** is equivalent to both **a** and **b**, which the arrow shows are non-equivalent. The diagram could only be consistent if one of the two dashed lines in ii was in fact an arrow.

determine relative size relations between every pair of elements represented? The answers reveal that far from there being a single type of 'resemblance' between a graphic and what it represents, there is a hierarchy of relations.

Consider the case in which we take the representation system of Figure 2 ii) and iii) but we not only interpret the arrow relation as meaning 'greater than', but also interpret its failure to hold as meaning 'not greater than'. Under this interpretation, Figure 2 iii) *does* determine size relations between all pairs of represented elements, specifically that $b = c$.

What determines whether information about size relations is enforced is the *exhaustiveness* of the interpretation of the representing relations. We define the interpretation of a representing relation to be *exhaustive* if the negation of the representing relation is interpreted as indicating the negation of the relation's interpretation. Under our original interpretation for Figure 2 iii), the absence of an arrow from b to c was not interpreted as meaning that the negation of the interpreting relation held. When this stronger condition is imposed, information enforcement appears.

Although this new interpretation, restores the system-level information enforcement noted in Figure 2 i), the new system retains a property of its old interpretation which contrasts with i). Whereas the interpetation of i) guarantees consistency for all representations drawn from the system, this is not true of either the old or the new interpretation of ii) and iii). A diagram such as Figure 3 has no model under this exhaustive interpretation, because $a = b$ and $a = c$, but $b > c$.

Information enforcement is a feature of interpretations which are exhaustive in the sense defined above. This condition is less restrictive than the full isomorphism of 2 i) which results from a match between the logical properties of representing and represented relations. The representing relation 'left of' is anti-symmetrical, anti-reflexive and transitive, just as is the represented relation 'greater than'. The result is that the system of Figure 2 i), under the interpretation described, cannot represent inconsistent sets of propositions. This property of the isomorphic style of interpretation we call *self-consistency* (see Barwise op. cit.). Mere exhaustive interpretation does not yield this property. For example, the representing relation of 'having an arrow from x to y' is non-symmetrical, non-reflexive and non-transitive. When this representing relation is exhaustively interpreted, far from being self-consistent, most of the system's diagrams have no model.

In emphasising the importance of distinguishing different types of correspondence, we have distinguished three grades of interpretation ranged between representing and represented relations. In the least expressive, there is a correspondence between the logical properties of the representing and represented relations which results in a self-consistency property of these systems; the next more expressive system results from relaxing the correspondence of logical properties but maintaining exhaustive interpretation in which the negation of a representing relation is interpreted as the negation of the represented relation; the third more expressive kind of system relaxes even this constraint. Below, when we consider linguistic semantics, we will make a further sub-division of the last category. So far we have only considered systems in which there is a single representing relation, relating a single category of elements. When links become typed (or equivalently nodes become typed), then even more expressive systems result.

5) Finally, our fifth difficulty with analysing graphical semantics through such global properties as similarity, is the problem that in complex real representation systems, these different types of correspondence coexist in the same system. This can be illustrated by further development of our current examples. We have so far tacitly supposed that the letters in the icons in the figures are logical constants interpreted according to a single-names axiom (things only have a unique name) ie. we have supposed that $a \neq b \neq c \neq d$. This tendency to interpret labels in this way in all the diagrams amounts to interpreting the sameness-of-position relation as isomorphic with the identity relation. Note that this is just as true of Figure 2 iii) when the arrow relation is *not* exhaustively interpreted, as it is of Figure 2 i) when it is. This interpretation of position as identity is a particularly pervasive characteristic of graphics which we return to below under the heading of *token referential* systems when we contrast linguistic and graphical semantics.

For the present, we note that representation systems sometimes interpret one relation in a way which enforces information, and another relation in a way that does not. We will see this illustrated in real graphical systems in the next section, and for languages later in this one. This is yet one further reason for understanding graphical semantics in terms of relations less global than resemblance or similarity.

3.1 Availability of constraints

Both information enforcement resulting from the exhaustive interpretation of representing relations, and the even stronger property of self-consistency resulting from correspondence between representing and represented relations, can be directly related to inferential tractability through such concepts as 'vividness' (Levesque 1988) and 'limited abstraction representation systems (LARS) (Stenning & Oberlander 1992).

As we noted in the introduction, establishing the weak expressiveness and inferential tractability of graphical representations is only half way to relating graphics to its cognitive characteristics. The very demonstration that the logical expressiveness of graphical systems is weak is achieved by showing a correspondence to a weak *sentential* system (see for example Levesque's definitions of vivid fragments of FOL). If weakly expressive graphical systems can be simulated in weakly expressive sentential systems having therefore the same resulting tractability, then the appeal to computational complexity has destroyed any account of *differences* between linguistic and graphical modalities. If computational tractability is the underlying explanation of why graphical systems are easily usable, then these weak sentential systems should also be easily usable, and the differentiation of graphical and sentential systems is lost.

An insightful response to this dilemma requires supplementing an account of computational complexity with an account of what we will call the *availability* of constraints on expressiveness. We must ask what information a user requires to identify and to exploit the constraints in a representation system? In a sentential system with abstract syntax, what weaknesses of expression exist do not arise from the way that representations map onto their semantics. A considerable amount of exploration and learning may be required to be able to tailor a theorem prover to exploit the weak expressiveness of the system. In graphical systems, the situation is more complicated, but at least sometimes, the weaknesses of expression of the system can be 'inspected' because they are consequences of a few relatively central properties of the graphical representations and their interpretations. In our example in Figure 2 i), just knowing that 'greater than' was represented by 'left of' is sufficient to know that only full orderings can be represented, because 'left of' shares logical properties with 'greater than'. As we shall see in more complicated examples described below (matrix graphics, Euler's Circles), quite intricate meta-logical properties of the graphics can be inferred from a few basic facts about the diagrams' geometry/topology and its interpretation.

What users with different background knowledge and learning opportunities can understand about the constraints on expressiveness of diagrams is potentially determined by several sources of knowledge about interpretations. We have distinguished several interpretations of figures 2 ii) and iii) which cannot be inferred from anything in the diagrams. Even when the graphical relation in these diagrams is interpreted exhaustively imposing an expressive weakness, exploiting this fact relies on being aware of it. Computational tractability can only predict usability when supplemented by an account of availability of constraints to users with different contextual knowledge. The discussion of example systems below will illustrate this point.

Providing we can substantiate this claim that there are differences between languages and some graphical systems in the availability of their constraints, the claim can offer a possible explanation of an often noted psychological fact about modality differences. Graphics are frequently used in the introductory teaching of abstract subjects but abandoned by students as they gain proficiency, often in favour of some algebraic generalisation of the topic. If the cognitive advantage of diagrams is that they make constraints of expression available to someone with little background knowledge of a new topic, then one might expect graphics to lose their advantage over sentential systems as the user learns to exploit the hidden weaknesses in sentential systems.

3.2 Linguistic expressiveness

Having described a hierarchy of expressiveness amongst interpretations of graphics we now turn to a comparison with the semantics of sentential languages. This comparison will make clear why constraints in the expressiveness of sentential representations are less available. We use the term *language* here to refer only to sentential representation systems even though it is often used more broadly. Our discussion of languages will assume the printed medium as most relevant to the present comparison with graphics. Though we believe the essence of our account will transfer to spoken languages there are certainly many interesting issues which arise but which will not divert us here.

Token language texts consist of sequences of tokens of alphabetical characters (including characters such as parentheses) drawn from a set of types of character. Natural languages are stratified at least into a lexical level and a sentential level. Sequences of characters combine to make 'words'. 'Words' combine to form sentences. Artificial languages may have different labels for the levels, or even different levels, but for current purposes the issues are the same. The only properties of *tokens* relevant to their interpretation are the *type* under which they fall. The only properties of texts relevant to their interpretation are the sequences of token words classified into types.

These are, of course, considerable simplifications. Actual texts use such features as spatial layout to define more than merely sequence of characters, and such features as font to carry significance beyond character-type information. But these are useful simplifications at this level of analysis. Examples of direct interpretation of characteristics of spatial layout which constitute exceptions to our generalisations about languages can lead to representation systems such as tables which are intermediate between the major modalities of pure text and pure graphic (see Section 4).

The essential property of pure linguistic modalities that sets them off from graphical ones is that the only inter-word relation which is interpreted is concatenation. Because concatenation is the only interpreted spatial relation its interpretation has to be differentiated by the types of elements which are concatenated. For there to be more than a single uniform semantic interpretation of concatenation there must therefore be an abstract syntax which provides a diversity of abstract relations to be interpreted. It is the paucity of interpreted temporal/spatial relations in language which requires an abstract syntax to express many semantic relations. Conversely, it is the abstract syntax which means that the representation of semantic relations is not direct.

There are one-dimensional representation systems, which are pretheoretically classified as sentential languages, which *do* interpret concatenation uniformly and directly. The sentences of finite state generated (regular) 'languages' are an example in which concatenation of characters representing event types is uniformly interpreted. Such 'languages' have no abstract syntax and do enforce the representation of total orderings between all pairs of represented events. In fact, concatenation is generally isomorphically interpreted as succession in time. Whether they should be regarded as languages is a moot point. From our point of view, their modality shares its style of interpretation with graphics, albeit in one-dimension, and they share with graphics the important property of token-referentiality (see Section 3, page 15).

With this account of the semantic bases of graphics and language, we can now see how it explains some intuitively gross differences between graphical and sentential representations. One key is the style of representation of identity relations which are distinguished as type- and token-referential systems. In *token* referential systems, multiple tokens of icons of the same type refer to *distinct* entities (generally with shared properties denoted by the features of the type). In *type* referential systems (such as sentential languages), names refer in virtue of their type—multiple token occurrences of the *same* name-type refer to the same entity. Different attributions are made to them by attaching different predicates to separate occurrences.

Token referential systems enforce the representation of identity information—it is not possible to represent two elements (say the evening star and the morning star) without representing whether they are one and the same, or distinct. Tokens of icons are typically identified by their positions at some suitable level of granularity and all identity relations are determined between the referents of all icons in all token representations. 'Having the same position as' is a relation which is interpreted isomorphically with identity in these systems.

Type referential systems generally do not enforce representation of identity relations because name-type identity is not generally interpreted exhaustively. Different names do not necessarily denote different things. In sublanguages with 'unique names' restrictions, the relation between name-type and reference is restored. But information enforcement only results if names are the *only* referring expressions. Only in that case is isomorphism established and every reference determines the identity relations of its referent to all other references. These successive conditions on languages bring them closer to the highly restricted systems Goodman (1968) defines as *notations* for which the main condition is that there be a unique representation for each state of affairs represented.

If a representation system is token referential, it can only use a 'name' (or icon) once within an individual representation since any recurrence of the name signals a distinct reference and simultaneously asserts non-identity of its denotation with the denotation of its previous ocurrence. This limitation to a single occurrence of each referring element in token referential systems also limits the number of properties of entities represented because each property must be represented by a feature of a single icon, and the number of dimensions on which icons can vary is limited. As has often been observed, graphics runs out of carrier dimensions (e.g. Barwise 1993, Hovy & Arens 1990) . While token-referential systems exhibit these limitations, they thereby purchase considerable facility of processing. For example, problems such as anaphor resolution can only occur in a type-referential system where mutiple occurrences of an icon which may share reference are possible.

Cartoon strips are an interesting hybrid of token- and type-referentiality which get around some of the limitations of token-referential systems by introducing a form of 'type identity' across frames. Multiple-frame graphics are a clear reminder that we have so far been addressing issues of interpetation of homogeneous graphical spaces. One symptom of the breakdown of graphical homogeneity is the inability to interpret spatial relations across frames. That a character appears in one frame to the left of another character in another frame is not interpretable as meaning they bear this relation in the domain. The terms of between-frame semantics are not the same as those of within-frame semantics (the sub-diagrams of Figure 2 are another example of this phenomenon).

We can now summarise our account of the contrast between linguistic and graphical modalities. In graphical representations, there are a hierarchy of interpretations which impose increasing constraints on the interpretation of spatial representing relations. In linguistic modalities, concatenation is the only graphical relation which is directly and uniformly interpreted, and it is generally interpreted as yielding an abstract syntax. Semantic interpretation then applies to relations within this abstract syntax. There is therefore no semantically uniform interpretation of concatenation, and therefore no constraint on expression arises from the nature of the interpretation of spatial relations.

This theory of the graphical/linguistic modality contrast departs from intuitive classifications of representation systems both for extremely weakly expressive languages (which it shows to share the critical property of information enforcement with graphical systems) and for highly expressive graphics (which it shows to share with languages the possession of abstract syntax—see Section 4 below where expressive semantic networks are considered). This conflict with pre-theoretical intuition is just what one expects to follow on the development of theory. We believe that the application of the approach to practical examples below fully justifies this departure from intuition by demonstrating unintuitive correspondences of usability with its analyses.

Having developed this framework for comparing modalities, the remainder of the paper considers three classes of common representation, ranging from the most weakly to the most strongly expressive, applying the analysis and reviewing its predictions against some of what is known empirically about the usability of each type of representation.

4 Three Example Families of Representation Systems

In order to test the usefulness of the conceptualisation we have proposed, we now apply it to three widely ocurring families of representation. In each case we first apply the theoretical framework. We then review some of what is known about the cognitive properties of the example and ask whether the observations are in accordance with the framework's predictions.

An important principle in selecting examples to analyse is that they should be representation systems for which there is a reasonable presumption of usefulness. Constructing hypothetical examples would be much easier, but a much less secure as a foundation for theory. Not much empirical evidence is available from careful experimental investigation of the usefulness of informationally equivalent representations, but at least analysing frequently used systems encourages the presumption that the examples are useful for something. We here take matrix graphics, logic diagrams, and semantic networks as our three example families of representations. We begin with the most constrained and progress to the more expressive.

4.1 Matrix Graphics

We use *matrix graphics* as a convenient term for graphical representations of populations of data cases which are usually multi-dimensional and in which all cases are specified on all dimensions. 'Business' graphics—graphs, pie-charts, histograms, Gant charts etc.—are common examples. The meaning of these graphics is determined by isomorphisms between representation and data.

Since Bertin (1963) presentational issues associated with this type of graphic have been extensively researched in the HCI literature. Most of this work has, however, concentrated either on the ergonomic aspects of the display of a particular type of graphic (e.g. Salvendy, 1987, Kolski & Millot 1991) or on the fine detail of alternative presentations of closely related graphical forms (Tufte, 1983). The latter type of work has been extended to automatic control of data presentation by Cleveland (1985) and Mackinlay, (1983, 1986).

Tables are one point at which textual and matrix graphical systems meet[2]. Simple algorithms can turn tables into running text by combining row and column headings with cell contents. From that perspective, tables are two dimensionally laid out language. Tables therefore represent a mixed modality in which some spatial relations are directly semantically interpreted in a way which intertwines them with the interpretation of their linguistic components. In the other direction towards the more graphical lies the transformation of tables to maps illustrated in Figure 4.

Both tables and matrix graphics are weakly expressive in that they cannot represent data cases which are not specified on all axis dimensions. Both can omit cases entirely but neither can include partially specified cases. There

[2]It is perhaps no accident that archeological evidence suggests that in every known case of the invention of writing systems, the initial use was to express tabulations of data, generally in the service of account keeping.

Alphabetical Table

	A	B	C
A	0	2	4
B	2	0	5
C	4	5	0

Long./Lat. Table

	B	A	C
C	5	4	0
B	0	2	5
A	2	0	4

Map

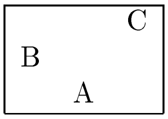

Figure 4: A table of intercity distances equivalent to a map

are some fine-grained differences of expressiveness within matrix graphics. For example, graphs are more weakly expressive than tables or histograms because graphs require an ordering on the categories ranged along axes, which then allows direct semantic interpretation of another spatial relation.

So, corresponding examples of tables, histograms and the tabular texts generated from tables, are different expressions of the same information, and as such each sub-family of the modality is equally expressive. The equivalences to be found between tabular language and graphical presentations is an example which underlines the dependence of analysis on the *system* level of description. The equivalence presumes that we understand tabular texts as ones generated from complete tables. As soon as texts are permitted which omit some but not all of the information about cells, the textual system becomes more expressive. The example also illustrates how the availability to users of constraints on expression will determine cognitive properties. A user faced with a text generated by translation from a whole table, can make no use of this constraint unless it is known to hold. If the same text is presented *as a table*, the spatial properties of the layout may be sufficient to allow immediate exploitation of the constraints in reasoning.

Our theory predicts that matrix graphics as a family of representation systems should have broadly similar cognitive properties. This view is supported by the empirical literature which mostly focusses on fine-grained interactions between specific tasks and different types of matrix-graphic presentation. Access mechanisms will differ in their facility according to the exact tasks at hand—judging whether a difference is positive or negative will be done more easily on the graph because it can be done by using only perceptual mechanisms. Retrieving an exact figure, or a figure for a difference should be easier from the table than the histogram. In terms of the MARS/LARS/UARS hierarchy described in Stenning & Oberlander (in press), matrix graphics and tables are MARSs. One graphic represents just one model.

4.2 Logic Diagrams

Diagrams use a wider range of spatial representing relations than do matrix graphics. Whereas the latter align categories along axes, diagrams exploit a wide range of spatial relations in their surface in representing their domains.

Work on the teaching of abstract subjects such as geometry and mechanics has both demonstrated the advantages for students of using diagrams and presented theories of why graphics have the effect they have. Larkin & Simon (1987) draw a distinction between *informational* and *computational* equivalence of representation. They see the advantages of graphics in terms of their computational tractability, citing searchability by parallel mechanisms and other properties of graphics as computational advantages of what they take to be informationally equivalent representations. Our approach can be seen as both a generalisation and a refinement of theirs. We agree that the central cognitive advantages of graphics come from their computational tractability. But we trace this tractability to their weak expressiveness, which in turn rests on the fact that at the representation *system* level, graphics are not informationally equivalent to languages. On this account, the computational advantages of graphics are more fundamental than architecture dependent features like parallel searchability (though these may well be the *result* of the possibility of implementing weakly expressive representations). To see where the accounts diverge, consider the case of processing a weakly expressive sentential language equivalent to a graphical system. It will be possible to construct a processor which makes this sentential language equally tractable (and therefore computationally equivalent). The difference between the graphical and sentential systems lies in the discoverability of the limitations on expression and the necessary methods of exploiting them in inference.

We can best illustrate our theory in a domain where the matching of the expressive power of the representation to the expressive demands of the task can be assessed more precisely than with mechanics problems. Here we take logic diagrams based on the analogy between spatial containment and set inclusion as an example where it is possible to make a rather deep analysis of relations between the limitations of the representing medium and limitations in the expressiveness of the logic represented. Logic diagrams are useful examples because they combine complete abstraction in some aspects (the relation between the respective circles and the sets represented is entirely arbitrary) with, in some cases, complete isomorphism of interpretation of topological relations to set relations. This example has the added attraction that a great deal is known about human reasoning ability in this domain and theories of mental process have been based both on graphical and non-graphical methods. Euler's circles have been the subject of well documented

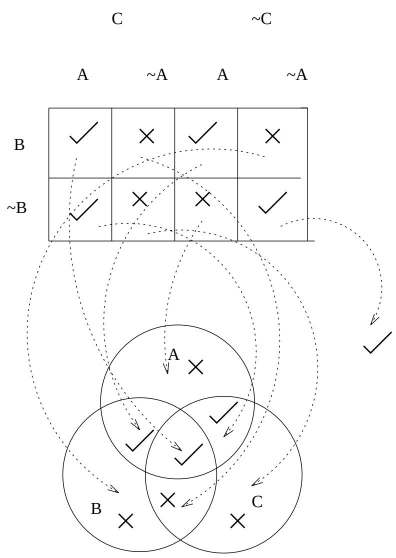

Figure 5: Venn Diagram as a variant of a table: dotted arrows indicate mapping from table cells to diagram sub-regions. Crosses indicate that a cell/sub-region is empty: ticks that it is non-empty.

misinterpretation in the literature, and this misinterpretation is of some interest to our cognitive theory (e.g. Erickson 1975, Johnson-Laird 1983, discussed in Stenning & Oberlander (in press)).

Some circle-diagram systems provide a link with the preceding section's discussion of tables since they start out as nothing more than tables organised according to a specialised method of spatial layout which facilitates search. Venn's system, for example, uses a single diagram of three circles intersecting as a 'table' of the eight possible 'types' of element defined by three properties. Figure 5 illustrates this correspondence between diagram and table.

The system annotates regions in this table to reflect presence or absence of types of individual in the model. There are different systems of annotation with different expressiveness. Shin (1991) shows how the system can be extended through a node-and-link notation for relations between cells to express much of monadic predicate logic.

The mapping of the several representing relations on to relations they represent illustrates the different kinds of mapping distinguished in Section 3. The fundamental idea of using spatial inclusion to represent set membership is fully isomorphic. A transitive, anti-symmetrical and anti-reflexive representing relation is mapped onto a represented relation with the same logical properties. In fact, the interpretation is such that spatial inclusion of points in a closed curve is actually an *example* of inclusion of members

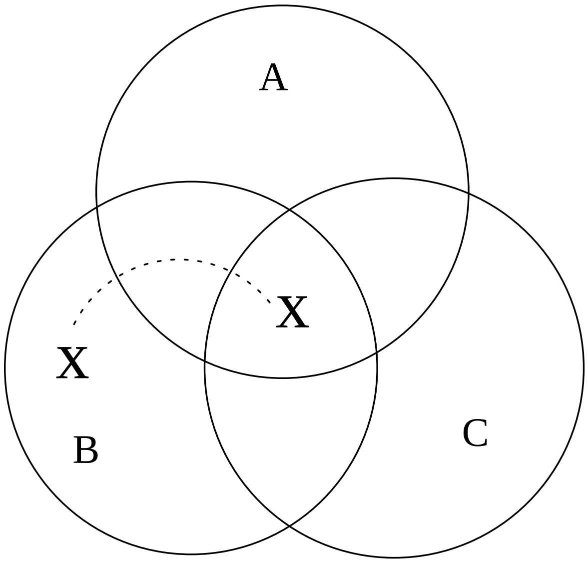

Figure 6: Interpretation of the link notation in Venn Diagrams. The arc connecting the Xs in the two regions asserts that *either* ABCs *or* $\tilde{A}B\tilde{C}$ exist. Absence of a link between regions is not interpreted as the negation of equivalent disjunctions.

in a set. But augmentation of the notation in which minimal sub-regions of the diagram are marked by crosses connected by links does not have this isomorphism (see Figure 6). The links are not interpreted exhaustively—absence of a link between two areas does not designate the negation of a disjunction between the linked regions. Adding the links notation is what takes the system beyond matrix graphics in expressiveness, but this addition intuitively makes the system much harder to understand. Anecdotal evidence from teaching with the system bears this out though we know of no systematic experimental test.

Venn diagrams are only one system of diagrammatic reasoning based on the same fundamental representing relation. Euler's circles exploit graphical constraints more fully than Venn diagrams, using many different arrangements of circle and incorporating a dynamic element into their use—reasoning can be modelled by mechanical constraints on movement of circles. Euler's Circles provide an interesting example of the complexities of analysing even apparently simple graphical systems and predicting their usability.

Euler's (1772) circles is a system of logic diagrams for doing syllogistic reasoning. Elsewhere we have developed a detailed reconstruction of Euler's algorithm and shown that when the system is correctly interpreted, it is a LARS rather than a MARS (Stenning & Oberlander (in press)). The reader is referred to that discussion for the detailed application of the current theory.

Briefly, a simple application of the analogy of spatial inclusion for set membership leads to a MARS interpretation and an explosion in the number of diagrams necessary to perform the abstract reasoning required. For example, the premiss *Some A are B* is satisfied by four of the five possible 2-circle diagrams representing the sets A and B. This means that to discover that the syllogism *Some A are B. Some B are C* does not have any valid conclusions using this concrete interpretation of the diagrams requires consideration of 16 pairs of diagrams. In each case it is necessary to show that there is no conclusion involving A and C which holds in all possible 'unifications' of these diagram pairs. This typically requires consideration of several arrangements of the three circles each consistent with the premisses, and leads to a total of about sixty diagrams to consider.

But no logic teacher, least of all Euler, ever proposed such exhaustive methods to a pupil. By strategically selecting what can be seen as 'weakest' cases it is possible to prune the search space. Using Euler's system the reasoner can solve any syllogism by algorithmic construction of a single diagram. A simple notational device, along with a subtle but far reaching reinterpretation of the ontology of the diagrams, provides a mechanically defined algorithm which allows this dramatic difference in effectiveness. Marks distinguish the sub-regions which designate *minimal* models of the premiss from regions designating types merely consistent with the premiss. A strategy of choosing unifications of diagrams with the most possible sub-regions ensures exhaustive search of the logical space. By adding a simple topological rule for 'cancelling marks' under some conditions after unification, the system is transformed from one which requires concrete examination of every single possible logical model, to an elegant one-pass algorithm which constructs a single diagram, and bases complete and valid reasoning on it.

Figure 7 provides an example syllogism solution using Euler's method. Unlike Shin's extension of Venn, this extension of Euler stays within a strongly graphical interpretation of its diagrams. Although it captures the whole of the syllogism, it is considerably logically less expressive than her system.

What do we know of the cognitive impact of this diagrammatic system? Although there are few systematic studies using this graphical reasoning technique its very longevity as a didactic technique speaks for its cognitive efficacy. Psychologists have mainly been attracted by using the system as the basis for a psychological model of 'mental' reasoning rather than an external teaching/reasoning aid. One teaching study Grossen & Carnine (1990), which compares Euler's Circles with sentential teaching does show quite radical facilitation of learning by graphics, bringing 'learning impaired'

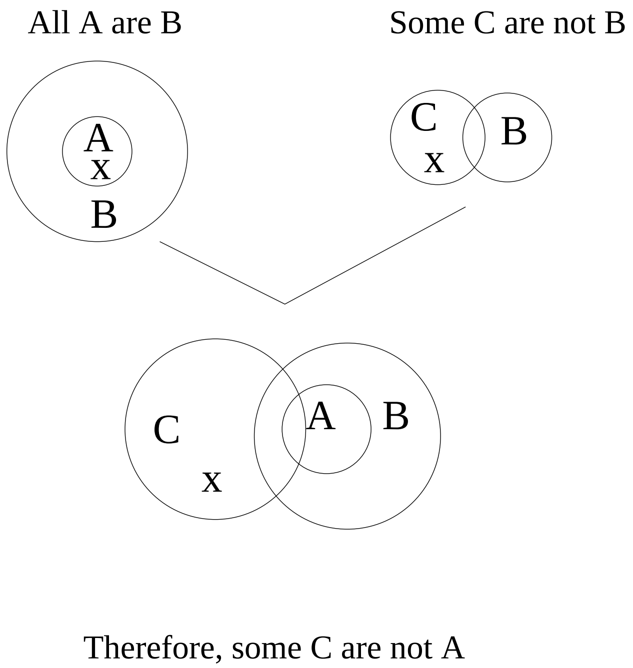

Figure 7: An example graphical syllogism solution. 'x' indicates a region corresponding to a minimal model. In unifying premiss diagrams, 'x'-marked regions bisected by the third circle become unmarked. Final 'x'-marked regions licence valid conclusions.

subjects up to superior performance levels, but only in a condition in which subjects actively manipulated the diagrams. The importance of manipulation may be due to temporal information enforcements.

When the diagrams are 'animated' by movements of the circles the resulting operations have certain computationally attractive properties of continuity. The movements transform models from one to another with minimal additions or subtractions of sub-regions denoting single types of individual. The 'relevant' models to consider in verifying a conclusion are always 'adjacent' in this space of models. Euler's spatial analogy is also a mechanical one. The logical constraints can actually be modelled by mechanical constraints. Marks representing minimal models can be simulated as 'pins' which prevent sliding movements of the circles Stenning & Oberlander 1994. This modelling of logical constraint as mechanical constraint, a possibility that stems directly from the isomorphism between representing and represented relations, may be one of the factors that leads to the contrast between active construction and passive inspection of the diagrams.

If Euler's system, correctly used, is a cognitively efficient representation for learning and performing a rather difficult fragment of reasoning, what does that say about the general theoretical framework? This diagrammatic system clearly takes us beyond matrix graphics in its expressive power. Syllogistic reasoning is inherently abstract, and a matrix graphic system requires exhaustive inspection of up to 60 diagrams to demonstrate the lack of any conclusion from some pairs of premisses. Shin's Venn-based system augments the power of the basic matrix-graphic core by notations which are not graphically constrained, capturing a larger fragment of logic, but, we suspect, losing some cognitive efficacy. The Euler system shows that diagrams can be extended in expressive power to be just adequate to capture the abstractions required for this fragment of reasoning *without giving up logically interpretable graphical constraints.*

Although the intuitions and anecdotal observations need empirical investigation, the framework encompasses some important observations. If the graphics are interpreted as only primitively expressive (MARS), they really are useless for this abstract task—viz the arguments that Euler's Circles cannot be the basis of mental syllogistic reasoning (Johnson-Laird 1983). It was Euler's genius to see that suitable strategies of selection and combination could remain within a graphically constrained interpretation and yield a tractable system for this limited domain of reasoning. Even minimal extensions of the logic to say, disjunctive syllogisms are beyond the expressive power of Euler's graphical formalism.

This graphical system provides a good example of the differential availability of constraints in graphical and sentential representations described in

Section 3. Once a diagram user understands the core of the interpretation of the diagrams, their weakness of expression immediately follows from their basic geometrical properties. Any point in a plane corresponds to a maximally specified individual. Spatial containment, like set inclusion, is a transitive, non-symmetric, reflexive relation like the monadic universal quantifier. Intersection models existential quantification; exclusion models ground negation. The five possible topological relations between two circles exhaust the possibilities and are each interpreted as the corresponding set relations. This exhaustiveness is particularly compelling because all the primitive topological relations are included. In contrast, deriving these properties from the sentential formulation of the syllogism takes a sophisticated grasp of model theory and considerable inference and enumeration of cases.

As an example of this facilitation of meta-logical reasoning, the graphical system makes available the abstract property *case-identifiability*. Case-identifiability is the name we give to the property of the syllogism whereby if, and only if, there is a maximally specified type of individual entailed to exist by the premisses (i.e. a type specified with respect to all three properties A, B and C) is there a valid conclusion (see Stenning & Oberlander (in press)). This abstract property is revealed by the observation that only final diagrams which have at least one remaining 'x' have a valid conclusion (see Figure 7). It is this model-theoretic property of the categorial syllogism that allows Euler's strategy of representation to work. For example, the disjunctive syllogism lacks this property and requires the unconstrained augmentation of Venn's system described above.

These important properties of ECs are available by inspection to a user with prior knowledge of only the core semantic concept behind the spatial-containment-as-set-inclusion interpretation. But there is still a big difference in ease of use between the primitively interpreted system with its explosion of diagrams to be checked, and the abstractive version presented here, even though both share the same core concept. Isomorphism is of little use if it blocks the expression of abstractions which are required for a tractable performance of the task at hand. We would still expect the usefulness of the system for reasoning by naive users to be determined by the pedagogical details of how the system is introduced.

In summary, Euler's Circles maximally exploit graphical constraints in capturing the domain of syllogistic reasoning. They are more constrained than Venn's system and the theory therefore predicts that for syllogistic reasoning they should be more effective. Our intuition and what observation is available supports this view but it stands in need of experimental test. Euler's system illustrates how constrained graphical representations make properties of reasoning systems available to users with only core knowledge of their semantic basis, but also reveals the need for a complete analysis of strategies of use and details of interpretation to fully predict usability. The usability of the system is determined by the combinatorics, and these are dependent on the subtleties.

4.3 Semantic Networks

In Section 3, we adopted node-and-link examples to illustrate interpretations of different expressiveness because this single formalism can bear interpretations of such a wide range of expressiveness. Semantic networks, one category of these formalisms, are therefore an important case for developing a theory based on constraints on expressiveness. Semantic networks are an insightful case for our purposes because much is known of their formal properties, and something of their usability. Networks are an important class of applied notation which has evolved in close association with computational technology. Much of their practical use is dependent on that technology. They have an intimate relation to linguistic representations in programming languages. The origins of semantic networks go back at least to Peirce (see Roberts 1973) who developed graphical notations for predicate calculus contemporaneous with Frege's development of sentential versions (see Sowa 1993) but did not achieve widespread use until the advent of computers.

A further interest of semantic networks is that an important motivation for their development has been cognitive theorising about human mental representation. Semantic networks have been used as the basis of models of human semantic memory and lexical representation (Collins and Quillian (1969); in simulation models of human text comprehension (Anderson and Bower (1972), Anderson (1983)); and as a notation for human memory representations in discourse (Kintsch and van Dijk, 1978). Even early formulations of mental models employs a variety of semantic network notation (Johnson-Laird and Steedman (1978).

We first review commonly ocurring examples of these formalisms of various expressive powers, and then review some evidence of their usability. We begin with the most expressive examples.

Semantic networks can be interpreted as logically highly expressive languages (e.g. Schubert 1976, Hendrix 1979, Sowa 1979). Schubert shows how simple nets can be extended in expressive power to equivalence with the lambda-calculus by progressively adding devices for capturing connectives, quantifiers, and abstraction operators. Schubert's point is to understand the semantics of networks by assimilating them to logical languages whose semantics is already understood.

These expressive interpretations of semantic networks are possible because an abstract syntax is imposed on the networks. Nodes appear which represent syntactic categories connected to their syntactic arguments by links. Sometimes the 'abstract' nodes are eliminated in favour of labellings on links. Either way the semantic interpretation is defined not directly on the representing relation (the links) but on a syntactical categorisation of the links based on configuration. Node-and-link formalisms can thus be seen as introducing multi-dimensional concatenation to the design of languages. Whereas sentential languages use one dimensional concatenation, networks, by allowing, indefinitely many arcs leading to and from any node, expand the dimensionality of concatenation available. As such our theory classifies these networks as essentially of a linguistic modality that exploits no graphical constraints.

The next less expressive grade of interpretation we distinguished in Section 3 was representations in which there was uniform interpretation of a representing relation (no abstract nodes), but this relation was not exhaustively interpreted (the initial interpretation of Figure 2 ii) and iii)). We saw that such diagrams were capable of expressing partial orderings. When these partial orderings are interpreted as expressing partial knowledge of a completely ordered world, the resulting representation system retains the discursiveness of a language—representation is selective.

Tree diagrams of taxonomies are an example of interpretation which applies the further constraint of exhaustive interpretation. Trees are probably the commonest representation of taxonomies. In a common interpretation, the transitive closure of their link relations are interpreted as ISA relations. Any two nodes which are not connected by a path of arrows are interpreted as being related by the negation of this relation. Whether the links are explicitly asymmetrical arrows, or whether this asymmetry is implicit in the vertical or horizontal layout, the interpretation is certainly asymmetrical. Because the logical properties of the representing arrows (non-symmetric, transitive) do not match those of the represented relation ISA (non-symmetric, transitive), trees do not have the guaranteed consistency property (for example, all cyclic graphs interpreted in this fashion represent inconsistent taxonomies).

Finally, an example of the most constrained interpretations of networks in which the semantic interpretation of the representing arrows does share all their logical properties is the circuit diagram. Here links are interpreted as physical connections, and their absence as the absence of physical connections. Nodes are interpreted token-identically as components, labelled according to their type. The result is a graphical MARS in which each graphic has a single model.

How does the empirical evidence of the usefulness of semantic networks map onto this classification of examples? The most systematic empirical studies of the cognitive properties of expressive semantic networks as compared to equivalent sentential counterparts come from the assessment of 'graphical' programming languages (see e.g. Myers (1990)). Results of studies of usability are mixed as our theory would predict, with some studies showing advantage but more showing roughly equal usability. But this review does not address expressiveness explicitly. At the most expressive end of the dimension, there is little indication that semantic networks are a useful formalism for lambda calculus (Schubert's intention was, of course, always analytical). Some have claimed that these formalisms are useful for teaching predicate calculus We know of no controlled studies, and our general intuition is that when sentential and network systems are carefully equated for expressiveness, they have roughly the same pedagogical efficacy. This empirical issue is worthy of investigation.

As we descend the expressivenes dimension, the first semantic networks for which we find solid evidence of wide employment are ones which stop short of employing variables. An examination of systems like ACT* which contain a semantic network representation as their declarative memory, certainly have typed arcs (or abstract syntactic nodes) but they signally restrict their entity nodes to logical constants, and are token-referential. Here there is real evidence of usefulness of the network formalism (relative to sentential equivalents) at least in theorising about human memory. Intuitively, much of the benefit probably stems from the ease of search which results from the token referential nature of the systems. Interestingly, the property of complete connectedness (having no unconnected 'islands' of nodes in a network representation) has been used to predict comprehensibility and memorability of texts the networks represent.

Our example of networks which impose exhaustive interpretations on their link relations were trees interpreted as representing taxonomies. The widespread use of this representation is sufficient testimony to its usefulness. The question arises as to how the limitations on their expression are understood since, as we have seen, they do not arise from any graphical constraint. So widely are these formalisms used that it is quite hard to answer questions about the origins of the knowledge required to interpret them. Just what does generate the inference that two classes of entity on different descending branches are mutually exclusive? What generates the related assumption that a class labelled 'a' at one node is a different class than one on another branch labelled 'b'? Is it the token referentiality of the system? Or the exhaustiveness of interpretation of the link relation? Or the 'formal' perception that cyclic graphs are not well formed members of the system? Or knowledge that what is presented is a taxonomy? Or ...? Whatever the right answer is for any particular example, predictions of usability from expressiveness will have to be relative to availability of contextual knowledge. At this level of expressiveness there is clear evidence of the usefulness of these graphical formalisms. When we take the final step to circuit diagrams usefulness can hardly be challenged. Sentential equivalents to these systems only arose in the process of computerisation of circuit design, and are rarely used in non-automated systems.

This brief journey along the dimension of expressiveness suggests that a thorough study of the interpretations of the many node-and-link formalisms used in computer science and elsewhere would be a valuable preliminary to a deeper study of the cognitive properties of these graphical representations. Although we can point to few systematic experimental comparisons of network and sentential systems, even the samples discussed here support the view that there is good evidence for usefulness at the inexpressive end of the continuum, but decreasing evidence as the formalisms' interpretation grows more expressive. Token referentiality appears to be an important watershed—there is little evidence of advantage over sentential systems of node-and-link formalisms so expressive as to jettison this feature. By and large, the complexity of the existing empirical literature on the usability of network representations bears out the theoretical framework developed here.

5 Conclusion

We believe that our proposals for the terminology of media and modalities, however terminologically distinct, capture concepts widely but implicitly held in the community concerned with the usability of representations. We believe they are the basis for a firmer foundation for research into cognitive properties of representations based on semantic analysis. Using semantic concepts to predict cognitive properties of representations turns theoretical analyses from logic and computer science back onto human information processing. Expressiveness of representation was first studied in connection with understanding processing by machine, but at a fundamental level the theory applies to the humans who devised it. This discussion of the family of graphical modalities is suggestive of the richness of classification this conceptualisation offers and of ways semantic analysis can contribute to the study of usability.

Previously we have proposed a semantically based hierarchy of expressivness, namely the MARS, LARS and UARS distinctions, based on the ability of representing systems to express partitionings of the space of all their possible models. Our present description of grades of interpretation function drawn in terms of constraints on the mapping between represented and representing relations is the beginnings of a 'syntactically' or formally motivated taxonomy. It is clear that the two classifications are not independent. Our hope is that in the future it may be possible to give an exact characterisation of their relations.

Our semantic conceptualisation of modalities focusses attention on the availability of information about interpretations to different classes of user. Degrees of expressiveness (and therefore tractability) of systems cannot predict usability unless combined with accounts of the availability to the user of constraints on expressive power. We have given some examples which illustrate how some graphical systems wear their weakness on their sleeves. As almost a side effect, we have emphasised what we all knew all along, that human cognition is as much about *selecting* representation systems to perform a task as about computation *within* a given representational scheme.

The current vogue for 'visualisation' which is witnessed in fields as disparate as programming methodology, the analysis of large data sets, human—computer interaction, mathematical proof, and education, is a powerful force for liberalising practice in many fields. But if we are to separate the value from the hype, it will take a more sophisticated distinction between modalities of presentation than the apparently simple one between graphics and language.

References

Anderson, J. R. and G. H. Bower (1972) Configurational properties in sentence memory. *Journal of Verbal Learning and Verbal Behavior* **11**, 594–605.

Bertin, J. (1981) *Graphics and Graphic Information Processing.* Berlin: de Gruyter.

Cleveland, W. S. (1985) *The Elements of Graphing Data.* Monterey, Ca.: Wadsworth.

Collins, A. M. and M. R. Quillian (1969) Retrieval time from semantic memory. *Journal of Verbal Learning and Verbal Behavior* **8**, 240–247.

Coutaz, J. (1992) Multimedia and multimodal user interfaces: a taxonomy for software engineering research issues. In *EWCHCI*, pp. 229–240.

Erickson, J. R. (1975) A set analysis theory of behaviour in formal syllogistic reasoning tasks. In R. L. Solso, ed., *Information Processing and Cognition: The Loyola Symposium.* Hillsdale, N.J.: Lawrence Erlbaum Associates.

Euler, L. (1772) *Lettres a une princesse d'Allemagne*, Vol. 2: *sur divers sujets de physique et de philosophie.* Letters 102–108.

Goodman, N. (1968) *Languages of Art.* Indianapolis: Bobbs-Merrill.

Grossen, G. and D. Carnine (1990) Diagramming a logic strategy: effects on difficult problem types and transfer. *Learning Disability Quarterly* **13**, 168–182.

Hendrix, G. (1979) Encoding knowledge in partitioned networks. In N. V. Findler, ed., *Associative Networks*, pp. 51–92. New York: Academic Press.

Hovy, E. and Y. Arens (1990) When is a picture worth a thousand words? — allocation of modalities in multimedia communication. In *AAAI Symposium on HCI.*

Johnson-Laird, P. N. (1983) *Mental Models.* Cambridge: Cambridge University Press.

Johnson-Laird, P. N. and M. J. Steedman (1978) The psychology of syllogisms. *Cognitive Psychology* **10**, 64–99.

Kintsch, W. and T. A.van Dijk (1978) Towards a model of text comprehension and reproduction. *Psychological Review* **85**, 363–394.

Kolski, C. and P. Millot (1991) A rule-based approach to the ergonomic "static" evaluation of man-machine graphic interface in industrial processes. *International Journal of Man-Machine Studies* **35**, 657–674.

Larkin, J. H. and H. A. Simon (1987) Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science* **11**, 65–99.

Levesque, H. (1988) Logic and the complexity of reasoning. *Journal of Philosophical Logic* **17**, 355–389.

Mackinlay, J. D. (1986) Automatic Design of Graphic Presentations. Technical Report STAN-CS-86-1138, Department of Computer Science, Stanford University.

Mackinlay, J. (1983) Intelligent Presentation: The Generation Problem for User Interfaces. Heuristic Programming Project Report No. HPP-83-34. Department of Computer Science, Stanford University.

Meyers, B. A. (1990) Taxonomies of visual programming and program visualization. *Journal of Visual Languages and Computing* **1**(1), 97–125.

Palmer, S. E. (1978) Fundamental aspects of cognitive representation. In E. Rosch and B. B. Lloyd, eds., *Cognition and Categorisation*, pp. 259–303. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Roberts, D. D. (1973) *The Existential Graphs of Charles S. Peirce*. The Hague: Mouton.

Salvendy, G. (1987) *Handbook of Human Factors*. New York: John Wiley and Sons.

Saussure, F.de (1966) *Course in General Linguistics*. New York: McGraw Hill. Translated by Wade Baskin, first published 1916.

Schubert, L. (1976) Extending the Expressive Power of Semantic Networks. *Artificial Intelligence* **7**(2), 163–198.

Shin, S-J. (1991) A situation-theoretic account of valid reasoning with Venn diagrams. In J. Barwise, J. M. Gawron, G. Plotkin and S. Tutiya, eds., *Situation Theory and its Applications*, Vol. 2, pp. 581–605. Stanford, Ca.: Center for the Study of Language and Information.

Sowa, J. F. (1979) Semantics of Conceptual Graphs. In *Proceedings of the 17th Annual Meeting of the Association for Computational Linguistics*, pp. 39–44, University of California at San Diego.

Sowa, J. F. (1993) Relating diagrams to logic. In G. Mineau, B. Moulin and J. F. Sowa, eds., *Conceptual graphs for knowledge representation*. Springer.

Stenning, K. and J. Oberlander (1994) Spatial containment and set membership: a case study of analogy at work. In J. Barnden and K. Holyoak, eds., *Analogical Connections*, pp. 446–486. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Stenning, K. and J. Oberlander (in press) A cognitive theory of graphical and linguistic reasoning: logic and implementation. *Cognitive Science*.

Tufte, E. R. (1983) *The Visual Display of Quantitative Information*. Cheshire, Connecticut: Graphics Press.