Chapter 2

Conversational Domains

In the previous monograph we described two basic procedures for constructing a conversational domain and its description ($D^1(R)$, $D^0(R)$) to represent a thesis about a subject matter.

One procedure is instrumented by a human interrogator/analyst who (given some mechanical "book-keeping" assistance) interviews a "source" or subject matter expert. The other procedure is a computer program, EXTEND, which performs a similar ritual. Operationally speaking, EXTEND replaces the interrogator/analyst but it does not "mechanise" the construction process. The fact is, only one human being, here the subject matter expert, is required. EXTEND uses him in an analytic role and provides the assistance needed to secure cyclicity and consistency (the essential properties of the relational network part of a conversational domain), as well as using him in the role of subject matter expert. This point was plainly exhibited by showing that EXTEND can be called as a routine by the tutorial operating system, CASTE, and *is* called whenever the student takes on the status of expert and (in an evolutionary system) enlarges the scope of subject matter by adding further topics.

Fig. 2.1 summarises the constituents of a conversational domain as it is produced by either of these methods. The labels *BG* (behavior graph) reflect the notation adopted to disambiguate the previous terminology (Task Structure, *TS*). Attached to each of the nodes, which stand for topics, there is a behavioural graph, *BG*(i), strictly a program graph. It is a class of programs of which any one, if executed in an appropriate modelling facility, will bring about and satisfy R_i the relation underlying this topic. Used

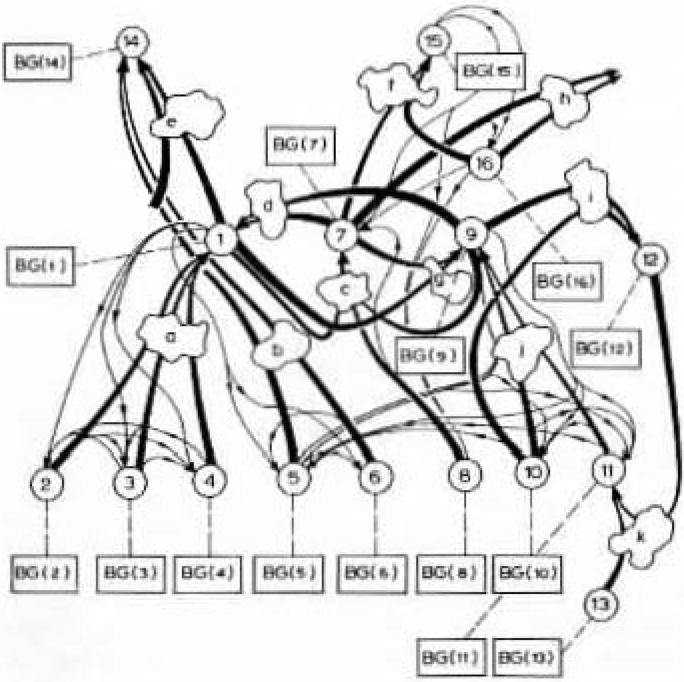


Fig. 2.1. Portion of a relational network. The nodes 1, 2, ... stand for topics. The arc bundles covered by a label represent a derivation of the topic on which the arcs are incident from the topics from which they emerge, by applying the relational operators specified in the labels a, b, ... The boxes attached by data links (not arcs) to each node specify the explanation of the topic in terms of a behavioural prescription or program graph (alias, Behaviour Graph, *BG*).

descriptively, *BG*(i) and its interpretation in the modelling facility *is* $D^0(R_i)$ of the previous monograph; used to prescribe a model-making behaviour which a student should carry out, it is *TS*(i). In either case, his (explanatory) model-making behaviour in the modelling facility ($Exec^0 i$) is compared for correctness with *BG*(i) and any correct model when executed in the facility also satisfies R_i.

It will be recalled that the relational network part of the conversational domain is processed to yield a structure such as Fig. 2.2 in which the relational operators, representing the derivation of one topic from others, are consigned to a data base, and the connective arcs depict entailment relations, i.e., derivations of any legitimate kind. The processing takes place (if and only if the original network is cyclic and consistent) at the point where the expert designates one topic or a cluster of strictly analogical topics as a head and specifies topics at a distance and direction from the head which he regards as subordinate to the head. Subsequently, the expert is required to *describe* the related nodes, using unary but many valued predicates, and the resulting mesh is embodied in a physical display, the *entailment structure*, in which each topic (or the node representing it) is associated with storage to accommodate tokens indicating its state during learning. It is possible to reduce the entailment structure to units of the type shown in Fig. 2.3 and it is important to notice that any legitimate network is an

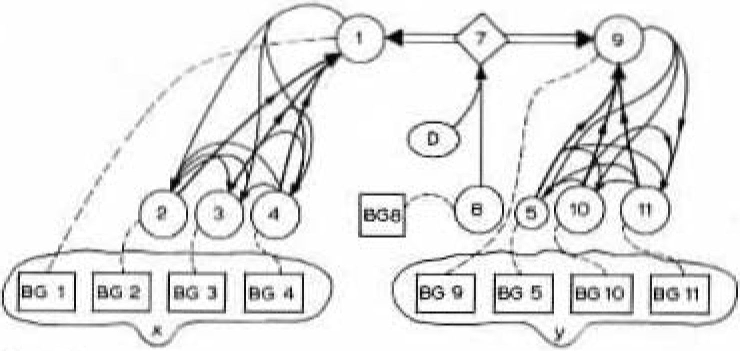
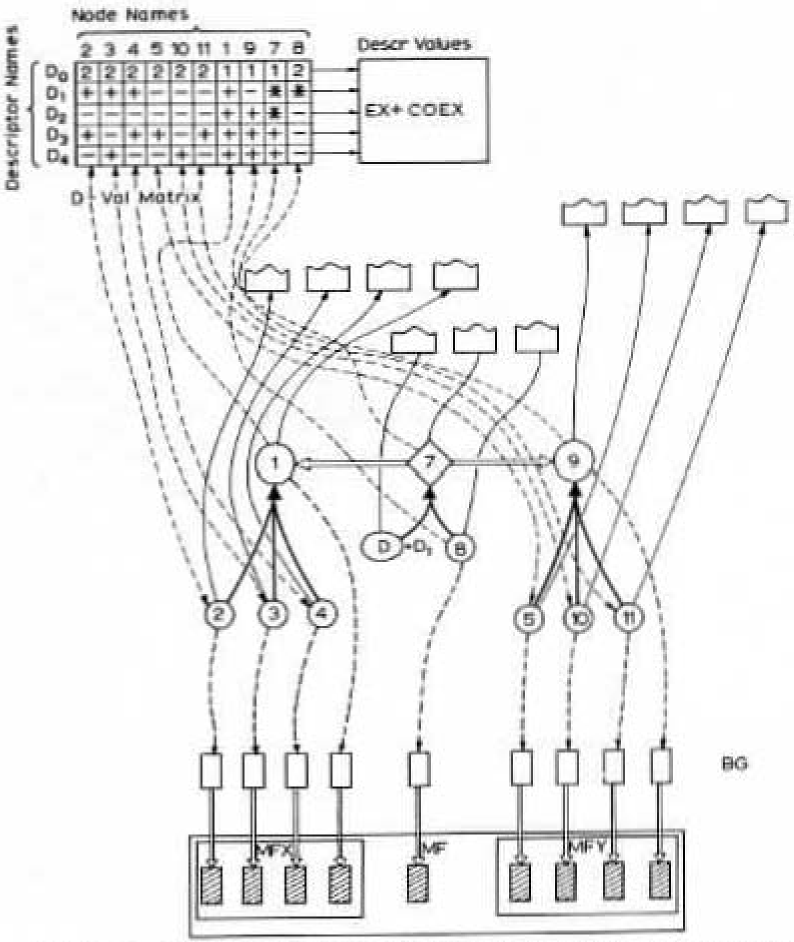


Fig. 2.2.

(a) The entailment mesh produced in preparation for pruning the network of Fig. 2.1, under the topic which is recognised as analogical, so that place holder node labelled D is introduced to accommodate the names of semantic descriptor(s), the values of which distinguish *topics* 1 and 9. Outgoing arcs from nodes 1, 7, and 9 are deleted except for those required to maintain the cyclicity of the structure (shown as thin arcs) and the cyclic component of the analogy relation (7) is represented by short hand ⇔ notation (*topic* 7 being itself distinguished as a ◊ node. The *BG* of *topics* 1, 2, 3, 4, are interpreted in a universe X, and *topics* 5, 9, 10, 11 are interpreted in a universe Y. X and Y are distinct, but as yet unspecified, and will be distinguished when D is named by the values of the D predicate.



(b) The entailment structure obtained if the mesh is pruned under the topic 7 (so that cyclic linkages are obscured *not* deleted), and its nodes are described by descriptors (unary many valued predicates or Fuzzy Predicates shown as $D_0D_1D_2D_3D_4$. Of these D_1 is D (the name of the distinguishing node) and D_0 is "depth from head in maximal arc distance". The D Val Matrix relates descriptor values to examples and counterexamples (the slide projected materials in INTUITION of Chapter 1) and to the name of nodes. attached to each node represent storage for "node state" markers (explore or aim or valid aim, or goal or understood). The *BG* of *topics* 1, 2, 3, 4 give rise to models in a independent modelling facility *MFX*: the *BG* of *topics* 5, 9, 10, 11 in *MFY*. Both *MFX* and *MFY* are part of a Lumped Modelling Facility containing several independent processors (for example, STATLAB of Chapter 1). *Topic* 8 may be realised in either part of *MF*.

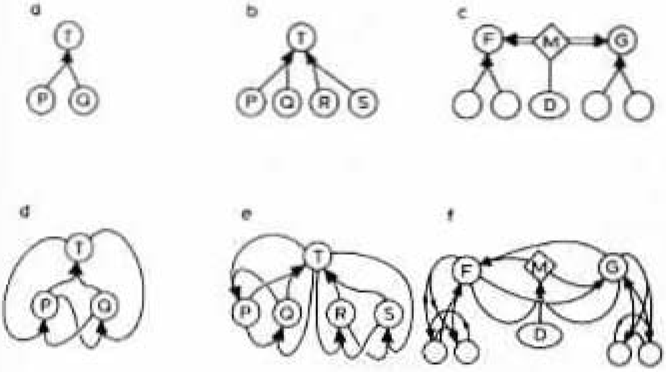


Fig. 2.3a, b, c, d, e, f. Typical structures. (a) Conjunctive, (b) Disjunctive, (c) Analogy relation with condensed symbolism, (d) Conjunctive, (e) Disjunctive, (f) Analogy meshes with cyclic derivations.

analogy relation in its own right. * Moreover, if the reconstructive derivation cycles of the original are reinstated, each substructural unit is cyclic (Fig. 2.3) unless it happens that it contains a node marked as primitive.

This essential property allows the mesh to be pruned under different head topics to yield completely different structures. For example, Fig. 2.4 shows a common construction in which a principle T, is reapplied to yield a topic relation A. On repruning in the most radical fashion, T is exhibited by examples (notice that these are *not* just aggregated under an arbitrary union. T is the join of A, B; or the join of B, C . . .; these topics may be *rederived*, as a result, from T). Other, intermediary prunings are illustrated.

These operations have been considerably refined since the previous monograph was written. Some of the refinements are of epistemological consequence and others of more pragmatic value; they will be described at appropriate points in the following discussion.

---

* Not any entailment structure. For example, an entailment structure can be, though seldom is, fully conjunctive. Even in that case, the network before processing contains cyclic derivations that are not eliminated by pruning (previous monograph).

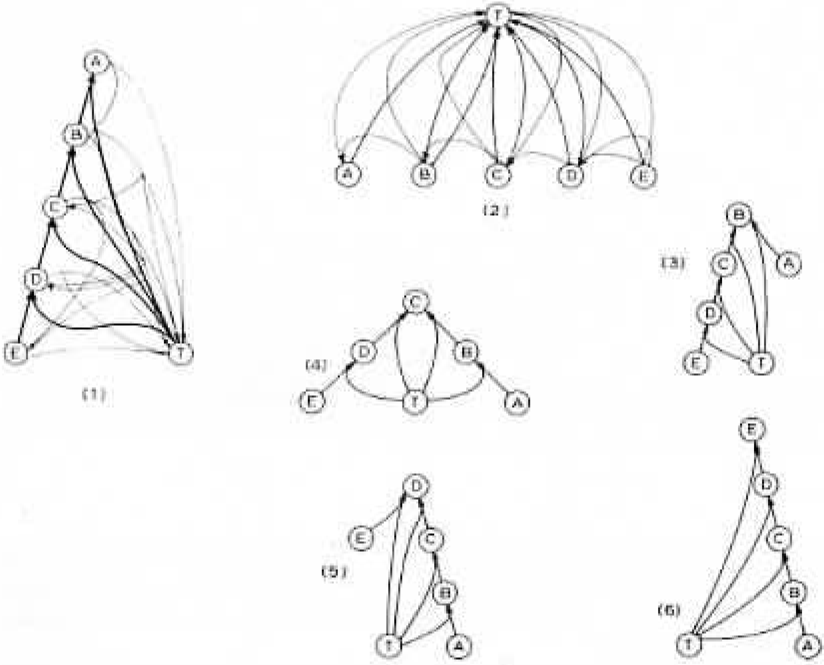


Fig. 2.4. Multiple head pruning. (1) Entailment mesh. Pruned under *topic* A (built up by reapplying a principle (T) with cyclic entailments shown as thin lines. (2) Converse pruning of the same mesh under the head topic of T. (3), (4), (5), (6) are other prunings (cyclic entailment connections are not shown).

1. SYNTACTIC AND SEMANTIC COMPONENTS OF A THESIS

It is expedient to discriminate between the syntactic, "5 is a prime number", and the semantic, "5 is a lucky number, or the numeral on your hotel room" aspects of a thesis and the structures representing it. The distinction is relative, "how do I know a purely syntactic entity, approximated by a logical text devoid of words?" But it is exceptionally useful.

Both of the construction procedures, and others introduced later in the book, are based on the idea that a thesis is a set of topics with syntactic relations between them and that a concept of a topic has a systemic (alias syntactic) core, roughly in Hartman's (1969) sense. Further, the syntactic component is output first, as a series of topic derivations, and later on is given semantic interpretation via the description. This is without prejudice to the fact that an expert or a student has a semantic interpretation in mind; it merely influences the order in which parts of his thesis are exteriorised.

For example, consider the unzipping (previous monograph) of the topic "efficiency" evoked by a question, "What does efficiency relate?" Clearly, the expert may be thinking of thermal/mechanical efficiency or some such interpretation, but the unzipping operation yields a syntactic derivation. For example, "Efficiency is a relation between work done and heat used, measured by a relation between source/sink temperature and the absolute temperature."

$$\text{Efficiency} = \frac{\Delta\text{Work}}{\Delta\text{Heat}} = \frac{T_{\text{high}} - T_{\text{low}}}{T_{\text{high}}}$$

All of the terms in this equation are discriminated upon syntactic grounds, as formally related symbols, and just this property renders them apposite as topics in a thesis which says, you can learn about *efficiency* if you understand "amount of work done and amount of heat" or "temperature difference and absolute temperature" or both. True, they also have semantic interpretations in a universe of heat engines, refrigerators, and the like; true also, these equipments are semantically related. But though the posited interpretations, or others, may be cognised, the mandatory feature of the derivation is a syntactic or formal relation.

Further, if the syntactic connection is pursued by successive unzipping until all of the subordinate topics are marked primitive, then these primitives are no more nor less than the constraints upon a modelling facility (a processor, not just a static entity) in which programs can be written to give imperative (temporally executed instruction) status to production rules. On execution, some of the possible programs, those that belong to *BG* (efficiency) and its subordinates, satisfy the "efficiency" relation and the relations "beneath" it.

Surely, any program, a syntactic entity, must be compiled as a model before it is executed; surely, also, the modelling facility (*MF*) in which it is compiled has a semantic description (it is a universe of possible actions). But this description appears later in the exposition of a thesis and it must do so in order to preserve the convenience of an "up to downwardly" directed derivation scheme (the thesis is the first and most global topic; further topics are differentiated as required), in contrast to the usual expedient of selecting sets of objects to begin with and using their members as building blocks (a "down to up" paradigm).

All this works satisfactorily except for analogy relations that are declared by the expert, in the simplest case, as isomorphisms. For an isomorphism (one to one correspondence) must be supported by a distinction between universes of interpretation (X, Y of Fig. 2.2), in practice, a distinction between modelling facilities designated *MF*(X) *MF*(Y). Lacking this support, the isomorphism would be confused with an identity and the derivation rejected as inconsistent. *

The topic that supports an analogy relation is one or more semantic predicate(s) (colour, texture, size, material, shape). The predicates supporting analogy relations (distinguishing *MF*(X) from *MF*(Y), for example) are the mandatory, and the only mandatory, semantic constituents of a thesis. The class of semantic properties named by these predicates includes time (execution time, order as determined by a processor clock). Recall from the previous monograph, that there are distinct clocks in the processors of *MF*(X), *MF*(Y).

One general point stressed in the previous monograph is worthy of repetition. Time, or precedence, is the least specific semantic interpretation given to syntactic productions, rewriting rules and implications. Moreover, any interpretation of such a (syntactic) sign involves time; though specific interpretations may entail specialised time orderings (realisable in the processor types of the previous monograph, L-Processors, the one clocked processors of modelling facilities, and so on).

The consequences of these observations ramify throughout the entire book. For example, they suggest a more systematic method

---

* *MF*(X) and *MF*(Y) figure as the "partitions of a modelling facility" in the previous monograph; for example, the "real" and "abstract" partitions of STATLAB. Henceforward, since analogy relations are considered in greater depth, we use the terminology *Lumped Modelling Facility* for the facility as a whole (for instance, all of STATLAB) and refer to its components or partitions, each with an a priori independent processor, as "modelling facilities" simpliciter: *MF*(X) or *MF*(Y) as the case may be.

for eliciting descriptions of the mesh depicting a thesis, which has been implemented and is described. They lie at the root of representing hypotheses/conditionals in a conversational domain. They are critical determinants of analogy relations. The class of analogies is far larger than isomorphisms (though the formal similarities can all be represented as morphisms of some kind). It includes, for instance, "analogies of analogies"; and the "of" ordering induces a hierarchy of descriptors. Finally, the distinction syntactic/semantic bears upon the issue of simplifying a thesis (a matter of practical consequence in course design).

2. DESCRIPTION METHODS AND THE SEMANTIC COMPONENT OF A THESIS

The expert's choice of a head topic and of a distance from the head at which topics are marked as primitive, is part of a description he gives to the entailment mesh. Any but specially contrived meshes permit the choice of several topics, and any such choice gives rise to a family of descriptors. Choice of a head topic extracts the thesis, under this head, from a potentially indefinite plexus of related knowables; it also imposes a quasi ordering (subordinate/superordinate) upon the structure which is isolated.

Under this ordering, the head topic(s) is (or are) *superordinate* to all others, and are assigned to a depth of zero. Several numbering algorithms may be used to convert entailment arc distances into values of the superordinate/subordinate descriptor. The algorithm currently employed in EXTEND (which is a refinement of the program in the previous monograph), is designed, so far as possible, to place the terms of all analogically related topics at the same superordinate/subordinate *depth* just as analogous head topics are at the same, zero depth.

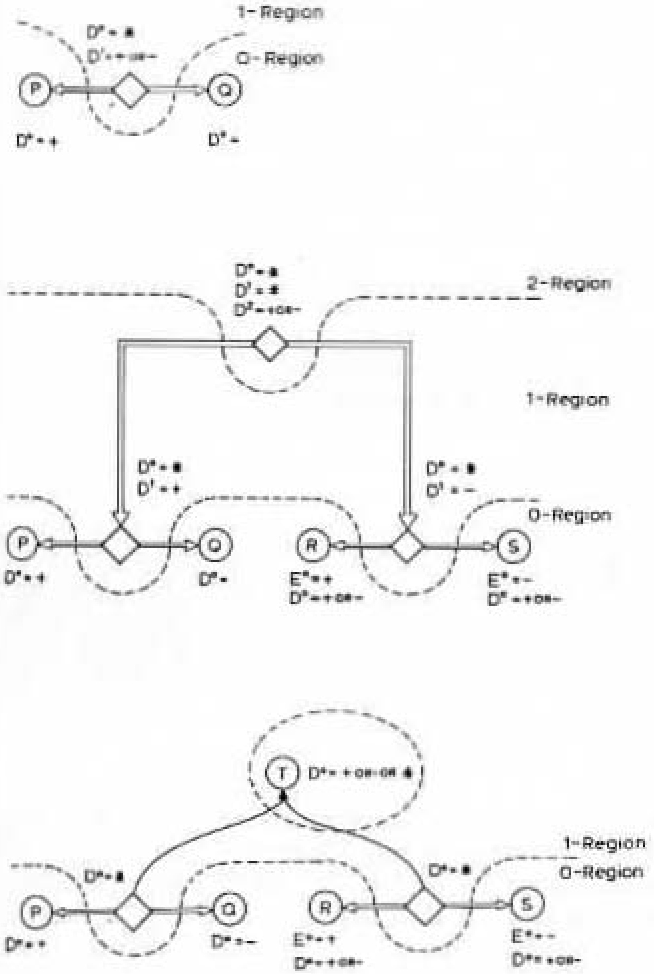
2.1. Forms of Analogy Relation

Suppose that a depth numbering scheme exists (one scheme will be described in Section 2.2), it is possible within the framework of a depth numbering to examine the analogies, if any, at a particular depth. Let us also anticipate the argument and suppose that semantic descriptors are to be chosen and given values on the nodes of one or more analogy relations and the topics which it/they relate. Semantic descriptors are unary but many valued predicates; for simplicity it is much more convenient at this stage to regard them as having the possible values "+" (meaning "has the property") or "—" (meaning "does not have the property") and "*" (either "irrelevant" or "undetermined"). In Chapter 4, it is noted that the semantic descriptors are really "Fuzzy Predicates" with more complex value sets and that the assumption of convenience delineates a limiting case. Descriptors (the predicate names) are symbolised D, E ,...; their values D = + or D = — or D = *.

Fig. 2.5(a) shows a standard analogy relation (for example between the real/abstract universes of "probability theory", relating topics P and Q. The central node represents the syntactic similarity between P and Q, the common rule or formal relation these share. Suppose the expert is required to discriminate P and Q (using one or more descriptors D for this purpose) so that the differences which refer to the analogy are delineated. Whatever D he chooses for this purpose, it is obligatory that if D = + on P, then to secure the discrimination, D = — on Q, and it will be intuitively evident that D = * on the node of the analogy relation; an analogy between topics cannot have the semantic interpretation of the topics, since it exists in a distinct analogical universe.

The rational justification for this intuitive statement is shown in Fig. 2.6(a); the semantic descriptor D itself (not its value) enters the analogy relation as the distinguishing predicate which captures the semantic difference component of the analogy. In general, the distinguishing predicate is a subset of an ordered set of semantic descriptors, and (Fig. 2.6b, Fig. 2.6c) any analogy based upon a similarity, U, may be reduced to an isomorphism between restrictions of the U similar topics.

Thus an analogy relation induces an hierarchical ordering amongst predicates. It could be expressed by a hierarchy of logical types, but, looking ahead to Chapter 4, it is more parsimonious to employ a property of Fuzzy Sets; namely, the elements of a Fuzzy Set may be Fuzzy Sets. Whichever notation is used, the hierarchical structure is represented as a series of regions, the 0 region and the 1 region of Fig. 2.5(a), with any node in the mesh belonging to a region. If the topics related by an analogy belong to region r, then the node of the analogy relation belongs to region r + 1. It is important to avoid any possible shade of confusion between depth



(a)

(b)

(c)

Fig. 2.5a, b, c. Analogy Relations, Descriptor Values, and Regions.

numberings, or levels, and the regions thus delineated. All nodes in Fig. 2.5(a) are at the same level and so are all nodes in Fig. 2.5(b), where the construction is iterated, as it may be indefinitely, by citing an analogy between analogies (alias, topics in Region 1 rather than Region 0) to generate a 2 region.

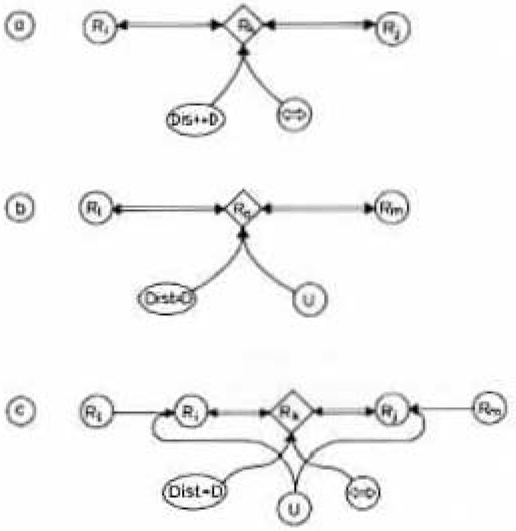


Fig. 2.6a, b, c. The distinguishing Predicate Dist on an analogy consists of one or an ordered set of predicate names that are used to indicate the *difference* between the analogous topics (here, R_i and R_j). The similarity of the analogical topic relation (R_k) is either an isomorphism (as shown in a) or a topic expressing the syntactic or systemic similarity (as shown in b) between R_l and R_m. This construction may always be reduced (as in c) to an isomorphism by restricting the analogous topic relations by U.

The region notation stems from the semantic descriptors and these are tagged by a superscript. For example, in Fig. 2.5(a), D^0 may have real values (+ or —) on nodes in the 0 region (and must have real values on the topics related by the analogy), but its value is $*$, by mandate, on nodes in the 1 region. Similarly, there is a descriptor, D^1, with real values (+ or —) in the 1 region and, in Fig. 2.5(b), a mandatory $*$ value on nodes in the 2 region.

Analogies between analogies are very common; especially so, it turns out, in physical science and other inherently compact subject matters. For example, the thesis on "energy conservation" used as a primary example in Chapter 7 is replete with them.

Another very common construction is a syntactic derivation involving the (syntactic parts of) two or more analogy relations of topic T in Fig. 2.5(c). The region convention clearly differs significantly; whereas an analogy between analogies with nodes in region r has a node in region r + 1, a derivation (like T) from analogies in region r has a node in region r. The model which is an interpretation of T exists in a distinct modelling facility. For example, in Fig. 2.5(c), if *topics* P and R are modelled in *MF*(X) and if topics Q and S are modelled in *MF*(Y), then *topic* T is modelled in *MF*(U) such that the models of T establish coupling relations between models built in *MF*(X) and in *MF*(Y). But notice that (though in the same region as the analogies) T is at a lesser depth.

Since the previous monograph was written, considerable effort has been devoted to analysing and representing analogies, motivated in part by the educational importance of analogies, properly used, as means for accelerating rapid comprehension of a subject matter. For example, though some analogies are isomorphisms (the type cited in Fig. 2.6) or isomorphisms valid for only some part of the related topics, others are generalisations. These varieties of analogy are amply discussed in subsequent chapters (notably Chapters 4, 6, 7 and 8) as they occupy a key role in innovative processes. Hence, generalisations are not examined at this juncture. It is, however, opportune to review one quite innocent complication which was mentioned in the previous monograph; namely, that analogy relations are not restricted to relating two topics.

Some of the more important many place analogies are shown in Fig. 2.7. Reading through the examples, Fig. 2.7(a) says that *topics* P, Q and R are analogous (their similarity represented in the central node, differences entering the central node as Dist). In Fig. 2.7(b) *topics* P, Q, and R are related by (possibly different) analogies. Fig. 2.7(c) asserts that the (different) analogies are themselves analogous. This construction is in register with Fig. 2.5(b), and Fig. 2.7(d), by the same token, is in register with Fig. 2.5(c). Fig. 2.7(e) expresses the existence of two analogies (x and y) between topics P, Q and R. For sensible discrimination x and y will be demarcated in terms of distinguishing properties that capture differences but also in terms of distinct (syntactic) rules (one to x and one to y). Even so, it often happens (Fig. 2.7(f)) that x and y have common features related by analogy between analogy relations (u). The constructions of Fig. 2.7(a), (b), (c) are all exemplified by the "real" department of "probability theory" (previous monograph) where P, Q, and R are topics in "games of chance", in "behavioural experiments" and in "genetics". The different con-

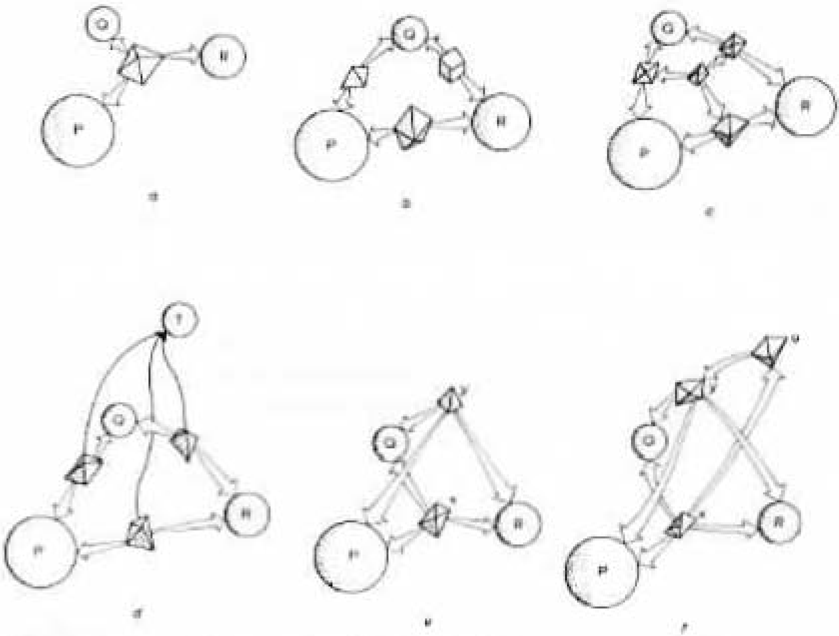


Fig. 2.7a, b, c, d, e, f. Complex Analogy Relations.

structions are appropriate to different levels and were deliberately glossed in the earlier treatment, as they may quite legitimately be, because the "real" nodes in this subject matter have the calibre of T in Fig. 2.7(d). The other constructions are more convincingly referred to generalised analogy relations of the kind we have promised to examine (in fact, any generalisation can be represented either in the fashion of Fig. 2.7(e) or else of Fig. 2.7(f)).

2.2. Depth Numbering

This preliminary discussion of analogy relations rested upon the idea of a depth numbering, the analogies being anchored to some depth. All depth numbering schemes rely upon the following types of process.

(a) A means for detecting the nodes of analogical topic relations.

(b) A means for determining the region of a node, using the 0 region nodes as a baseline.

(c) Some numbering arrangement that orders the nodes in a mesh from a head node (or a cluster of analogical head nodes), assigned a depth of 0, that are located in the 0 region.

Analogies are detected in syntactic terms by noting that they differ in establishing some kind of morphism. Hitherto, only the isomorphism operation was seriously employed; since the mechanics of generalisation have been studied, there is a general morphism (a mapping between relations that preserves some formal relation). In the scheme we employ, isomorphism $\Leftrightarrow$ appears as a relational operator; so, now, does a general morphism. If he employs the isomorphism, the expert is provided with a place holder node (Dist = ?) to accommodate the distinction between universes of interpretation required to maintain the integrity of isomorphism in contrast to equality; a similar distinction is needed if a general morphism is invoked. The nodes associated with these operators and placeholders (Dist = ?) are marked, mechanically, by an analogy detection algorithm. They are listed together with nodes, like T in Fig. 2.5(c), that represent derivations from analogy relations, *provided they are not part of a derivation re-entering nodes in the 0 region* (if the italicised condition is false, they will be numbered from their 0 region entailments). Call this list the *analogy list*.

A further algorithm is applied to the union of the original node list and the analogy list. Nodes that are not members of the analogy list are assigned to Region 0. The analogy list is now searched for analogies between nodes in Region 0 and these, together with nodes corresponding to immediate derivations (like T), are assigned to Region 1. The process is iterated, at the next stage finding analogies (between analogy nodes) in Region 1 which are assigned to Region 2, and continues until all the analogy list entries have been exhausted (for Regions 0, 1, ..., r, ..., r_{max}).

Finally, a depth numbering algorithm is applied to the original mesh and the distinguished (and region assigned) analogy list. This algorithm operates from the head downwards, first, with nodes in the 0 region. So far as possible, it satisfies the condition that the nodes related by an analogy and the analogical node itself are placed at the same depth. It is not always possible to satisfy this condition, and the expert is given the option of deleting an analogy he has previously inserted or of permitting analogies that cross between depths. Such analogy relations are not necessarily pathological and can be handled. As they are rare and because handling them greatly complicates the description process (to follow), it will be supposed that all 0 region topics related by one analogy are at the same depth. Having exhausted the 0 region nodes and assigned them depth numbers, the algorithm next addresses any nodes in the analogy list that are derivations from analogies and have no direct derivational link to 0 region nodes, again operating from depth 0 downwards.

2.3. Improved Method for Eliciting Descriptions and Their Values

Meshes are numbered as they are isolated from their surrounding (the pruning of the first monograph), and in practice pruning and depth assignment are carried out automatically before the current mesh is displayed to the expert. Since a mesh cannot be a simple chain of nodes, it is evident that the superordinate/subordinate descriptor does not uniquely name each node and the onus is placed upon the subject matter expert to select and assign values to further descriptors ("unary but many valued predicates of the nodes") so that:

(a) Statements of the conjoint values of the descriptors uniquely ostend one node (there is at most one node, standing for a topic, in each "cell" of a grid made up from descriptor values).

(b) Some "cells" are empty.

However, no restriction is placed upon the number of descriptors employed, and the description scheme may be as redundant as desired.

From the student's point of view, the descriptors, or some subset of them which he can show that he understands, furnish the means for exploring, gaining access to, and learning about the topics.

From the expert's point of view, it is useful to separate descriptors into the categories, syntactic and semantic.

The values of a syntactic descriptor, such as superordinate/subordinate, say nothing (except perhaps to the expert) about interpretation. They are properties (in this case a "depth" or "arc distance" property) of all the nodes in a mesh. The entire mesh could be described in these terms as an abstract graph and, for that matter, the syntactic component of this thesis, revealed in the derivation structure, could be described as an uninterpreted and formal system. Under these circumstances, however, it is difficult to see how a student could make sense of it; at any rate, since the incorporation of aim validation (Chapter 1), a student would not be allowed to use only syntactic descriptors when specifying his aim and starting to learn.

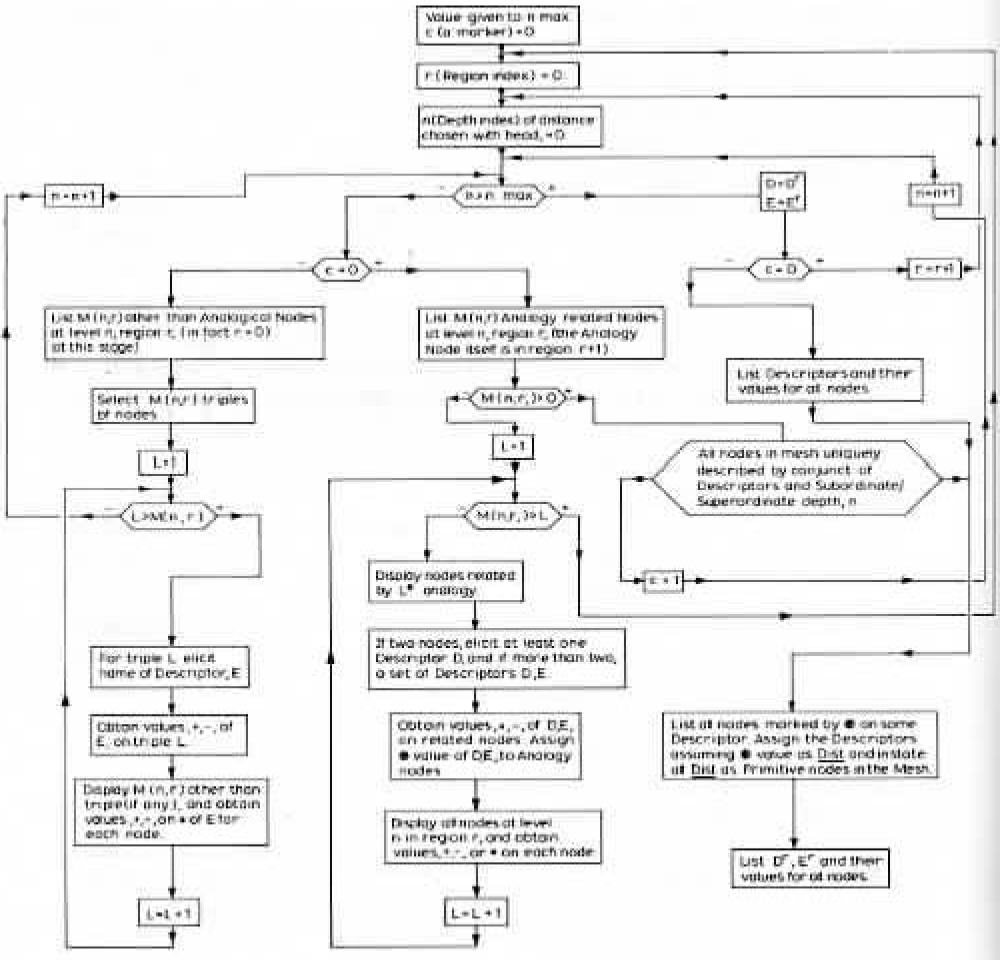
Semantic descriptions have values that refer to the universes of interpretation in which explanatory models for topics are realised as programs. One semantic descriptor is the head name (notice, this name is the value of a semantic descriptor, though the values, 0, 1, ... of subordinate/superordinate depth are values of a syntactic descriptor). Other semantic descriptors carve up the topics in various ways. For example, "steam engines" and "heat pumps", or "turbines" and "piston impulsion", in the "energy conversion" thesis of Chapter 7, or electrical/mechanical in physics. The current recommendation is that large numbers of semantic descriptors are specified.

Apart from the superordinate/subordinate descriptor, which is derived automatically once a head topic is chosen, the remaining descriptors are systematically elicited as "personal constructs" (Kelly 1955) using a modified repertory grid technique (Bannister and Mair 1968). The objects over which the personal constructs are elicited are the nodes in the mesh.

However, insofar as the expert is really evaluating interpreted explanations (models) of the topics which the nodes stand for, the constructs are semantic descriptors and convey substantive meaning. Even so, they are treated uniformly as unary (many valued) predicates of the nodes. For expository convenience we limited the values in the last section to +, —, and * (irrelevant). This limitation is inessential, but whatever values are permitted, the value * (irrelevant) must be preserved.

The names and values of the descriptors are elicited mechanically by a program akin to Thomas's (1971) DEMON. The chief peculiarity lies in the way that nodes are sorted and presented to the expert (as the objects having, or not having, a property).

The descriptor eliciting procedure is outline charted in Fig. 2.8. It accepts as an input a mesh with depth numbering (n) and regions (r) already specified, and its output is a described mesh to which is adjoined a set of primitive nodes representing the descriptors D, E, which figure as the distinguishing predicates (Dist) of analogy relations. The remaining descriptors, (d, e, ...) if any, that are eli-



● is machine marked value of ● and node so marked is recognisable

Fig. 2.8. Outline Flow Chart for Descriptor Elicidation Process.

cited to safety condition (a) and (b) for other than analogical topics are listed but are not represented by nodes.

Several points are usefully kept in mind whilst reading this flow chart. First, when the expert is asked to choose the name and values of a descriptor (alias a personal construct, or a property) with respect to a set of nodes, he is really being asked to contemplate the models which will, on execution in an appropriate modelling facility, satisfy the topic relation. Semantic descriptors are properties of this interpretation.

Next, the "model" of an analogy relation between two (or more) topics is a coupling between two (or more) models, distinguished by execution in a priori (without the coupling) independently clocked processors and by the distinguishing predicate (Dist) which is specified by way of the selected descriptors.

Finally, although the program which realises this flow chart can be interfaced with the expert using a teletypewriter terminal, this expedient is completely impracticable except for the simplest meshes. All practical systems employ a display of the mesh which is continually accessible to the expert and an "interrupt" which provides the expert with the displayed values of the descriptors he has so far chosen, superimposed upon the nodes in the mesh. One interface of this kind is described in Chapter 7, but most graphic consoles will provide the required facilities.

2.4. *Tutorial Materials*

The described and pruned mesh is transformed into an entailment structure (Fig. 2.2) by encoding (either in computer storage or the hard wired form of Chapter 1), each node being associated with storage locations to indicate its state as learning proceeds.

Tutorial materials are based upon demonstrations constructed from the $BG(\mathrm{i})$ as task structures $TS(\mathrm{i})$ (previous monograph), together with the "How" questions ($\underline{\mathrm{EQuest}^0}$ and $\underline{\mathrm{Comm}^0}$ and their qualified forms). "What" questions ($\underline{\mathrm{PQuest}^0}$) span the topic relations, again as described in the previous monograph.

In Chapter 1, we noted that experience with both operating systems, CASTE and INTUITION, has underscored the necessity of providing rich semantic data in response to explore transactions, and shown, also, that an *aim* must be validated before it is accepted by the system. The data provided when a topic is explored (by citing a conjunct of descriptor values that ostends and uniquely identifies the topic) consist in one or more slides. The artwork is important (some examples are shown in Chapter 1), but is generated systematically as a series of illustrations that exemplify the topic and a series of counterexamples that differ in one or more descriptor values.

Aim validation questions (of the form PQuest[1] in the previous monograph, since they refer to subsets of nodes) are multiple choice questions having one and only one (correct) response alternative that illustrates the descriptor values conjoined to identify the topic. The remaining response alternatives (incorrect) show counterexamples differing in one or more descriptor value.

3. SOME USEFUL OPERATIONS UPON ENTAILMENT MESHES

Insofar as the derivations of a thesis are retrievable, it is always possible to generate a binary decomposition of any conjunctive or disjunctive (but not analogical) structure in a given mesh. Each kernel in the binary decomposition has exactly two members.

Labelled clusters of relational operators (Fig. 2.9), reduced to a kernel in the entailment mesh, are replaced by sequences of the complete subset (Natural Join, Projection, Union) of operations. These sequences are arranged in order and further nodes are introduced (Fig. 2.9). These nodes stand for topics which were not made explicit in the original thesis (and which in general need not be made explicit), but which are needed to satisfy the requirement that each kernel has two members.

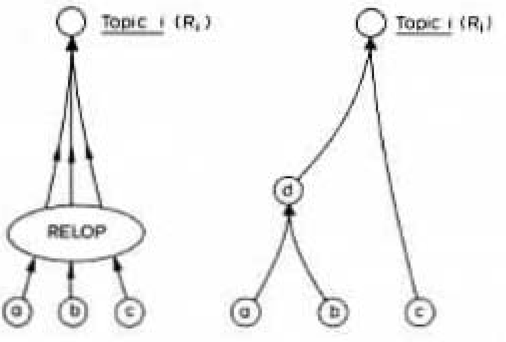


Fig. 2.9a, b. Binary decomposition. (a) A conjunctive substructure (kernel) in which *topic* i with formal relation R_i is obtained from a, b, and c. In the original thesis the derivation was labelled by a complex of relational operators Relop. (b) One Binary Decomposition. The components of Relop are replaced by sequences of {Natural Join, Projection, Union} and nodes, such as d are introduced to represent intermediary relations.

3.1. Trade Off Methods

The binary decomposition of a structure showing the derivation of a topic relation R_i (at its head) together with all Behaviour Graphs, *BG*, of its primitive nodes (*BG*(a), *BG*(b), *BG*(c)) has as much information or specificity as the relation R_i and its Behavioural Graphs *BG*(i).

It is also true that an undecomposed structure representing the same topic, R_i, with the same task structures attached satisfies this condition; in fact, if B (Fig. 2.10) is a binary decomposition of A (Fig. 2.10), then A and B contain the same amount of information or specificity.

The information or specificity is differently arranged. In A, it is relatively localised, since most of it is packed into the Behaviour Graph or, tutorially speaking, the task structure, *TS* of R_i. In B, it

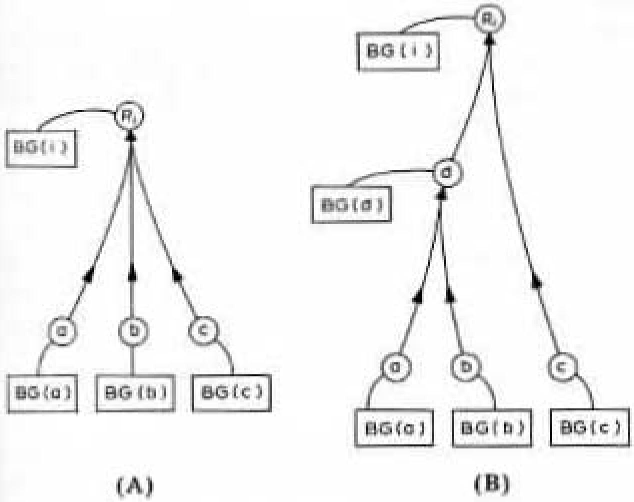


Fig. 2.10A, B. Trade off and the distribution of specificity or information between the entailment mesh and the Behaviour Graphs/Task Structures, connected to its nodes. The redundancy in any conversational domain (even with purely conjunctive mesh) should not be confused with redundancy in disjunctive mesh representing alternative derivations of the same topic. Equalities: If Sp = Specificity, then Sp(*BG*(i), R_i) = Sp(*BG*(a), *BG*(b), *BG*(c), Derivation R_i from a, b, c) = Sp(*BG*(d), *BG*(c), Derivation R_i from c, d).

is distributed over the network. We comment that a trade off is always possible. Though a behavioural specification *BG* or *TS*, and a cognitive (relational network) specification are distinct, and though they are both needed in a tutorial system, their combination is also fundamentally redundant.

Hence, within limits, there is a systematic method for deploying the information in a thesis in an educationally desirable manner. It may be conveyed primarily by demonstrations and the tutorial materials attached to them, or primarily by an entailment structure display, or, redundantly, in both ways.

There are restrictions upon the kind of information which is traded off in this manner and upon the amount of trade off which is possible; namely:

(1) Kind. The traded off information is in the syntactic (not in the semantic) content of a thesis; the semantic information is conveyed by descriptor values and in exemplary data, accessed by explore transactions.

(2) Amount. The distribution which maximises the information in the entailment structure is obtained by constructing and displaying a binary decomposition of the underlying relational network (as in B of Fig. 2.10). The distribution which minimises the information in the entailment structure is obtained by maximising the number of arcs that contribute to the derivation of a topic (as in A of Fig. 2.10). The limit is set by the following rule: "no essential precedence ordering may be omitted." Thus, in A, there is only one precedence requirement (a, b, c must all be understood before R_i is understood, but a, b, c may be studied in any order, or simultaneously). In general, this is not the case, though it *is* possible to eliminate precedence orderings that are not required on syntactic or computational grounds.

Binary decomposition and trade off work for disjunctive structures, but some care is needed to avoid confusion. Any disjunctive structure represents the fact that the same topic may be derived in several ways, or that the thesis is redundant. This redundancy is quite distinct from the redundancy immanent even in conjunctive structures, due to the fact that the entailment structure and the task structure have information in common. So long as this distinction is appreciated, disjunctive structures may be reduced to the set of all possible conjunctive components and dealt with as before.

3.2. Simplification

A locally cyclic (conjunctive or disjunctive) structure of topic relations stands as an understandable topic. This is emphasised by drawing a line around a structure headed by the topic in question; for example, R_j in Fig. 2.11(a).

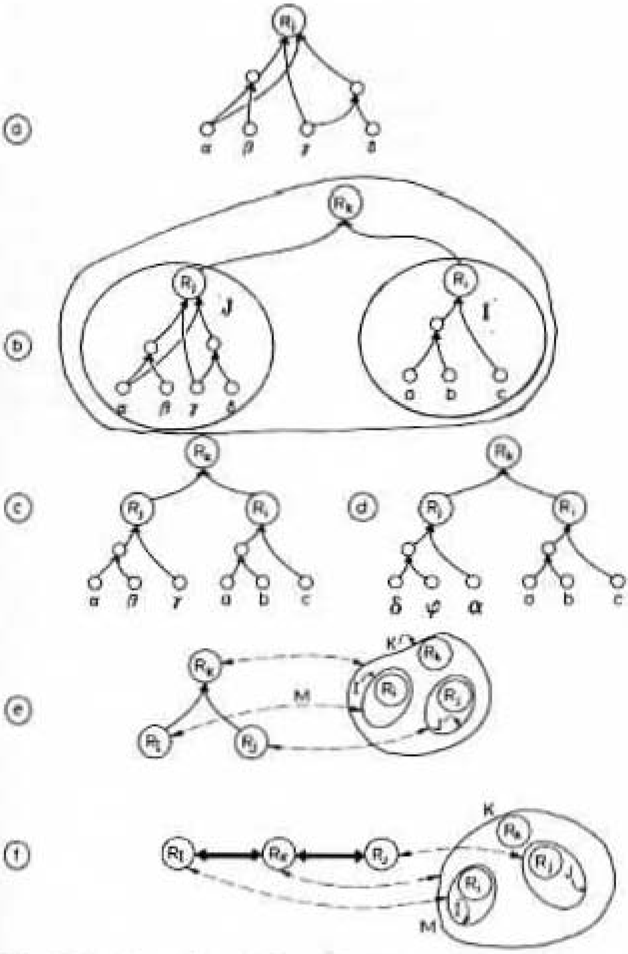


Fig. 2.11a, b, c, d, e, f. Simplifications. The circular regions in (e) and in (f) are those delineated in (b).

Call the circumscribed region J (since it is headed by R_j). J forms part of a system, insofar as the circumscribing lines are nested with respect of superordinate topics naming hierarchically arranged subclasses such as I (headed by R_i), J (headed by R_j), and K (headed by R_k), in Fig. 2.11(b).

"What is the simplification of R_i (or of R_j) in the context of R_k?"

One answer to this question is that a simplification is any irredundant or conjunctive structure, compatible with the original, and yielding the same derivation. For example, the structures in Fig. 2.11(c) and Fig. 2.11(d) are simplifications (in this sense) of R_j; there is no simplification (in this sense) of R_i. This sort of simplification (by "selection") implies that since there is less content to a course representing an irredundant thesis than there is to a course representing a redundant thesis, the "selected" irredundant representation is "simpler". Though of dubious utility (since the irredundant representations are rarely easier to learn), there is an algorithm for extracting all such "simplifications" from a given structure.

A very different kind of simplification (by consistent "smudging") maps the circumscribed regions I, J, and K of the original picture onto points representing nodes in a distinct network (Fig. 2.11(e)).

The mapping (M in Fig. 2.11(e)) is plausible enough. What must be ascertained is the precautions needed to ensure that M gives rise to a coherent simplification rather than a mess.

There is no difficulty in convincing oneself that simplifications exist, that they are widely employed in practice, and that they are used to good effect. For example, let R_k represent a statement of the gas law $P^* \times V = \text{Const} \times T^*$ as conceived by an elementary student for whom P^* is pressure and T^* is temperature, taken as matters of experience (how much "push" there is, how "hot" it is), though being, of course, susceptible to measurement. V, the volume of a container, and Const (the gas constant) are understood as thoroughly as required at any point in the course of studies for which the entailment structures have been devised.

Conversely, let R_k represent the gas law $P \times V = \text{Const} \times T$, as conceived by a fairly sophisticated student, for whom P and T are known in terms of the motion of idealised molecules and the mean kinetic energy of these idealised molecules, the volume V having the meaning it has for the elementary student. If Boltzman's constant is S, the P and T terms are defined for the advanced student by equations such as:

$$P = \frac{1}{3}\frac{N \times m \times Z^2}{V}$$

and

$$T = \frac{2}{3S}\left(\tfrac{1}{2}m \times z^2\right)$$

where $(\frac{1}{2}m \times z^2)$ = Mean Kinetic Energy, m = Mass of an idealised molecule, N = Number of idealised molecules in gas, Z = the mean velocity of idealised molecules.

The mapping M is legitimate since it may be maintained (by a physics master, for example) that if the elementary student used the prescribed measuring methods on objective reality to reach (an obviously simple minded) understanding, it would still be the case that (numerically) $P = P^*$ and $T = T^*$.

The relevant psychological requirement is that in the context of a course up to R_k (which determines, for example, the uniform connotation of volume V), no statement made in teaching R_K and understanding R_I and R_J as its prerequisites shall contradict or falsify any statement made later (when more complex material is presented) in teaching R_k and understanding its prerequisites, R_i and R_j. Of course, more "true" statements appear in understanding the "enriched" or detailed course materials.

Mappings, M, that satisfy these requirements exist if the primitives of R_k, R_i and R_j belong to (are modelled in) the same universe of interpretations, say U. It is also *possible* that topic R_K is an analogy and that its separate terms are modelled in distinct universes of interpretation, R_I in X and R_J in Y (Fig. 2.11(f)).

In general, analogy relations cannot be simplified by consistent smudging, though all of the conjunctive or disjunctive subtheses that are analogically related may be simplified. The particular example of Fig. 2.11(f) is exceptional insofar as there is a thesis containing some conjunction, to which the analogy is subordinate. Such a structure unifies the distinct universes X and Y.

For example, let I stand for elementary physics and X for a universe where Temperature (T^*) is "hotness" and pressure (P^*) is "push". Let J stand for advanced physics and Y for a universe where temperature and pressure have the other meanings T and P. There is a thesis about science which unifies X and Y in the sense that all X measurements or actions are open to expression as clusters (homomorphic images) of measurements or actions in Y. The analogy, in this case, is a cognitive reality but is not epistemologically essential.

3.3. Discussion

The educational uses of trade off and of simplification by consistent smudging are fairly obvious, though the merit of simplification becomes most obtrusive for really large scale subject matters. The following notes are an attempt to augment the concept and to exhibit the advantages in terms (as usual) of realisable operating systems. It is not too difficult to bridge the gap between quasi mechanical (but definite) realisations and classroom practice.

Just as a topic is described, so may a class of topics be afforded a coarser grained description. For example, the class named I is described by subsets of the values of the descriptors of the topics within class I, and such subsets are readily pointed out more economically by the values of additional descriptive predicates; call them *attributes*, for reference.

Using explore transactions in the coarse grained attribute space (in contrast to the fine grained descriptor space), a student can locate I or J and determine its properties. Moreover, he can establish his aim on I; meaning "on the head node of R_i in I".

At this point, supposing the operating system accommodates the underlying fine grained structure, he can mechanically "zoom in" on the detail; for example, to engage in a fine grained exploration or to relocate his aim at some node (other than the head node).

A coarse grained display of a large structure in an attribute description, circumscribing regions like I and J, is generally desirable, provided it is possible to retrieve the underlying fine grained structure and its descriptors. Practical implementation involves an interactive graphic display, the structures in question being represented in computer storage.

Under these circumstances, there is no objection to storing the entire derivation as a relational network together with its cyclic components, and it is possible, as a result, to realise an identity

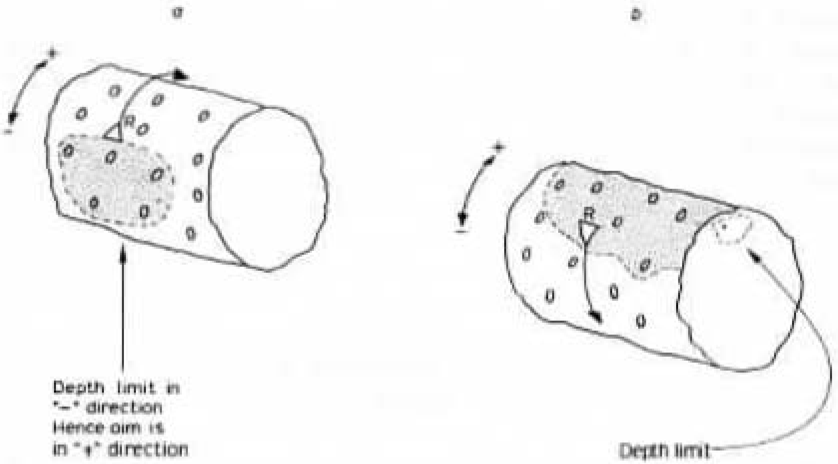


Fig. 2.12a, b. The student's aim as an oriented or directional marker on a par with the expert's head and depth. For simplicity the cyclic mesh of the stored derivation is wrapped round a cylinder (primitive topics are thus at the edges) (a) shows an aim oriented from node of topic R in one direction (+); (b) in the other direction (—). Only nodes in shaded region are displayed to the student, but he may vary area or depth.

between the aim topic chosen by a student and the head topic chosen by a subject matter expert. The student's aim of necessity becomes a vector, corresponding to the expert's "head and depth", naming the aim topic itself and a lower boundary, which may be established in several directions.

Fig. 2.12 shows two such directional aims which reverse the orientation of the syntactic depth descriptor (subordinate/superordinate). For all that, the underlying derivation is unchanged and the values attached to semantic descriptors are unchanged whichever of the two (or more) aims is selected.

4. DERIVATIONS

As noted at length in the previous monograph, the syntactic or derivational component of a thesis is represented in terms of formal topic relations (subsets of a product set) and relational operators that transform relations into other relations. The calculus of relational operators was introduced into data base design by Codd (1970) and the originality, if any, of the present approach resides in how the topic relations and derivations are specified ("from up downwards" rather than from "basic unit upwards"), certainly not in how the relations are manipulated.

Even in the field of education, other researchers have independently developed comparable schemes with their own peculiar advantages; the differences are chiefly notational. For example, Scandura's (1973) "Structural Learning" Techniques represent topics (Scandura calls them "Concepts") as sets and functions rules and "higher order" rules. Bunderson and Merrill (1973), together with their colleagues working on the TICCIT computer aided instruction system, have much the same approach. The topics appear as sets, functions and relations abutted by compositions and set theoretic combinations that either are, or are equivalent to, relational operators.

These and similar spirited schemes referenced in the previous monograph have proved useful and flexible. The present work deviates only in respect of how the topic relations and derivations are elicited (as noted already) and in the emphasis placed upon analogy relations. Though very comprehensive in most respects, the other schemes are not primarily intended to uncover the structure of analogies (as this scheme *is*).

There is nothing sacred about the choice of relational operators as a canonical means for representing derivations. The calculus is used metalinguistically and by programs like EXTEND which sort out derivation paths and determine legality. Any other competent calculus would serve just as well. In particular, the "axiomatic" schemes due to Steltzer and Kingsley (1974) are more appropriate, more amenable to manipulation by a subject matter expert, and more clearly exhibit the distinction between the syntactic (formal, axiomatic, derivational) part of a thesis and its semantic content. A good deal of our recent work has employed this axiomatic scheme in place of our augmented relational operator scheme.

As in the present discussion, Stelzer and Kingsley distinguish between what may be known (the theses represented in a GCN or General Cognitive Net) and what may be done (a set of *BG*s or Task Structures). Only the derivational component (the GCN) will be discussed.

An axiomatisation of a thesis about a subject matter (represented as a GCN) is rooted upon the following categories of objects called constituents: primary notions, derived notions, basic principles (axioms), and established principles (theories). The constituents x, y enter into two relations F (x, y) ("y is formulated in terms of x") and E(x, y) ("y is established in terms of x"), and these relations may hold as follows:

F(x, y)
Possible constituents in x, y

x	y
Primary Notion	Derived Notion
Derived Notion	Basic Principle

E(x, y)
Possible constituents in x, y

<table>
<tr><th>x</th><th>y</th></tr>
<tr><td>Primary Notion</td><td rowspan="4">Established Principle</td></tr>
<tr><td>Derived Notion</td></tr>
<tr><td>Basic Principle</td></tr>
<tr><td>Established Principle</td></tr>
</table>

The GCN may be expressed as the complex of relations type F and E holding between a set of constituents. Since the intention is to obtain an axiomatisation, the GCN will be minimally redundant, but there is no necessary restriction upon the order in which parts of the complex relation are spelled out, nor upon the order in which the final constituents are chosen. It is evident, on inspecting Steltzer and Kingsley's examples, that GCNs correspond to generally conjunctive derivations which exhibit the kernel structure of a subject matter.

The GCN rules (for using F, E, and so on) are designed to prohibit loops; hence, analogy relations (which surely hold between the task structures; for example, the course on photography, one instance in the 1974 paper, has several universes of interpretation) are not made explicit. The prohibition is computationally convenient as well as axiomatically defensible but is unacceptable (on psychological grounds) from the present point of view.

Several kinds of compromise are possible. Our present approach is to obtain conjunctive substructures as GCNs, to adjoin an extra-axiomatic postulate that any established notion or principle is cyclic (consistency is guaranteed), to form disjuncts of GCNs after they are constructed, and to add on analogies between the F, E relations of the GCNs by an independent process. In other words, GCN rules are used locally in course assembly and the local products (GCNs or conjunctive structures) are unified by the methods already outlined or to be described.

5. THE SIGNIFICANCE OF ANALOGY RELATIONS

In hindsight, it was fortunate that conversational domains were first constructed for theses dealing with applied science. As a result, we were forced to take the representation of analogies seriously from the beginning.

In particular, analogies are non-verbally explained by executing two or more models that are built in two or more a-priori-independent processors or universes of interpretation (*MF*(X) and *MF*(Y)) together with a coupling that establishes their dependency. Though at first sight this looks like an overly complicated technique, and at the next glance seems to be a statement of the obvious, it turns out to be one starting point for a theory of innovation.

Any thesis represented in an entailment mesh is a justifiable hypothesis expounded by someone, the subject matter expert. He may remain anonymous until more than one thesis is represented in the mesh, for example, more than one scientific theory or an overall thesis about several rival hypotheses. In this case, it is necessary to name the advocates or protagonists as people, schools of thought or whatever. Call them A and B. Now A's thesis is justified insofar as A can model it in some universe and B's insofar as B can do the same in another universe, and there is a sense (to be developed in chapter 4) in which these universes are a-priori-independent.

The basic transaction between A and B, regarded as dynamic entities in conversation, is an agreement over their theses, including an agreement to differ. This agreement may sometimes be founded upon an additional act (a constituent of verification and falsification methods) whereby A's thesis and B's thesis are modelled in a common or reference universe. But, prior to that, A and B must agree upon or accept a reference universe (as a student does when he subscribes to an experimental contract). In either case, the microstructure of agreement may be complex as it will entail A's hypotheses about B (and B's hypotheses about A), in addition to the theses, alias hypotheses, to which they overtly adhere. If the act of agreement is frozen and the result inscribed in an entailment mesh, then it is an analogy relation.

Conversely, any analogy relation represented in a mesh is the inscription of a petrified agreement between people, and there is a sense in which the dormant and possibly unnamed participants (A, B) are resuscitated when the analogy is understood. This may illuminate the obscure, even cryptic, remark in the previous monograph that the basic utterances in an L Conversation are agreements, the basic statements are L Metaphors designating analogy relations. Any thesis contains such a unit, explicit or not. Most theses of interest contain many.