Chapter 4

Theoretical Developments

Only in recent years have mathematical logicians seriously considered systems that are constructed from the point of view of a participant; that are, in the non vicious sense of the first monograph, "subjective" and "reflective". The most comprehensive development is due to Andreka, Gergeley and Nemeti (1973a,b); but several, seemingly quite different pieces of work complement the picture. In toto, these mathematical systems lend credibility to the "string and sealing wax" formulation of conversation theory. These additions could be advanced independently, as systemic propositions which are supported by empirical data. But, since their otherwise peculiar form fits the larger and axiomatically respectable framework, a more convincing case is made if they are viewed as instances of this general and well-formulated system.

The logico-mathematical advances bearing upon reflective systems belong to the following areas of study:

(a) A non classical and model theoretic treatment of languages and logics (Andreka, Gergeley, Nemeti 1973a,b).

(b) A general formulation of Fuzzy Predicates (Goguen 1968).

(c) A theory of Fuzzy Algorithms, and Fuzzy Sets (Zadeh 1971).

(d) Coherence Theory (Rescher, 1973). *

* Since this book went to press, certain important theoretical developments have taken place and the list should be updated by two additions, namely, (e) Varela's logic of self reference and (f) work, chiefly due to Goguen, on the category theoretic foundations of General System Theory. These recent developments are briefly outlined in a footnote at the end of the chapter on p. 162.

These developments are discussed in the context of conversation theory (not always in the order listed above), to provide the general framework promised in the last paragraph.

In this chapter we review (a), (b), (c), and (d) as they apply to conversation theory and, where necessary, generalise the dicta and definitions of the first monograph. At this juncture, attention is still concentrated upon strict and *one-aim-at-once* conversations (though the underlying mental operations are often more liberally conceived). Even within this framework, it is possible to advance the notion of a *common meaning* reached by an agreement having a syntactic and a semantic component. In particular, we develop the idea that topics in a conversational domain which stand for analogy relations between other topics, are static inscriptions of a *common meaning.*

Analogy relations have a curiously central position because of their educational significance (only by using them accurately, can the student genuinely accelerate his learning of a subject matter), and because the appreciation of analogy relations is at the root of innovation and discovery. In order to introduce these ideas, it is convenient to describe certain augmentations of the transactions permitted in a conversational operating system such as CASTE or INTUITION. The augmented transactions were mentioned in the first monograph, but have been incorporated since it was written (in 1973). The understanding of analogy relations is discussed in these terms, and some ideas about the construction of analogy relations are sketched by way of introduction to the next chapter. Little more can be done until the "one aim at once" condition is relaxed.

1. GENERAL AND NON-CLASSICAL LANGUAGE, LOGIC AND MODELS

A language has a semantic (or interpretative) as well as a syntactic (or formal) status. The conversational language L is necessarily of this type, so are the languages in Barralt-Torrijos and Chiaraviglio's formulation, referenced in the first monograph. But the notion contrasts quite sharply with an "uninterpreted formal language", a purely syntactic construct of symbolic logic. Notationally, a language is a triple:

Language $\triangleq$ ⟨Set of Sentences, Interpretation Function, Universe of Interpretation⟩

or, for brevity

$\mathcal{L} \triangleq \langle S, I/F, \underline{Univ}\rangle$

These entities and their relations are shown in Fig. 4.1. There are indefinitely many languages. Labelling any one of them by an index i:

$\mathcal{L}_i \triangleq \langle S_i, I/F_i, \underline{Univ_i}\rangle$

and they may differ in any or all, of their component terms.

A logic is conceived as any pair:

Logic $\triangleq \langle \mathcal{L}_i, \text{Calculus}\rangle$

where the calculus is capable of expressing algorithms or programs, themselves syntactic entities which generate sentences in a set, S.

Clearly, a calculus could degenerate to one program or a certain class of programs (Prog). Generally, we equate the notion of calculus with this degenerate form.

$\langle \mathcal{L}_i, \underline{Prog}\rangle = \langle\langle S, I/F, \underline{Univ}\rangle, \underline{Prog}\rangle$

where Prog produces some, or all, members of S.

Further, the conversational language L is held in mind as $\mathcal{L}_i$ or a class of $\mathcal{L}_i$ accommodating full (or degenerate) logics.

At the cost of stratification (as in the conversational language $L = L^1, L^0$), or any other trick which discriminates between a description and an instruction to bring about whatever is described, it is possible to incorporate goal descriptions. Thereby, a pragmatic of goal satisfaction is adjoined to the syntactic and semantic system to form a semiotic system.

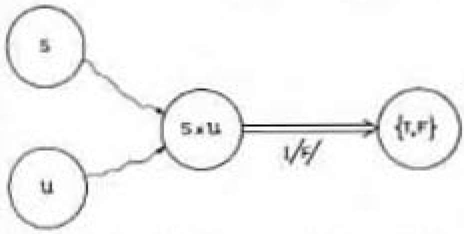

Fig. 4.1. Graphical representation of a language L, considered as a triple, consisting in a set S of syntactically admissible statements, a universe of interpretation U, and an interpretation function which maps the product of S and the universe onto values in a truth set (here {T, F}).

The motivation behind this development is that a sentient being (unspecified, but aware) is able to have indefinitely many languages, $\mathcal{L}_i$, and in them to imagine as possibilities and contemplate an indefinitely large number of realities (universes). In any universe, the interpretation of those sentences of S that are true (in this universe) is a model.

For a classical logic, the truth set is {True, False} (Fig. 4.2). But the overall scheme is designed to accommodate non classical logics. For example, a logic of action or of command execution (Rescher 1966), as sketched in the first monograph, is non classical. Statements prescribe operations. The statement of a procedure is true in a certain universe, if this procedure satisfies a mooted goal in this universe. For another example, a logic of Fuzzy Predicates in a non classical logic: the truth set is values in {Interval 0, 1; Meaningless}. A logic of Fuzzy Programs that compute the values of Fuzzy Predicates is also non classical and is of special interest.

One noteworthy aspect of the logic and language under discussion is its systemic orientation. Most treatments of model theory are constructed according to canons of parsimony and are directly applicable only to the simplest situations. So, for example, a universe is generally regarded as a set of elements, objects, or at the most, events. In the present case the restriction is waived, as it must be in the interpretation of a logic of action and operation. The universe can have the characteristics of a processor. Using the terminology of the first monograph, universes of interpretation are M-Individuals (one or other sort of processor). The *model* for an action engendering statement (Prog, for example) is a compilation of Prog. Moreover, time is implicit in the universe (perhaps only in

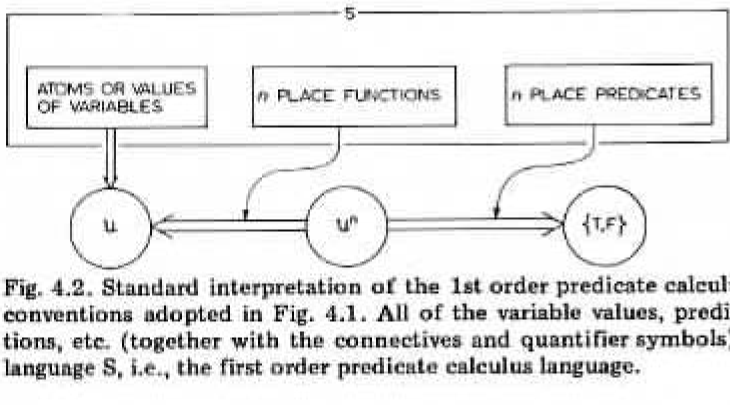

Fig. 4.2. Standard interpretation of the 1st order predicate calculus using the conventions adopted in Fig. 4.1. All of the variable values, predicates, functions, etc. (together with the connectives and quantifier symbols) are part of language S, i.e., the first order predicate calculus language.

the weak sense of order and the injection of negentropy to set a process in motion). But it is neither necessarily nor usually the case that time is uniform, so that different parts of the processor are a priori synchronised. If synchronicity exists, this is a special constraint built in as part of the syntactic statement which is given an interpretation either as a compiled and executable program, or as a result of productions manifest when instructions are taken in the imperative form, or as a special condition (for example, in the first monograph, the characterisation of modelling facilities, as "one clocked" or "many clocked" processors).

It is possible to view a scheme or system of this type from two equally legitimate perspectives, by considering the various processes that are licenced by the scheme from either an outward or an inward looking stance.

Of these two perspectives, the outward looking is less familiar and more definitely subjective or reflective (though in a sense, both of them have a reflective component). The notion underlying the outward looking perspective is that languages and, a fortiori, universes can be generated constructively in some medium which will be identified (as suggested already) with a *processor*. That is, the interpretation function I/F is regardable not only as a mapping (Fig. 4.1) between truth values and the product of statements and universes, but also, given certain statements and a truth criterion, as a process in the stipulated *processor* which constructs universes as imagined possibilities. Under these circumstances, the interpretation function I/F is itself a constructive process and it will be distinguished as such by writing (with processor given).

$\mathcal{L}_i \triangleq \langle S_i, \underline{\text{Inter}}_i \rangle$; *where* $\underline{\text{Inter}}$ i is a compiler that produces a specific universe, $\underline{\text{Univ}}$ i, as an interpretation in the processor.

Logic $\triangleq \langle \mathcal{L}_i$, Calculus$\rangle$; or, degenerately, $\langle \mathcal{L}_i, \underline{\text{Prog}}$ i$\rangle$ where $\underline{\text{Inter}}$ i produces a compilation of $\underline{\text{Prog}}$ i (and an interpretation of its input and output domain) in the processor.

We shall identify the processor with an L-Processor, the most general kind of M-Individual considered in the first monograph. At least, an L-Processor is an indefinitely sized ("inexhaustible") collection of a priori independent and asynchronous, programmable machines; of course, these machines are brought into local dependency and synchronicity when a program is executed.

The psychological interpretation of this construction is obvious in experience. If the L-Processor is identified as a human brain, then the compilation and subsequent execution of any program gives rise to an imagined world in which the input and output variable of the program range over sets of imagined objects. These may be abstract objects (for example, the set of real numbers) or they may be concrete objects (as in the case of a visual image, or the apparitions of any other sense modality) or they may have an undetermined status in this respect (the "imageless thought" of the Würzburg School, or simply unclassifiable impressions). In any case, these universes of interpretation (the input and output sets of the compilation of a program) are dubbed "imaginary", because they are constructed in a processor which the participant has described, for tenable but all the same arbitrary reasons, as *his own*. Apart from this, and to a lesser extent, the peculiarities of compilation in an L-Processor, the objects are no more "imaginary" and no more nor less "real" than the objects of everyday sensation and perception. *

* If an L-Processor is equated with a human brain, then this proposal is no more outrageous than Muller's 19th century doctrine of "specific nerve energy"; the notion that modalities of sensation, and ultimately of perception, are determined by patterns (of "specific nerve energies" to sustain the archaism), rather than being direct consequences of physically distinctive stimulation. Conversation theory takes "L-Processor" more generally (though a human brain is an L-Processor, so are many other systems).

Judging by everyday experience, internal compilations exist for different sense modalities and compilations that are not identifiable with any known sensory organ. Such introspections are well supported experimentally; for example, in the work of Wallach and Averbach (1955) or Posner (1966) on the existence of distinct visual and verbal memory traces, the informational value of which is stressed in Atkinson and Shiffrin (1965) "copy trace" scheme. As a matter of fact, there is no serious dispute about the existence of dedicated sensory buffer stores (which is of little immediate concern) or of distinct internal compilations of whatever process represents a memory in the theory at issue. It is also empirically obvious that sensory traces are translated from one modality to another at the least provocation either in short-term storage or long-term storage (Atkinson and Shiffrin (1967), so that only under special circumstances will different compilations of the same process (or a process engendered by the same stimuli) remain unrelated. But these special circumstances can be engineered (as witnessed by the results referred to above). The resulting interactions and occasional independences are clearly compatible with the present theory under the caveat that we refer to "what is remembered" as a concept (compiled procedure) and reserve the name

A fortiori, L-Processors are able to accommodate several languages, $\mathcal{L}_i$, $\mathcal{L}_j$, simultaneously, or as a special but important case, several degenerate logical systems (with "Calculus" set equal to "Prog"), for example, the systems:

$$\langle \mathcal{L}_i, \underline{\text{Prog}}\ \alpha \rangle = \langle\langle S_a, \underline{\text{Inter}}\ x\rangle, \underline{\text{Prog}}\ \alpha \rangle$$

$$\langle \mathcal{L}_j, \underline{\text{Prog}}\ \beta \rangle = \langle\langle S_b, \underline{\text{Inter}}\ y\rangle, \underline{\text{Prog}}\ \beta \rangle$$

where S_a contains goal descriptions proper to Prog α (relations satisfiable if Prog α is compiled and executed), S_b contains goal descriptions for Prog β. Inter x, Inter y are processes that realise models, generally distinct models, that are compilations in the L-Processor of Prog i, Prog j that *do* satisfy the goals described and, in this sense, are *true* valued.

"Truth" in this internal organisation need only refer to the possibilities of compiling and successfully executing a certain class of programs (all with an associated goal description) in an L-Processor, the existence of which is surely affirmed. These possibilities depend indirectly upon the program classes, Prog, already compiled and under potential or current execution. Hence, "truth" is tantamount to a statement that a system of inferences, hypotheses or beliefs is *coherent*, that it "sticks together" and (first monograph) is "conflict free". Contradiction is not excluded, provided it is conditional and thus hypothetical; for example, the system may contain programs that are modelled and interpreted in *distinct* parts of the universe which compute statements that *would be* contradictory, *if* the distinction were obliterated.

There is nothing in the outward looking perspective, sketched in the last paragraph, to preclude an inward looking perspective. From this latter point of view, certain universes of interpretation *exist*, usually outside the boundaries of an L-Processor, each with its own structure; for example, a molecular view of chemistry, a wave mechanical view of chemistry, the mechanics of a quite different part of the real world. If so, it is possible to reinstate the interpretation function as a mapping (I/F) between existing universes, truth sets and program statements. This is an external observer's image of things. Or, as a more pertinent alternative, interpretation

memory for "a procedure that compiles and reconstructs this concept". In other formulations (chiefly directed towards laboratory sized tasks) our "memory" is more often a "retrieval search".

may still be regarded constructively except that it leads to an identification with some pre-existent reality whereby, for instance, an imagined and "coherent" model is "tested" empirically to establish "correspondence truth" or veridicality.

2. FUZZY PREDICATES

Just as an ordinary predicate, or adjective, names a set of entities having the named property, so a Fuzzy Predicate names a Fuzzy Set. A Fuzzy Set is a function from a universe (its own particular universe) to a truth set. Though several possibilities exist, our immediate concern is with Fuzzy J Sets (Goguen 1968) for which the truth set is $\{0, 1, *\}$ or, verbally, "The interval $\{0, 1$: meaningless$\}$. Some Fuzzy Sets F, G, ... are shown in Fig. 4.3, named by Fuzzy Predicates. It is crucial to notice that the Fuzzy Set itself *is* the *function*. However, someone in a position to select an element x in the domain of F may refer to the value picked out in F's range (the truth set) as x's "grade of membership" in F; written, in Zadeh's (1973) notation, as μ_F/x. Similarly, x (of Fig. 4.2) has a grade of membership μ_G/x in G and the pair $\langle x, y\rangle$ has a

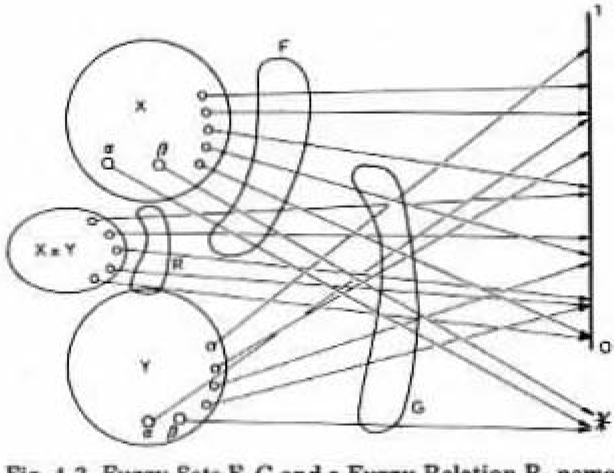

Fig. 4.3. Fuzzy Sets F, G and a Fuzzy Relation R, named by Fuzzy Predicates considered as functions from universes, X, Y and their Cartesian product X × Y onto a many valued truth set.

grade membership $\mu_R/\langle x, y \rangle$ in the Fuzzy Relation, R. In general. just as 1-ary Fuzzy Predicates name Fuzzy Sets or properties, so also, n-ary $(n > 1)$ Fuzzy Predicates name relations; as usual, a property is a unary relation. Quite possibly, the elements in the domain of a Fuzzy Set are Fuzzy Sets; so hierarchical organisations are perfectly permissible.

The algebra of Fuzzy Predicates differs somewhat from the algebra of Non Fuzzy Predicates (see, for example, Goguen 1968). Union and intersection are defined, so are *various* forms of complementation. But the behaviour of subsets of Fuzzy Sets is atypical and interesting.

Goguen, explicitly in the 1968 paper, has proposed Fuzzy Sets as the semantical or interpretative images of inexact concepts; that is, real concepts as entertained by minds beset by ambiguity and vagueness to a greater or lesser extent. This point of view is consonant with the position taken in this book and in the first monograph, but it is not identical with it. We maintained previously that a concept is a procedure under execution in an L-Processor which does in fact compute some property or relation named by an L-Predicate. In the generalisation, a procedure is identified with a Fuzzy Program (to be specified below; but a term which encompasses the various programs and non deterministic programs of the first monograph). Surely, the Fuzzy Program (alias procedure) will, if it undergoes execution, produce (stabilise, compute the values of) a Fuzzy Relation or property. This relation or property is given, in extenso, by a Fuzzy Set named by a Fuzzy Predicate. Our (entirely compatible) usage remains: that the concept *is* a procedure undergoing actual or potential execution. The *Fuzzy Predicate* is identified (in proper context) with a *topic* designating a (generally fuzzy) *topic relation.*

Goguen's major insight (which is used in Section 11) is that the universe of interpretation for a *natural* language consists in a set of Fuzzy Sets and that *natural* languages are distinguished from other languages primarily because this is so. The proposal is compatible with conversation theory. Natural language interpretations, especially the analogy relations that are the interpretations of natural language metaphors, serve as a peculiarly flexible modelling facility. The degree of freedom so obtainable may, in principle, be approximated in a physical modelling facility, made in the metal external to the user, and would be a processor able to accept and execute Fuzzy Programs. Any L-Processor is such a thing, paradigmatised by a brain which, we argue, is the (internal to the user) modelling facility for thought.

3. FUZZY PROGRAMS

Just as an ordinary program may be represented by a series of "instructions" which reduce to assignment statements and conditional imperative statements, so a Fuzzy Program (Zadeh 1973) may be represented as a series of Fuzzy or deterministic "instructions" * which reduce to assignments and Fuzzy Conditional Imperatives. A Fuzzy Conditional Imperative specifies a Fuzzy Relation, and that the execution of such a step (for example, using Zadeh's 1973 rule of compositional inference) usually results in a range of values or elements.

Fuzzy Programs have been characterised as algorithms by Santos (1970) and by Zadeh. But they yield Fuzzy or "approximate" results within a certain "tolerance" (see, for example, Cin Dal 1974). Broadly, a Fuzzy Program is a "heuristic". This amounts to slightly more than a mating of nomenclature; something is added to the idea usually conveyed by "heuristic" (even used carefully, as in the context of problem manipulation, by Polya (1954) and others). In fact, the multiplicity of values (or elements of sets pointed out by values of a variable) which generally results from an execution step may either be *perpetuated* or *resolved.* For execution on a serial machine, resolution is almost mandatory. Of the several values generated by execution, one is selected as the representative value to be carried forward into subsequent stages in the computation. Any defensible resolution rule can be employed for this purpose; for example, to select the maximum value or the numerically mean value as the representative. On the other hand, there is nothing in the formulation of a Fuzzy Program to suggest resolution, and given an other than serial processor (notably, an L-Processor accommodating several a priori asynchronous operations), the program itself calls for operations that may either be parallel or, in the sense of the first monograph, concurrent and only local-

* Generally, *Go To*, and *Start* and *Stop* are deterministic instructions, but these structures may be relaxed.

ly synchronised. In the sequel, perpetuation (no resolution) and parallel or concurrent execution are taken for granted.

Mechanically speaking, nothing remarkable is involved but the resulting computation (a heuristic operation) is far richer than the serially resolved process which merely simulates it. The heuristic or Fuzzy Program is, if permitted this fuller meaning, a class of programs for achieving the same Fuzzy Result (for computing values of a Fuzzy Predicate), together with those communicative or locally synchronising interactions required for the execution of these programs.

4. SPECIALISATIONS AND NOTATION

Henceforward, Prog stands for Fuzzy Program.

The term S Prog is reserved for a serial representation of Prog which computes the same relation as Prog but may be compiled in a serial processor.

Similarly Inter (given an L-Processor) stands for a (Fuzzy) Program that compiles a Prog in the processor and assigns the values needed if it is to be executed.

Further, as a special case, Inter is the generation of the mapping in Fig. 4.2 (the constructive realisation of an interpretation, as in Fig. 4.1). *

* The price is that we are committed to a view of the world of conceivable realisations; namely, an L-Processor containing any required number of asynchronous programmable machines (the loci of control of the first monograph) in which an indefinite number of independent dynamic systems may be specified and brought into local synchronicity and/or dependency by instruction, and in which the least element is a system. Realisations of a more restrictive kind are characterised, as they are needed, by the expedients of the first monograph; for example, by *stipulating* that a modelling facility is a one clocked processor, or a collection of one clocked processors permitting, as the case may be, serial or parallel execution of programs. Precisely the same commitment is fairly characteristic of general system theory though it is differently voiced. For example, Beer (1966) refers to the richness of fabric (nature, the unrestricted case of an L-Processor) insofar as fabric accommodates a diversity of process types that are obscured in a tractable abstract representation. Ackoff (1973) makes the statement differently. Elements or "atoms" are systems; their interaction is implicit unless specifically "excluded"; inferences are multiple causal, rather than causal, Singer's producer product relation and its refinements.

In order to deal with analogy relations, it is convenient to develop the very terse definitions employed in the previous monograph (L^0 Procedures, $\underline{\text{Proc}}^0$; L^1 Procedures, $\underline{\text{Proc}}^1$; and so on). The notation introduced in the last section is used for this purpose. Since, in the last monograph, all procedures were qualified as undergoing execution in an L-Processor, no disparities exist. The main objects of the exercise are (a) to distinguish between a class of programs that compute the same *abstract* relation and a class of interpretations for whatever is computed. Due to the difference in interpretation, the results may be called different *topic relations*, even if they have an abstract relation in common; (b) to establish a correspondence between procedural representations and images depicting states of an L-Processor or an external processor (the modelling facility of the previous monograph).

Consider the notion of an L-Procedure (undergoing execution in *any* L-Processor, but some processor always at hand). Henceforward,

$$\text{Procedure} \triangleq \underline{\text{Proc}} \triangleq \langle \underline{\text{Prog}}, \underline{\text{Inter}} \rangle.$$

Thus, observing the artificial stratification of L, the L^0 procedure is

$$\underline{\text{Proc}}^0 i \triangleq \langle \underline{\text{Prog}}^0 a, \underline{\text{Inter}}^0 x \rangle$$

which computes, stabilises, or brings about a topic i.

Notice (as an important feature of the generalisation) that change in either right hand term may determine a fresh $\underline{\text{Proc}}^0$; thus, if $i \neq j$

$$\underline{\text{Proc}}^0 j \triangleq \langle \underline{\text{Prog}}^0 a, \underline{\text{Inter}}^0 y \rangle$$

which computes a topic $j \neq$ topic i, though since the $\underline{\text{Prog}}$ is identical ($\underline{\text{Prog}}$ a), topic i and topic j may share the same abstract topic relation, i.e., the relations in question (as in the first monograph, R_i and R_j) are isomorphic. This possibility is precluded by the other variation for $i \neq k$, namely,

$$\underline{\text{Proc}}^0 k \triangleq \langle \underline{\text{Prog}}^0 b, \underline{\text{Inter}}^0 x \rangle.$$

The term "concept" has the same meaning as it has before (a procedure under actual, or potential, execution) but is more con-

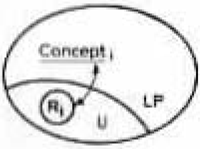

Fig. 4.4. Shorthand notation for action of concept i, compiled in an L-processor (LP) and bringing about relation, R_i in an outcome set interpreted in universe U.

veniently specified as follows:

Concept i ≜ Stable compilation of Proc^{0}i.

We use the shorthand notation of Fig. 4.4 to indicate that concept i on execution brings about relation R_i in its interpretation set $\mathcal{U}$.

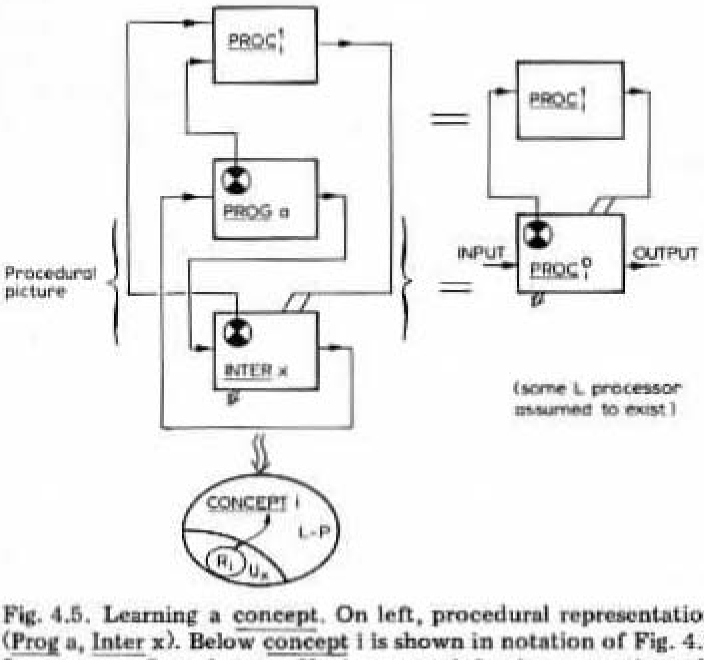

Fig. 4.5. Learning a concept. On left, procedural representation of Proc^{0}i = (Prog a, Inter x). Below concept i is shown in notation of Fig. 4.3. LP denotes L-processor. Set of states U_x is reserved for interpretation of outcome domain in which concept i brings about R_i if it is executed. On right: notation as used in previous monograph where *some* L-processor is assumed to exist.

An L^1 procedure is

$\underline{Proc}^1 \triangleq \langle \underline{Prog}^1, \underline{Inter}^1 \rangle$

where $\underline{Inter}^1$ need not be made explicit. The form of constraint is discussed in Chapter 5, Appendix 2.2; it consists in any workable structure existing in the L-Processor and possibly the repertoire of existing $\underline{Proc}^0$. As in the previous monograph, the limiting case in which $\underline{Proc}^1$i acts upon and reconstructs $\underline{Proc}^0$i is a memory, the simplest form being merely a recompilation of $\underline{Proc}^0$i. This is shown in Fig. 4.5, together with the stable compilation (in an L-Processor) to which it gives rise "℘" (that is the symbol "℘" links the procedural representation to the notation employed in Fig. 4.4.

More generally, $\underline{Proc}^1$ carry out *constructive* as well as reconstructive operations and they must do so in the case of an *understanding* (the primary condition detected in a strict L Conversa-

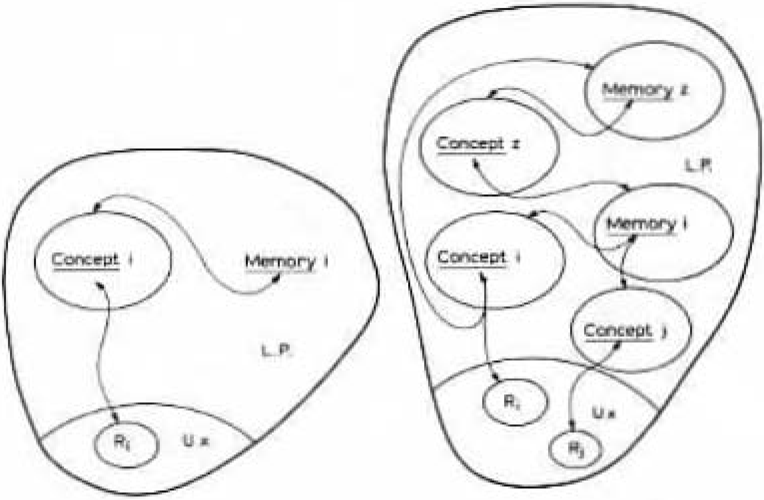

Fig. 4.6. On left, construction of memory comparable to construction for a concept; on right, bifurcating and looping constructions are permissible, insofar as stratification of $L = L^1, L^0$, (an artificial distinction, in any case) is preserved. Here, for example concept z is not allowed to bring about a relation R_z and neither concept i nor concept j can, as one of their immediate products, reconstruct memory i. This is a matter of edict, however, not of fact.

tion). The minimal form of an *understanding* is shown in Fig. 4.5.

Fig. 4.6 depicts the actual liberality of bifurcating and cyclic connections and reveals the frequently stressed fact that the stratification of L into L^1, L^0 is conventional, not factual. If imposed for convenience, the stratification disallows many cyclic organisations which could otherwise exist.

6. GENERALISATION OF ENTITIES IN THE CONVERSATIONAL DOMAIN

For notational clarity, the programs extensionally equivalent to (that do the same thing, by computing the same relation as) Prog^0i are represented as behaviour graphs (Chapters 1 and 2) denoted *BG*(i): meaning "descriptions of and precipitations for *model-building* behaviour". As noted before, all behaviour graphs are thus program graphs (for example, Chang and Lee 1973). The (many) programs exhibited in one behaviour program graph, *BG*(i), only represent $\underline{\text{Prog}}^0$i (Section 4) since *BG*(i) is non Fuzzy. These representations are designated S $\underline{\text{Prog}}$ i (Section 4).

A modelling facility to accommodate non verbal explanations as compilations and interpretations of programs in *BG*(i) is a (restricted) universe of interpretation, or a set of a priori independent universes of compilation and interpretation. In other words, it is one or more processors, together with interpretation sets for the input and output domains of programs that may be compiled and executed. If there are several a priori independent processors, we use the neoglism "Lumped Modelling Facility" to denote the aggregate.

In either case, the modelling facility executes compiled programs as models to yield results in an interpreted input-output set (more usefully, in the product of *input-set* × *output-set* = *outcome-set*). Any correct model for *topic* i is such that the execution yields one or more outcomes (all of which belong to $R_i \subset$ *outcome-set*. Since a "Lumped Modelling Facility" is described by L-Predicates, the models that may be constructed in the facility form a model space. The facility is more restricted than an L-Processor due to the "clocking" restrictions (first monograph) upon the constituent processors. The graphical notation of Fig. 4.7 is used to represent a model. In this pictu·e, which is intended to

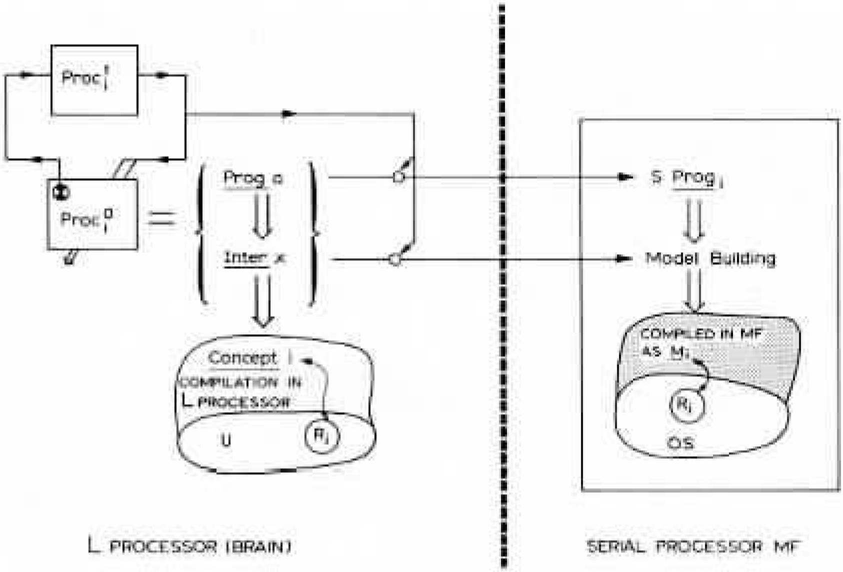

Fig. 4.7. Model building.

clarify the distinction between a model *external* to the L-Processor and a concept as a compilation *in* an L-Processor, the modelling facility *MF* is based on a serial or one clocked processor with an interpreted outcome set distinguished as *OS*. Some S Prog i, representative of Proc i, is compiled in the modelling facility as a model M_i. It is important to realise that whereas Prog a is a Fuzzy Program and Inter x is its Fuzzy Compilation (in an L-Processor, the student's brain in this case), as a concept the representative program S Prog i is serial and M_i is its compilation in the serial processor of *MF*. Model M_i is correct if the result of its execution is equivalent to the result of executing some program (S Prog i) in *BG*(i), and if it secures R_i in the interpreted outcome set *OS*. (Models for analogy relations requiring lumped modelling facilities with several independent processors are discussed in Section 9 and shown graphically in Fig. 4.8.)

One other feature of Fig. 4.7 is of importance. Just as the stable compilation of $Proc^0 i$ depends upon the operation of an L^1 process, $Proc^1 i$, so the selection of a representative S Prog i and its compilation in *MF* as M_i depends upon $Proc^1 i$.

The task structure *TS*(i) is an imperative form of the program graph *PG*(i). It represents all the demonstrations that can be given to a student using the modelling facility, and (as in the first monograph) is tantamount to a class of behavioural prescriptions for achieving the behavioural objective of satisfying R_i.

The entailment structure, *ES*, figures, as it did in the first monograph, in a dual capacity. On the one hand, it represents legal derivations of topics and thus *what may be known* (in the same way that *TS*(i) stipulates *what may be done* if the ith topic is selected). On the other hand, *ES* constitutes a modelling facility at the cognitive level in which the student exteriorises his actual derivation of a topic as a state marker distribution or learning strategy. In this capacity the entailment structure and the storage locations for marker placements (for the aim topic for the goal topic and so on), *ES* is an L^1 analogue for the L^0 modelling facility *MF* in which explanatory behaviour is exteriorised.

Finally, the *conversational domain* is the entire collection: entailment structure and the operator data base (first monograph) that back it up; for each *topic* i in the entailment structure either *BG*(i), or *TS*(i); the syntactic and semantic descriptions $D^1(R)$ of the derivations in *ES* and $D^0(R_i)$ of the compilations of each *BG*(i) in *MF*.

7. DIFFERENT TRUTH CRITERIA AND TRUTH VALUES

The following types of "truth" are generally recognised; correspondence truth, consensual truth and coherence truth.

Of these, correspondence truth is concerned with the result of testing that something has a mooted property, or that a given relation holds and is qualified by "in such and such (or *all*) worlds or universes of interpretation". If this qualifier is rescinded by supposing that any person or entity able to make a test is looking at the *same* world, then empirical evidence is obtained and a hypothesis based upon this evidence may be conditionally verified. (The contingency is present because things change, because the assumption of similarity is doubtful, because the relevance of data is never completely determinable, and because tests are fallible.)

Consensual truth is a form of gross accord between observers. In its naive form, consensus (over the admissibility of evidence, for example) is the outcome of a voting match between the observers.

But the refined versions of consensus admit for discussion in the course of reaching agreement, and in this case, consensual criteria are really being treated as the coherence criteria of the next paragraph.

The coherence truth of a proposition, p, is a question of the extent to which p forms part of a system of cogent inference with respect to some other corpus of propositions; for one example, those entailed by a prevailing thesis or a body of convictions, beliefs or (even) dispositions; for a further example, those propositions apposite to different possible worlds. Advocates of coherence truth include Bradley (1914); many of the notions are presaged in the writings of Leibnitz (especially in the sense of the "further example") and can be traced back as far as the ancient philosophers. The field is reviewed and an up to date coherence theory is developed by Rescher (1973). This recent theory is of peculiar interest insofar as one goal is to extract the maximum possible coherent content from a set Θ of generally inconsistent propositions, $\{p, q, ...\}$.†

Let Θ^* be Θ devoid of p. Now p is maximally coherent with Θ^* (thus, is a "strong" member of Θ), if p is a deductive consequence of the propositions in Θ^* (so that the negation of p is incompatible with Θ^*); p is coherent (to some extent) with Θ^*, if p is not incompatible with the deductive consequences of Θ^* and is thus a possible member of Θ. Now, given a set, Θ, it is possible to specify a family of non empty maximally consistent subsets of propositions (*mcs*) of Θ, such that any *mcs* is consistent, and such that the addition of any q in Θ to an *mcs* devoid of q renders that subset of propositions inconsistent. The coherently true content of the original collection might be specified as "that which is a deductive consequence of all the maximally consistent subsets" (Rescher's "I consequence"), or "that which is a deductive consequence of any of them" (Rescher's "W consequence"). In fact, Rescher recommends the use of intermediary criteria. A preference (an alethic or truth oriented preference) is employed to determine a set of eligible maximally consistent subsets of Θ, and the coherent content is whatever is a deductive consequence

† Our set "Θ" is Rescher's set S; as usual, the limitations of the alphabet make it impracticable to maintain a concordant notation, and we have used "S" for other purposes.

(Rescher's "P consequence") of any subset in this preferred set.

The theory is primarily concerned with working out the truth about a phenomenon based upon a set of observations and, perhaps, some existing observations. Hence, the propositions p, q, ... are data; they are candidates to be accepted or rejected according to whether they (and prior propositions) form a system, with the caveat that as much content as possible be extracted from the data. In order to count as a datum, however, the propositions (results of observation, for example) must have an extra logical claim to datahood and must also be sufficiently comprehensive to cover all possibilities relevant to the phenomenon under scrutiny. Similarly the preference criterion (in the original, an alethic preference unrelated to desire and attuned to objectivity) is also extra logical, and in the province of epistemology.

8. AGREEMENT AND COMMON MEANING

Our concern in this book is certainly not "logical" in the technical sense. It is psychological and epistemological. Consequently, our motives in mustering notions of coherence are distinct from Rescher's, and it is prudent to stress the differences at the outset. Except indirectly, the argument has little bearing upon rational assessment or even upon "necessary" or absolute truth. Nevertheless, the truth conditions of correspondence, coherence and consensus (as a form of coherence) hinge upon various kinds of agreement which implicate (at least) provisional and idiosyncratic truth.

Correspondence truth values (albeit local to a universe of interpretation) appear in adjudicating the "correctness" of a model; of whether or not a relation, R_i, is satisfied when the compiled model is executed, and whether or not the syntactic component (S Prog) of a model matches some other program or a class of programs, such as *BG*(i). In general, the logic is "non classical" both in the sense that it is a logic of action and in the sense that its truth sets are many valued (the valuations are of Fuzzy Predicates).

An external interpretation of $\underline{\mathrm{Proc}}^0 i$ is the explanatory model preferred by a participant who is learning *topic* i. The correctness of this model (Section 6) depends upon whether or not its execution satisfies R_i. Correctness is thus, amongst other things, an index of correspondence truth, local to the universe of interpretation furnished by the modelling facility. Indirectly, correctness also implies that the representative program can be compiled and that its compilation as a model can be executed. Similar remarks hold good if the model itself is matched against the class of models *TS*(i) obtained by interpreting any of the programs in *BG*(i), all of which satisfy R_i. Moreover, both correctness and matching (against models in *TS*(i)) are special cases of a semantic or interpretative agreement; a participant agrees that the model (or the result of its execution) tallies with a canonical form.

The general case of semantic agreement involves two or more participants. That is, some other participant, often in a dominant and judiciary role (for example, a teacher), makes a demonstrative model in the same modelling facility. The result of executing the authority's demonstrative model is compared with the result of executing the submitted explanatory model, and the two participants agree that these results do, or do not, satisfy the *same* interpretation of a topic relation.

Such a semantic agreement is severely limited. It says nothing, of necessity, about general empirical "truth" or absolute rationality; nor does the related canon of correctness. For example, if the original thesis propounds a falsified theory, correctness means "correct with respect of some part of this false theory, or with respect of it all". Participants are obviously able to reach agreement upon irrational, or empirically refuted, propositions.

But, to do so is not pointless. Though the status of a semantic agreement is limited, it does mean more than a vague accord. The participants who agree have been able to interpret a relation (and a program which computes it) in some world, perhaps a very bizarre world, and they agree that these interpretations (of the relation) are the same, or are within tolerance. Moreover, in this universe, the compilations of their programs work to bring about the given result.

The companion notion of coherence is also essential to the idea of agreement. The main point is that coherence between the statements entertained by two or more participants implies a basically *syntactic* agreement, though depending upon the circumstances, more than syntactic agreement *may* be involved.

In the first monograph, we specified the mediator of cognition as a Psychological-Individual or P-Individual. Any P-Individual is the replication (or self-stabilisation) of a repertoire consisting in units ⟨Proc^1, Proc^0⟩ (Section 5). The construct is essentially dynamic; the procedures making up the P-Individual are undergoing execution in some L-Processor. However, we do not insist that a P-Individual is localised, geographically, in a particular brain. Nor do we exclude the possibility that several P-Individuals cohabit the same brain, provided it is an L-Processor and thus is able to execute L-Procedures. As a matter of fact, both kinds of distribution of cognition are commonplace and are necessary features of a strict conversation, in which *understandings* are observable.

Having insisted that a P-Individual is a dynamic system, it is plausible to characterise it, alternatively, as some consistent and self-replicating system of hypotheses or beliefs, and thus to liken it to the sociological construct of a role. In this specification, "hypotheses" and "beliefs" are regarded as active cognitive processes "entertaining hypotheses" or "subscribing to beliefs", so that this picture of a P-Individual is quite similar to Kelly's (1955) picture of "man as an experimenter" or even, at a different and broader grain of theorising, Lewin's (1936) view in this matter.

Consider the artificial and imaginary expedient of freezing the P-Individual into momentary stasis. Under this imaginary assault, the "hypotheses" and "beliefs" make an appearance as "L Propositions". Call the set of L Propositions Propset. Manifestly, "any p, q, ... in Propset is coherent with Propset", i.e., the set of propositions representing the hypotheses or beliefs of the P-Individual at a particular instant are (L) coherent. If this were not the case, the P-Individual would not be (as asserted) self-replicating (though the converse of this contention to the effect that "if Propset is coherent then the system is self-replicating" is, clearly, not valid).

Coherence of Propset, in this sense, may have no greater status than a personal and private "truth"; the P-Individual's set of "beliefs" are amongst the deductive consequences of {p, q, ...} in Propset. To be more discriminating, we invoke instruments analogous, on the one hand, to Rescher's alethic preference ordering (so that only the deductive consequences of the preferred *mcs* are followed up) and, on the other hand, to the criteria of datahood (that the {p, q, ...} are truth candidates, both relevant and worth having). Lacking such an augmentation, a P-Individual may be nothing but a dreamer or a solipsist or a system that regurgitates the ultimately tautologous verb and adjective chains of an internal dictionary. In the extremity, a coherent Propset is a syntactic construction, and the further assertion that this Propset characterises a P-Individual leads only to the semantic inference that an L-Processor exists and is able to execute it. Perhaps the creature can do nothing except to say "I" repeatedly, like the bleating "point" of Flatland, in Abbot's (1884) geometrical fantasy.

Suppose there are two or more P-Individuals, A, B, in conversation, and their Propsets are constructed and symbolised as Propset A and Propset B. If the propositions in Propset A and Propset B are mutually coherent (so that Propset A, B is coherent), then the mutual coherence is an index of syntactic agreement between A and B. By the same token, there may be a syntactic agreement between factors of one P-Individual (A, for example) in respect of a conversational domain. This agreement is a statement of consensus (between A and B, or the factors of A) in terms of coherence. Consensus, in the sequel, is identified with such an agreement. But the statement, as it stands in its syntactic form, is minimal. Much more can be said if the conversation is *strict* and based upon *understanding* (in the technical sense of explanation conjoined with derivation) or the construction of $\underline{\text{Proc}}^1$i, $\underline{\text{Proc}}^0$i, as in Section 5.

Let a P-Individual engage in a strict conversation anchored upon a fixed conversational domain, taking place over occasions 0, 1, ... n ... N. Upon each occasion some topic in the domain is understood.

As in the first monograph, let $\pi(n)$ stand for the repertoire of pairs $\langle \underline{\text{Proc}}^1\text{i}, \underline{\text{Proc}}^0\text{i} \rangle$ that are learned, reproduced, and stabilised at occasion n. Due to the construction of the conversational domain and the characterisation of any P-Individual, it is possible to order the repertoires π (n), as follows:

$$\pi(0) \subset ... \subset \pi(n) \subset \pi(n+1) \subset ... \subset \pi(N).$$

With each π (n) associate a Propset (n); it will contain propositions asserting, or hypothesising, topics in the domain and relations between these topics. Each assertive proposition is rooted upon an *understanding*, and explanation and a derivation of some topic, held to evidence $\langle \underline{\text{Proc}}^1\text{i}, \underline{\text{Proc}}^0\text{i} \rangle$. The hypotheses concern topics up to and including the aim topic current at occasion n. The act of pointing out a topic to learn (issuing a command or asking a question, as the case may be) introduces at least one further candidate, and in general, the coherence of Propset is reduced by adding this candidate with hypothetical status. On the other hand, the act of *understanding* restores coherence and may increase it. That is, *Coherence* (Propset (n + 1)) $\geqslant$ *Coherence* (Propset (n)) and, in general, *Coherence* (Propset (n + 1)) $>$ *Coherence* (Propset (n) (The Gestalt property, claimed for the conversational domain). *

Further, if the topic is correctly explained (as it must be for *understanding*), then the resulting proposition is credited with the weak correspondence truth, to which we previously alluded. If the topic is legally derived, then a similar credit is given to at least one proposition affirming a relation between topics. In brief, the coherent propositions are, by virtue of *understanding*, assigned a (weak) *semantic* truth value.

For a consensual externalisation, suppose two or more P-Individuals (A, B) or two or more factors of one P-Individual engaged in a strict conversation on a fixed domain. We recognise the following types of consensual agreement between one and the other.

(a) A syntactic agreement of degree depending upon the coherence of their Propset.

(b) A semantic agreement regarding interpretations or models at level L^1 as well as at level L^0).

(c) If the participants have *both* syntactic and semantic agreement in respect of one or more topics, then these topics have the same *meaning* to the participants.

9. COMMON MEANING AND ANALOGY RELATIONS

Consider the conversations proper to learning about a given conversational domain. Such conversations may be of several types (reviewed in the next section), but the simplest kind amounts to a student engaging in "conversation with himself"; that is, a student represented as a pair of internal participants, one teacherlike and one learnerlike, who is "learning on his own".

* In practice, it is possible to determine the style of learning by examining the magnitudes of coherence values. For example, someone who recalls topics in a conversational domain by deriving them in many ways from other topics has a greater coherence, associated with his Propset, than someone who learns and uses just one derivation.

Now ask, "Is there an inscription in a conversational domain which may be learned of a *common meaning* (Section 8(c) above) agreed between participants?" The reply to this question is affirmative, and the desired inscription is an Analogy Relation.

To see this clearly, distinguish between the syntactic and semantic components of a subject matter thesis represented in a conversational domain. The syntactic component is an expression of derivations of topic relations and the uninterpreted program graphs attached to each topic. The semantic component is made up of the modelling facility (the compilation/interpretation set) and the *description* of the entailment structure afforded by the descriptor values assigned in $D^1(R)$ (from which a description of the compilation/interpretation set is derived).

Specifically, an analogy relation, (Fig. 2.6) is distinguished from the derivations in a disjunctive or conjunctive substructure by the fact that a semantic component is essential to its *cyclicity* (first monograph) and is represented by the semantic predicate Dist which distinguished between universes that are related in the analogy by a *morphism;* in the limit by an *isomorphism.* At the risk of labouring this point, notice that conjunctive and disjunctive substructures are also cyclic, Fig. 2.3, but the cyclicity of the analogy alone depends upon Dist. In terms of the first monograph, this fact demarcates *isomorphism* (where there is one to one register between topic relations, but no identity) and the other relational operations able to preserve specificity all of which secure relational identity. More generally, the similarity part of an analogy relation is syntactic and the difference part is semantic, as minimally indicated by the predicate Dist.

This general statement is in complete accord with Hesse's (1963) elegant-analysis of the analogy relations of science. The similarity is expressed by a morphism (and ideally, an isomorphism) between rules or abstract systems or scientific laws; the difference is expressed by a possibly incomplete list of properties characterising universes of interpretation (for example, "optics" and "sound").

Hesse's argument is peculiarly germane to the current theory, since it stresses that material analogies (those which can be modelled in a modelling facility) are based upon similarities of a causal or functional sort; as a result, upon rules that can be expressed by means of programs executed in a serial processor (the compilations of which are finite state machines). Our terminology "syntactic" covers this case quite adequately but will, later on, allow access to less well structured analogy relations. The analogy between similar rules (programs) that are compiled and executed in different universes of interpretation may be expressed as an isomorphism between some of the properties of each of the universes (X, Y); namely, those properties which enter into a specification of the *outcome set* (Section 6). The list of properties pertinent to each universe of interpretation, for example, the lists:

Optics	*Sound*
Colour	Pitch
Intensity	Loudness
⋮	⋮

are isomorphically related, if the given rule relates those in "optics" and the same rule relates those in "sound" (for instance, a simple wave propogation equation involving these terms when it is interpreted). But the list may either be complete: each of the indefinite number of properties that might be cited can be given a "positive" (i.e., member of the list) or a "negative" value. Or it may be incomplete: some properties have undetermined relevance at the moment the analogy relation is stated; for example, the "medium" in which sound waves travel may or may nor correspond to a "medium" (the aether, historically) in which light waves travel.

If it happens that the syntactic rule corresponds to a Program Prog a, then (for colour and intensity relevant or positive optical properties, and pitch and loudness relevant sonic properties) the isomorphism in the analogy relation is:

Colour ⇔ Pitch
Intensity ⇔ Loudness

Given Prog a

Similarly, if x and y are outcome sets characterised by these properties, this is in accord with the formulation of Section 5, namely, for an analogy relation R_k between R_i and R_j

$\underline{\text{Proc}}\ i = \langle \underline{\text{Prog}}\ a, \underline{\text{Inter}}\ x \rangle \Leftrightarrow \langle \underline{\text{Prog}}\ a, \underline{\text{Inter}}\ y \rangle = \underline{\text{Proc}}\ j$

where Proc i, Proc j satisfy R_i, R_j.

The only significant difference is that Proc k is seen as a procedure which computes the value of a distinguishing predicate Dist x, y which determines in what respect the universes of interpretation differ. But it is surely the case that any student having Proc k in his repertoire is in a position to test any property that comes to mind, or is observed, as being relevant or irrelevant to the analogy relation.

All the analogy relations considered in the first monograph can be expressed in these terms; notably, those holding between the "real" and the "abstract" universes of interpretation in "probability theory". Several other specific examples, culled from our work, are discussed in connection with conversational domains.

The immediate point of emphasis is that R_i and R_j, whatever they represent, are presented to a student as distinct; they are the relations of different topics in the conversational domain. Insofar as the student regards them as distinct and consequently views them from a different perspective, he is, at any rate in a momentary way, represented as two distinct entities. Consensually, these may oscillate so that R_i is learned (or thought of, or attended to) at one moment and R_j at the next. Insofar as R_i and R_j both occupy his attention (or are learned about, still as distinct topics, simultaneously), the entities are two participants. Finally, if it happens that the student assimilates an analogy relation between *topic* i and *topic* j, then the (albeit transient) participants reach a *common meaning* which, if it tallies with the analogy relation inscribed in the conversational domain, is the specified analogy R_k between R_i and R_j. An understanding (explanation and derivation) of R_k is evidence to this effect.

Evidently, R_k is also a *common meaning* to comparable "participants" inside the subject matter expert who inscribed it as a topic in the conversational domain, and we return to the question of *constructing* analogy relations after detailing the act of *understanding* an analogy relation. Some groundwork is needed in order to update the exposition in the previous monograph so that the account adumbrates certain revised transactions in the operating systems (either CASTE or INTUITION as described in Appendix A) and the, now explicit, distinction between the Prog and Inter components of a procedure.

10. UNDERSTANDING OF AN ANALOGY RELATION: EXPLANATION

Consider first, the non verbal explanation or modelling of an analogy relation which is one necessary component of its understanding. Suppose a Lumped Modelling Facility consists of two a priori independent serial processors connected to one or more outcome sets.)Generally, there are very many possible outcome sets, but they are invariably partitioned by the semantic descriptors into parts proper to each processor.) For example, in STATLAB (of the first monograph and the Appendix) there are two processors, one attached to the "real" universe of interpretation, and one attached to the "abstract" universe of interpretation. Their outcome sets consist in "simple events" or "composite events" or "measures" (in one case), and "simple results" or "composite results" or "frequency ratios" (in the other). Call one processor and its outcome sets X, and the other processor and its outcome sets Y.

In order to explain the analogy relation R_k between topic i and topic j, the student must ultimately do as follows:

(a) Build a correct model, M_i, of concept i which on execution in the modelling facility realises R_i in X.

(b) Build a correct model, M_j of concept j which on execution in the modelling facility realises R_j in Y.

(c) Couple X and Y so that the isomorphism between R_i and R_j is realised, and execute the models *simultaneously* to satisfy R_i and R_j. This coupling, a model M_k, satisfies the analogy relation R_k.

To summarise, clause (a) is evidence that $Proc^0 i$ exists; clause (b) that $Proc^0 j$ exists; and clause (c) that $Proc^0{}_k$ exists. If backed up by evidence for derivations, this provides evidence for a stable concept, and thus for an *understanding*.

Now clause (c) has greater content than it seems to have. In general, the simultaneous and successful execution of M_i and M_j implies more than a coupling between their outcome sets. It is possible if, and only if, the pair of a priori independent processors in the X and Y parts of the Lumped Modelling Facility are partially synchronised, either by interruption signals or by other methods. The crux of this requirement is not well illustrated by the examples in the first monograph (where, for the most part, the reality of concurrent execution of M_i and M_j was not encouraged and analogy relations were modelled formally). In fact, the subject matter employed permitted this glossing, though we noted persistent student demands to "compare models". As promised, the defects have now been remedied and concurrent execution is rendered mandatory. Its impact is easily imagined in the context of the "optics" and "sound" example, and the matter is pursued in Chapter 7 using the subject matter of energy conversion and simple thermodynamics.

With these comments in mind, Fig. 4.8 shows the structure built up in a modelling facility. M_i is the *compilation* in X of S Prog i (representing Proc i) with its outcome set (OSX) distinguished and having R_i as a subset. M_j is the compilation in Y of S Prog i (representing Proc j) with its outcome set (OSY) distinguished and having R_j as a subset. For correctness, M_i and M_j are matched against *BG*(i) and *BG*(j) and the satisfaction of R_i, R_j, is determined. This operation is not shown. The coupling and partial synchronisation appear in the picture as the connections between OSX and OSY, together with those between X and Y. If, and only if, the models can be jointly executed to satisfy R_i and R_j, the analogy relation R_k is correctly modelled.

The sense in which the entire model, M_k constitutes an index of

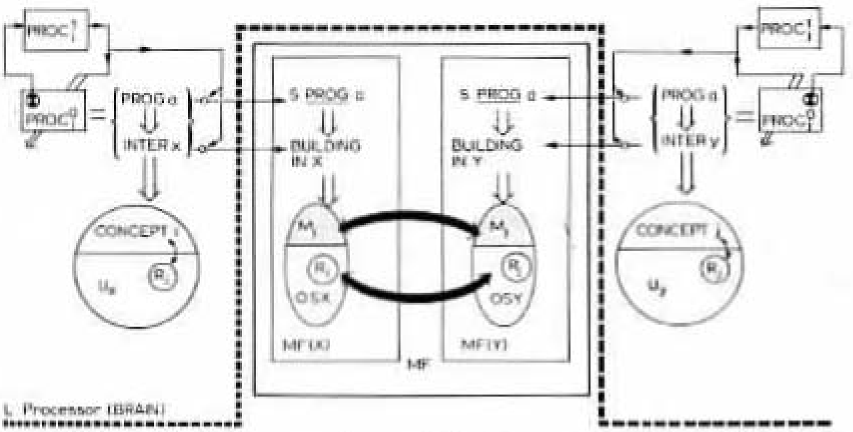

Fig. 4.8. The model for an analogy relation is a coupling (signified by ⟳ between models M_i and M_j) as a result of which the execution of M_i and M_j in *MF* is synchronised.

common meaning is shown in Fig. 4.8, which is no more than an outline sketch for the cognitive organisation we suppose to be responsible for the analogy relation model. As before, Proc i and Proc j are L^0 procedures in the cognitive repertoire. They consist in Prog and Inter components; namely, ⟨Prog a, Inter x⟩ and ⟨Prog a, Inter y⟩. Insofar as the student selects S Prog i (compiled as M_i in X of the external facility) and S Prog j (compiled as M_j in Y of the external facility), Inter x and Inter y are generating distinct compilations in the students brain (an L-Processor) for Prog a. These are sketched as Concepts i, j, conceived as internal representations on a par with the external models M_i in X and M_j in Y. Insofar as the student places x and y in register with the correct (in the sense of relevant) properties of X, Y, and to the extent that correctness is betokened by the successful (joint) execution of M_k he also has in his repertoire a further procedure Proc Dist = ⟨Prog Dist Inter x × y⟩ which is internally compiled in the product $U_x \times U_y$ and distinguishes between $U_x \times U_y$ appropriately. (And must do so, since under an operating system, Dist (x, y) is a semantic descriptor and is already marked as being understood.) That is, Proc Dist computes the distinguishing descriptor Dist (x, y).

The "internal" (or imaginary) participants said to reach "agreement" over a *common meaning* are centered upon Proc i and Proc j; they are held distinct by the action of Proc Dist; they have Prog a in common; their agreement amounts to a recognition of this communality, even though Proc Dist exists. The semantic (or correspondence) component of the agreement is the model for the analogy relation. Its syntactic (or coherence) component is the isomorphic register between Progs in Proc i, Proc j. We refer to the internal participants as "imaginary" because we are concerned with experiments or tutorials in a one aim at once facility, such as CASTE or INTUITION. Hence, although the foci of attention of the "participants" may be real enough to a student (and common experience suggests that they are), the transactions are not distinctly observable as exteriorised stretches of behaviour.

11. UNDERSTANDING OF AN ANALOGY RELATION: DERIVATION

Now, consider the other aspect of understanding an analogy relation: its derivation, which is exteriorised as a learning strategy traced out on an entailment structure. For a two term analogy * just four basic configurations are possible (though these give rise to innumerable variants). Assuming that the student's aim (his "focus of attention" or the "topic that he appreciates") is *at* or superordinate to the analogy relation R_k (Fig. 2.6), these configurations (Appendix A) are as follows:

(A) *Topic* i is understood, *topic* j is understood and the analogy relation is marked as goal which is a legal member of workset.

(B) *Topic* i is understood, the subordinates of *topic* j are understood and the analogy relation is marked as goal which is a legal member of workset.

(C) Vice versa, but *topic* j is understood instead of *topic* i.

(D) Neither *topic* i nor *topic* j is understood but the subordinates of at least one of them are understood. The analogy relation is marked as goal. Under the conditions discussed in the first monograph, this placement of markers does not admit to goal as a member of workset. However, in the revised operating systems that are currently in use, it does (and *may* do so because of the possibility of concurrent modelling).

Configuration A obtains if the student intends to learn the analogy relation as a relation between existing concepts for *topic* i and *topic* j. As a practical consequence, the student may (if he wishes) receive a demonstration of the isomorphism and of Dist, and he must model the analogy, as in Fig. 4.8, if this *topic* R_k is to be marked as understood. Notice, however, that M_i and M_j both exist.

Configuration B obtains if the student intends to learn the analogy relation in terms of *topic* i and to derive an explanation of *topic* j in terms of R_k. As a practical consequence, the student may (if he wishes) receive a demonstration of R_k as a path to *topic* j, and for understanding of R_k he must model R_k which involves constructing M_j (since M_i already exists).

Configuration C is the reverse situation in which *topic* i is accessed through R_k. The student may (if he wishes) receive a

* That is, an analogy relation between two topics. Similar comments apply to analogies involving many terms or other analogy relations of the type exhibited in Chapter 2, but the configurations are much more complicated and are difficult to represent graphically.

demonstration of R_k as a path to *topic* j. He must model R_k if R_k is to be marked as understood; this entails the construction of M_j, but M_j exists.

Finally, the configuration D leads to a conditional transaction. The student may receive demonstrations of *topic* i and *topic* j (if he wishes). But in explanatory modelling, he can essay the construction of a coupling between unspecified and not-yet-understood topics. However, a model such as this is accepted conditionally. The analogy relation is marked as understood unconditionally if, and only if, M_i and M_j are produced (to be united by the coupling), as a result of which *topic* i and *topic* j or both of them will be marked as understood. In the process R_i or R_j or both of them will be marked as goals at the same moment as R_k. Since this implies that workset has more than one member, the manoeuvre is necessarily part of a holistic learning strategy and is, in fact, adopted by holistic students.

One psychological interpretation (which we favour as by far the most plausible) is that conditions A, B, C involve learning an analogy relation when one (condition B or C) or both (condition A) of the terms of the analogy are known already. In condition D, on the other hand, the analogy relation appears first of all and the terms (*topic* i or *topic* j or both) are understood because the analogy is known. For example, using the subject matter "energy conversion" of Chapter 7, the student in condition A discovers a relation ("heat conservation" cycle) between "heat engines" and "refrigerator" both of which are known to begin with; in condition B or C, he knows about "heat engines" or about a "refrigerator" and derives "heat conservation cycle" because of that. Of course, we may not exclude a global looking and comprehensive approach in these cases, since any student could fail to exteriorise his mental gambits. But in condition D, either "heat engines" or "refrigerator" or both are understood as a result of knowing about "heat conservation cycle" and in this case the student *must* be adopting a global method.

All of the conditions for learning an analogy relation are consonant with Fig. 4.8 and with the notion that the analogy relation is just the inscription of a common meaning (recall that each term is modelled, though it is only marked as understood if it was marked as a goal). On the other hand, the order of events and the type of interaction between Procs differ radically according to the condition selected. In particular, conditions A, B, and C involve the existence of Proc i or Proc j or both before there is an L^1 operation (a Proc^1) that places these concepts in register; whereas in condition D, this operation is performed over Dist (x, y) before Proc i and Proc j are constructed. This we can only construe (mechanically speaking) as implying the existence of a hybrid procedure, neither Proc i nor Proc j, which is differentiated to yield Proc i and Proc j.

12. THE ACT OF CONSTRUCTING AN ANALOGY RELATION

So much for *learning* an isomorphic analogy, as it is inscribed in the conversational domain. When it comes to *constructing* an analogy (during course assembly, or under the control of an evolutionary heuristic), the participants we dismissed as "imaginary" may be very real. These participants could be members of a team of subject matter experts, or equally they could be distinct cognitive organisations that are parts of one subject matter expert.

Since the course assembly heuristic EXTEND, considered in the first monograph, is (like CASTE) restricted to one aim at one, these interesting segments of cognition cannot be exteriorised in the system. But other heuristics to be described in Chapter 7 (as part of an operating system called THOUGHTSTICKER) allow for many aim topics.

ADDENDUM

Two recent papers by F. Varela ("A calculus for self reference", *Int. J. Gen. Syst.* 2, p. 5, 1975 and "The extended calculus of indications interpreted as 3 valued logic", *Notre Dame Journal Formal Logic*, 1976) provide a respectable means for talking of reflective and self reproducing systems (a fortiori, P Individuals) *within* the language L. The basic idea is to reserve a truth value for the condition of recursive or vicious circularity and the possibility of doing so stems from the calculus of distinctions and indicators (Spenser Brown, G. 1969, *Laws of Form*, George Allen and Unwin, London), to which the first monograph owes so much. The difficulty that different kinds of circularity are inadequately distinguished is resolved in a further paper, "The Arithmetic of Closure", which will be part of the proceedings of the 3rd European Conference on Cybernetics and Systems Research, Vienna, 1976; with this augmentation, it becomes possible to speak, similarly, of interactions between several distinct P Individuals.

Almost simultaneously J. Goguen ("Objects", *Int. J. Gen. Syst.* Vol. 1, p. 237, and "Complexity of Hierarchically Organised Systems and the Structure of Musical experience," *Technical Report* in Department of Computer Science, UCLA, 1976) has rooted general system theory in "objects" that depend (in a sense) upon observation and has shown how systems are amalgamated by dependency/independence, or synchronicity/asynchronicity, to create further systems.

The relevance of this work is evident; it is compatible with our informal argument, though more elegant. These significant innovations are currently being incorporated (under the notion of categories of "objects" that are P Individuals) and tangibly implemented, by Robert Newton, at this laboratory.