Collections, Systems and Mathematical Metaphors

Ray Paton
Department of Computer Science
The University of Liverpool
Liverpool L69 3BX
UK

Abstract

The purpose of this paper is to consider a number of conceptual and linguistic issues associated with the nature and occurrence of collection concepts in the biological sciences. Collections are very common in biology and may be conflated with the idea of a system. The relations between collections (in the linguistic/psychological sense) and systemic metaphors in the biosciences will be examined. Two important systemic constructs that are relevant to both collections and systems are verbs and 'glue'. Examination of certain aspects of these ideas leads to a consideration of potentially valuable insights into the construction of descriptive biosystemic concepts using the source metaphor of category theory.

Keywords: biosystem, verb, glue, collection, systemic metaphor, category theory

1 Introduction

"If we try to squeeze science into a single viewpoint ... we are like Procustes chopping off the feet of his guests when they do not fit on the bed".

Freeman Dyson (1995)

The aim of this paper is to develop meaningful tools of thought for biology within integrating and integrative conceptual frameworks. Some ideas associated with category theory will be used to enhance the conceptualisation of collections and systems in the biosciences. Tools of thought for dealing with the complexities of biosystem organisation and adaptation are incomplete. To partially accommodate this incompleteness a pluralist approach to modelling and description is followed. This is the sentiment reflected in the quotation that opens this section of the paper. Although the paper blends mathematical and scientific ideas and sources, it must be noted that neither a Pythagorean nor Aristotelian stance is being emphasised. This is an important point to make as neither pattern (mathematics) nor matter (materials) are taken as singular fundamentals underlying the whole approach.

The underlying philosophical basis of this paper follows Harré (1986) and many others who expound a semantic/cognitive view of scientific theory. A theory is viewed as the evolving cognitive complex that enables us define the objects in the real world that we seek to model, represent, explain and understand. Theory provides the conceptual environment by which models can be constructed, predictions and explanations made and hypotheses generated. Many theories are more than ordered collections of statements. They also contain an iconic component. At the core of a scientific theory is a conception of a mechanism or structure at work.

As argued more fully elsewhere, the language that is used to articulate scientific models - in that it is used represent one thing in terms of another - is metaphorical (e.g., Paton, in press). Thus, a biological cell or its parts have been described in terms of machine, network, computer, factory, laboratory, society, ecosystem and text (Paton, 1993). Scientific metaphors are not ornamental properties of language that can be replaced by literal descriptions. They fulfil central roles in the development of a scientific theory including catechresis - supplying new terms to the theoretical vocabulary, and ontology - being involved in the formulation of hypothetical entities. There is also a didactic function in that they facilitate dialogue between a teacher and pupil. Within the scope of the didactic function there is an important *caveat mentor*: metaphors have limits and can just as easily misinform a subject as illuminate it.

Three kinds of integrative structure will now be examined viz.: collections, systemic metaphors and categories. Each of these will provide a number of metaphorical source ideas for clarifying our thinking about biological systems.

2 Concepts and Collections

The purpose of this section is to consider one group of nouns that allow the components of whole to be collected together. Collective nouns include terms like family, army and forest as well as more anonymous terms such as group, pile and stack. The former are related to specific component nouns whereas the latter are much less specific. Collective nouns are relatively rare when compared with class terms. Collections may be organised to produce hierarchies or networks and many verb types may be used. This contrasts with classes which produce *is_a* or *is_part_of* hierarchies and use few verb types. An important property for the purposes of the present discussion is that collections require interrelationships between parts (Markman *et al*, 1980).

It is relatively uncommon in English for single words to be used to incorporate relational structures and events (Markman and Hutchinson, 1984). The relational structures of events and themes are a common ways of organising information/knowledge but we do not often have single words for them. Investigations of the understanding of biological systems among high school showed that although they may find it easy to collect terms together, it is often difficult to assign an organising collecting concept (Paton, 1988).

Even simple examples of systemic relations raise problems about collections. For example, a consideration of whether blood is a part of the circulatory system highlights a number of semantic issues. Not only is there is a multiplicity of terms for the system such as vascular, blood, cardiovascular, haemodynamic, circulatory, etc, there are problems regarding the inclusion/exclusion of other components such as lymph vessels, bone marrow, thrombus, etc. In part these differences are related to representation (e.g., whether the system is viewed as an anatomical conduit or as a blood processing machine). Hunt (1978) noted that if an object is defined as a member of a collection and is related (by a verb) to another object then the second object is also a member of that collection. This relation does not always hold for systems, but the criteria for making distinctions are far from clear.

A number of collective nouns are used for monopspecific groupings of animals such as brace of grouse, brood of hens, cete of badgers, colony of ants, drove of cattle, gaggle of geese, plague of locusts and pride of lions. Collective nouns can be used in a figurative/comical sense. Consider the following example, a brace of professors was summoned by the dean but a pride of professors appeared to the students who brooded over their examination results before going to the bar in a gaggle. In this case the collective concepts are used to help organise a sequence of events.

Some collective nouns have a rich semantic structure. Examples involving human social encounters include: conference, congregation, congress, consortium, delegation, meeting and symposium. Figure 1 provides a summary of some internal relations regarding a meeting. The representation is far from complete and is not based on any particular representational formalism. Prepositions associated with "meeting" can be used to identify a number of internal relations. The richness of this internal organisation provides a cohesive binding on the conceptual whole.

It is now possible to examine some molecular biological collections that also have important internal organisation. Many examples could be considered, we compare operon, regulon and genome. An operon is a controllable unit of transcription consisting of a number of structural genes that are transcribed together and at least two distinct regions: operator and promoter. In terms of internal organisation it has been viewed as a one-dimensional array of interactions. A regulon consists of a group of genes or operons that are regulated together but may be located a good distance from each other in the genome. This time the internal organisation is a network of interactions.

A genome is another collecting concept that we now examine in more detail. Comparisons between older and more recent ideas about this structure reveal an evolving view that makes increasing use of computational ideas about parallel distributed systems, emergence and cognition (Paton, 1998). Shapiro (e.g., 1991) contrasted a number ways in which views or models of the genome have changed.
Firstly, there has been a shift from a constant to a fluid genome in which storage is dynamic rather than rigid. With regard to internal organisation, the genome is no longer viewed as a bag of isolated genes but rather as multigenic interacting networks. There

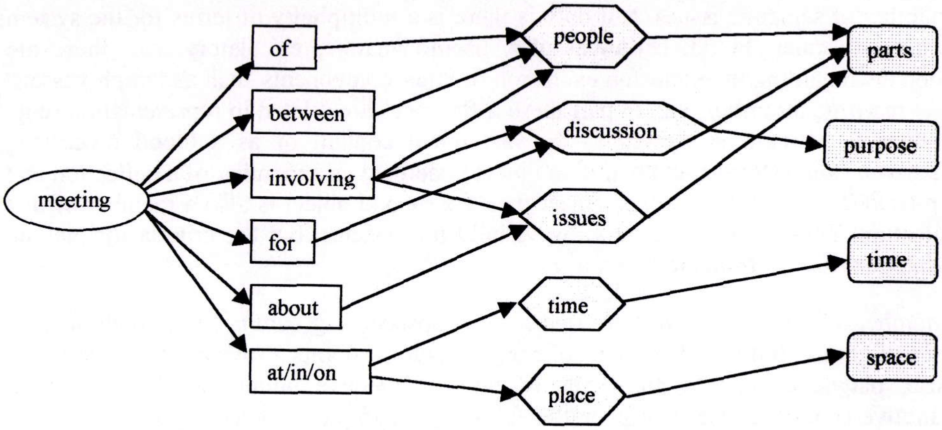


Figure 1 - Schema for Meeting

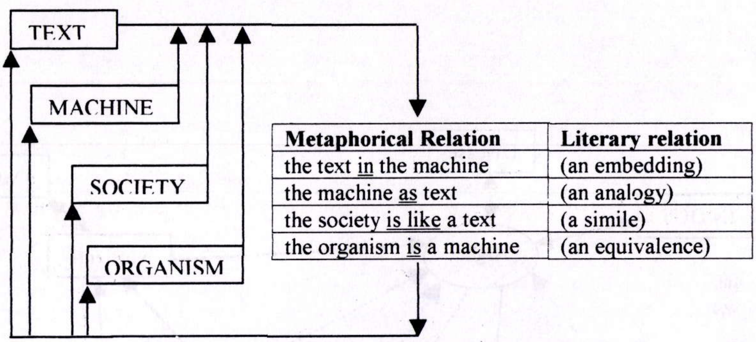
has also been a move from a mechanical-chemical view to hereditary organisation reflecting an information-rich or 'smart' thinking. Rather than information utilization taking place in an automatic/mechanical fashion there is an integrated, coordinated and complex information system. The components of collections may be a homogeneous or heterogeneous, ranging from small number to large number and at one or several organisational levels. Numerous relations may be described such as spatial, temporal, conceptual, etc. The internal organisation of a collection may be attributed to certain systemic properties.

3 Systemic Metaphors

A system is a collecting concept that has a number of distinguishing properties such as: interactions among parts, organisational form or structure resulting from interactions and whole-system functionality or the emergence of behaviours of the whole that are greater than the summing together of the parts. Systemic metaphors are general constructs which can be associated with general tools of thought for organising knowledge. Examples of systemic metaphors include: machine, text, organism and society. We briefly examine some of these issues in relation to the notion of an ecosystem.

Ecosystems' thinking has been subject to many displacements from various systemic sources. For example, a number of ecological machines can be described including chemical (i.e., matter), thermodynamic (i.e., energy) or cybernetic (i.e., information). Circuit thinking was used for example by H. T. Odum when applying equivalent circuits from electrical systems to deal with energy flows (for historical review see Golley, 1993). Associated terms in relation to machines and circuits include, balance, input-output, feedback, regulation, *etc.* A range of organismic concepts have been displaced to ecological ideas including open system, growth, development, sickness and health, self-maintenance, individuality, *etc.* The source idea of a society includes terms like agents, context, interaction, exchange, co-operation and competition could be displaced. A theatre metaphor might include stage, play, performance, actors, roles, setting, scenario, script, etc. Closely linked to some of these theatre ideas is the text metaphor which in terms of ecology relates to displacements regarding natural *history,* script, interpretation, meaning, context, as well as societies-as-texts. It is not only nouns that can be associated with particular systemic metaphors, verbs also follow certain types of usage pattern.

An example of the application of this general approach is the organisations of hereditary information and the development of a metaphor related to the idea that life is a dance (for source see Singer, 1959; for application Paton, 1998). Within the framework of this metaphor the script or score (what tends to be called the hereditary material), the cast (the metabolic agents and processes) and the stage (the cellular structure) co-exist and pre-exist the phenotypic life history which inherits them. The above discussion also indicates how recursive relations can operate between source ideas. Figure 2 provides a summary scheme for how multiple displacements can be achieved.



Metaphorical Relation	**Literary relation**
the text in the machine	(an embedding)
the machine as text	(an analogy)
the society is like a text	(a simile)
the organism is a machine	(an equivalence)

Figure 2 - Some Recursive Relations between Some Systemic Metaphors

It was noted in an earlier section that theories often have an important iconic or visual component at their centre. A simple case study will now be considered to illustrate the relation between visual representations and systemic metaphors. We focus on an ecological example that will lead to more general displacements in cell biology in the next section. Consider the diagram of a 'factor complex' shown in Figure 3. This type of representation shows factors that influence a particular variable within an ecosystem (in this case algae) and also how these factors influence each other. Multiple interacting components are depicted in this diagram which summarises an autecological point-of-view. The society metaphor can be a dominating source here covering issues like agency, context, interaction, exchange and competition. A shift from society to circuit can result in a diagram like Figure 4. Here the autecological focus has shifted (with a concomitant reduction in adjacent arcs to the algae node) and a network of flows with cycles and loops can be seen. The model depicts openness by using "IN/OUT" boxes.

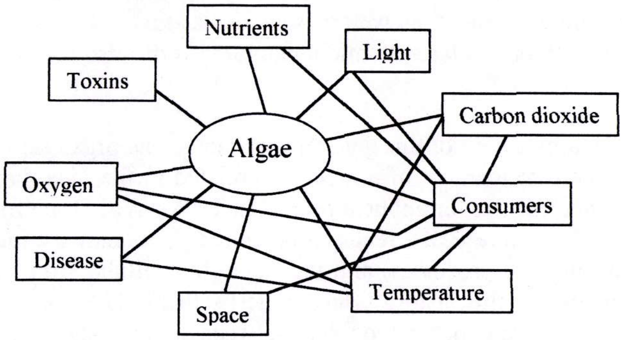


Figure 3 - A Factor Complex

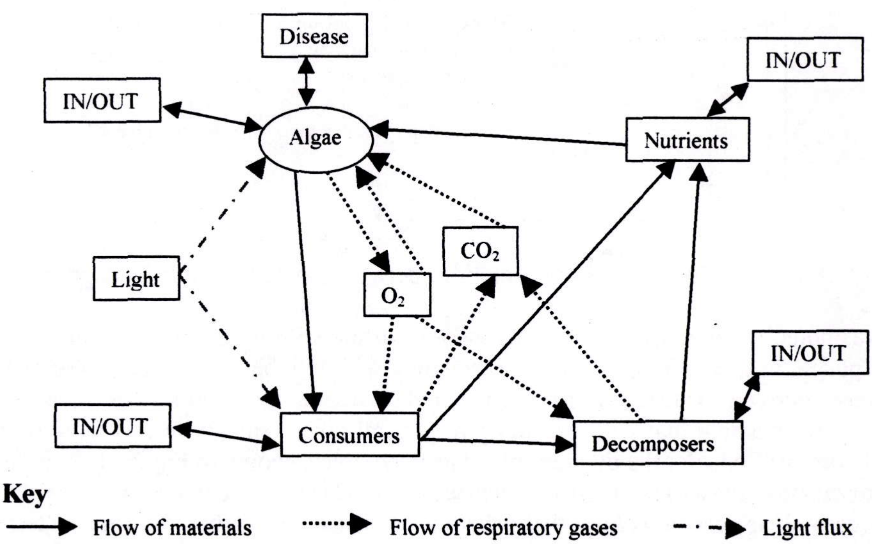


Figure 4 - Network of Flows

4 From Glue and Verbs to Categories

Many ideas in general systems thinking have been inspired by biology (e.g., Bertalanffy, 1973; Capra, 1996). In this section we explore some conceptual bridges between biological systems and mathematical categories. A problem for conceptual organisation is the potential multiplicity of systems that may be generated as well as the danger of making system an empty concept (Marchal, 1975). Each time a verbal relation joins component(s) of a system with another object, a new collection is formed. Sometimes the new collection is also a candidate system. One partial solution to the conceptual organisation problem would be to manage candidate systems in terms of their relationships with systemic metaphors. A complementary approach is to utilise systemic properties of categories. We start to approach the latter through the metaphorical ideas of 'verbs' and 'glues'.

The application of category theory to systems thinking in general and relational biology in particular has a long history (e.g., Rosen, 1991). Some recent biologically-orientated approaches have been developed by biologists (e.g., Chandler, 1998). The approach presented here has a different focus in that it seeks to develop integrative tools of thought for dealing with the conceptualisation of biological systems (at all levels of the organisational hierachy) through the application of certain systemic sources. The emphasis is on the displacement of concepts rather than the reduction of biology to mathematics. The connection between biosystems and category theory will be made using ideas about 'glue' which is used here in a metaphorical sense and covers such terms as adhesion, combination and cohesion. The starting point will be the following quotation from Ehresmann & Vanbremeersch (1987):

> *"... an object in a category has such a complex structure when it is composed of a family of more elementary objects, 'glued' together, the gluing depending on some specific links between the components"*
>
> Ehresmann & Vanbremeersch (1987)

This idea of 'glue' is non-trivial and although its common usage relates to adhesives, we use it here to cover a range of related verbs such as adhere, attach, bind, bond, cohere, combine, connect, fuse, join, liase, link, paste, tie and weld (Paton, 1997). The elaboration of 'glue' ideas will begin by considering a number of network representations of a system. This kind of representation tends to make the most use of circuit/machine thinking and we shall endeavour to keep this in mind as some verb/'glue' issues are discussed.

Consider the 6-node (vertex), 7-arc (edge) network at the bottom of Figure 5. This graph can be transformed in number of ways as shown by the various steps labelled with letters. Given that a transformation can be interpreted as the representation of one thing in terms of something else, and following the discussion of metaphor and model in section 1, we may say that the process of transforming the various graphs applies certain

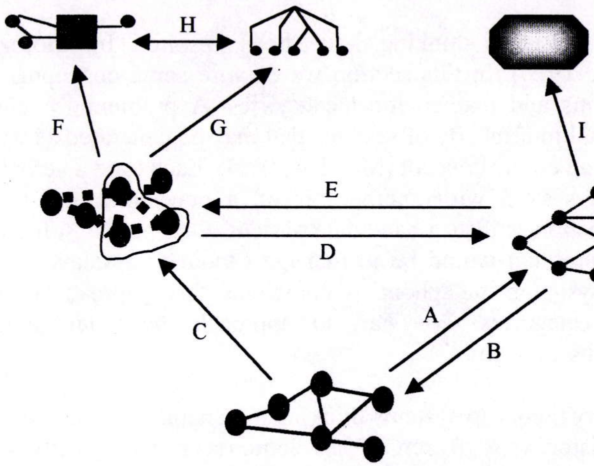


Figure 5 - Some Abstractions on a Network

metaphors. For example, step A represents the production of a new network in which the relations of the original network become nodes and the relations in the new network are implied as relations between relations in the original. Hence, process becomes object in this representation (a *line graph*). We can say that this type of transformation adds meaning to the network which can be related to the increase in arcs, and resonates with the generation of a factor complex to account for an autecology (as discussed previously). Step I relates to higher order extensions of the line graph (e.g., a line graph of a line graph). The biological significance of the higher order relations in these graphs may become problematic. Step C presents the process of abstracting a subgraph or clique of nodes in the network to produce (for example) a compartment. Step F depicts the abstraction produced in step C as a compartmental model and Step G as a hierarchical tree of objects. Given certain rules it is also possible to step between the compartmental model and the tree (Step H). The search for preserving/conserving 'glue' within these transformations can be examined from the context of the cell-as-text metaphor (see also Paton & Matsuno, 1998).

(a) Adhesion via a binary relation

(b) Adhesion via multiple relations

Figure 6 - Adhesions

We shall illustrate the important idea of adhesion with a couple of visual representations (Figure 6). The shaded node can be thought of as a verb(process) that takes a number of cases or as an object that adheres to a number of other objects. In the latter situation the arcs(edges) represent the 'glue' and as shown with Figure 5 there can be a switch between verbs as arcs and verbs as nodes. Both verb and noun emphasise the society metaphor. Elsewhere, the idea of enzyme-as-verb is discussed in connection with the role of verb as glue (e.g., Paton, 1997; Paton & Matsuno, 1998). The adhesive properties of enzymes can be seen by using the example of Figure 6(b) in which the central (shaded) node is the verb (enzyme) and the other nodes relate to the cases the verb takes. Such enzyme cases would include: substrate(s), regulatory molecules, targetting domains and association domains. A good example is Calmodulin-dependent Kinase II (CaM Kinase II). This is a large multimeric enzyme acts on upwards of 50 substrates and four functional domains: catalytic, regulatory (it has both inhibitory and CaM binding regions), variable (for targetting and localisation) and association (with other subunits). Other examples of this viewpoint are discussed in Fisher, Malcolm and Paton (in press). Enzymes can be described as 'smart' materials in the sense that they integrate chemical, thermodynamic and electrochemical signals in a context sensitive manner. Indeed, examples like CaM Kinase II also exhibit a memory capacity. Transcription factors also satisfy this case based approach (see later). The reader should note a similarity between this adhesion by cases and the factor complex discussed above. A shift from the society to circuit metaphor will shift the focus to a network of fluxes. However, it is important to keep in mind that society is not being replaced by circuit. Both metaphors are important but emphasise different views on the system as we see with the potential nestings of systemic metaphors in Figure 2.

So far, the discussion about verbs and 'glues' has focussed on adhesion and demonstrated that even at this level we are not simply talking in terms of single binary relations or the sum of independent binary relations. This has been reinforced by the notion of a factor complex and how it relates to the society metaphor, autecology and enzyme behaviour. We now shift attention from a local (adhesion) view of 'glues' to more global perspectives. In particular we shall distinguish between local (worm's eye), semi-local (kangaroo's eye) and global (bird's eye) accounts. Figure 7 summarises these three views in relation to a trefoil knot. A number of points can be made regarding these viewpoints and 'glues':

- a local worm's eye cut can bring about a global change to the whole structure. The whole nature of 'knotiness' is lost when certain cuts are made.
- semi-local interactions are important to the configuration of the whole (a biochemical analogue would be non local interactions for stabilising 3D protein structures). Again loss of certain semi-local interactions affect the whole conformation.
- re-gluing a global change (e.g., putting the pieces together after the knot has been in a blender) may not be possible.

Interestingly, the de-gluings of a trefoil knot can act as a simile for the de-gluings of biosystems and also of the conceptual schemes we employ to model and think about them. Table 1 provides various analogical stances in which glue, verb and viewpoints are compared. Biological hierarchies need vertical and horizontal 'glues' (Paton, 1997).

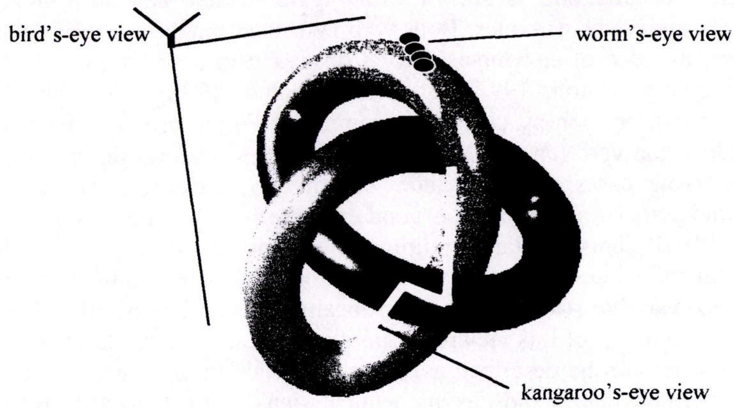


Figure 7 - Some Viewpoints on a Trefoil Knot
(source of knot image, MATLAB™)

Glue Analogue	Textual Analogue	Topographical Disposition	Viewpoint
Adhere	Verb	Local	Worm's eye
Combine	Connector	Semi-local	Kangaroo's eye
Cohere	Subnector (theme)	Global	Bird's eye

Table 1 - Glue from Various Viewpoints

The 'gluing' properties of enzymes were previously discussed in terms of case relations and adhesion. We now examine further insights into the 'glue' relations with some other proteins that have catalytic and other activities. CBP and p300 are large multi-functional proteins (2441 amino acids each). They participate in various basic cellular functions, including DNA repair, cell growth, differentiation and apoptosis. CBP/p300 are focal points for multiple protein-protein interactions and co-activate many other transcription factors including CREB, nuclear receptors, signal transducer and activator of transcription (STAT) proteins, p53, and the basal transcription proteins. Members of the CREB grouping bind to cAMP response elements (CREs) in the promoters of the genes they induce which include somatostatin, enkephalin and α-gonadotrophin.

Giles, Peters & Breuning (1998) review the importance of these two proteins in some detail. Within the context of the present discussion, CBP/p300 act like 'glue' in a number of ways that relate to:

- molecule-molecule bindings and interactions,
- enzymatic processes,
- as a physical bridge between various transcription factors and the basal transcriptional machinery,
- acting as histone acetyltransferases (HATs) - linking transcription to chromatin remodelling and,
- mediating negative and positive crosstalk between different signalling pathways.

It is important to note that from our point of view, this 'glue' is not just that the molecules have intrinsic adhesive properties, they also provide the cell with combinatorial and cohesive properties at a functional level. As was noted above, adhesion is not only through a binary operation but rather can involve more than two interacting parts.

Figure 8 summarises some of the relations between 'gluings' and function for CBP/p300. Object and process can be subsumed as one general term 'glue' with respect to the multi-functionality of these proteins. What is more, it is possible to introduce the topological thinking of local ↔ semi-local ↔ global into the context in which the 'glue' functions. This relates to the terms in the boxes on the left of the diagram. The general issue of 'glue' in relation to object-verb, structure-function and system-category can now be brought together.

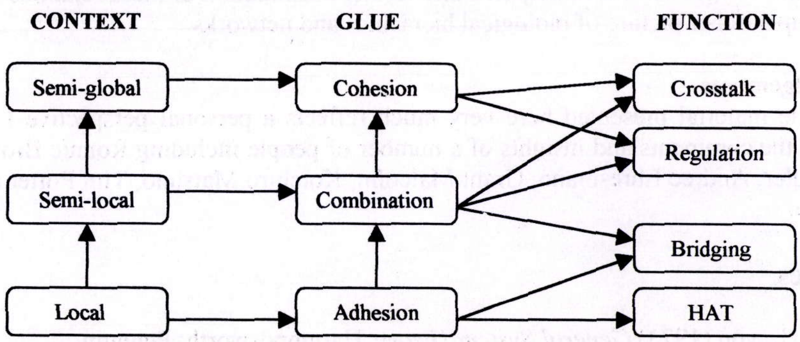


Figure 8 - Some 'Gluings' Associated with CBP/p300

Ehresmann & Vanbremeersch (1987) note that a system can be modelled as a category whose objects are its components and whose arrows (morphisms) are the interrelations. A number of categories may be considered for example, paths in a graph, neurons in a network, neuronal paths in a network, cells in an organism and organisms in an ecosystem. This section has emphasised the importance of verbs and 'glue' in relation to biological systems. Within this representational framework of a system verbs can be likened to the arrows and functors of category theory whilst the product of 'glues' have similarities with pattern, colimit and limit. A **pattern** (diagram) is a collection of cooperating objects. For example, the internal organisation of a protein can by modelled by a pattern of atoms in which links represent chemical relations. A **colimit** (cohesive binding) glues a pattern into a single unity in which the degrees of freedom of the parts are constrained by the whole. A **limit** represents the relationship between whole (i.e., the single unity) and its components. In the protein example, the singular molecule is represented by the colimit of the pattern. The colimit models the integration of the pattern into a single unity. Ehresmann & Vanbremeersch (1996) give another example of the operation of a colomit with the case of an ambiguous figure in which there is simultaneously the colimit of two patterns into which it can be decomposed. The three 'glues' discussed above each have colimits. The nestings of such colimited objects presents a valuable tool of thought for organising and characterising biological hierarchies in terms of intra- and inter- relationships.

5 Concluding Remark

This paper has sought to bring together ideas from biology, linguistics, mathematics and psychology in order to address issues concerned with integrative and intergrating systems. It is hoped that this kind of conceptual analysis will help reveal the conceptual complexity and subtlety of this area of knowledge. It must be noted at this stage that this paper has not explored Ehresmann & Vanbremeersch's evolutive systems that are based on state categories. It is anticipated that further examination of these relations will help build up a richer picture of biological hierachies and networks.

Acknowledgements
Although the material presented here very much reflects a personal perspective I am grateful for the comments and insights of a number of people including Ronnie Brown, Jerry Chandler, Andree Ehresmann, Grant Malcolm, Koichiro Matsuno, Tim Porter and Jim Shapiro.

References

Bertalanffy, L. von (1973) *General System Theory*, Harmondsworth: Penguin.

Chandler, J. L. R. (1998), "Semiotics of Complex Systems: A Hierarchical Notation for the Mathematical Structure of a Single Cell", in Holcombe, M. & Paton, R.C. (eds), *Information Processing in Cells and Tissues,* Plenum, NY, 185-195.

Dyson, F. (1995), "The scientist as rebel", in *Nature's Imagination The Frontiers of Scientific Vision*, Cornwell, J. (ed) Oxford University Press, Oxford, pp 1-11.

Capra, F. (1996). *The Web of Life*. London: HarperCollins.

Ehresmann, A.C. & Vanbremeersch, J-P. (1987), "Hierarchical evolutive systems: mathematical model for complex systems", *Bull. Math. Biol.*, 49, 1, 13-50.

Ehresmann, A.C. & Vanbremeersch, J-P. (1996), "Multiplicity principle and emergence in memory evolutive systems", *SAMS,* **26**, 81-117.

Fisher, M.J., Malcolm, G. & Paton, R.C. (in press) "Spatio-logical Processes in Intracellular Signalling", *BioSystems*.

Giles, R.H., Peters, D.J.M. & Breuning, M.H. (1998), "Conjunction dysfunction: CBP/p300 in human disease", *Trends in Genetics*, **14**, 5, 178-183.

Golley, F.B. (1993), *A History of the Ecosystem Concept in Ecology*, Yale University Press: New Haven.

Goguen, J. A. (1991), "A Categorical Manifesto", *Math. Struct. in Comp.Sci.*,**1**, 46-67.

Harré, R. (1986), *Varieties of Realism: a Rationale for the Natural Sciences*, Blackwell, Oxford.

Hunt, E.B.(1978), In J. W. Cotton and R. L. Klatzky (eds) Semantic Factors in Cognition (Chapter 1). New Jersey: Lawrence Erlbaum.

Markman, E.M., Horton, M.S. & McLanahan, A.G. (1980), "Classes and Collections: Principles of Organisation in the Learning of Hierarchical Relations", *Cognition*, **8**, 227-241.

Markman, E.M. & Hutchinson, J.E. (1984), "Children's Understanding of Constraints of Word Meaning: Taxonomic Versus Thematic Relations", *Cognitive Psychology*, **16**, 1-27.

Paton, R.C. (1988), "Some Investigations into the Meaning and Use of 'System' in Biological Education", Unpublished PhD Thesis, University of Leeds, UK.

Paton, R.C. (1993), "Some Computational Models at the Cellular Level", *BioSystems*, **29**, 63-75 .

Paton, R.C. (1997), "Glue, Verb and Text Metaphors in Biology", *Acta Biotheoretica*, **45**, 1-15.

Paton, R.C. (1998), "The Ecologies of Hereditary Information", *Cybernetics and Human Knowing*, 5, 4, 31-44.

Paton, R.C. (in press), "Systemic Metaphors and Integrative Biology", *Theoria et Historia Scientiarum*, in press.

Paton, R.C. & Matsuno, K. (1998), "Some Common Themes for Enzymes and Verbs", *Acta Biotheoretica*, 46, 131-140.

Rosen, R. (1991), *Life Itself*, Columbia University Press: New York.

Shapiro, J.A. (1991), "Genomes as Smart Systems", *Genetica*, **84**, 3-4.

Singer, (1959), *A Short History of Scientific Ideas*, Oxford: University Press.