*Indiana Undergraduate Journal of Cognitive Science* 4 (2009) 42-53


# Fluid Learning

Jordan Barnes

*Cognitive Science*

*Simon Fraser University*

---

## Abstract

This paper attempts to clarify some of the challenges associated with high-level learning and provide a context for future research directions in this area. A moderate to advanced level of familiarity with the fluid-analogy systems developed by Douglas Hofstadter and his team of researchers at Indiana University is presupposed. These are models of cognition and perception that are characterized by their use of analogies and distributed agents to solve problems. Fluid-analogy systems like Metacat (Marshall, 1999) and Letter Spirit (McGraw, 1992; Rehling, 2001) have received distinguished interest in Cognitive Science for their unique and psychologically plausible approaches. It is with the contributions of these systems in mind that a serious discussion about the obstacles and directions for a comprehensive theory of learning can begin to be fleshed out.

## Introduction

Dampening the wider acceptance of fluid-analogy systems in Cognitive Science, is the difficulty posed in explaining how these systems might be used to learn new concepts. Learning is central to what the mind does, and as such, complex structured systems that make vague claims about learning are easily, and perhaps rightly, viewed as being out of step with the practical problems of the field. This being said, many interesting learning mechanisms already exist within most fluid-analogy systems. Metacat in particular, is endowed with several sub-systems that allow processing runs to be stored and used as influence in future or ongoing workspace activity. Further to this, a good argument can be made that fluid-analogy systems require a functioning working memory, and by virtue of this fact, are *learning* how to appropriately represent a given problem or percept within this working memory.

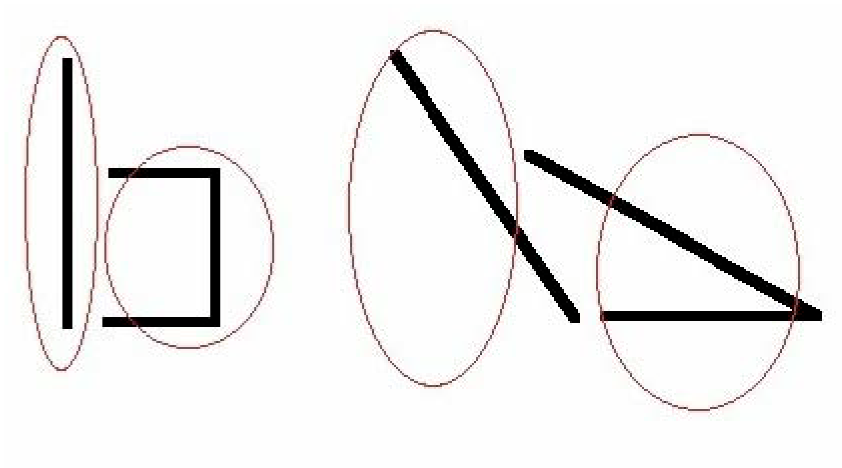
The problem is that the learning that underlies these memory systems is, at present, restricted to the specific domain for which the program was created. This has been a longstanding criticism of fluid-analogy models (Forbus et al, 1998), and one that researchers (Michael Roberts, 2001) at the Center for Research on Concepts and Cognition (CRCC) have indicated that they are working to address. At issue with respect to learning are two critical points. A concept, such as the abstract notion of the crossbar in the letter "t", must be learnable if a system like Letter Spirit is to have any kind of automated extensibility. Letter Spirit currently requires hand-tailored descriptions for its role concepts, which are developed from a combination of statistics and human intuition (Hofstadter et al, 1995). Additionally, learned concepts must have properties that make them domain independent. In other words, what properties does the crossbar concept share with other objects in the world that would allow someone to make it look 'electronic', 'medieval', or 'mountainous'? It is not at all clear what level of analysis is needed for these questions but looking at the problem from a low-level point of view as well as discussing general concept learning seems like the only appropriate way to begin.

For a problem as complex as bridging the gap between low-level, neurobiologically plausible learning, and high-level structured concept learning, two simultaneous research programs are needed. One stream of thought should be devoted to expanding on methods of developing abstract, "role" based, connectionist representations for objects and problems. This kind of work was advanced in a significant way by Douglas Blank in his Analogator (1997) neural networks. These experiments show how simple, abstract representations, for things like the parts of letters (or 'roles' in the terminology of Letter Spirit), can be learned in very basic analogical situations. This work appears to be under-appreciated in Cognitive Science, given that Analogator has provided an interesting way of developing genuinely structure sensitive representations that could be useful in a variety of circumstances. The research potential in this area is wide-open and an example of what can be done with this work is detailed here. The other stream of thought should focus on how fluid-analogy systems can be integrated, such that they can share their individual talents and fill-out, in rich detail, the attributes of learned concepts. This is an immensely complex task, and one that will only be foreshadowed here.

*"[N]aive approaches that assume that backpropagation (or some similar algorithm) will cause roles to emerge in very simple three-layer networks without specific training subtasks which, if well-designed, might force the system to process roles or without specific architectural features (which likewise might force a network to process roles) is misguided."*
- Gary McGraw (pp. 359)

The roles that Gary McGraw refers to in the above quote have to be learned somehow. The discussion here is about what a well-designed system that forces a network to process roles would have to look like. A good first step to understanding the scope of the problem of role-learning would be to work with the learning system that has so far yielded the best results in this domain, Analogator, and see what it can do in terms of discovering roles as opposed to just identifying them.

Analogator's training procedure works as follows. Roles like the kind pictured below in the letter "b" are supplied to a standard backpropagation neural network.



*Two instances of the letter "b" broken down in to its constituent abstract parts of left-post and right-bowl.*

One input bank of the network might accept the right-bowl, and the other would accept the entire letter. Blank refers to the right-bowl in this case, as the "figure". There is one hidden layer and two output banks. The first output bank would be trained through backpropagation to identify the figure and the second would be trained to identify what would be the left-post in this case, or the "ground". Once the network has had sufficient training iterations to identify each piece, the activations of the hidden layer are passed along with a new "b" in a different typeface, as the input to a new network. This network is again trained to dissociate the figure and ground of this new "b", in effect, learning the abstract similarities between the figure and grounds in both letters. This process is repeated many hundreds of times until Analogator is able to identify the figure and ground components for a "b" it has never seen.

Importantly, Analogator has not discovered the abstract roles in the letter "b" on its own. It was a human that decided how a "b" should be broken up, and a human that decided where the breaking point is in all of the training examples. This does not address what McGraw is getting at when he says that backpropagation networks are unlikely to learn new roles. What Analogator needs to be trained to do is discover what the roles are for a letter it has never seen before. This is a much trickier task.

To see if there are such a thing as *analogous* roles in different letters, that can be discovered by a network of this type, Analogator should be trained on questions of the form: 'What is the "|" of "d", in "p" or 'what is the "c" of "d" in "b". The hope is that the abstract roles of the ascender left-post or right-bowl may be learned as a principle of letter construction in general.

## The Analogator Role Learning Procedure

1. Provide the letter "d" in the source scene.
2. Backpropagate the ascender bar as the figure and the bowl or closure as the ground.
3. In a new network, provide the letter "b" with the hidden layer from step 1 as input.

4. Backpropagate the ascender bar as the figure and bowl as the ground.

5. Train on a variety of typefaces, with several different letter combinations such as "b" and "d", "m" and "w", and "z" and "n".

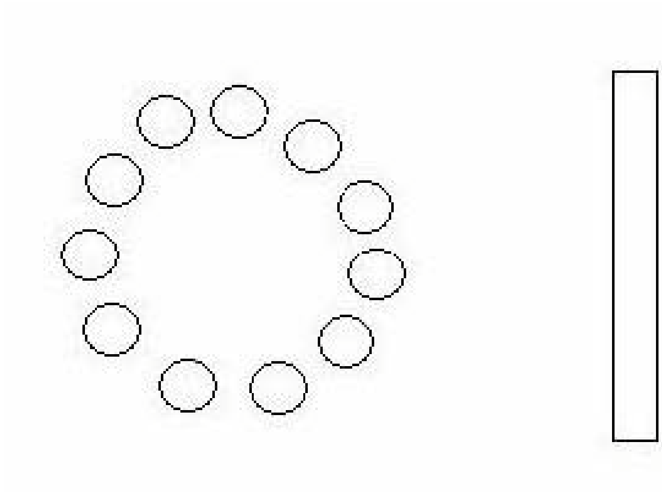
6. Measure the success of the network in parsing a variety of different "p" styles in to parts that correspond to its abstract role categories.

A successful identification should be evidence that the network was able to learn roles for a new category of letter without exposure to it. Of course the problem of how the seed roles were learned is an open question, but one that can be addressed at a later time should the procedure prove successful.

The results of this investigation are not yet conclusive or ready for publication. For reasons that will be elaborated on in the coming discussion, I am skeptical that this approach can possibly have a very high degree of accuracy. The way people develop and use letters seems to be sensitive to all kinds of cultural and instrumental nuances that are beyond the scope of this simple kind of analogy making. At best it seems that the network may learn some rather arbitrary type-design rules about using one letter as inspiration for another. Nevertheless, this approach should be exhausted, as something akin, but likely much more complex, must be taking place in some way in our minds. Before a discussion about high-level learning can take place in earnest, there must be a closer look at some seemingly unrelated issues in high-level perception research. Without being able to presuppose some significant advancements in fluid-analogy architectures, high-level learning really isn't possible. To this end, I will be looking at some of the integration and improvements that will need to take place before detailing a theory of high-level learning.

## Letter Spirit is not good at recognizing "thickness"

Letter Spirit does not fare well at all with "bubble" letters or "thick" parts. Metacat however, has a potential solution to this limitation. Instances of concepts and relationships can be linked to their cardinality in Metacat's Slipnet. For instance, three "E"'s in a row would first be individually labeled by *1-Group* codelets, then a neighboring pair might get a *2-Group* label and finally a *3-Group* codelet might bond them all together. Codelets designed to discover groups permeate the workspace allowing for the re-application of concepts like "successor" to higher-order groups. For example the problem "ABC --> ABD; QBBTTT --> ?" is solvable in Metacat by allowing codelets to discover the relation between letter successorship and numerical successorship of groups. A likely answer from Metacat to the problem posed might be "QBBTTTT" (which relies on seeing QBBTTT as a 1,2,3 group). Grouping is the concept needed to understand "thickness" in Letter Spirit. Consider these two letters:



*An "o" and an "l", composed of lower-level groups of o's and l's.*

By grouping the "o"s together and reapplying the same codelets used in the identification of the low-level "o"s, the high-level figure emerges. Analogously, a standing rectangle can also be seen as a 2-Group of ascenders in the right circumstances. [1]

**Letter Spirit is unable to see a style involving a sequence over several letters**

Not all of the elements of style in a particular typeface are intrinsic to isolated letters within that typeface. What this means is that a stylistic pattern may emerge at a higher-level than the themes contained in individual letters themselves. One could imagine a typeface designed to be taller and taller as letters approach the middle of the alphabet and then shorter and shorter as they approach the end.[2]


*A typeface that grows and shrinks.*

Clearly this is an extrinsic pressure on the style of the alphabet that requires knowledge of the alphabet's successorship relations, and a concept for middle. There is no way Letter Spirit could currently perceive a pattern like this and it does not come up as a topic of discussion in any of the work done on the program.

**Metacat's concepts should be graded**

Metacat can also be improved by incorporating aspects of Letter Spirit. Metacat should

---

1 The assessment of the "right" circumstances gets complicated but the point remains that Letter Spirit's role model could benefit from the use of a concept like successorship.

2 Referring to immutable, a priori information

have varying degrees of relational activation. The cleanness of problem solving in Metacat is a luxury of having "Platonic"[3] (Hofstadter, 1995) concepts. "B" is the predecessor of "C" and "C" is the successor of "B". These relationships are explicitly coded in to Metacat and are either true or false. The relation itself can not reflect the strength to which those objects may or may not activate their respective categories. Letter Spirit's self-determined confidence in its classifications are a natural fit for influencing the degree to which Metacat should pursue a particular path of processing in a given letter string analogy problem.

## A Gestalt-Driven, High-Level, Learning Hypothesis

From a high-level standpoint, the system would work by grouping information together at every level that it can and then refining those groupings over time with feedback from the environment. Said in this way, this theoretical architecture doesn't sound much different than any other common sense view about learning. What makes it unique are the specific details regarding hybridization of existing fluid-analogy systems, in order to *bootstrap* concepts from one level to the next and how to appropriately store and retrieve items from memory.

## Methods

Stimuli to be learned would be provided on a grid meant to be a kind of retinotopic map. Activated points (or quanta, to borrow the term from Letter Spirit) on the map would initially be chunked together into groups based on very coarse sensory discriminations such as color, distance, and connection (Wertheimer, 1938). The topology of these chunkings would be used to try and activate any Gestalts that may have similar topologies, in precisely the same way as it is done in Letter Spirit.[4] If no gestalts are adequately activated to provide top-down biasing, the chunk may be held with temporary activation in something like an iconic memory store. Pressure would mount to re-parse the chunk, which may result in a separate stream of processing spawning two worlds of possibilities[5] with varying interpretations about the nature of the chunks in the workspace. [6]

---

3 Referring to immutable, a priori information

4 Rehling (pp. 168) describes Gestalt activity in Letter Spirit this way: "An Examiner run begins with the segmentation of the gridletter into parts ... The Coderack is then initialized with one Gestalt codelet, which looks for hints about the possible letter category of the gridletter, and, for each part in the segmentation, two looker codelets, which begin the process of identifying what roles the parts may be fillers of [sic]."

5 The language here would seem to invoke ideas from Dennett's Multiple Drafts Theory of the mind (1991). In this theory, people may have several competing versions of reality being argued for and compositionally spliced together in a continuous flow of thought.

6 This kind of parallelism would have to be very carefully constructed however. Allowing new worlds of thought based on every interpretation and reinterpretation of parts could easily end up being combinatorially explosive; as was suggested of Hearsay II's parallelism (McGraw, ibid). I don't pretend to know how to solve this difficulty without first having a model to try different approaches with.

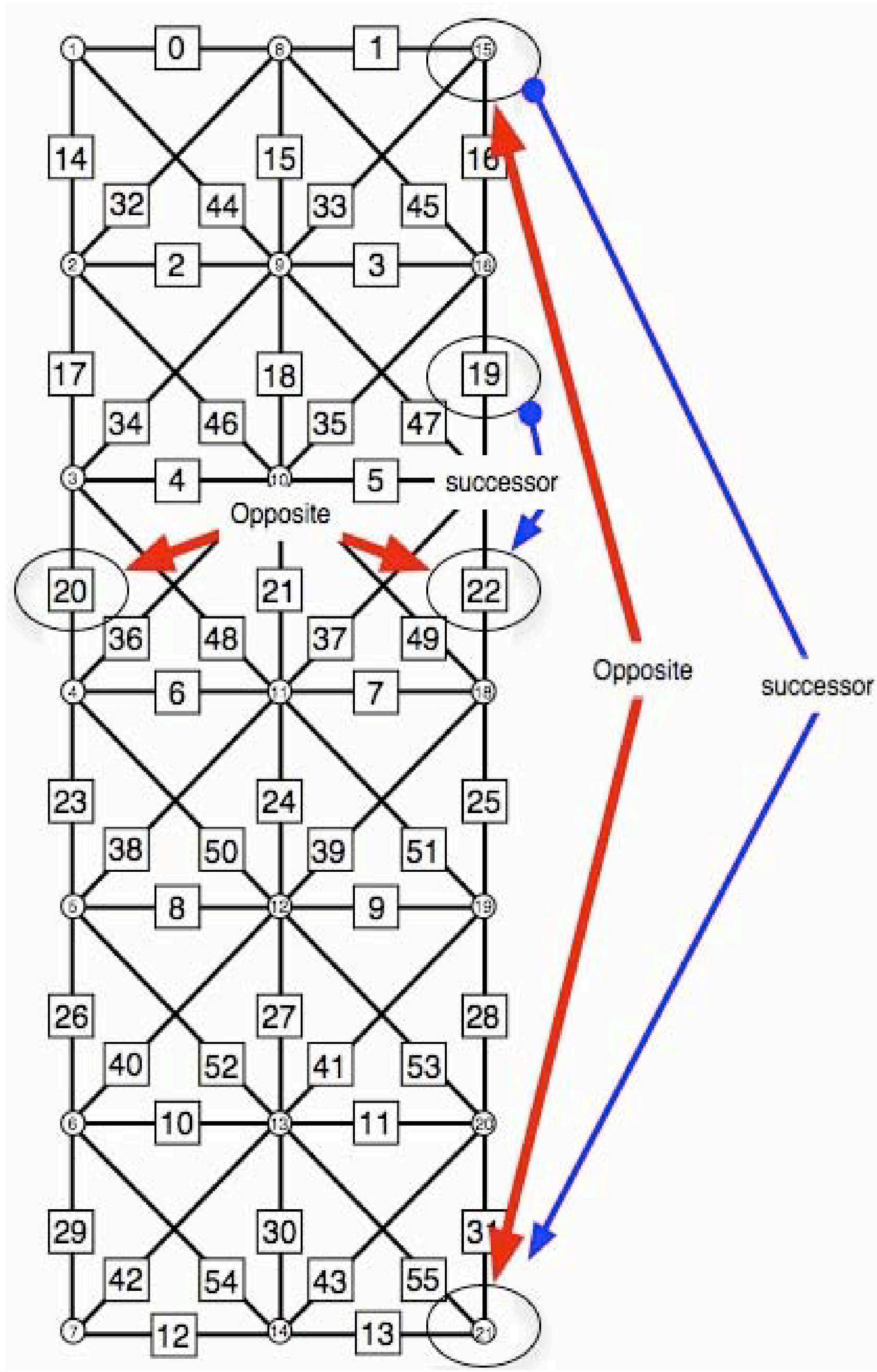
Gestalt codelets corresponding to the topologies of all the different chunks in the workspace would be, at a very minimum, placed in some kind of iconic memory with connections to the other chunks that were developed contemporaneously. It is very important that every chunk in the workspace gets a gestalt codelet with descriptions about its shape, use and context, even if it is never seen or used again. Whether or not a Gestalt codelet gets placed in longer term memory depends on whether or not it gets used and how relevant that use was to the processing of the scene in general. These kinds of abstractions have been shown to be possible as in the case of Metacat's self-determined assessments about critical events in the course of its processing, which build up the contents of the temporal trace and episodic memory. Gestalt codelets are to be thought of as cheap and expendable in this model. Most of what we see only needs to be remembered at the highest levels and without focused attention, unused gestalt codelets should simply be discarded after a brief period of time, in order to clear up processing resources. Evidence for this assumption can be found in work done by Kikuno and Haruo (1991), who showed that people have terrible memories for features that exist at levels lower than what is normally needed for classification, as in the case with the features of everyday coins.

Chunks that do activate existing gestalt codelets, whether in some kind of iconic or longer term memory, would begin to bias the processing similarly to how it is done in Letter Spirit. Perhaps the chunk activates a simple closure, or a more complex figure. Whatever the case may be, these activations initiate top-down processes that begin to terrace the processing landscape in favour of certain pathways over others. These pathways may affect the course of attention over time and unrecognized gestalts may experience pressure to reorganize their parsings in a new way, or they may become the locus of attention because of their poor fit, or both may happen simultaneously in competitive interpretations.

Where this process begins to diverge from Letter Spirit is when chunks are also given inter-chunk relations in something like Metacat's Sipnet. Overlaying concepts from the Slipnet with grid information from Letter Spirit's conceptual network in the same workspace makes inter-domain spreading activation possible.

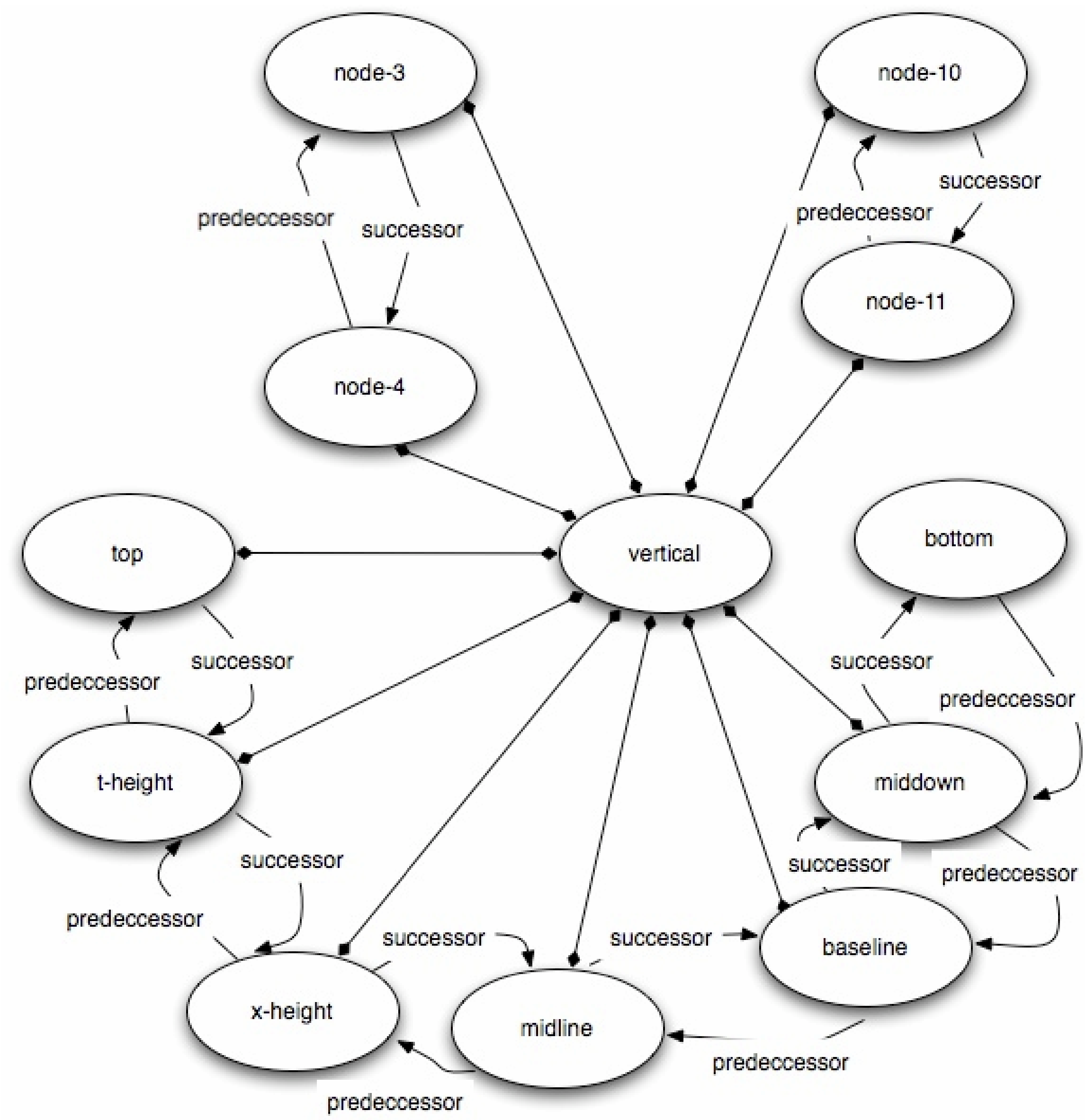
In the previously discussed example of the growing and shrinking alphabet, the concepts of alphabetic-successorship and increasing-relative-size, would need to be connected together in order to form a high-level theme.[7] It is not entirely clear however, how the concept of "increasing" can be represented over the concepts currently in use in Letter Spirit. There is no concept called "taller" for instance, that would allow Letter Spirit to translate zone information relationally. This is an indication that the concepts used to describe the grid in Letter Spirit are not in fact primitives. They are labels for more fundamental sets of relations that have not yet been described but would have to be in order to make the system truly general.

7 At least for the first half of the alphabet




*The Letter Spirit grid (Rehling, 2001, pp. 25) with a subset of example concepts from Metacat's Slipnet (Marshall, 1999, pp. 19) overlaid on top.*

If this is a little unclear, the concept of "tall" in this case should be thought of as a certain group length of successor relations relative to everything else in the grid. There are two advantages to describing areas of the grid in these terms. One is that the concept of "taller" in the example has an actual meaning. It is the successor of a certain magnitude of other successors along the vertical dimension. A potentially useful idea for converting Letter Spirit's norms into sets of more primitive descriptions is to think of everything as being relative: tall is tall, relative to other members of the same kind of category. The second advantage of this terminology is that it becomes yet another topology of the concept upon which analogies can be seen between domains.




*Another subset of examples: this time of conceptual connections using pairwise connections between some of the nodes from Letter Sprit's grid and their slipnet relations.*

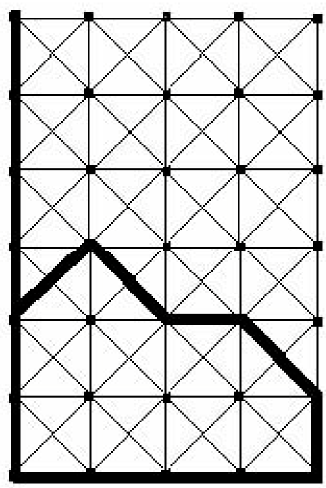
In "To Seek Whence Cometh a Sequence", Hofstadter (1995) provides sequences of numbers for what he calls a "mountain-chain" (pp. 57). If the y-coordinates <1,2,3,4,5,4,4,3,2,1> are plotted in succession over the x-axis you get a pattern that looks similar to a mountain with a peak and a plateau. Similarly, a norm description in these numeric terms would look something like this for the letter "v":

<9,8,7,6,5,4,3,2,1,2,3,4,5,6,7,8,9>

or in successorship terms:[8]

<s(s(s(s(s(x))))), s(s(s(s(x)))), s(s(s(x))), s(s(x)), s(x), x, s(x), s(s(x)), s(s(s(x))), s(s(s(s(x)))), s(s(s(s(s(x)))))>.

What this provides is a neutral way of describing objects such that relationships between them can be identified at higher-levels of perception and *inter-domain* similarities discovered by virtue of the descriptions of other concepts in the same terms. Earlier in this paper, it was asked how you might describe a part of a letter such that it could be considered "mountainous". Domain neutral, topological descriptions are the way to do it.



*A crude attempt at a "mountain b", demonstrating the usefulness of relative descriptions for inter-domain analogies.*

An entangled hierarchy involving the machinery of fluid-analogy systems like Letter Spirit and Metacat, could form the basis for shifting levels of abstraction from quanta to parts to roles to wholes to parts to roles to wholes *ad infinitum*, until a representation is understood for what it truly is in context. Learning then, is the storage in memory of those chunks or gestalts that contributed to a higher-level picture, along with how they were used. Groups and their contextual relations become the roles or *active concepts*[9] through which semantics are constructed. Their existence is necessitated by the whole they became a part of, and by the gestalt that was created for that whole. It is a role *hypothesis* because their validity is relative. Should the role cease being useful in its current form, it can change or be discarded over time.

---

8 s(x) being the Peano notation for "successor of x".

9 *Active concepts* is a term coined to describe references that are not merely empty symbols that get shunted around in a computer, but animate and dynamic processes that contribute to a non-deterministic semantics.

## Development

Of course, this model is only a rough approximation of an actual implementation. It is meant to pump the intuitions of those interested in learning. The kinds of codelets, their parameters and the precise structure of the system would all have to be tweaked over hundreds if not thousands of trials, mirroring the development of the other fluid-analogy models. Although it may not be achievable in full at the moment because of processing constraints, an idealized learning system of this type is very much a possibility[10].

## Works Cited

Blank, D. “Learning to see analogies: a connectionist exploration”. PhD thesis, Indiana University, Bloomington, 1997.

Dennett, D. “Consciousness Explained” Little, Brown & Co. 1991.

Forbus, K., Gentner,D., Markman, A., and Ferguson, R. “Analogy just looks like high-level perception: Why a domain-general approach to analogical mapping is right”. Journal of Experimental and Theoretical Artificial Intelligence, 10, 231–257. 1998.

Foundalis, H. "Phaeaco: A Cognitive Architecture Inspired by Bongard’s Problems". 2006.

Hofstadter, D. “Fluid Concepts and Creative Analogies: Computer Models of the Fundamental Mechanisms of Thought”. Basic Books, Inc. 1995.

Kikuno and Haruo. “Memory for distinctive and common features of coins”. Psychological Reports, Vol69(3, Pt 1) pp. 867-870. 1991.

Marshall, J. "Metacat: A Self-Watching Cognitive Architecture for Analogy-Making and High-Level Perception". 1999.

McGraw, G. “Letter spirit: Recognition and creation of letterforms based on fluid concepts”. TechnicalReport 61, Center for Research on Concepts and Cognition. 1992.

Rehling, J.”Letter Spirit (Part Two): Modeling Creativity in a Visual Domain”. Doctoral Thesis. Indiana University. 2001.

Roberts, Michael. ″A Cognitive Magnifesto". 2001. http://www.cogsci.indiana.edu/farg/michael/proleg.html

10 Readers interested in these ideas may find the recently published dissertation of Harry Foundalis (2006) entitled "Phaeaco", to be a much more articulated and complete vision for fluid learning and cognition. I was only made aware of this work after finishing this paper - unfortunate, given its high degree of relevance.

Wertheimer, Max. First published as Untersuchungen zur Lehre von der Gestalt II, in Psycologische Forschung, 4, 301-350. Translation published in Ellis, W. A source book of Gestalt psychology (pp. 71-88). Routledge & Kegan Paul.[1]. 1938.