Relating language, logic, and imagery

Arun K. Majumdar, John F. Sowa*

Kyndi, Inc., San Mateo, California

Abstract

The world is a continuum, but words are discrete. Sensory organs map the continuous world to continuous mental models of sights, sounds, actions, and feelings. Those mental models, which represent a moving 3-D virtual reality (VR) are the semantic foundation for all versions of language and logic. A common model for cognition must be able to process and relate all modalities. Kyndi technology represents all information in graphs. They include conceptual graphs for symbolic information and arbitrary graphs for 2-D icons or 3-D VR. All graphs are stored in Cognitive Memory, which can find approximate mappings for arbitrary graphs in logarithmic time. Those mappings include formal unification for logics, and informal analogies for case-based reasoning. The analogies can even map conceptual graphs to the graphs derived from imagery or VR simulations. For reasoning, Peirce’s rules of inference for existential graphs can support operations on arbitrary icons, such as the diagrams in Euclid’s geometry. Those rules, when adapted to conceptual graphs, can map symbolic languages to and from Euclidean style geometrical reasoning. With two new rules of inference, called observation and imagination, the Standard Model of Cognition can support mental models without making any current software obsolete.

© 2018 The Authors. Published by Elsevier B.V.

This is an open access article under the CC BY-NC-ND license (https://creativecommons.org/licenses/by-nc-nd/4.0/)

Keywords: Deep Learning; Computational Creativity; Cognitive Science

* Corresponding author.

E-mail address: sowa@bestweb.net
1. Generalizing the standard model to include imagery

The mechanisms for relating language, logic, and imagery are fundamental for cognitive science and for any general AI system. Long before they speak their first words, infants learn to relate their bodies to everything they see, feel, and hear, especially their parents and siblings. That knowledge is the semantic foundation for language and reasoning. To illustrate the issues, consider sentences spoken by a child named Laura shortly before her third birthday [11]:

   Here’s a seat. It must be mine if it’s a little one.
   I want this doll because she’s big.
   When I was a little girl I could go “geek-geek” like that. But now I can go “this is a chair.”

Laura’s content words express concrete images, directly related to her actions. But her syntax and function words express a surprising amount of complex logic: possibility, necessity, tenses, indexicals, conditionals, causality, quotations, and metalanguage about her own language. As another example, a mother was talking with her son, who was about the same age as Laura [12]:

   Mother: Which of your animal friends will come to school today?
   Son: Big Bunny, because Bear and Platypus are eating.

   The mother looked in his room, where the stuffed bear and the platypus were sitting in a chair and “eating”. The boy had imagined a situation, built a model of it, and based his reasoning on it: The bear and the platypus are eating. They can’t eat and go to school at the same time. Big Bunny isn’t doing anything. Therefore, Big Bunny is available.

   The semantic aspects of imagery, action, and feelings appear to be more important than syntax. This point is supported by studies of infants with one deaf parent and one speaking parent. At every stage of development, they have equal ability to express themselves in one-dimensional speech or in moving, three-dimensional gestures [30]. In fact, infants with two deaf parents babble with their hands, not with vocal sounds. The neuroscientist Antonio Damasio [6] summarized the issues:

   The distinctive feature of brains such as the one we own is their uncanny ability to create maps... But when brains make maps, they are also creating images, the main currency of our minds. Ultimately consciousness allows us to experience maps as images, to manipulate those images, and to apply reasoning to them.

   The maps and images form mental models of the real world or of the imaginary worlds in our hopes, fears, plans, and desires. They provide a “model theoretic” semantics for language that uses perception and action for testing models against reality. Like Tarski’s models, they define the criteria for truth, but they are flexible, dynamic, and situated in the daily drama of life. The logician, mathematician, scientist, and philosopher Charles Sanders Peirce would agree. Although he invented the algebraic notation for predicate calculus [22], Peirce claimed that all reasoning is based on a “concrete, but possibly changing, mental image” that may be aided by “a drawing or a model”:

   All necessary reasoning without exception is diagrammatic. That is, we construct an icon of our hypothetical state of things and proceed to observe it. This observation leads us to suspect that something is true, which we may or may not be able to formulate with precision, and we proceed to inquire whether it is true or not. For this purpose it is necessary to form a plan of investigation, and this is the most difficult part of the whole operation. We not only have to select the features of the diagram which it will be pertinent to pay attention to, but it is also of great importance to return again and again to certain features. [26, 2:212]

   The word diagram is here used in the peculiar sense of a concrete, but possibly changing, mental image of such a thing as it represents. A drawing or model may be employed to aid the imagination; but the essential thing to be performed is the act of imagining. Mathematical diagrams are of two kinds; 1st, the geometrical, which are composed of lines (for even the image of a body having a curved surface without edges, what is mainly seen by the mind’s eye as it is turned about, is its generating lines, such as its varying outline); and 2nd, the algebraical, which are arrays of letters and other characters whose interrelations are represented partly by their arrangement and partly by repetitions. If these change, it is by instantaneous metamorphosis. [26, 4:219]

   Methods of reasoning with and about symbols, diagrams, and models have been analyzed and debated since early days of AI. The simplest way to resolve the debates is to change the terminology and the diagrams that describe the
issues. Figure 1 compares the diagrams for the Standard Model [10], and DCT, the Dual-Coding Theory by Paivio [20, 21].

More software has been developed for the Standard Model than for DCT, but the two models are complementary. With revised terminology and a reorganization of the boxes in the diagrams, it would be possible to design an integrated version that is upward compatible with both. In DCT, the double arrow labelled “referential connections” suggests a mapping. To define that mapping, Sowa [34] proposed two rules of inference called observation and imagination, which map images to and from symbolic propositions that describe those images. To illustrate the mappings, Figure 2 shows a statement of Euclid’s Proposition 1 in controlled English and a diagrammatic representation based on Peirce's existential graphs (EGs) combined with Euclidean diagrams.

The referential connections between the DCT symbols and images may be represented by lines, pointers, variables, labels, or names. Peirce called them indexes or selectives. As Figure 2 shows, Euclid’s labels correspond to pronoun—\( \sim q \). That corresponds to an if-then statement, which may contain any kind of symbols, graphs, or images inside the ovals.

The first step of mapping images to symbols is a translation of Figure 2 to the Existential Graph Interchange Format (EGIF), which is a subset of the Conceptual Graph Interchange Format (CGIF). The ovals in Figure 2 are a hybrid
notation, but EGIF itself is purely linear. The ovals are represented by a pair of brackets with a tilde in front: \(~[~]\); for readability, a pair of negations may be represented by the keywords 'If' and 'Then'. Following are three symbolic translations:

\[
\text{EGIF: } [\text{If } (\text{Line } *AB) [\text{Then } (\text{Triangle } *ABC) (\text{Line } *AC) (\text{Line } *BC) (\text{HasSides } ABC \text{ AB AC BC}) (\text{Congruent } AB \text{ AC BC})] ]
\]

\[
\text{CLIF: } (\text{not exists } ((\text{AB Line})) (\text{not exists } ((\text{ABC Triangle})) (\text{and } (\text{Line AC}) (\text{Line BC}) (\text{HasSides } ABC \text{ AB AC BC}) (\text{Congruent } AB \text{ AC BC}))))
\]

\[
\text{English: } \text{If there is a line AB, then there is a triangle ABC; there are lines AB, and AC; ABC has sides AB, AC, and BC; AB, AC, and BC are congruent.}
\]

Inside the negations, EGIF can assert three kinds of sentences: (1) something of some type exists; (2) there is a relation or function among two or more things; (3) two or more names refer to the same thing. CGIF and CLIF are two dialects of ISO standard 24707 for Common Logic, which can be translated to many versions of logic, including controlled English and the many notations for the Semantic Web. The asterisk * in EGIF or CGIF maps to the keyword in predicate calculus, and the phrases ‘there is’ or ‘there are’ in English.

Although Peirce did not insert images inside EGs, he wrote that any image marked with one or more indexes can state a proposition. As an example, he mentioned a portrait with an attached name of the person portrayed. Peirce also wished he had the funds to purchase equipment for making motion pictures and stereoscopic images. He would have loved the software for virtual reality. He invented the algebraic notation for predicate calculus, but he called EGs his “chef d’oeuvre” and claimed that their rules of inference generate “a moving picture of the mind in thought.” After a detailed comparison of Peirce’s EGs to current theories about mental models, the psychologist Johnson-Laird [8] agreed:

Peirce’s existential graphs are remarkable. They establish the feasibility of a diagrammatic system of reasoning equivalent to the first-order predicate calculus. They anticipate the theory of mental models in many respects, including their iconic and symbolic components, their eschewal of variables, and their fundamental operations of insertion and deletion. Much is known about the psychology of reasoning... But we still lack a comprehensive account of how individuals represent multiply-quantified assertions, and so the graphs may provide a guide to the future development of psycho-logical theory.

2. Cognitive memory

With billions of neurons working in parallel, the brains of humans and other animals can access an immense memory in milliseconds. By contrast, path-based methods for searching and retrieving graphs take N-cubed time for a database of N graphs [7]. Even for a thousand graphs, \(N^3\) is a billion, and the time for a million graphs would be longer than the age of the universe. Word vectors can be used to find graphs with similar labels in logarithmic time, but they ignore the graph structure. Chemists have developed algorithms for finding graphs of organic molecules in logarithmic time [31]. But they are designed for a highly specialized ontology, and the algorithms cannot find all subgraphs of the graphs that are indexed. To enable all subgraphs of any graph to be found in \((\log N)\) time, each subgraph would require a separate entry in the index. That would cause an exponential increase in the space requirements.

In The Emperor’s New Mind, Roger Penrose [28] suggested that human intelligence results from quantum mechanical operations in the neural cells. But he later admitted “In my view the conscious brain does not act according to classical physics. It doesn’t even act according to conventional quantum mechanics. It acts according to a theory we don’t yet have” [29]. For their theory of quantum cognition, Busemeyer and Bruza [4] observed that human beliefs don’t jump from one state to another. Instead, people feel an ambiguous superposition of all the options or “eigenstates.” Although Peirce didn’t use the terminology of quantum mechanics, he aptly characterized the uneasy feeling:

Doubt is an uneasy and dissatisfied state from which we struggle to free ourselves and pass into the state of belief; while the latter is a calm and satisfactory state which we do not wish to avoid, or to change to a belief in anything else. On the contrary, we cling tenaciously, not merely to believing, but to believing just what we do believe. [25 5.372]
As Penrose observed, a theory that predicts the probability of one option or another is useful. But a computational system for language understanding and reasoning requires a theory that can represent the details of those options. Some method of quantum knowledge representation (QKR) is necessary. Some authors propose a cognitive version of quantum mechanical wave functions: an infinite dimensional Hilbert space of complex functions whose eigenstates represent alternative beliefs [5, 1, 36]. The “fixation of belief,” as Peirce called it, would correspond to a measurement that resolves the superposition by forcing the QKR for a confused state of doubt into an eigenstate that has a definite answer.

To find and retrieve graphs in (log N) time, Majumdar [14] developed a quantum-like knowledge representation (QKR), which encodes graphs in a numeric form. For a knowledge base of N graphs, the retrieval time scales as (log N). For any graph g, its QKR encoding is called the Cognitive Signature™ of g, and the system that stores and retrieves them is called Cognitive Memory™ (CM). The cognitive signature is called a QKR because it exhibits the key properties of superposition, entanglement, and uncertainty. The operations of searching and graph matching, when performed on the QKR, are analogous to the measurements in quantum-mechanical systems. A quantum computer would be ideal for processing them. But even with today’s digital computers, searching and graph matching with a QKR can be performed with floating-point computations that scale in logarithmic time. For random access of “Big Data” during language understanding and reasoning, CM enables an ordinary laptop to outperform a supercomputer.

As a QKR, the cognitive signature encodes conceptual graphs or similar graph notations in a numeric form. For any conceptual graph g, such an encoding is called the cognitive signature™ of g. Figure 3 shows Cognitive Memory™ as a system for storing the signatures of all CGs in a knowledge base. The lower level of Figure 3 shows how a query CG q is mapped to a cognitive signature and compared to a database of stored signatures. That DB would also store links back to the original CGs or to other notations from which the CGs were derived.

The mapping shown in Figure 3 involves two kinds of vectors or multivectors: topological vectors that encode the structure of the graphs, and ontological vectors that encode the labels on the nodes and arcs. The topological vectors represent the connectivity, branching, and cycles of the graphs. The ontological vectors represent the types of concepts and relations. The type labels are organized in a hierarchy (partial ordering) that may be specified by the axioms and
definitions of some ontology. The details of how the hierarchy was specified are not encoded in the cognitive signatures. An informal taxonomy such as WordNet or a partial ordering derived from a set of axioms would be encoded in the same way.

The function labelled $f$ in Figure 3 combines the two vectors to form a cognitive signature. But $f$ may be any member of a family of encodings that could emphasize different aspects of the structure or the ontology. A function that ignores the ontology would encode a structure of unlabelled graphs. A function that ignores the structure would encode a bag of concepts and relations. For most applications, a combination of the two is best, but $f$ may give greater weight to certain aspects of the structure or the ontology. During the development of Kyndi technology, many algorithms for representing $f$ were implemented and tested. In the best versions, the encoding of a graph $g$ is a superposition of the encodings for each of its subgraphs. As in quantum mechanics, the subgraphs are eigenstates that are virtually present in the encoding. They can be detected by a measurement (i.e., a search with a query graph), but they do not occupy space in any physical medium. As a result, a QKR encoding has the effect of making all subgraphs accessible to a search, but without increasing the time or space requirements.

The time to encode a graph scales as a polynomial in the size of the graph. The time to find matching graphs scales logarithmically with the number of graphs in the database. For encoding a single graph, the exponent of the polynomial depends on the complexity of the interconnections. In encoding a document, for example, the connections within a sentence may be complex, but the references that link sentences within a document or collection of documents are much sparser. The observed encoding time for the CGs derived from a document or a corpus of documents is proportional to $(N \log N)$, where $N$ is the total number of sentences.

For finding analogies and case-based reasoning, some similarity measure is necessary to evaluate the semantic distance between any two graphs. But similarity is highly context dependent. In a biological classification, for example, the number of levels that separates molluscs from vertebrates is less than the number that separates cats from dogs. In general, semantic distances at the lower levels should be considered much shorter than distances at the upper levels. But context is also important: a guppy, for example, is not a typical fish nor a typical pet, but it is a typical pet fish. When considered as food, salmon would be closer to clam than guppy. For an orthodox Jew, however, clams are not considered food.

When people hear or read language, they take context into account at every stage of interpretation and reasoning. Some context can help resolve syntactic ambiguity during the initial stages, but more complex issues are usually processed during a later reasoning stage. As examples, consider the following four sentences:

1. Bob bought fish for dinner.
2. Bob bought fish for his aquarium.
3. Yoko Suzuki bought seafood for dinner.
4. Isaac Levi bought seafood for dinner.

By finding relevant semantic information while a parser is analyzing a sentence, Cognitive Memory (CM) can improve performance by resolving syntactic ambiguities as soon as possible. That contribution to computational efficiency also emulates human performance. For the simple sentences above, the syntax is unambiguous, and the parser would not need context-dependent information. It would just use a lexicon to find the syntactic features and a suitable concept type for each word. The words 'fish' and 'seafood', for example, would be mapped to concept nodes [Fish] and [Seafood] in conceptual graphs.

At a later stage, somebody might ask for more detail about the fish or seafood. To find relevant information, the Kyndi processor would use the conceptual graphs as query graphs to find matching graphs in CM. It’s possible that some matching graph might state the exact kind of fish or seafood. More likely, CM would find graphs about the typical kinds of fish for dinner or for an aquarium. To determine the kind of seafood for Ms. Suzuki or Mr. Levi, several steps would be required. Kyndi software would not take those steps unless somebody asked more detailed questions about the food choices by people from different ethnic groups.

To compute a semantic distance that takes the levels and the context into account, Majumdar [11] specified a method of conceptual relativity that fixes the distance from the top to the bottom of any ontology at 1.0. Distances decrease at the lower levels, and they are affected by qualifiers such as pet or food. But the semantic distances are not affected by new levels that may be added with methods of learning or knowledge acquisition.
For finding all subgraphs of a given graph, the QKR aspects of cognitive signatures are essential: the cognitive signature of a subgraph is a function of the cognitive signature of any graph that contains it. For example, the CG for “fish for dinner” would be a subgraph of the CG for “Bob bought fish for dinner”, which would be a subgraph of the CG for an entire paragraph. When the cognitive signature for the paragraph is stored in CM, the cognitive signatures for all subgraphs are accessible in logarithmic time. There is no need to store and index every subgraph separately. For the detailed algorithms, see the patents by Majumdar [12, 13, 14]. For the use of CM in various applications, see the examples by Majumdar et al. [15, 16]. Although the CM algorithms are very different from anything that happens in the human brain, the examples show that CM can make NLP more human-like.

3. Society of mind

The human brain is a complex system of diverse components with different kinds of representations interacting at various stages during the processes of perception, action, learning, reasoning, and life. The psychologist Lawrence Barsalou [3] maintained that cognition is not a single, unified method or mechanism. Instead, he claimed that cognition is *coordinated non-cognition*: Cognition is “embedded in, distributed across, and inseparable from” the “processes of perceiving, acting, and emoting.” When people view a static object, they anticipate working with it. When people view food, they anticipate its taste when eating it. Musicians identify their own performances by recognizing the judgments are fundamental to all aspects of reasoning and decision making. No single aspect is cognition, but all of them together are cognition. Social interactions facilitate learning by stimulating more aspects.

After a dozen years of research in AI, Marvin Minsky [18] reached a similar conclusion: no single mechanism, by itself, can adequately support the full range of functions required for a human level of intelligence. Two decades later, he developed those insights in the book *Society of Mind* [19]:

What magical trick makes us intelligent? The trick is that there is no trick. The power of intelligence stems from our vast diversity, not from any single, perfect principle. Our species has evolved many effective although imperfect methods, and each of us individually develops more on our own. Eventually, very few of our actions and decisions come to depend on any single mechanism. Instead, they emerge from conflicts and negotiations among societies of processes that constantly challenge one another.

These observations imply that human-level cognition requires an open-ended variety of coordinated modules or agents. Section 1 of this article proposed a merger of the Standard Model and Dual-Coding Theory with the rules of observation and imagination to relate symbolic methods and mental models. Section 2 introduced Cognitive Memory for high-speed storage and retrieval of any kind of information that could be mapped to and from graphs. And this section shows the need for a framework that supports and coordinates high-speed message passing among heterogeneous modules. Kyndi technology uses the Flexible Modular Framework (FMF), which was proposed by Sowa [32] and developed further by Majumdar, Sowa, and Stewart [15].

Four important features of the FMF: First, no module needs to know anything about the internal representations or mechanisms of any other module. Second, a module can discover other helpful modules just by broadcasting a request and checking which modules, if any, respond. Third, any module may use cognitive memory to store and retrieve any information about anything, including itself. Fourth, any hardware or software system of any kind may become an FMF module just by putting a wrapper around it that responds to FMF messages. Those features imply that any module may be replaced or upgraded at any time without disrupting any previous services. They also imply that no module ever becomes obsolete. If one of the newer modules breaks or fails for any reason, one or more of the old modules could still respond to the same kinds of messages it handled in the past.

Finally, The FMF would enable a system designer to combine AI components of any size. The simplest method is to choose one system as the master and to put an FMF wrapper around other components, including complete cognitive systems. But a deeper level of integration would subdivide each components into one or more FMF modules that could automatically form new interconnections among modules from different sources. That method would enable a cognitive system to grow with the aid of a human designer or to reorganize itself, as it responds to the patterns of usage in typical applications. For examples, see Majumdar et al. [15, 16].
References


