Non-Well-Founded Set Theory and the Circular Semantics of Semantic Networks



Robin K. Hill
Computer Science Department
226 Bell Hall
State University of New York at Buffalo
Buffalo, New York 14260
hill@cs.buffalo.edu

Abstract

The theory of non-well-founded sets-that is, sets that are potentially circular in hereditary membership—can be applied to a semantic network knowledge/belief representation, SNePS. It provides a particular type of node, the base node, which represents a primitive concept, with a semantics that is both influenced by and influences its dominating compound nodes. A semantic function is defined that interprets each node in the network as a non-well-founded set over the sensory input nodes (words, visual stimula, etc.).

Under certain axioms governing SNePS structure, results show that the semantics does not allow meanings that are circular to the point of vacuity, and that the semantics is inherent in the graphical structure of the SNePS network itself. The semantics supports SNePS principles such as that the meaning of each node (concept) is distinguished from all others, and that the meaning of a node is internal, dependent on its location with regard to other nodes in the network (rather than on external phenomena). An enhanced semantics, which incorporates into the hyperset semantics the relations used to label the arcs, is also developed.

1 Circularity in Semantics

Several researchers in artificial intelligence and knowledge representation note that some real-world phenomena seem inherently circular [Barwise, 1989, pages 194–198], [Nebel, 1991], [Smith, 1991, pp. 265 ff.]. The semantics of SNePS is intended to be circular, notwithstanding its acyclic graph representation, in the sense that the meanings of certain directly-connected nodes influence each other. We address the provision of circularity in semantics using the theory of non-well-founded sets the tool.

sech node

のでは、10mmのでは、

Non-Well-Founded Set Theory

Set theory provides a rigorous environment in which to reason about objects. Set theorists adopt a homogeneous typing under which all members of sets are regarded

5Ner.

375

A. Yantis (ed.), Intelligent Systems, 375–386.

2 1995 Kluwer Academic Publishers. Printed in the Netherlands.

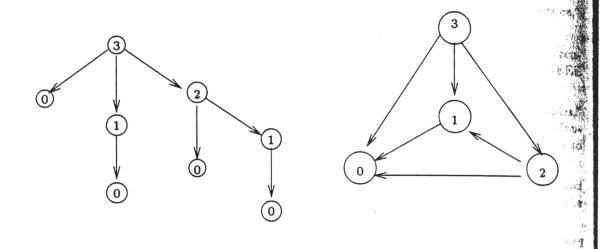


Figure 1: Graphical pictures of 3 as a (hereditary) set

as sets themselves; sets are hereditary. A member of a set S might also be a subset of S. When we need concrete objects that are not sets, we will call them 'atoms'. This simple mechanism permits the standard set-theoretic definition of the natural numbers \mathcal{N} as $\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \{\emptyset, \{\emptyset\}\}\}, \dots$ Each $n \in \mathcal{N}$ is represented by the set $\{m \in \mathcal{N} \mid m < n\}$, or, equivalently, $n+1 = n \cup \{n\}$, where \emptyset is identified with zero. For example, $\emptyset \subset \{\emptyset, \{\emptyset\}\}\$ and $\emptyset \in \{\emptyset, \{\emptyset\}\}\$; 0 is both a member and a subset of 2. The hereditary nature of these sets allows them to be meaningfully depicted as rooted directed graphs, with the arrows showing membership. Figure 1 shows the set we call '3' in two ways, on the right with unique occurrences of the member nodes.

Various axiomatic systems of set theory have been developed; one standard is that called Zermelo-Frankel Set Theory with the Axiom of Choice, abbreviated ZFC. It has nine axioms, with the Axiom of Foundation, disallowing any sets that contain themselves, either directly or indirectly, like the sets below:

$$a = \{a\} \tag{1}$$

$$a = \{a\}$$

 $b = \{s,t\}, \text{ where } s = \{t\} \text{ and } t = \{s\}$ (2)

In his theory of non-well-founded sets, Peter Aczel negates the Axiom of Foundation, retaining the others [Aczel, 1988]. The statement of his axiom relies on the graphical depiction of sets, where an accessible pointed graph or apg of a set is a directed graph with a distinguished node called the point from which all other nodes are reachable, which has a decoration, an assignment of sets to each node such that the children are members of their parents' sets. (A node with no children is assigned the empty set.) Non-well-founded set theory is ZFC with the Axiom of Foundation replaced by a negation of it called the Anti-Foundation Axiom.

The Anti-Foundation Axiom: Every graph has a unique decoration.



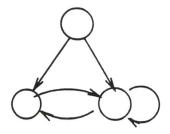


Figure 2: Other pictures of Ω

In other words, even cyclic graphs represent sets. All acyclic graphs, such as that of Figure 1, still represent sets, too. A simple example of one of the new non-well-founded sets, or *hypersets*, is given by the picture below.



Aczel calls the hyperset depicted Ω . Like other hypersets, this one has many pictures—including those in Figure 2. Note that writing down its contents in standard set notation is problematic. Because it is its own child, the solitary node would have to be decorated with something like this: $\{\{\{\ldots\}\}\}\}$.

We now bring in a set of atoms A; a decoration then assigns to each childless node either an atom or the empty set, with the decoration of other nodes made up of the sets assigned to their children, as before. The notation V_A stands for all hypersets over the set of atoms A, all hypersets that may (but are not required to) include elements from A or other sets that include elements from A. The range of the decoration is V_A . The thrust of the Anti-Foundation Axiom is that every picture, even if it has cycles, has a membership-relative assignment of sets to each node.

To put it another way, an equivalent of the Anti-Foundation Axiom called the Solution Lemma that any system of membership relations can be satisfied by hypersets. The notions of a system of equations and its solution are built up as in albegra, with the equations providing a definition for each "indeterminate", expressed as a hyperset over atoms and indeterminates (and so possibly circular), and the solution being an assignment to each indeterminate of a hyperset over atoms only, not other indeterminates, such that all of the original relationships given by the equations still hold when its assignment is substituted for the indeterminate.

The Solution Lemma: Every system of equations in a collection \mathcal{X} of indeterminates over $\mathcal{V}_{\mathcal{A}}$ has a unique solution.

This gets us non-well-founded set theory without reliance on graph notation or concepts. We will use both forms.

SNePS and its Principles

The particular knowledge representation used is SNePS, developed by Shapiro dialia. SNePS is a semantic network representation, with nodes and directed

arcs, and it adheres to principles that promote intensionality, such as relative belief spaces; it has other significant aspects, such as inference systems, that do not bear directly on this work. SNePS networks are acyclic and can be defined compositionally, as new nodes and arcs are added to networks in certain patterns, thereby creating new larger networks.

A SNePS network is a propositional semantic network, that is, one in which every proposition represented in the network is represented by a node (rather than an arc). Arcs are best regarded as punctuation, having no conceptual semantics. For this reason, it is forbidden to add an arc between two existing nodes. Certain arc labels come with SNePS, while others necessary for a particular implementation are to be defined by the user. Together they form the set of relations of the implementation.

Arcs are directed; the node at the origin is called the 'tail' node and the node at the arrowhead, the 'head' node. There would be no point in connecting two nodes with multiple instances of the same arc (arcs with the same label), but there may well be multiple arcs with the same label emanating from the same tail node but terminating at different head nodes, or multiple arcs with different labels connecting two nodes. Nodes with no arcs emanating from them are called 'atomic'. They include sensory nodes, which represent the real-world interface, being associated with printed text strings, other visual or tactile or aural data, etc., and base nodes, which represent individual concepts, and hence have in-arcs from other nodes.\(^1\) Nodes that do have arcs emanating from them, i.e., that dominate others through their out-arcs (and may themselves be dominated with in-arcs), are called molecular nodes. They are defined to be sets of wires, structures consisting of a labeled arc and the node at its head.

A wire is an ordered pair (r, n), where r is a SNePS relation, and n is a SNePS node. [Shapiro, 1991, page 145]

A guiding force in the theory behind SNePS is the Uniqueness Principle, which states that each concept in the modeled "mind" of a cognitive agent is represented by a unique node in the SNePS network. In SNePS, the FIND/BUILD mechanism creates networks, performing one of two operations when some new concept is submitted to it: (1) if the concept already exists (a node with exactly the right connections is already in the network), then the new information is added to it, or (2) if such a node does not exist, it is created and assigned an unused unique identifier.

4 Circularity through Non-well-founded Sets

Along with the degree and nature of the circularity postulated for SNePS networks, another outstanding question of current SNePS research is the semantics of base nodes, which seem heavy with meaning that doesn't "go" anywhere. Non-well-founded set theory is the stone that can kill both birds. Let the meaning of a node, informally, be the set of meanings of subordinate nodes, except that base nodes have circular meaning, participating also in the meaning of the parent node.

¹ Variable nodes, which represent arbitrary concepts (individuals or propositions) are not in the scope of this work.

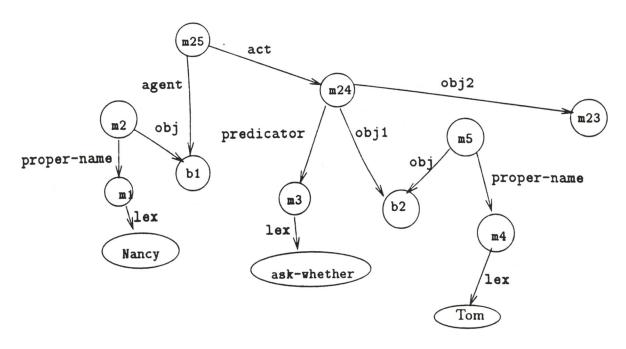


Figure 3: The SNePS network S

Sensory nodes will be the atoms A. To focus on the graphical structure, arc labels will be stripped off for now. Ergo, construed as hereditary sets, base nodes will be members of their parent nodes, and vice versa, which can be shown graphically by an additional edge from a base node to each of its parents. The semantics $\mu(n)$ of a node n will then be the hyperset, over the sensory atoms, that is depicted by the subgraph with n as its point.

An example of a SNePS network (taken from a larger SNePS network) is shown in Figure 3 and called S. S is the SNePS representation of the sentence fragment "Nancy asks Tom whether (m23)"—the rest of the structure, rooted at m23, is omitted. Base nodes include b1, representing the concept of Nancy, and b2, Tom; m25, m2, m24, etc., are molecular nodes; Nancy, ask-whether, and Tom sensory nodes representing the respective written words. The relations are { proper-name, obj, lex, agent, act, predicator, obj1, obj2 }. The new graphical structure that includes circularity is shown as the form called S* in Figure 4.

To formalize the suggested semantic function from nodes to hypersets, we partition the set of nodes in S into the subsets BASE, SENSORY, and two types of molecular nodes—those being treated as atoms due to circumscription of the network, the subset MOLATOM (in the example, $\{m23\}$), and those treated as molecular, MOLFULL (including m1, m2, m25, m24, and so forth). Each sensory node $s \in SENSORY$ is tagged with its lexeme or other sensory datum, and each atomic" molecular node $a \in MOLATOM$ with its label. The decoration and the semantic function μ are formally derived to allow for circularity by making the meaning of a node the hyperset assigned to it, respecting the meanings of surrounding nodes. No semantics is assigned to nodes from SENSORY or MOLATOM, on the principle that they are best regarded only as sources of input (actual or po-

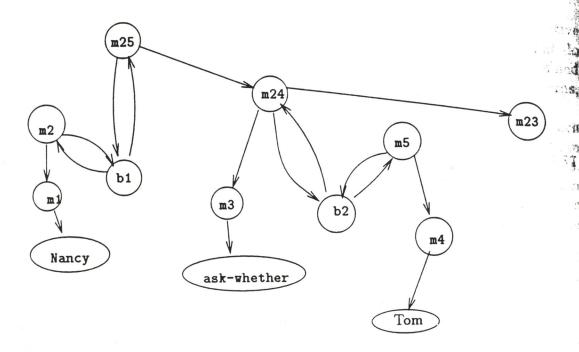


Figure 4: The derived graph S^*

tential), not as meaningful in their own right. (Of course, any molecular node in MOLATOM may be interpreted fully under μ whenever desired.) Under this definition, the semantic value of a base node b is influenced by its parent nodes, as well as vice versa, as desired.

For example, the meanings under μ of the base nodes b1 and b2 in the given S (Figure 3) are shown in Figure 5 as pictures, where the function f turns out to be the hypserset assignments given by μ . We can develop this solution rigorously using the Solution Lemma. Since they comprise the indeterminates,

$$\mathcal{X} = \{b1, b2\},\tag{3}$$

者はなる かき

and

$$A = \{\text{Nancy}, \text{ask-whether}, \text{m23}, \text{Tom}\}. \tag{4}$$

As noted, the node m23 is taken as an atom, rather than as a set with its own subordinates/members, for convenience. The universe of hypersets, then, is

$$V_{\mathcal{A}} = V_{\{\text{Nancy, ask-whether, m23, Tom}\}}$$
 (5)

The solution sought will be an assignment f of sets from $\mathcal{V}_{\mathcal{A}}$ to b1 and b2. The hyperuniverse $\mathcal{V}_{\mathcal{A}'}$ is all hypersets over $\mathcal{A} \cup \mathcal{X}$. We need a system of equations defining b1 and b2, where each set on the right-hand side is in $\mathcal{V}_{\mathcal{A}'}$:

$$b1 = \{m2, m25\} = \{\{\{\{nancy\}, b1\}, \{b1, \{\{ask-whether\}, b2, m23\}\}\}\}$$
 (6)

$$b2 = \{m24, m5\} = \{\{\{ask-whether\}, m23, b2\}, \{b2, \{Tom\}\}\}$$
 (7)

The function μ is an assignment of hypersets to b1 and b2 such the equations that defined them in terms of their membership still hold when those hypersets are

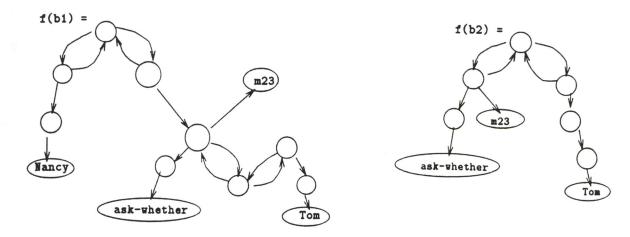


Figure 5: Assignments to b1 and b2 under μ

used in place of the indeterminates b1 and b2. Indeed, that is the case when the hypersets pictured in Figure 5 are substituted for occurrences of the indeterminates in Equations 6 and 7. For the definition of μ as the solution to the system of equations given by the network, the computation of the sets, and verification that they are indeed the solution to the defining equations, see [Hill, 1994]. We have a function μ with the desired scope:

$$\mu : MOLFULL \cup BASE \longrightarrow V_{SENSORY \cup MOLATOM}$$
 (8)

In a full network with no molecular nodes treated as atoms, all molecular and base nodes would be assigned hypersets over sensory nodes only, that is, hypersets from $V_{\rm SENSORY}$. The crucial point is that the semantics given to nodes are hypersets over the atoms, not over other ungrounded hypersets. No references to intermediate nodes, such as m24 or b2, are left.

5 Results and Implications of μ

Results derived from the foregoing bolster the respectability of μ as a semantics for SNePS. For proofs, see [Hill, 1994].

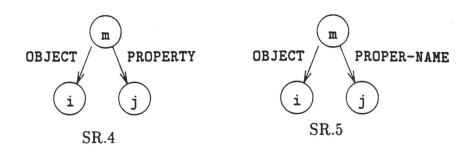
Theorem 1 states that no node is circular to the point of "vacuity".

Theorem 1 \forall nodes $n \in S, \mu(n) \neq \Omega$

Shapiro employs a definition of domination of one node by another that is analogous to hereditary set definition, providing Theorem 2 to the effect that, for the meaning of a molecular node, what you see is what you get. The approved at a node $n \in S^*$ is $\mu(n)$.

Theorem 2 (Node-Picture Principle)

 $\forall m \in MOLFULL, \mu(m) = \{\mu(c) \mid m \text{ dominates } c\}$



An analog to the Uniqueness Principle holds for the semantic function μ as Theorem 3. Nodes with different meanings under μ are different nodes and therefore represent different concepts (except in the case where uniqueness depended on distinction of arc labels).

Theorem 3 (Uniqueness Principle under μ) In any $C' \in SNets'$ derived from a full network C, unless nodes n and m dominated exactly the same subordinate nodes (in which case the arc labels differed),

$$n = m$$
 if and only if $\mu(n) = \mu(m)$

6 A Richer Semantics, with Relation

So far in this development, the arc labels, taken from a set of relations \mathcal{R} supplied by SNePS and by the user, have not been considered, and the semantic function μ has been defined over the skeletal structure of a SNePS network only. What is the role of arc labels, and what exactly is the nature of their "punctuation" function [Shapiro and Rapaport, 1991, pages 221-222]? In [Shapiro and Rapaport, 1987], the two networks given by syntactic rules SR.4 and SR.5 differ only in that one has an arc labeled PROPERTY and the other has an arc in the same position labeled PROPER-NAME; the two networks have different semantics, as expressed in the semantic rules SI.4 and SI.5.

- SI.4 m is the Meinongian objective corresponding to the proposition that i has the property j.
- SI.5 m is the Meinongian objective corresponding to the proposition that Meinongian objectum i's proper name is j. (j is the Meinongian objectum that is i's proper name; its expression in English is represented by a node at the head of a LEX-arc emanating from j.)

If arcs have no semantic import, but the meaning of a node is the entire network in which it is embedded, in what principled way can the "structural" contribution of an arc be distinguished from the "semantic" contribution of a node? If arcs make fixed contributions to the meanings of molecular nodes, they should be involved in the semantic function μ . For an exploration of several alternative solutions, see [Hill, 1994, Chpater 7].

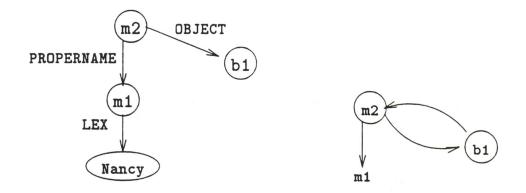


Figure 6: Z

Figure 7: Z^*

Here we present a solution that respects nicely both the definition of a molecular node and the treatment of atoms as the sole grounding of decorations. Recall that a molecular node is a set of a wires. Since wires are ordered pairs (relation, head-node), they can be expressed as sets in the usual way:

$$\langle a, b \rangle = \{\{a\}, \{a, b\}\}\$$

These properties can be used to enhance the original μ semantics to provide a much richer value for the semantics $\mu(n)$ of a node n, which treats the set of relations \mathcal{R} as *atoms*, along with the sensory data.

Let us return to the previous example, S and its S^* of Figures 3 and 4, and (because the computations become quite complex) define an even smaller network context, the very restricted Z, shown in Figure 6, from which Z^* is derived. Instead of the semantically sterile labels for molecular nodes, we use their definitions as sets of wires, which are ordered pairs, and convert the ordered pairs to sets.

$$m1 = \{\langle LEX, Nancy \rangle\}$$
 (9)

$$= \{\{\{LEX\}, \{LEX, Nancy\}\}\}$$

$$(10)$$

$$m2 = \{\langle PROPERNAME, m1 \rangle, \langle OBJECT, b1 \rangle\}$$
 (11)

$$= \{\{\{PROPERNAME\}, \{PROPERNAME, m1\}\}, \{\{OBJECT\}, \{OBJECT, b1\}\}\}\}(12)$$

For this small example, the set of atoms, extended to include the relations that participate in the semantics, is $\mathcal{A} = \{\text{Nancy}, \text{LEX}, \text{PROPERNAME}, \text{OBJECT}\}$. To apply the Solution Lemma to the same task as before—finding assignments to the selected set of indeterminates, $\mathcal{X} = \{\text{m2}, \text{b1}\}$ —we need a system of equations expressing them as hypersets over $\mathcal{A} \cup \mathcal{X}$.

$$m2 = \{\{\{PROPERNAME\}, \{PROPERNAME, \{\{\{LEX\}, \{LEX, Nancy\}\}\}\}\}\}, \\ \{\{OBJECT\}, \{OBJECT, b1\}\}\}\}$$

$$(13)$$

$$b1 = \{m2\} \tag{14}$$

Compare these equations to the system that would have been used under the original μ , where the set of atoms was $A = \{m1\}$:

$$m2 = \{m1, b1\} \tag{15}$$

$$b1 = \{m2\} \tag{16}$$

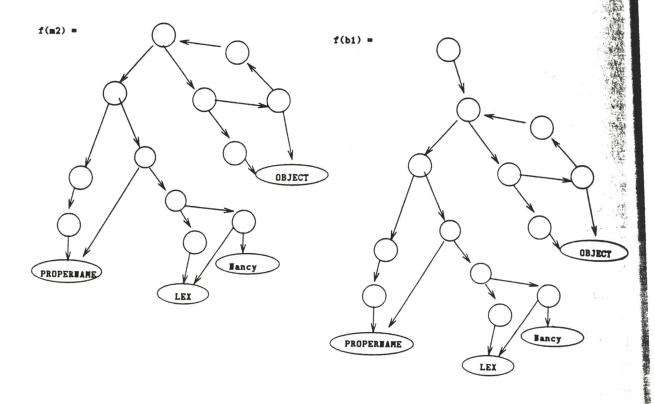


Figure 8: Assignments to b1 and b2 under μ_r

In Equations 13 and 14, m1 no longer exists as an object, having been superseded by its definition as a set of wires.

As the solution f, of course, we want hypersets over A, that is, hypersets from the universe V_A , such that the following relationships are maintained:

$$f(m2) = \{\{\{PROPERNAME\}, \{PROPERNAME, \{\{\{LEX\}, \{LEX, Nancy\}\}\}\}\}, \\ \{\{OBJECT\}, \{OBJECT, f(b1)\}\}\}$$
 (17)
$$f(b1) = \{f(m2)\}$$
 (18)

The solution is shown graphically in Figure 8. Compare the hypersets assigned here (via f, which is the enhanced semantic function μ_r made manifest), to those given under the plain μ , shown in Figure 5. For the definition of μ_r , computation of the solution, and its verification, see [Hill, 1994, Chapter 7].

We have computed the semantics of nodes in a standard SNePS network Z, incorporating rather than neglecting the arc labels (and even allowing for multiple arcs between the same pair of head and tail nodes, a configuration that can exist in SNePS where the arcs are labeled with distinct relations). Given a SNePS network S, the set of molecular nodes MOLFULL is not unstructured primitives, but consists of wires, sets of ordered pairs $\langle r, n \rangle$, and the relations are included in

$$\mu_r : MOLFULL \cup BASE \longrightarrow V_{SENSORY \cup MOLATOM \cup R}$$
 (19)

7 Results and Implications of μ_r

Ty

古地域を通過される

Though complex, μ_r is significant for reasons of the integrity of the semantics. It seems obvious that the two arcs labeled PROPERTY and PROPERNAME in SR.4 and SR.5 above have something to do with the establishment of distinct meanings for their respective dominating molecular nodes. In fact, they could both occur in the same cognitive agent, as discussed above, since the BUILD command of SNEPSUL would not judge them to violate the Uniqueness Principle. In other words, the new version of Theorem 3 would state that $n = m \Leftrightarrow \mu_r(n) = \mu_r(m)$, no longer qualified by the exclusion of the case where n and m dominate exactly the same structure but have different arc labels. Theorem 2 does not hold for μ_r , however, since the SNePS network itself does not show the hyperset structure rooted at nodes if relations from the arc labels are to be atoms along with the sensory nodes.

The treatment given above distinguishes SNePS from other semantic network approaches that have explicitly-named relations between nodes, but no way to build them into nodes themselves at a fundamental level. The definition of the SNePS object "wire" as a node and relation is the key here. (Of course, any semantic network treatment could have such a definition added to it.) On the other hand, the original semantic function μ , which ignores are labels in favor of node identifiers and connectivity, shows what participates in the meaning of a node (that is, what other nodes) without making a commitment as to how, and could be applied (with its handling of circularity) to any graphically structured knowledge/belief representation—even those that do not allow propositions modelled to have multiple arguments in a single position (see [Shapiro, 1991, page 138ff.] for comparison).

8 Contributions to Semantics of Representations

The semantic function μ expresses the meaning of a concept, such as grandmother, in terms not of other concepts, such as grandparents' house, lilac eau de toilette, and blood being thicker than water, but in terms of the discrete sensory stimula involved—the voice, the sight of the house, the scent of lilacs, and the myriad other components of feeling that contribute to the cognitive agent's notion of grandmother. Motivations and principles of SNePS are also supported. The meaning of a node is highly dependent on its location within the surrounding network, rather than on some external property of the concept itself.

We have seen, as an improvement on the basic idea, an enhanced semantics μ_r incorporates the relations used as arc labels into the semantics as atoms, treating them also as primitives. It is also straightforward to define a "measured" μ^{δ} that provides semantics to some degree of elaboration δ , useful for computational and algorithmic analysis [Hill, 1994]. A further use of μ would be to provide the semantics of the entire "mind" of a cognitive agent, the union of meanings of all

nodes constituting the point basis, those nodes from one of which all other nodes are reachable.

Since non-well-founded sets must rely on apgs for their finite depiction, so a knowledge representation that embraces circularity will naturally be graphical, providing an advantage exclusive to graphical representations.

References

- [Aczel, 1988] Peter Aczel. Non-Well-Founded Sets. CSLI Lecture Notes; Number 14. Center for the Study of Language and Information, Stanford, California, 1988.
- [Barwise, 1989] Jon Barwise. The Situation in Logic. Center for the Study of Language and Information, Leland Stanford Junior University; Stanford, California, 1989.
- [Hill, 1994] Robin K. Hill. Issues of semantics in a semantic-network representation of belief. Technical Report 94-11, Department of Computer Science, SUNY at Buffalo, 1994.
- [Nebel, 1991] Bernhard Nebel. Terminological cycles: Semantics and computational properties. In John F. Sowa, editor, Principles of Semantic Networks, chapter 11, pages 331-361. Morgan Kaufmann Publishers, Inc., San Mateo, California, 1991.
- [Shapiro and Rapaport, 1987] S. C. Shapiro and W. J. Rapaport. SNePS considered as a fully intensional propositional semantic network. In N. Cercone and G. McCalla, editors, The Knowledge Frontier, pages 263-315. Springer-Verlag, New York, 1987.
- [Shapiro and Rapaport, 1991] Stuart C. Shapiro and William J. Rapaport. Models and minds: Knowledge representation for natural-language competence. In Robert Cummins and John Pollock, editors, *Philosophy and AI: Essays at the Interface*, pages 215–259. MIT Press, Cambridge, MA, 1991.
- [Shapiro, 1991] Stuart C. Shapiro. Cables, paths and "subconsious" reasoning in propositional semantic networks. In John F. Sowa, editor, Principles of Semantic Networks, chapter 4, pages 137-156. Morgan Kaufmann, San Mateo, CA, 1991.
- [Smith, 1991] Brian Cantwell Smith. The owl and the electric encyclopedia. Artificial Intelligence, 47:251-258, 1991.