Associating Semantics with Graphs

David Israel
Artificial Intelligence Center
SRI International

So, What's Special About Graphs?

- Is there something specially problematic about graphs?
- How about *associating semantics* with:
  - (x)(Fx --> (Ey)(Gy & Rxy))
  - (DEFUN (Factorial *x*) (COND (*x* = 0), 1, (Times *x* (Factorial (*x* -1)))
  - The quick red fox jumped over the lazy brown dog

Or with the analysis (parse) tree for any of the above?

Note: Sure, trees are a special kind of graph, but sufficiently special that we shouldn't kid ourselves into thinking that work on trees is *just* work on graphs.

And What Do We Have in Mind When We Speak of Graphs??

- Graphs, understood abstractly?
  - Two-sorted relational / algebraic structures: Nodes (Vertices, Elements) and Edges (Links, Arrows)
  - Of many, many different kinds

- Graphs, understood, at least in part, essentially *graphically* (= diagrammatically)
  - Topologically
  - Geometrically
  - Geo**metrically**
  - Visually

A Digression on Euclidean Geometry

- Is diagram-based reasoning essential?
- Euclid (and Kant and Frege): YES
- Hilbert (and Pasch, and before them Descartes et al., and after them, almost all of 20th Century Mathematical Logic, and Automated Reasoning -- Chou et al; Zao et al): NO WAY!
- Small body of work on diagrammatic / hybrid reasoning
- Taking diagrams seriously: What general inferences can one draw from a specific individual diagram in a Euclidean proof?
  - Paradoxically, only those that depend on abstract topological features of the diagram, not on geometric features

Categories as Graphs, with Extra Structure

- Category Theory: Abstract algebra of maps between mathematical structures
  - *Graphs can be treated as categories as well*
- In the end, all about existence of and conditions on commuting graphs
- Graphs: Two sorted structure, edges (arrows) and nodes, with two mappings:
  - source (domain) and target (codomain)
- Add for each node A, a distinguished arrow, 1_A: A ---> A
- For any two arrows: f: A -->B ; g: B --> C, an arrow, gf: A --> C
- Governed by the equations:
  - $f1_A = f = 1_Af$
  - $(hg)f = h(gf)$
- Specific kinds of categories make more demands

No problem here about associating semantics with these (highly non-random!) graphs

Caveat Lector/Auditor!

- Model-Theoretic treatments of languages (constructions) are not necessarily contributions to a useful account of *meaning*
- *Model-Theoretic semantics are not necessarily semantics*
- Model-Theoretic treatments are not necessarily/usually/often contributions to semantics at all!

A Little Ancient History

- Semantic Nets (a la Ross Quillian, 1968)
- Intended to model semantic memory … and *inference* (?)
  - Lexical concepts
  - Word senses (word + sense)
  - “representational format [that would] permit the `meanings' of words to be stored, so that humanlike use of these meanings is possible”
- Edges represent the (symmetric?) relation of “being (psychologically) associated with”
  - Edges might be labeled with weights (strengths)
  - Semantic priming phenomena
- As for inference:
  - No dynamics!
  - Semantic priming effects: Is this inference?

A Little Less Ancient History

- Collins & Quillian (1969) stress the hierarchical organization of memory
- With hierarchy comes direction: many edges now understood as expressing class / subclass relation … or concept /subconcept relation
- This latter idea goes over big in AI!
  - Over time much less interest in the psychology of semantic memory
- And the semantics of these trees? Lattices?
- Note (again): yes, lattices can be represented as graphs, but let's not kid ourselves…..

Classical Interpretations

- First-order model theory:
  - a domain **D** (a set)
  - 1-place relation symbols interpreted as *arbitrary* subsets of **D**
  - No further constraints on interpretation of relation symbols
  - So all possible interpretations of the *n* or infinitely many 1-place relation symbols means *all possible independent assignments* of subsets of **D**

Admissible Interpretations

- To capture the idea that some relations between the *meanings* of predicates (= common nouns = nodes in taxonomic hierarchies) are built into (analytic in) the language
  - As opposed to empirical generalizations, expressed, e.g., as axioms in the language.
- Suppose *M(x)* is interpreted as “x is a mammal” and *D(x)* as “x is a dog”. In any admissible I, I(*D*) is a subset of I(*M*).
- Changes the model-theoretic metatheory
- Hence, changes the requirements on sound and complete proof-systems as well

Real Applications of Graph Theory: Probabilistic Graphical Models

- A family of probability distributions defined in terms of a graph -- directed or undirected
- Nodes are identified with random variables
- Joint probability distributions defined by taking products over functions defined on connected subsets of nodes
- By exploiting graph-theoretic representation, such formalisms provide general algorithms for computing marginal and conditional probabilities
- And in providing computational control over complexity of such operations

Simplest Case: Bayes Nets

Directed acyclic graphs

Edges represent evidential dependence

Edges associated with probabilities of a child (target), conditioned on its parents

Markov condition: in DAGs, equivalent to d-separation, a graph-condition relating 3 disjoint sets of nodes. If in G, S_1 is d-separated from S_3 by S_2 then, S_1 is prob. independent of S_3, conditional on S2 in G.

From Bayes Nets to Causal Nets

- Simple!
- Interpret the arrows as causal dependencies and treat the Markov Condition as the Causal Markov condition

A variable X is independent of every other variable (except X's effects) conditional on all of its direct causes.

- – Amounts to assuming that d-separation reflects conditional causal independence
- – Assume also that all common causes of variables are included
  - a kind of closed-world assumption about common causes