

Visualizing Inference

Henry Lieberman and Joe Henke

MIT Media Lab

Cambridge, MA 01239 USA

Abstract

Graphical visualization has demonstrated enormous power in helping people to understand complexity in many branches of science. But, curiously, AI has been slow to pick up on the power of visualization.

Alar is a visualization system intended to help people understand and control symbolic inference. Alar presents dynamically controllable node-and-arc graphs of concepts, and of assertions both supplied to the system and inferred. Alar is useful in quality assurance of knowledge bases (finding false, vague, or misleading statements; or missing assertions). It is also useful in tuning parameters of inference, especially how “liberal vs. conservative” the inference is (trading off the desire to maximize the power of inference versus the risk of making incorrect inferences). We present a typical scenario of using Alar to debug a knowledge base.



Figure 1. An Alar visualization, centered on the assertion “Orange is a food”. Inferred assertions (green) use related knowledge about food to infer new assertions, e.g. “Orange AtLocation grocery store”.

Visualizing concepts and assertions in AI

We present Alar, a visualization system for a large commonsense ontology and knowledge base, ConceptNet, and its associated heuristic inference technique, AnalogySpace [Speer et al 08]. Alar can visualize both graphs of concepts, and also graphs of assertions. Alar is based on display of dynamic node-and-arc graphs, dynamically adjusting using the force-directed layout of the visualization toolkit D3JS [Bostock 14].

In the Concept view, nodes represent Concepts (like elements of an ontology), and links represent similarity between concepts. Link thickness represents the degree of similarity, and lines exert a proportional spring-like “force” in the dynamic graph pulling its nodes closer (working against a repelling force that spaces out the nodes). Concepts with similar meanings will be seen to cluster together. Words with, say, more than one meaning, will find themselves pulled between clusters that represent each of their meaning contexts.

In the Assertion view, nodes represent assertions, a triple of Concept-Relation-Concept (e.g. “Fork UsedFor Eating”). Links represent similarity of assertions (not necessarily that one assertion logically implies another, although for a link between an assertion in the knowledge base (black) and an inferred assertion (green), it is usually the case that it exerts a strong influence. The size of the dot indicates its truth value.

We believe the visualization of assertion graphs to be particularly novel. Visualization of related sets of assertions can be a powerful tool in debugging knowledge bases and inference. The assumption is that the inference space has a kind of smoothness characteristic – inference about similar concepts should have similar truth values. When incorrect assertions are inferred, they are often connected to incorrect assertions in the knowledge base, so can be easily spotted. Discontinuities in the assertion space can also be a clue that some important knowledge is missing from the knowledge base. Finally, incorrect assertions can also appear when inference is “too liberal” – it concluded something without sufficient evidence.

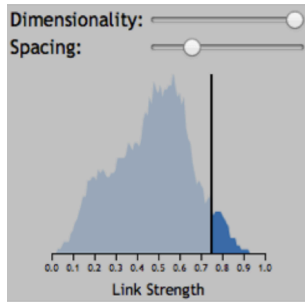


Figure 2. Interactive control over permissiveness of the inference, layout, and level of detail

Similarity of concepts and assertions are computed using the analogical heuristic reasoning technique AnalogySpace [Speer et al 08]. It works by making a matrix of concepts vs. “features” (relation + concept), and taking the principal components of this matrix. Such components often represent important semantic distinctions, such as “good vs. bad”. A previous visualization for ConceptNet displayed these components as axes in a multidimensional space [Speer et al 10]. Alar’s visualization technique should be applicable to other systems based on concepts and inferred assertions, and that have a “liberal vs. conservative” control parameter.

Debugging inference with Alar

There are three interactive controls over the visualization, shown in Figure 2. First, *dimensionality*, which controls how “liberal” or “conservative” the inference is. For concepts, liberal inference results in more similarity links; for assertions, more inferences. *Spacing* supplies “negative gravity” making semantic clusters more readable. The *link*



Figure 3. An assertion network showing that “Cup IsA Drink” was inferred from a bogus assertion in the KB, “Glass IsA Drink”.

strength is a movable slider on a histogram of number of links vs. strength. Only links to the right of the slider are displayed, giving control over the level of detail. The interface is seeded with one or more initial concepts (e.g. “Orange”) or assertions (“Orange is a food”). An operation, “Add Related Nodes” finds the most similar concepts (or assertions) to the seeds and expands the graph.

Figure 3 shows a typical situation where there is a bug in the knowledge. The incorrect assertions, “Cup IsA Drink” and “Bottle IsA Drink” were inferred from the incorrect KB assertion, “Glass IsA Drink” (probably a failure of our parser on something like, “A glass of water is a drink”). Assertions around it like “Wine IsA Drink”, “Beer...” etc. are unaffected.

Related Work

There is surprisingly little work on visualization of inference in AI. Visualization tools are often used on concept ontologies [Protégé 14], [Katifori 07] (but not on associated inference of assertions). Specialized inference algorithms provide visualization of their internal data structures [Cossalter et al 11], and/or sets of examples [Amershi et al 11], usually unique to that algorithm. Logical systems visualize proof trees [Eisenstadt et al 91], but not large sets of related assertions that are not part of the same inference chain.

We hope that this work will spark more investigation of the use of creative and beautiful visualizations to help people understand what AI programs are doing.

References

Amershi, S., Fogarty, J., Kapoor, A., and Tan, D.(2011) Effective End-User Interaction with Machine Learning. In AAAI Conference on Artificial Intelligence (AAAI 2011), pp. 1529-1532.

Bostock, M. Data Driven Documents, D3JS. <http://www.d3js.org>, February 2014.

Cossalter, Michelle, Ole J. Mengshoel, Ted Selker, Visualizing and Understanding Large-Scale Bayesian Networks, AAAI Workshop on Scalable Integration of Analytics and Visualization, San Francisco, California, USA, 2011.

Eisenstadt, Marc, Mike Brayshaw, Jocelyn Paine, The Transparent Prolog Machine: Visualizing Logic Problems, Springer, 1991

Katifori, A., Halatsis, C., Lepouras, G., Vassilakis, C., and Gianopoulos, E. 2007. Ontology visualization methods—A survey. ACM Comput. Surv. 39 , 4, Article 10 (October 2007).

Protégé, accessed February 2014.

<http://protegewiki.stanford.edu/wiki/Visualization>,

Speer, Robert, Catherine Havasi and Henry Lieberman, AnalogySpace: Reducing the Dimensionality of Commonsense Knowledge, AAAI-08, Chicago, July 2008.