**Presentation Paper Title**: *Designing metaschemas for the UMLS enriched semantic network*
**Date of Presentation**: 02 February 2004
**Full Citation**: Li Zhang, Yehoshua Perl, Michael Halper, and James Gellera (in press). Designing metaschemas for the UMLS enriched semantic network. *Journal of Biomedical Informatics*.
**Keywords**:  Unified Medical Language System (UMLS); Metathesaurus; Semantic Network (SN); Enriched semantic network (ESN); Directed Acyclic Graph (DAG); Schema; Metaschema; Multiple inheritance; Visualization
**Slides**: http://ils.unc.edu/bioinfo/docs/20040202-Herron.ppt

**Summary of Presentation:** *Zhang et al.* was presented starting with an overview of the Unified Medical Language System, the Metathesaurus, metaschema, the Semantic Network, and the proposed Enriched Semantic Network. Special care was taken to explain what a DAG is and why it helps to give the ESN some advantages over the SN.  A definition of metaschema was introduced during the presentation followed by a brief exploration of both the Q-metaschema and C-metaschema described in the current paper.  The paper was then evaluated as a matter of class discussion; some fundamental questions relating to the current paper were explored pertaining to the utility of metaschemas.

The Unified Medical Language System was created by the National Library of Medicine in 1986 in order to provide an ontology for medical texts that allows for resolving terminological differences between publications.  The UMLS organizes terms by concept.  The UMLS terminologies are represented in the *Metathesaurus*, one of three fundamental knowledge sources for the UMLS.  Another knowledge source, the *Semantic Network*, provides a generalized hierarchical (double tree with root nodes "event" and "entity") representation of the contents of the Metathesaurus.   The ESN is a proposed re-representation of the Semantic Network.  Instead of the SN's strictly hierarchical double tree representation, the ESN uses a directed acyclic graph that allows for multiple inheritance of semantic information.  The ESN representation is essentially a network representation of the semantics of the Metathesaurus.

The SN and the ESN are representations designed for machines.  They are highly complex and thereby difficult for people to use as a tool for manually exploring semantic relationships between medical sources.  A metaschema eases human exploration of the contents of the ESN; for present purposes a metaschema is an abstraction, a simplified representation, of the ESN.  Practically speaking, a person would use a metaschema visually.  Both the Q- and the C-metaschemas successfully reduce the complexity of the ESN.  While the Q-metaschema is more compact than the C-metaschema, the C-metaschema represents the ESN as a DAG and is therefore more "true" to the structure of the DAG-based ESN.

*Zhang et al.* adequately demonstrated methods for constructing two different types of metaschemas in order to ease human comprehension of the ESN.  It is clear from the paper that the Q- and C-metaschemas provide means for visually representing complex ESN information, and that they may be useful for people working with the ESN to eliminate redundancies in the representation.

It is clear that the ESN provides a significant advantage over the SN because, unlike the SN, it allows for multiple semantic inheritance.  However *Zhang et al.* does not seem to make clear whether the C-metaschema lends itself to being constructed algorithmically.  It does seem transparent that the Q-metaschema does not lend itself well to automated construction because its rule set is essentially qualitative.  Given that the ESN is dynamic, if a metaschema cannot be constructed in an automated fashion, then the metaschema may be difficult to maintain. Building a metaschema representation may be a cumbersome process if done manually; such metaschemas may require a great deal of manual labor to construct if they cannot be constructed in an automated fashion.  Given any change in the ESN, how much of a metaschema would have to be reconstructed?  Do we just alter a subsection of the metaschema, or does the metaschema need to be reconstructed from the ground up?  Finally, the paper does not explore the difference between auditing the ESN and eliminating redundancies in the ESN; the paper left us wondering whether they were the same thing.

While it seems unlikely that audits can be performed automatically, it does seem that a metaschema at the very least is a useful tool for manually eliminating redundancies.  However, whether the ESN will be adopted by the National Library of Medicine is unknown.  Without adoption of the ESN, the present paper lacks a certain practical utility.