

Project title

Interface for Web-like Databases

Supervisor

Vladimir Sazonov, Logic and Computations Group,
<http://www.csc.liv.ac.uk/~sazonov/>

Brief description

The goal of this Project is *to implement interface convenient for browsing and observing semistructured or Web-like Databases (WDB)*. For a student working on this project this would be a good and challenging opportunity to apply his/her programming skills and to learn new ideas on unstructured databases — a fresh and promising direction of research.

More detailed description (cf. also [2, 1])

A Web-like or semistructured database under hyperset approach is a generalisation of the ordinary approach to relational databases. Any data under this approach is considered as a finite set of labelled elements

$\{label_1 : x_1, \dots, label_n : x_n\}$, for example,
 $\{\text{Student} : s, \text{Department} : d, \text{StartAndEndOfStudy} : t\}$.

The order and repetition of the labelled elements $label_i : x_i$ does not matter. Labels also correspond to *attributes* of relational databases (actually, strings of symbols — not sets). Labelled elements x_1, \dots, x_n or s, d, t are themselves complex data (sets). Say,

$s = \{\text{Name} : n, \text{Birthdate} : b, \text{Address} : a\}$.

For uniformity, *atomic data* are also represented as a labelled *empty set* $label : \emptyset$ where $\emptyset = \{\}$. We also abbreviate $label : \emptyset$ as $label$. Then, for example,

$n = \{\text{FirstName} : \text{David}, \text{LastName} : \text{Beckham}\}$.

Admitting an analogy with WWW, a set is an analogue of a *Web page*, labelled elements of a set are *hyperlinks* (labels can be ‘clicked’ to ‘download’ corresponding Web pages—sets). Then, as in WWW, cyclic (hyper) sets like $\Omega = \{\Omega\}$ are allowed. We assume here an ‘empty’ label \square before the element Ω , which is usually omitted. *Pure* hypersets are those which contain no labels at any depth (or, equivalently, only the empty label \square).

Any *database state* (i.e., a hyperset) is represented in a computer as a system of equations, as above or as a directed edge labelled graph. Thus, s considered as a vertex of the graph has three outgoing labelled edges: $s \xrightarrow{\text{Name}} n$, $s \xrightarrow{\text{BirthDate}} b$, and $s \xrightarrow{\text{Address}} a$. Vertices also correspond to URLs of a Web page, and the labelled sets, such as $\{label_1 : x_1, \dots, label_n : x_n\}$, correspond to HTML files. In a browser we see only (clickable) labels $label_1, \dots, label_n$ with corresponding URLs x_1, \dots, x_n *hidden*.

In the case of WDB (not WWW), such kind of an interface is too primitive (like observing only the attributes of a relational table, instead of the whole table). We would like to see a deeper structure than only $label_1, \dots, label_n$, like in the ordinary relational databases. There could be alternative ways to do that, taking into account that WDB has no rigid structure as Relational DB.

Thus, the essence of the Project consists, in

1. Choosing a way how a WDB will be presented in a computer (or in many computers — distributed case like WWW).
2. (the main clause) Inventing and implementing a possibility of visualising WDB by creating a suitable interface.
 - This may be based, for example, on a transformation of any given system of set equations into a graph represented in a flexible way, convenient to the user.

- Vice versa, there should be a possibility to update this visualised graph with automatic updating the system of set equations.

3. Inhabiting an interesting WDB.

4. Demonstrating that this interface is working on this WDB.

If done well, the resulting system could be used as demonstration tool in the module COMP311 “Semistructured or Web-like databases”.

Background requirements

Familiarity with traditional approaches to Relational Databases, elementary set-theoretic concepts, and having sufficiently advanced programming skills. It is very desirable to attend the Module COMP311 especially devoted to semistructured or Web-like databases.

References

- [1] Lisitsa, A., and Sazonov, V., Bounded Hyper-set Theory and Web-like Data Bases. *Computational Logic and Proof Theory, 5th Kurt Gödel Colloquium, KGC'97*, Springer LNCS Vol. 1289, 1997, pp. 172–185.
- [2] Sazonov, V.Yu.: 1993, ‘Hereditarily-finite sets, data bases and polynomial-time computability’, *Theoretical Computer Science* **Vol. 119**, Elsevier, pp. 187–214.
- [3] V.Yu.Sazonov, Querying Hyperset/Web-Like Databases, *Logic Journal of IGPL*, 2006; 14(5): 785-814. doi:10.1093/jigpal/jzl010