mizar-items: Exploring fine-grained dependencies in the Mizar Mathematical Library

The Mizar Mathematical Library

The Mizar Mathematical Library (MML) is one of the largest collection of formalized mathematical knowledge that has been developed with various interactive proof assistants. It comprises more than 1100 "articles" summing to nearly 2.5 million lines of text, each consisting of a unified collection of mathematical definitions and proofs. Semantically, it contains more than 50000 theorems and more than 10000 definitions expressed using more than 7000 symbols. It thus offers a fascinating corpus on which one could carry out a number of experiments. Our system, mizar-items, computed **fine-grained dependencies** among the contents of the MML. For an overview of Mizar, see (Grabowski et al., 2010); for a discussion of some successful initial experiments carried out with the help of mizar-items, see (Alama et al., b, a).

Dependence analysis

We say that a definition, or a theorem, ϕ **depends** on some definition, lemma or other theorem ψ, (or equivalently, that ψ is a **dependency** of ϕ) if ϕ "needs" ψ to exist or hold. The main way such a "need" arises is that the well-formedness or the justification of provability does not hold in the absence of ψ. We are interested in computing what minimally accounts for the success of a **specific** mathematical proof that has been formalized in the Mizar language.

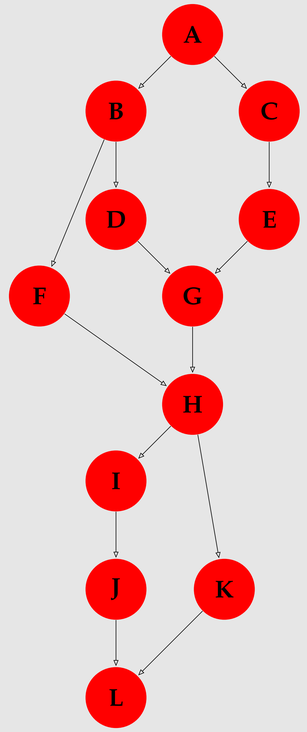
The ultimate aim of mizar-items is to facilitate the foundational study of those items of the MML that have substantial mathematical or foundational value, such as the Jordan curve theorem, the axiom of choice, the existence of strongly inaccessible cardinal numbers, or Euler's polyhedron formula. The fine-grained dependency data exposed by mizar-items could also be used in theory exploration and reverse mathematics (Simpson, 2009) or Lakatos-style (Lakatos, 1976) investigations of necessary and sufficient conditions for mathematical theorems.

Figure 1 illustrates that kind of analysis that mizar-items facilitates: it shows the precise sense in which the axiom of infinity in Tarski-Grothendieck set theory (TG), on which the MML is based, depends on the axiom of foundation (which is a proper axiom of TG—the axiom of infinity is a theorem of TG, unlike in ZF).

Site

mizar-items is accompanied by a website, http://mizar.cs.ualberta.ca/mizar-items/, for exploring these dependencies. With the site one can view any particular Mizar item and see precisely what it depends upon (and what depends on the item). At the site one can ask such queries as:

- **Is there a path between two given items?**
- **Do all paths from one item to another pass through a given intermediate node?**
- **Are there any paths between two given items that do not pass through a given node?**



Item	Meaning
A	The axiom of infinity: ω exists and is unique
B	Every ordinal belongs to a limit ordinal
C	Every non-empty class of ordinals has a least member
D	Limit ordinals are those ordinals closed under ordinal successor
E	Every non-empty set included in an ordinal has a $\subseteq$-minimal element
F	X is an ordinal if $\forall x \in X$, x is an ordinal and $x \subseteq X$
G	For ordinals A and B, $A \subseteq B$ iff every ordinal in A is in B
H	For ordinals A and B, either $A \in B$, $A = B$, or $B \in A$
I	Members of ordinals are themselves ordinals
J	For all sets X, Y, and Z, either $X \notin Y$, $Y \notin Z$, or $Z \notin X$
K	If x is an $\in$-transitive proper subset of an ordinal A, then $x \in A$.
L	Every non-empty set has a member with which it is disjoint

Figure 1 A path of dependence in the MML

To facilitate exploration, one can start by visiting a list of selected interesting entry points into the vast Mizar library. Our site tracks, for example, F. Wiedijk's famous "100 Theorems" site.

Literature

1 Grabowski, A., Kornilowicz, A. and Naumowicz, A. (2010). Mizar in a nutshell. *Journal of Formalized Reasoning*, *3*(2), 153–245.

2 Alama, J., Kuehlwein, D., Tsivtsivadze, E., Urban, J. and Heskes, T. Premise selection for mathematics by corpus analysis and kernel methods. . Submitted.

3 Alama, J., Mamane, L. and Urban, J. Dependencies in formal mathematics. . Submitted.

4 Alama, J. Generalizing theorems of the Mizar mathematical library by type promotion and property omission. To be presented at *ITP 2011*.

5 Simpson, S. G. (2009). *Subsystems of Second Order Arithmetic*. 2nd edition Springer.

6 Lakatos, I. (1976). *Proofs and Refutations*. Cambridge University Press.