

Enhancing MML via Isomichi's Classification of Domains

Adam Grabowski

Magdalena Jastrzębska

Institute of Mathematics, University of Białystok
ul. Akademicka 2, 15-267 Białystok, Poland
adam@math.uwb.edu.pl, magja@osiedle.net.pl

Abstract – The purpose of the work described in this paper is to test the formalization of the Kuratowski closure-complement problem already solved in the Mizar language as well as to further develop a selected discipline formalized in the Mizar Mathematical Library. We aimed at the possibly faithful encoding of a regular mathematical paper while at the same time reflecting some features of the language chosen for this translation. As a result, we also obtained better granulation of the knowledge stored in the MML.

Keywords – general topology, closure operator, topological domain, subcondensed set, supercondensed set.

1. Introduction

The Mizar system is developed as a proof-assistant for mathematicians and computer scientists. Researchers can benefit from this assistance at all stages of their work and computer-managed knowledge can be browsed and learnt much easier. Also new results in the field can be encoded within a repository as long as it is developed well enough and provides proper mathematical background.

Roughly speaking, there are two ways to test the strength of the repository. One of them is to encode a chosen textbook completely as in the case of the project to formalize *Compendium of Continuous Lattices* [4] in Mizar. The other is to choose a set of papers on a topic and obtain their full encoding. Our primary aim was to completely formalize a paper published in a renowned mathematical journal, namely *Pacific Journal of Mathematics*. Earlier codifications of various papers in the MML were incomplete and we can imagine the cause of this to be because authors chose the easier parts of mathematical papers to be translated and left those sections which cannot be fairly reflected in the repository language due to the lack of necessary auxiliary facts.

The choice of Y. Isomichi's work in general topology [7] is by no means accidental – some Mizar articles by Z. Karno [9] also used some ideas mentioned in the paper. Topology is also one of the better developed disciplines in the MML. It is also very close to the variant of Zermelo-Fraenkel set theory on which the library is based.

The formalization of a paper is also less work-consuming and it can be done by a narrower group of people because of the smaller number of interdependencies. Usually however papers are self-contained and much more compact, so they are in our opinion a better testbed for proof-assistants than handbooks.

The paper is organized as follows. In the following section, we present some basic topological notions needed to understand the whole problem. Section 3 mentions the formalization of general topology in the Mizar Mathematical Library. In the following sections, we discuss the language features which made our work more readable and most of them are illustrated by examples taken from our Mizar article [8]. A description of activities aiming at the enhancement of the MML can be found in Section 6, while at the end we draw some concluding remarks.

2. Regular Subsets

There are two main approaches to the fundamental notion of topology considered in the literature. One of them, by Kuratowski, is based on four axioms of a closure operator acting from 2^X into 2^X defined as follows:

$$\begin{aligned}\overline{\emptyset} &= \emptyset \\ A &\subseteq \overline{A} \\ \overline{\overline{A}} &= \overline{A} \\ \overline{(A \cup B)} &= \overline{A} \cup \overline{B}\end{aligned}$$

where A and B are arbitrary subsets of a space X . Closed sets are the fixed point of the above operator.

Another definition of the closure of a subset A of a topological space $(X, \mathcal{T})$ is as follows:

$$p \in \overline{A} \iff \forall_{G \in \mathcal{T}} (p \in G \Rightarrow A \cap G \neq \emptyset)$$

(here, $\mathcal{T}$ is a family of open subsets of X , i.e., a family which contains $\emptyset$ and X and is closed for arbitrary unions and binary intersections.)

Furthermore, in terms of closure and set-theoretical complement we can define the interior of a subset A of a space $(X, \mathcal{T})$ as

$$\text{Int } A = X \setminus \overline{X \setminus A}.$$

Kuratowski poses a question in [10] of how many distinct subsets can be obtained from a single subset applying closure and interior operators (closure and complement, originally) and the answer is 7 (resp., 14). (see [5] for a discussion of this problem and [1] for its Mizar encoding.)

In looking for a concrete example of a set from which exactly fourteen distinct subsets can be obtained, we can try to formulate conditions under which such objects may be distinct.

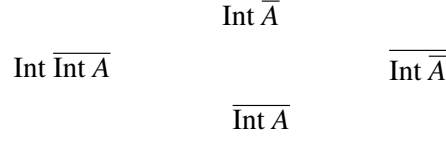


Figure 1. Diagram of a closure-interior diamond.

The arrows in Fig. 1 denote set-theoretical inclusions between the objects.

The following equalities characterize well-known important classes of subsets (by Kuratowski):

- open domains (also called regular open subsets): $A = \text{Int } \overline{A}$
- closed domains (or regular closed subsets): $A = \overline{\text{Int } A}$
- domains: $\text{Int } A \subseteq A \subseteq \overline{\text{Int } A}$

Isomichi's work introduces another classification of subsets with respect to the shape of this part of the diagram. As can be easily observed, if the inclusion relation holds between $\text{Int } \overline{A}$ and $\overline{\text{Int } A}$, this diamond can flatten:

- supercondensed sets: $\text{Int } \overline{A} = \overline{\text{Int } A}$
- subcondensed sets: $\overline{A} = \overline{\text{Int } A}$
- condensed sets: $\text{Int } \overline{A} = \overline{\text{Int } A} \wedge \overline{A} = \overline{\text{Int } A}$

Isomichi divides all subsets into three classes:

- I $\text{Int } \overline{A} \subseteq \overline{\text{Int } A}$
- II $\overline{\text{Int } A} \subseteq \text{Int } \overline{A}$
- III $\text{Int } \overline{A}$ and $\overline{\text{Int } A}$ are $\subseteq$ -incomparable.

Clearly, the classes are disjoint and their union is the whole 2^X .

3. General Topology Codified

Following successful (but still unfinished) work on the encoding of [4] (57 Mizar articles in total written by 16 authors), the Białystok team started formalizing Engelking's handbook on general topology [3]. An article series with MML identifiers `TOPGEN` [6], which includes five articles at the moment and another four being prepared, aims for a complete encoding of this book. Japanese MML developers work on an alternative approach for finite topological spaces (`FINTOPO` series). Formal topology is also one of the key points in the TYPES FP6 grant sponsored by the European Union. One of the recently emerged subfields of research in MML was algebraic topology which allowed us to increase the speed of codification of the proof of the Jordan Curve Theorem.

Topological spaces are defined as structures with descendants of `1-sorted`, with two fields, the second – `topology` – is the family of subsets of its carrier, with an additional adjective `TopSpace-like` reflecting the sense of the family of open subsets as described in Section 2. Also the definitions of closure `Cl` and interior `Int` are very close to their informal originals contained in [12] and [13].

4. The Classification

The core of Isomichi's paper is the classification of subsets of a given topological space into three disjoint classes. In the Mizar language it was encoded as follows (we quote here only one of the definitions):

```
definition let T be TopSpace, A be Subset of T;
  attr A is 1st_class means
:: ISOMICHI:def 6
  Int Cl A c= Cl Int A;
end;
```

Obviously, this class is non empty because all super- and subcondensed sets are its elements. This is understood automatically by the checker via the conditional cluster registrations:

```
registration let T be TopSpace;
  cluster supercondensed -> 1st_class Subset of T;
  cluster subcondensed -> 1st_class Subset of T;
end;
```

Examples are very important in proving the correctness of existential registrations of clusters. Without this, all mathematics encoded in Mizar would not fully reflect reality and clusters would not work properly. To assure that we do not consider the same empty class, we can take into account the empty set which is both super- and subcondensed.

On the other hand, the existence of the elements of the second class is a rather non-trivial fact (the space T is not simple TopSpace):

```
registration let T be anti-discrete non trivial TopSpace;
  cluster 2nd_class Subset of T;
end;
```

Obviously, we did not prove the existence of elements of the third Isomichi's class and it was not done by the author either (since it is just false for arbitrary topological spaces). To complete the formalization we decided to make a workaround. First, we defined an attribute for a topological space which can be read as "containing subsets belonging to the class III":

```
definition let T be TopSpace;
  attr T is with_3rd_class_subsets means
:: ISOMICHI:def 11
  ex A being Subset of T st A is 3rd_class;
end;
```

As an example we took a subset $A = (-\infty, 1) \cup \mathbb{Q}(2, 4)$ of a real line with natural topology ($\mathbb{R}^1$ in MML).

```
registration let T be with_3rd_class_subsets TopSpace;
  cluster 3rd_class Subset of T;
end;
```

After this registration, we can write the functorial cluster which together with two similar ones for class I and II can be understood as one of the basic properties of classes: *Both an arbitrary subset and its complement belong to the same class.*

```
registration let T be with_3rd_class_subsets TopSpace;
              let A be 3rd_class Subset of T;
              cluster A' -> 3rd_class;
end;
```

Our formalization took 1451 lines of Mizar code (1325 excluding blank lines and comments), which comparing with eleven pages of informal original text make the de Bruijn factor much below 2.

5. Mizar Language Features and Formalization Problems

The Mizar language is very close to the mathematical vernacular (as considered by de Bruijn) and is relatively well-readable. The resulting database is available to mathematicians with no additional difficulties. Although the library is not devoted to the formalization of a specific discipline, especially general topology can be fairly reflected there.

5.1 Features

Some language features which can be of special interest for the working mathematician were also used during Isomichi's formalization:

Redefinitions. Used if an author prefers an alternative definiens over the traditional one. This is the case in the definition of boundary which is already introduced in MML, but with another definiens.

```
definition let T be TopStruct, A be Subset of T;
  redefine func Fr A equals
  :: ISOMICHI: def 4
    Cl A \ Int A;
end;
```

Obviously, compatibility with the original has to be shown. In our case, we had to prove that $Cl R \setminus Cl R' = Cl A \setminus Int A$.

Synonyms. This is a kind of syntactic sugar. Using this construction we can write `Bound A` instead of `Fr A` as Isomichi did. As it appeared to be very readable when solving the Kuratowski 14 set problem, we also introduced postfix notation for closure in this way (`A-` instead of `Cl A`).

```
notation let T be TopStruct, A be Subset of T;
  synonym A is regular_open for A is open_condensed;
  synonym A is regular_closed for A is closed_condensed;
end;
```

Also notions of regular open and regular closed subsets were in fact just open and closed condensed sets already introduced in Mizar.

Attributes instead of classes. This substantially changed the quality of our work by turning on many automatizations provided by the Mizar type checker. Specifically, we could use all three types of cluster registrations: existential, conditional, and functorial. Frankly speaking, the hardest proofs were necessary when proving correctness of the first type. We would very much like to have this useful and powerful mechanism of adjectives work also for sentences of the type “ A is open” and have additional attributes added without a type changing statement (via `reconsider`).

5.2 Problems

The main drawback of the formalization was the need to construct concrete examples; sometimes this work was way above the level of the rest of the effort. From Isomichi’s paper, only Remark 2 was not formalized. It mentions T_0 topological spaces as an interpretation of the results and is given without the proof.

Also many of the theorems from the paper were accepted by the checker without any proof after registering proper clusters. Such corollaries, if left in the source, can be detected and removed automatically by auxiliary MML software. In such cases, an appropriate note should be provided by the author.

6. Library Revisions

The creation of a large library at the hands of many authors is a great challenge. At the same time however, it could be hard to achieve a sufficient level of homogeneity.

As of the time of this writing (March 2006), the Mizar Mathematical Library available for downloading together with the Mizar system is 4.60.938. The first number originally denoting the version of the syntax of the library files is now kept for technical reasons and is to change when the number of revisions will be too large.

The Library Committee of the Association of Mizar Users accepts new articles and also enhances old ones. In fact, both activities are closely connected because many repetitions can be removed on the author’s side with his strong support. In such a manner, an author can keep an eye on how his article evolves and at the same time learn how to achieve better quality in his future submissions. Otherwise the changes are done independently; all major revisions are numbered (the second number of the MML version).

The Committee claims the policy of the last number to be equal to the number of `.miz` files in the distribution, in the `mml` subdirectory. Some of the articles are deleted from the library when the net of structures is tightened (e.g., functions and relations originally developed independently from scratch, but now functions are relations with an additional property).

Instead of `open condensed` and `closed condensed` adjectives we would like to have super- and subcondensed sets, so a revision will be necessary to make the apparatus of clusters work in their full strength. This will certainly result in a weakening of the assumptions of many theorems.

7. Conclusion and Further Work

The main achievement of this work is the full encoding in the Mizar language of the regular mathematical paper [7] in [8].

With the completion of the formalization of the proof of the Jordan Curve Theorem (in September 2005, as a result of a long-lasting cooperation between Shinshu University, Nagano, Japan and University of Białystok, Poland), we can concentrate on proof restructuring and enhancement processes to make them more understandable for humans. We believe that our work can be a step towards this direction because as of now, we can hardly measure what was really needed to complete the proof of JCT.

Although at first sight the paper seems to be of rather minor interest (as may be concluded by using any of the popular Internet searching services) and not so fresh, some parallel works can be observed recently.

The work can be applied to the generalization of Kuratowski's 14 problem as well as to the improvement of clustering and reusability of notions collected in the library due to the extensive use of attributes. Similar policy may be observed recently – the Library Committee revises the MML to replace types like “Element of COMPLEX” by “complex number” (similarly with “real numbers”).

A more thorough revision of the repository reusing our work is not being done at the moment because we would like to keep the translation from the natural language in a single Mizar article, so a reasonable solution would be to place it somewhere earlier in the list of MML articles than it is now. However, because of examples which use rather advanced techniques and hence relatively late articles, this cannot be done easily without moving other items of the library.

Following the work described in this paper, we would also like to translate completely into Mizar a representative collection of other papers in topology and to reach research frontier in the nearest future.

References

- [1] L. K. Bagińska and A. Grabowski, *On the Kuratowski Closure-Complement Problem*, Formalized Mathematics, Vol. 11, No. 3, 2003, pp. 323–329, MML Id: KURATO_1.
- [2] P. Cairns, J. Gow, and P. Collins, *On dynamically presenting a topology course*, Annals of Mathematics and Artificial Intelligence, 38(1–3), 2003, pp. 91–104.
- [3] R. Engelking, *General Topology*, PWN – Polish Scientific Publishers, Warsaw, 1977.
- [4] G. Gierz, K.H. Hofmann, K. Keimel et al, *A Compendium of Continuous Lattices*, Springer, 1980.
- [5] A. Grabowski, *On the computer-checked solution of the Kuratowski closure-complement problem*, Mechanized Mathematics and Its Applications, Vol. 4, No. 1, 2005, pp. 25–34.
- [6] A. Grabowski, *On the boundary and derivative of a set*, Formalized Mathematics, Vol. 13, No. 1, 2005, pp. 139–146, MML Id: TOPGEN_1.
- [7] Y. Isomichi, *New concepts in the theory of topological space – supercondensed set, subcondensed set, and condensed set*, Pacific Journal of Mathematics, Vol. 38, No. 3, 1971, pp. 657–668.
- [8] M. Jastrzębska and A. Grabowski, *The properties of supercondensed sets, subcondensed sets and condensed sets*, Formalized Mathematics, Vol. 13, No. 2, 2005, pp. 353–359, MML Id: ISOMICHI.
- [9] Z. Karno, *The lattice of domains of an extremally disconnected space*, Formalized Mathematics, Vol. 3, No. 2, 1992, pp. 143–149.
- [10] K. Kuratowski, *Sur l'opération $\bar{A}$ de l'Analysis Situs*, Fundamenta Mathematicae, 3, 1922, pp. 182–199.

- [11] Y. Nakamura and G. Takeuchi, *On the Jordan curve theorem*, Technical Report, Shinshu University, Nagano, Japan, 1980.
- [12] B. Padlewska and A. Darmochwał, *Topological Spaces and Continuous Functions*, Formalized Mathematics, Vol. 1, No. 1, 1990, pp. 223–230.
- [13] M. Wysocki and A. Darmochwał, *Subsets of topological spaces*, Formalized Mathematics, Vol. 1, No. 1, 1990, pp. 231–237.