

Tarski's Geometry Modelled in Mizar Computerized Proof Assistant

Adam Grabowski

Institute of Informatics

Department of Mathematics and Informatics

University of Białystok,

ul. Konstantego Ciołkowskiego 1 M, 15-245 Białystok, Poland

Email: adam@math.uwb.edu.pl

Abstract—In the paper, we discuss the formal approach to Tarski geometry axioms modelled with the help of the Mizar computerized proof assistant system. Although our basic development was inspired by Julien Narboux's Coq pseudo-code and is dated back to 2014, there are significant steps in the formalization of geometry done in the last decade of the previous century. Taking this into account, we will propose the reuse of existing results within this new framework (including Hilbert's axiomatic approach), with the ultimate future goal to encode the textbook *Metamathematische Methoden in der Geometrie* by Schwabhäuser, Szmielew and Tarski. We try however to go much further from the use of simple predicates in the direction of the use of structures with their inheritance, attributes as a tool of more human-friendly namespaces for axioms, and registrations of clusters to obtain more automation (with the possible use of external equational theorem provers like Otter/Prover9).

I. INTRODUCTION

FOR YEARS, foundations of geometry attracted a lot of interest of researchers from various areas of mathematics. From the very beginnings, human thought was stimulated by geometrical objects, to take Thales as the prominent example of an ancient philosopher. Classical geometry involved illustrative examples and construction problems instead of a building strong axiomatic basis. However, from the modern viewpoint of automated theorem-provers, diagrams can deliver some really tough problems. Here an important example is the possibility of ruler-and-compass construction: impossibility of trisecting the angle and doubling the cube (as two out of four problems of antiquity), where the treatment of constructible numbers is way more efficient from the formal point of view.

Euclid and his *Elements* are often recalled as one of the first successful uses of an axiomatic method in mathematics, and such an approach can be formalized efficiently with the use of computer proof-assistants. Then, changing even simple notions with obvious (at least at the very first sight) properties, as parallel postulate (or Playfair axiom), gave rise to various geometries (e.g., Bolyai-Lobachevskian hyperbolic geometry). The same work can be modelled with machine formalizations, using various sets of axioms (creating *types*). Now, apart from the discussion whether the non-emptiness of types in Mizar is real difficulty (because from informal point of view one can consider an object with any properties, even mentioning their coherence), and how much more can

be attained if the reimplementing of the Mizar type system will be done in the foreseeable future, we have to cope with the limitations of the existing type structure. On the other hand, type checking allows some errors to be caught early – when making mathematical definitions. In this context the requirement of constructing at least one object of the desired type is quite natural, as it prevents contradictory types. Similarly, the appropriate model had to be constructed either to assure that proposed axioms are correct (which is not very hard as they can be parsed by ordinary mathematician even straight from its corresponding Mizar source code), or (and this is probably even more important) to bind the fresh formal apparatus with the existing Mizar developments. Some of them are written in a language which is not as expressive as contemporary Mizar language is; in the time of the beginnings of the Mizar Mathematical Library (MML) as a tight collection of Mizar articles covering various branches of mathematics, geometry was an area which was developed quite dynamically.

The language of the Mizar system was influential for other systems for formalization of mathematics, e.g. `miz3` is a proof interface built on top of HOL Light interactive theorem prover, with the declarative language compatible with the Mizar language [47]. Recently, William Richter used `miz3` tool to formalize Tarski geometry axioms, with the ultimate aim to incorporate it in HOL Light, but of course part of his pseudo-code could be also treated both as a case study and as a good starting point for further work.

The choice of the topic is not accidental – recent code available in Coq [7] and the use of automated equational provers caught an eye of researchers and, as a by-product, some results, which shed some new light on the axiomatization of geometry, were published. One of the bright milestones was also the publication of the new issue of the classical textbook *Metamathematische Methoden in der Geometrie* by Wolfram Schwabhäuser, Wanda Szmielew and Alfred Tarski [40] (to which we refer by the acronym SST) with the foreword of Michael Beeson.

In this paper, we are not focused on any of geometric challenges known in the community, as proving Hilbert axioms from Tarski [7], or formalizing full SST in Mizar, although it can be definitely good starting point as in [38] we prove some Hilbert's axioms. What we were trying to do was to

increase of the integrity of (the geometrical part of) the MML as pointed out in the paper of Piotr Rudnicki and Andrzej Trybulec [39] and this could be a kind of partial realization of their postulates. This was done mainly via the mechanism of *revisions* of the repository – stepwise refinement of items already included in MML, done not necessarily by authors themselves [19]. An alternative approach – focusing on computations instead of proofs, and work in the analytic framework of Euclidean spaces \mathbb{R}^n is also well-represented in the repository of Mizar texts, with the recent examples of Morley trisector theorem or Ceva theorem. The process of formalizing geometry within Mizar Mathematical Library started years before first Coq geometry formalizations; furthermore, constructive logic behind the Coq proof assistant naturally forces program extraction from proofs and intuitionistic setting of the reasoning. In Mizar the stress is put on three main issues:

- writing readable proofs using classical logic,
- possibility of cooperation with external theorem provers, and
- knowledge reuse (increasing possible connections between various developments, called *integrity of a repository*).

The outline of the paper is as follows. In Section II, we describe the history of formalizing geometry in Mizar, Section III presents the discussion on abstract and concrete mathematics; the next one outlines some basic constructions needed to understand our work, that is concrete translation of chosen properties straight from Tarski's axioms (A1)–(A7). In Section V and VI we give some insight on knowledge reuse and on its readability, respectively, while in the last part we describe related work, then we draw some concluding remarks and propose some future work.

II. MIZARING AFFINE GEOMETRY

In Tarski's system of axioms [44] the only primitive geometrical notions are points, the ternary relation B of “soft betweenness” and quaternary relation \equiv of “equidistance” or “congruence of segments”. The axioms are reflexivity, transitivity, and identity axioms for equidistance; the axiom of segment construction; reflexivity, symmetry, inner and outer transitivity axioms for betweenness; the axiom of continuity, and some others. The original set consisted of 20 axioms for two-dimensional Euclidean geometry and was constructed in 1926–27, submitted for publication in 1940, and finally appeared in 1967 in a limited number of copies. There are many modifications of this system, and Gupta's work in this area [20] offers an important simplification. The strict betweenness was studied even before: it gave rise to “betweenness geometry” by Veblen in 1904.

Another notable axiomatization, proposed by Hilbert [21] in 1899, has three sorts: planes, points, and lines, and three relations: betweenness, containment, and congruence. In this sense it is a little bit more complex than Tarski's (but not necessarily in terms of numbers of axioms as it has also 20 of these). The two approaches establish a geometrical framework,

in which theorems can be proven logically (remember Euclid's *Elements* proofs are mainly pictures or graphical constructions and rely heavily on the intuition). This allows to use a computerized theorem prover in order to find the proof or proof checker to check the theory for its correctness. In the paper, we deal with the proof checker Mizar, based on classical first order logic and Tarski-Grothendieck set theory (a variant of Zermelo-Fraenkel). Using this tool we describe, how Tarski's theory was built formally.

There are two significant connections of Andrzej Trybulec, the founder of the Mizar project, with persons involved in the area of Tarski's geometry. The first one and very influential was the cooperation with Lesław W. Szerbera, the author of [42], in the early 80s of the previous century. Szerbera, who also submitted to the Mizar Mathematical Library twice, was at that time a head of the Institute of Mathematics in Białystok, Poland; the place the Mizar system was mainly developed (and that was also the affiliation of Trybulec). These contacts resulted in the research on the theory of interpretation and semantic foundations of logic in the sense of Epstein. At that time, Trybulec himself was not very active in formalizing geometry. The other connection was that after finishing his study in mathematics at the University of Warsaw under the guidance of Karol Borsuk (famous Polish topologist), Andrzej Trybulec took the position of an assistant in the Chair of Geometry, where Wanda Szmielew was a professor.

Although the very first approach to formal geometries we can consider the work of Wojciech A. Trybulec INCSP_1 [45], at the state of its writing it was not tightly connected with the rest of the formal approach to geometry in Mizar. Krzysztof Prażmowski (who is currently the head of the Institute of Mathematics, University of Białystok, Poland) created quite an active research group of fifteen people in the field of automated deduction in geometry, with the extensive use of the Mizar system. Main authors of these contributions were Krzysztof Prażmowski, Henryk Oryszczyszyn, and Wojciech Leończuk, all from former University of Warsaw, Białystok Branch. Together with the other authors, e.g., Kusak, Skaba, Muzalewski, and others, they wrote 43 Mizar articles on geometry (5 were removed later in the process of revisions). At the very beginning, there were many independent paths of formal development of various geometries (which expressed in various Mizar structures: directed vs. undirected parallelity relation, orthogonality, etc.).

Statistical data from Table I are not very impressive with respect to the current state of MML. The number of authors is quite big (six percent of all the authors of MML), hence the style of geometry in Mizar is not very uniform. The writing style can be measured by the percentage of zipped kBytes, which is higher than for the whole MML. This specific ratio (zipped instead of unpacked bytes) is usually taken when computing the so-called de Bruijn factor, showing how much additional code we should write comparing with informal proof to be fully understandable by computers. Furthermore, hierarchy of geometric objects is based on eight various structures, so it raises communication issues between various

TABLE I
STATISTICAL DATA ON THE FORMAL DEVELOPMENT OF GEOMETRY
WITHIN MML

	Geometry	MML	Percentage
files	38	1,254	3.0
authors	15	255	6.0
kBytes	1,981	93,324	2.1
zipped kBytes	406	17,642	2.3
definitions	294	11,606	2.5
theorems	1,319	56,547	2.3
attributes	135	3,079	4.4
clusters	113	13,071	0.9
structures	8	157	5.0

approaches. The last article was PROJPL_1 dated back to 1994, and the series was not very actively developed (with the exception of the paper of the combinatorial Grassmannians COMBGRAS). But in 1990 Mizar articles on geometry were about 30% of the whole Mizar repository (out of 140 files). Of course, after that time revisions of this specific area were quite active: one of the main streams (done also by the current author) was to get separate axioms of selected properties instead of a large block for *the mode* (i.e., the constructor of the type in Mizar). As we can read from the table, the use of clusters (relatively new feature of Mizar language) is still very low comparing with the rest of MML, so there is a lot of work to be done. Affine approach to geometry was less important as there was another big challenge which for fifteen years stimulated geometry (in analytical setting, however): the proof of the Jordan Curve Theorem [23], which will be recalled in Section V.

As notable affine geometrical facts already formalized in Mizar we can enumerate:

- Hessenberg's theorem – HESSENBE;
- Desargues theorem (present in "Top 100 mathematical theorems") – ANPROJ_2;
- Pasch configuration axioms – PASCH;
- Fano projective spaces – PROJRED1;
- Desarguesian projective planes – PROJRED2;
- Pappus, Minor, Major and Trapezium Desargues axioms – AFF_2;
- Minkowskian geometry – ANALORT.

The "Top 100 mathematical theorems" list of important theorems in mathematics, proposed by Paul and Jack Abad at the end of previous century, is the popular collection of challenges for contemporary computerized proof-assistant systems; the list maintained by Freek Wiedijk described in [48] is available at <http://www.cs.ru.nl/~freek/100/>. Currently, the Mizar system holds the third position with 64 items formalized so far.

In our opinion, the unifying approach in Tarski's spirit could be quite useful to bind all of these geometrical results together. Among another significant facts in geometry we can point out Morley trisector theorem [8], Ceva and Menelaus theorem [41]. These facts however deal with Euclidean plane, so the proofs are in the area of analytic geometry; they were developed more than ten years later than the foundations

TABLE II
THREE MAIN SECTIONS OF THE MIZAR MATHEMATICAL LIBRARY

Part of MML	Files	%	kBytes	%
classical part	323	25.8	20,013	21.8
abstract part	866	69.0	67,996	74.0
SCM	65	5.2	3,837	4.2
Total	1254	100.0	91,846	100.0

of geometry in Mizar, when Euclidean spaces were more thoroughly explored in MML.

III. THE CHOICE OF FORMAL FRAMEWORK: CLASSICAL VS. ABSTRACT MATHEMATICS

The distinction between classical and abstract mathematics (i.e. the one based on ordinary axioms of set theory, and all those using the notion of a structure, respectively) is important from the viewpoint of the organization of the Mizar repository. We had to choose between two paths:

- it is possible to formulate Tarski's axioms without the use of a structure, and also set theory could be meaningless for that framework, only the classical logic with Mizar predicates is enough (equality plays a special role in the system based on set theory [18]);
- the use of Mizar structures forces us paradoxically to use basics of set theory – defining a signature of Tarski's plane needed to give a type of congruence of segments and betweenness relation, which was set-theoretic (Mizar language is typed, and in the earlier case one should also give a type at least to points, but it can be defined as Mizar object).

The latter was also chosen by us as the whole geometry in MML is written in abstract style (as the majority of MML, as you can read from Table II) as structures in Mizar are present for a long time. Even if in ordinary mathematical tradition they are considered as ordered tuples, in the implementation in Mizar they are treated rather as partial functions, with selectors as arguments, and ordinary inheritance mechanism (with polymorphic enabled, which will be extensively used in our formalizations). The details of the implementation can be found in [16], we give here only general syntax of declaring Mizar structure:

```
struct (Predecessor_List) Structure_Name
  (# selector_1 -> type_1,
   selector_2 -> type_2,
   ...
   selector_n -> type_n #);
```

To every argument, its type should be declared, which corresponds to ordinary definition of the signature of an algebra.

In the contemporary MML, the basic Mizar article devoted to structures is [26]; it is dated as for 1995, earlier than the first article from MML on structures [45], i.e. 1989, because it was created much later as a result of revision. Its main step was introducing common predecessor of all structures – `l-sorted`, and the name of the selector carrier was

chosen. Wojciech A. Trybulec proposed in [45] the formalization of Hilbert axioms (hence we have three sorts: points, lines and planes). Interestingly, the original name `Points` remained untouched (although many similar approaches were later unified).

```
definition
  struct IncProjStr
    (# Points, Lines -> non empty set,
     Inc -> Relation of
       the Points, the Lines #);
end;

definition
  struct (IncProjStr) IncStruct
    (# Points, Lines, Planes -> non empty set,
     Inc -> Relation of the Points, the Lines,
     Inc2 -> Relation of the Points, the Planes,
     Inc3 -> Relation of the Lines, the Planes #);
end;
```

```
definition
  let S be IncProjStr;
  mode POINT of S is
    Element of the Points of S;
  mode LINE of S is
    Element of the Lines of S;
end;
```

Introducing two different structures to host only points and lines in the first case was a result of revision allowing to use simpler structures to cope with planar geometry. Note however, that this disagrees with Tarski's single sort universe (of points); and this is quite a basic structure in MML, 1-sorted, as we already mentioned. Based on this, other descendant objects can be defined:

```
definition
  struct (1-sorted) AffinStruct
    (# carrier -> set,
     CONGR -> Relation of
       [:the carrier,the carrier:] #);
end;
```

This was the very first approach, and in fact the article with MML identifier `ANALOAF` offered quite simple structure; here, `CONGR` is a relation on the Cartesian square of the carrier, and it is used mainly in the context of parallelity predicate. Merging it with another relation, the orthogonality, MML offers a variety of affine geometries. The outline of this universe is shown in Figure 1. It can be observed that the hierarchy is not very deep, as many approaches were done originally in parallel, without reusing notions or theorems between various paths of development. Furthermore, incidence structures are not descendants of 1-sorted, which should be definitely corrected in the future.

```
definition
  struct (AffinStruct, OrtStr) ParOrtStr
    (# carrier -> set,
     CONGR, orthogonality -> Relation of
       [:the carrier,the carrier:] #);
end;
```

In the original form of [45], the number of axioms was

introduced in the form of a single big formula introducing mode `IncSpace`. What is interesting, after the revision, and evolution of the Mizar system, the file is 10 kB smaller.

```
definition
  struct (1-sorted) TarskiPlane
    (# carrier -> set,
     Betweenness -> Relation of
       [:the carrier, the carrier:], the carrier,
     Equidistance -> Relation of
       [:the carrier, the carrier:],
       [:the carrier, the carrier:] #);
end;
```

Observe that the choice of the arity of the relation is really meaningless here: the betweenness relation can be treated as a ternary relation; but the choice of this concrete model was quite arbitrary as the difference between dealing with ternary relations and relations between ordered pairs and elements will not cause any major problems later (we use mainly predicates).

```
definition
  let S be TarskiPlane;
  mode POINT of S is Element of S;
end;
```

```
definition
  let S be TarskiPlane;
  let a, b, c be POINT of S;
  pred between a,b,c means
  :: GTARSKI1:def 1
  [[a,b],c] in the Betweenness of S;
end;
```

```
definition
  let S be TarskiPlane;
  let a, b, c, d be POINT of S;
  pred a,b equiv c,d means
  :: GTARSKI1:def 2
  [[a,b],[c,d]] in the Equidistance of S;
end;
```

One can mention possibly misleading type definitions for the betweenness relation (it can be considered as three-argument relation than as it is now, i.e. relation between an ordered pair and an element), and also – pretty technical – predicate between a,b,c for arbitrary points a,b,c . Of course, the notation could be changed easily into well-readable candidate `b is_between a,c` (maybe with the need for replacing arguments), but we have chosen this notation to be closer to SST. Brackets can be used as `between (a,b,c)`, but they can also be omitted, there is no strict obligation for using them.

IV. THE ORIGINAL SELF-CONTAINED APPROACH

In the very first version of Mizar formalization of Tarski's axioms done by William Richter with the help of `miz3`, we can find the remark:

In Mizar it isn't possible to define such a type (or model) without proving that some model exists. Trybulec's existence proofs runs over 450 lines. So we define a predicate 'S Tarski-model' which means that the plane S satisfies the axioms A1–A7, and

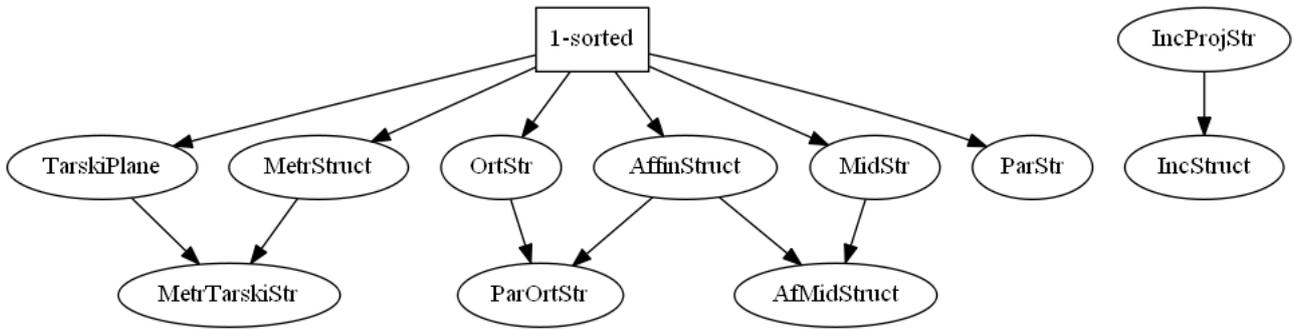


Fig. 1. Geometrical structures in the Mizar Mathematical Library

then prove trivial theorems A1–A7 which say that if S Tarski-model, then S satisfies an axiom A1–A7. The extra clutter involving the predicate Tarski-model, and the label TarskiModel which stands for the statement ‘ S Tarski-model’ could be avoided by loading all our results into one gigantic theorem. Our approach seems preferable.

This motivation essentially caused the lack of the appropriate Mizar type. Indeed, Trybulec’s existence proof gets even more lines (over 500), but it should be taken into account that [45] was one of the very first Mizar articles submitted to MML (numbered #25) and auxiliary set of handy lemmas or models was really modest at that time.

One can consider having predicates, but attributes instead of them seems to be better idea: we can have modular building of a complex structure, all other can be reused; hence in our Mizar article we focus on pure betweenness-equidistance part, not really mentioning the question of dimensions. We prove 44 theorems (properties of the predicates), with the Gupta’s proof of Hilbert’s I1 axiom (that two distinct points determine a line).

```

theorem ::: I1:
  a <> b & x <> y &
  a on_line x,y & b on_line x,y
  implies x,y equal_line a,b
  proof
  assume
H1: a <> b & x <> y; then
P2: b,a equal_line a,b by LineEqA1;
  assume
H2: a on_line x,y & b on_line x,y;
  per cases;
  suppose x = b; then
    x,y equal_line b,a by H1, H2, I1part2;
    hence thesis by P2;
  end;
  suppose
    x <> b; then
H4: x,y equal_line x,b by H2, I1part2; then
    x,b equal_line a,b
      by H1, I1part2Reverse, H2;
    hence thesis by P4;
  end;
end;
end;

```

TABLE III
THE STATISTICS OF THE CONTENT OF [38]

Items	Numbers
attributes	10
lines	1522
kBytes	50
theorems	47

Final Gupta’s proof of I1 (GTARSKI1:46) above uses additional auxiliary predicates, e.g. `on_line`, which shortens the notation and provides future connections with Hilbert’s axioms. Of course, at the end we are left with the proofs that the set of formulas (obtained by the so-called *definitional expansions* [24]), chosen formula can be deduced, but our attribute-steered approach seems to be better, not in terms of the efficiency of proving, but readability and usefulness for human mathematician.

Our Mizar versions of Tarski’s axioms have descriptive names, and follow the ones from SST (using \equiv for congruence of segments and B for betweenness relation):

- CongruenceSymmetry (A1):
 $\forall_{a,b} ab \equiv ba,$
- CongruenceEquivalenceRelation (A2):
 $\forall_{a,b,p,q,r,s} ab \equiv pq \wedge ab \equiv rs \Rightarrow pq \equiv rs,$
- CongruenceIdentity (A3):
 $\forall_{a,b,c} ab \equiv cc \Rightarrow a = b,$
- SegmentConstruction (A4):
 $\forall_{a,q,b,c} \exists_x B(q, a, x) \wedge ax \equiv bc,$
- BetweennessIdentity (A6):
 $\forall_{a,b} B(a, b, a) \Rightarrow a = b,$
- and Pasch (A7):
 $\forall_{a,b,p,q,z} B(a, p, z) \wedge B(b, q, z) \Rightarrow \exists_x B(p, x, b) \wedge B(q, x, a).$

One attribute has the form which substantially differs from SST version: in order to shorten the notation, we introduced technical predicate

```

definition
  let S be TarskiPlane;
  let a, b, c, x, y, z be POINT of S;
  pred a,b,c cong x,y,z means
  :: GTARSKI1:def 3
    a,b equiv x,y &

```

```

a,c equiv x,z &
b,c equiv y,z;
end;

```

denoting essentially SSS predicate for triangles. Using this notion, SST axiom (A5) could be encoded as follows:

```

definition let S be TarskiPlane;
  attr S is satisfying_SAS means
:: GTARSKI1: def 9
  for a, b, c, x, a1, b1, c1, x1
  being POINT of S holds
  a <> b & a,b,c cong a1,b1,c1 &
  between a,b,x & between a1,b1,x1 &
  b,x equiv b1,x1 implies
  c,x equiv c1,x1;
end;

```

that is,

$$\forall a,b,c,x,a',b',c',x' (a \neq b \wedge ab \equiv a'b' \wedge bc \equiv b'c' \wedge ac \equiv a'c' \wedge \wedge B(a,b,x) \wedge B(a',b',x') \wedge bx \equiv b'x') \Rightarrow cx \equiv c'x'.$$

It should be noted that the sequence of internal Mizar utilities can be run on the code, and hence proofs can be modified or even shortened. The remarkable case are the three assumptions from auxiliary lemmas added in the original `miz3` code, which was automatically removed as unnecessary. The usual running cycle of auxiliary programs like `relprem`, `relinfer`, `chklab`, and `inacc` is as follows: first, unnecessary premises are marked, then irrelevant (obvious for the Mizar checker) proof steps are marked as unused and eliminated, and finally the whole structure of the proof can be pretty-printed. Such “cleaning” cycle [16] caused the reduction of proofs of properties by more than 300 lines, not affecting readability that much. The net of connections between definitions, lemmas, and theorems obtained in this way can be further studied in order to get better refactoring of knowledge via dedicated techniques [17].

Having separate attributes for distinct axioms had already shown its usefulness in various geometrical settings. It could also allow later for defining equivalent axiom systems for Tarski geometry (and due to mechanism of clusters this equivalence will be obvious for the checker, once proven).

V. INTRODUCING METRIC STRUCTURE

It is well known fact that every metric space can be equipped with the natural topology. This informally obvious mathematical property brings some unexpected difficulties when dealing with structures if automated proof-assistants play a role. Namely then, if one considers topological spaces in a quite natural way, that is $\langle U, \tau \rangle$, and metric space as $\langle U, d \rangle$, respectively, one can rather naturally merge both structures into common $\langle U, \tau, d \rangle$.

During the formalization of the Jordan Curve Theorem in Mizar [23], however, another approach was chosen (and pushed consequently until the successful finale): `Euclid 2` denoted metric space concerned with the Euclidean plane, and then special functor converting any metric space into the topological space was applied to obtain `TOP-REAL 2`

(of course the conversion can be made for arbitrary natural number n , not necessarily 2, but Jordan curves deal with two-dimensional case). The basic signature for metric spaces are `MetrStruct`, where distance is a function defined on the Cartesian square of the carrier with the real values. Then, metrics (or pseudo-, quasi-, semimetrics, etc.) can be defined in terms of attributes [22], that is properties of the distance function.

```

definition
  struct (MetrStruct, TarskiPlane) MetrTarskiStr
  (# carrier -> set,
  distance -> Function of
  [:the carrier, the carrier:], REAL,
  Betweenness -> Relation of
  [:the carrier, the carrier:],
  the carrier,
  Equidistance -> Relation of
  [:the carrier, the carrier:],
  [:the carrier, the carrier:] #);
end;

```

Then we have two worlds merged: affine, where we have two Tarski’s relations, and Euclidean, where in terms of distance function, we can have betweenness relation and the measure for segments. We argue that, regardless of all the complications caused by merging structures [12] (which increases the number of selectors, hence the chain of notions gets more complicated), such approach – not converting between two contexts, but rather to make reasoning in the world which is successor of both – allows for more flexible reuse of knowledge from two original areas.

```

definition let M be MetrTarskiStr;
  attr M is naturally_generated means
:: GTARSKI1: def 15
  (for a, b, c being POINT of M holds
  between a,b,c iff b is_Between a,c) &
  (for a, b, c, d being POINT of M holds
  a,b equiv c,d iff
  dist (a,b) = dist (c,d));
end;

```

In merged structure, we want to have two segments congruent, if they are equal in terms of distance, and betweenness relation uses the predicate `is_Between`, also given in terms of the sum of distances. To construct a model of the Tarski’s space, such “one gigantic theorem” mentioned by Bill Richter in the quotation from the previous section can be somehow avoided. In Trybulec’s paper [45], the massive construction of the appropriate space satisfying all incidence axioms was hidden under the existence proof of the following Mizar type:

```

registration
  cluster strict IncSpace-like for IncStruct;
end;

```

Recalling the previous discussion on the structure merging, the construction of such a space from scratch (essentially more or less modified copy-and-paste work on the proof from [45]) can be avoided with the help of some useful tricks, investigating knowledge already present in MML with its possible reuse. For example, based on the geometry on the real

line, appropriate geometrical structure from metric structure can be just an extension with properly defined distance.

```

definition
  func TarskiSpace -> MetrTarskiStr equals
  :: GTARSKI1:def 22
    the naturally_generated TarskiExtension
      of RealSpace;
  coherence;
end;

```

Then, we can show that in such extension the metric is well defined, i.e. this space is reflexive, symmetric, and discerning. Furthermore, if we take into account “geometrical part”, the following axioms can be proven true:

```

registration
  cluster TarskiSpace ->
    satisfying_CongruenceSymmetry
    satisfying_CongruenceEquivalenceRelation
    satisfying_CongruenceIdentity
    satisfying_SegmentConstruction
    satisfying_BetweennessIdentity;

```

VI. HUMAN-ORIENTED REPRESENTATION OF KNOWLEDGE IN *Formalized Mathematics*

Of course, detailed proofs of all the work are already present in the Mizar Mathematical Library and were published in the special issue of *Formalized Mathematics* devoted to 25 years of MML. We tried to find subtle compromise between strict mathematical notation and a fluent text looking quite as good as that written by human – as the author has the opportunity of choosing the translation formats for newly introduced definitions (in XML stylesheet); the example of two automatically translated theorems are given on Figure 2 (versions straight from MML and FM).

```

theorem :: GTARSKI1:31 ::: GuptaEasy:
  a <> b & between a,b,c & between a,b,d &
  b <> c & b <> d implies not between c,b,d;

theorem :: GTARSKI1:32
  a,b,c cong a9,b9,c9 &
  between a,x,c & between a9,x9,c9 &
  c,x equiv c9,x9 implies b,x equiv b9,x9;

```

VII. RELATED WORK

Although the history of the development of the axiom system for geometry by Tarski is not very clear from the beginnings (as the early works by Tarski seem to be postponed by the World War II), and even if the ultimate source of information is SST which was badly printed (the book was recently reissued with the foreword of Michael Beeson), now it attracted a lot of focus from automated deduction systems.

The remarkable item here is of course Julien Narboux’s formal development of Tarski’s system with the use of Coq [33]. We are planning to include some interesting results from GeoCoq in the Mizar system; Nakasho’s et al. MML symbol reference system [32]¹ offers quite user-friendly interface for

¹The system can be browsed at <http://webmizar.cs.shinshu-u.ac.jp/mmlfe/current/> with the official version of the Mizar system.

browsing appropriate content. Of course, GeoCoq provides ready-to-use list of notions, which are connected only with the Tarski’s system – that makes the browsing more convenient. Of course, after rescaling all definitions and theorems can be browsed within MML providing a look in the style of GeoCoq project.

Similar web-browsing system giving practical insight into Tarski’s geometry is offered by the aforementioned Michael Beeson’s *Tarski Formalization Project*. It offers linkages with underlying items from SST; the correctness relies on the Otter prover. Here the readability is not the main issue, however a few open questions was answered. Tim Makarios used Isabelle to formalize elliptic geometry in order to provide independence results for the parallel postulate [28]. Accidentally, this is one the items from the “Top 100 mathematical theorems”, which is also challenging for us. It can be treated as a further development of axiomatic approach to geometry in Isabelle started by Meikle [31] (in Hilbert’s *Foundations of Geometry* [21] style).

As we are concerned with the integrity issues for the Mizar Mathematical Library, we do not want to make basics from scratch, especially if the strong fundamentals are already done. There a need for further careful study of how much can be done in the direction of unveiling the knowledge included in the repositories of the Mizar system. Due to mechanism of revisions it is not the case that this work is just useless as it is very old. Mizar geometry just used the other way.

There are many simplifications of the original system by Tarski; e.g. by Tarski himself and his collaborators, Gupta’s [20], or Makarios [28]. Also showing the correspondence between various axiom systems (with Hilbert’s at the very beginning) for geometry is quite influential and here provers can be quite useful. Of course, one of the well-advertised areas is the topic of axiom system for various equationally-defined classes of algebras, with the leading problem by Robbins and its solution by William McCune obtained in 1996 with the help of EQP/Otter equational theorem prover. But the question remains, how to cope with the space between obviousness for the human and for the machine, in the time before *QED Singularity*, as Michael Beeson (at QED+20 workshop in Vienna, July 18, 2014 [4]) called this time when formal proofs will be the norm in mathematics.

VIII. CONCLUSIONS AND FUTURE WORK

In the paper, we have shown how already quite well established repository can be enhanced to cope with new capabilities of the Mizar system. We mention here a new approach to structures, including their merging, extensive use of attributes, and implemented automatization of definitional expansions. Also the issues of the integration with external provers should be taken into account.

We have created in [38] (and recent GTARSKI2 [9]) complete formal axiomatization of Tarski’s geometry which, at least in our opinion, has the advantage of higher readability for ordinary mathematicians than, e.g., Coq or Prover9 proof objects. In the same it is tightly connected with another

- (31) If $A \neq B$ and B lies between A and C and B lies between A and D and $B \neq C$ and $B \neq D$, then B does not lie between C and D . The theorem is a consequence of (30), (21), and (18).
- (32) Suppose $\triangle ABC \cong \triangle A'B'C'$ and X lies between A and C and X' lies between A' and C' and $\overline{CX} \cong \overline{C'X'}$. Then $\overline{BX} \cong \overline{B'X'}$. The theorem is a consequence of (5), (11), (8), (6), (12), (25), (7), (16), and (19).

Fig. 2. Screenshot of an excerpt from *Formalized Mathematics*

axiomatization of Euclidean plane, due to Hilbert, already available in MML. We have shown also that the real Euclidean plane satisfies all Tarski's axioms.

Of course, one of our aims is to do similar work to that of Michael Beeson's – the use of Prover9 to formalize practically everything from SST. But taking into account that Prover9 proof object can be relatively easily converted into working (and correct) Mizar proofs, it is important to construct appropriate background in order to have something which can offer a kind of help for a mathematician by the following requirements:

- it will reflect all theorems from SST;
- it will offer relatively high readability;
- the proofs can help to catch the idea of a proof, not necessarily will be just *l'art pour l'art*.

Translations from OTTER (at that time, currently Prover9) were helpful in the process of building various complemented lattices: Sheffer-stroke based and those needed in the Robbins algebras [13]. Also encodings of various short axiomatizations for Boolean algebras were easier with that tool; also in the area of rough sets [14] some results were obtained in that way [11]. On the other hand, fuzzy sets, which are much closer to set theory than rough sets, can be also nicely formalized [15].

We are also interested in building a Mizar model of elliptic geometry; Japanese team did some introductory work in [10], but it is not quite feasible now. This could help in showing that parallel postulate is independent from the other ones, which will solve another item from "Top 100 mathematical theorems", which is also an important issue. Further research on logical reasoning systems (as in the case of Tarski's system formulable in first-order logic with identity, set theory is not required) formalized by means of automated proof-assistants can be fruitful not only in the area of pure or applied mathematics, but also, e.g., in the area of argumentation theory, including legal expert systems [49].

We should definitely also have in mind the usefulness of the approach in the context of didactic use of the Mizar system. In order to achieve this goal, the existing state of geometry in MML should be made definitely more coherent, without the need of the creation of new special encyclopaedic articles which could serve better for students.

REFERENCES

- [1] J. Alama, M. Kohlhase, L. Mamane, A. Naumowicz, P. Rudnicki, and J. Urban: Licensing the Mizar Mathematical Library, Proceedings of Mathematical Knowledge Management 2011, *Lecture Notes in Artificial Intelligence*, 6824, pp. 149–163 (2011)
http://dx.doi.org/10.1007/978-3-642-22673-1_11
- [2] J. Avigad, E. Dean, and J. Mumma: A formal system for Euclid's Elements, *Review of Symbolic Logic*, 2 pp. 700–768 (2009)
<http://dx.doi.org/10.1017/S1755020309990098>
- [3] G. Bancerek, Cz. Byliński, A. Grabowski, A. Kornilowicz, R. Matuszewski, A. Naumowicz, K. Pał, and J. Urban: Mizar: state-of-the-art and beyond, *Intelligent Computer Mathematics, Lecture Notes in Computer Science* 9150, pp. 261–279 (2015)
http://dx.doi.org/10.1007/978-3-319-20615-8_17
- [4] M. Beeson: Mixing computations and proofs, *Journal of Formalized Reasoning*, 9(1), pp. 71–99 (2016)
<http://dx.doi.org/10.6092/issn.1972-5787/4552>
- [5] M. Beeson and L. Wos: OTTER proofs in Tarskian geometry, in Proceedings of IJCAR 2014, *Lecture Notes in Computer Science*, 8562, pp. 495–510 (2014)
http://dx.doi.org/10.1007/978-3-319-08587-6_38
- [6] K. Borsuk and W. Szmielew: *Foundations of Geometry*, North-Holland (1960)
- [7] G. Braun and J. Narboux: From Tarski to Hilbert, in *Automated Deduction in Geometry*, T. Ida and J. Fleuriot (eds.) Proceedings of ADG 2012 (2012)
http://dx.doi.org/10.1007/978-3-642-40672-0_7
- [8] R. Coghetto: Morley trisector theorem, *Formalized Mathematics*, 23(2), pp. 75–79 (2015)
<http://dx.doi.org/10.1515/forma-2015-0007>
- [9] R. Coghetto and A. Grabowski: Tarski geometry axioms – part II, *Formalized Mathematics*, 24(2) (2016).
<http://dx.doi.org/10.1515/forma-2016-0012>
- [10] Y. Futa, H. Okazaki, D. Mizushima, and Y. Shidama: Operations of points on elliptic curve in projective coordinates, *Formalized Mathematics*, 20(1), 87–95 (2012)
<http://dx.doi.org/10.2478/v10037-012-0012-2>
- [11] A. Grabowski: Automated discovery of properties of rough sets, *Fundamenta Informaticae*, 128(1–2), pp. 65–79 (2013)
<http://dx.doi.org/10.3233/FI-2013-933>
- [12] A. Grabowski: Efficient rough set theory merging, *Fundamenta Informaticae*, 135(4), pp. 371–385 (2014)
<http://dx.doi.org/10.3233/FI-2014-1129>
- [13] A. Grabowski: Mechanizing complemented lattices within Mizar type system, *Journal of Automated Reasoning*, 55(3), pp. 211–221 (2015)
<http://dx.doi.org/10.1007/s10817-015-9333-5>
- [14] A. Grabowski: On the computer-assisted reasoning about rough sets, in *Monitoring, Security and Rescue Techniques in Multiagent Systems*, B. Dunin-Kępczyk, A. Jankowski, M. Szczuka (Eds.), *Advances in Soft Computing*, 28, 215–226 (2005)
http://dx.doi.org/10.1007/3-540-32370-8_15
- [15] A. Grabowski: On the computer certification of fuzzy numbers, in *Proceedings of 2013 Federated Conference on Computer Science and Information Systems, FedCSIS 2013*, M. Ganzha, L. Maciaszek, M. Paprzycki (Eds.), pp. 51–54, IEEE (2013)
- [16] A. Grabowski, A. Kornilowicz, and A. Naumowicz: Mizar in a nutshell, *Journal of Formalized Reasoning*, 3(2), 153–245 (2010)
<http://dx.doi.org/10.6092/issn.1972-5787/1980>
- [17] A. Grabowski and Ch. Schwarzweiller: Towards automatically categorizing mathematical knowledge, in *Proceedings of 2012 Federated Conference on Computer Science and Information Systems, FedCSIS 2012*, M. Ganzha, L. Maciaszek, M. Paprzycki (Eds.), IEEE, 63–68 (2012)

- [18] A. Grabowski, A. Kornitowicz, and Ch. Schwarzweiller: Equality in computer proof-assistants, in *Proceedings of 2015 Federated Conference on Computer Science and Information Systems, FedCSIS 2015*, M. Ganzha, L. Maciaszek, M. Paprzycki (Eds.), IEEE, 45–54 (2015) <http://dx.doi.org/10.15439/2015F229>
- [19] A. Grabowski and Ch. Schwarzweiller: Revisions as an essential tool to maintain mathematical repositories, *14th Symposium on Towards Mechanized Mathematical Assistants*, Calculemus'07/MKM'07, Lecture Notes in Computer Science, 235–249 (2007) http://dx.doi.org/10.1007/978-3-540-73086-6_20
- [20] H.N. Gupta: Contributions to the axiomatic foundations of geometry, PhD thesis, University of California, Berkeley (1965)
- [21] D. Hilbert: *The Foundations of Geometry*, Chicago: Open Court, 2nd ed. (1980)
- [22] S. Kanas, A. Lecko, and M. Startek: Metric spaces, *Formalized Mathematics*, 1(3), pp. 607–610 (1990)
- [23] A. Kornitowicz: Jordan curve theorem, *Formalized Mathematics*, 13(4), pp. 481–491 (2005)
- [24] A. Kornitowicz: Definitional expansions in Mizar, *Journal of Automated Reasoning*, 55(3), pp. 257–268 (2015) <http://dx.doi.org/10.1007/s10817-015-9331-7>
- [25] A. Kornitowicz: On rewriting rules in Mizar, *Journal of Automated Reasoning*, 50(2), 203–210 (2013) <http://dx.doi.org/10.1007/s10817-012-9261-6>
- [26] Library Committee of the Association of Mizar Users, Preliminaries to structures, *Mizar Mathematical Library*, MML Id: STRUCT_0 (1995)
- [27] M. Lombard and R. Vesley: A common axiom set for classical and intuitionistic plane geometry, *Annals of Pure and Applied Logic*, 95, pp. 229–255 (1998)
- [28] T. Makarios: A mechanical verification of the independence of Tarski's Euclidean Axiom, MSc thesis (2012)
- [29] R. Matuszewski and P. Rudnicki: Mizar: the first 30 years, *Mechanized Mathematics and Its Applications*, 4(1), pp. 3–24 (2005)
- [30] W. McCune: Prover9 and Mace4. Available at <http://www.cs.unm.edu/~mccune/prover9/>, 2005–2010.
- [31] L. Meikle and J. Fleuriot: Formalizing Hilbert's Grundlagen in Isabelle/Isar, in Proceedings of TPHOLS'03, *Lecture Notes in Computer Science*, 2758, pp. 319–334 (2003) http://dx.doi.org/10.1007/10930755_21
- [32] K. Nakasho and Y. Shidama: Documentation generator focusing on symbols for the HTML-ized Mizar library, Proceedings of the Conference on Intelligent Computer Mathematics, CICM 2015, *Lecture Notes in Computer Science*, 9150, pp. 343–347 (2015) http://dx.doi.org/10.1007/978-3-319-20615-8_25
- [33] J. Narboux: Mechanical theorem proving in Tarski's geometry, in *Automated Deduction in Geometry*, F. Botana and T. Recio (eds.), Lecture Notes in Computer Science, 4869, pp. 139–156 (2007) http://dx.doi.org/10.1007/978-3-540-77356-6_9
- [34] A. Naumowicz and A. Kornitowicz: A brief overview of Mizar, in *Theorem Proving in Higher Order Logics 2009*, S. Berghofer, T. Nipkow, Ch. Urban, M. Wenzel (Eds.), Lecture Notes in Computer Science, 5674, 67–72 (2009) http://dx.doi.org/10.1007/978-3-642-03359-9_5
- [35] K. Pał: Methods of lemma extraction in natural deduction proofs, *Journal of Automated Reasoning*, 50(2), 217–228 (2013) <http://dx.doi.org/10.1007/s10817-012-9267-0>
- [36] A. Quaife: Automated development of Tarski's geometry, *Journal of Automated Reasoning*, 5(1), pp. 97–118 (1989)
- [37] W. Richter: Hilbert Axiom Geometry in HOL Light, <http://github.com/jrh13/hol-light/tree/master/RichterHilbertAxiomGeometry/>
- [38] W. Richter, A. Grabowski, and J. Alama: Tarski geometry axioms, *Formalized Mathematics*, 22(2), pp. 167–176 (2014) <http://dx.doi.org/10.2478/forma-2014-0017>
- [39] P. Rudnicki and A. Trybulec: On the integrity of a repository of formalized mathematics, in Proceedings of Mathematical Knowledge Management 2003, A. Asperti, B. Buchberger, O. Caprotti (eds.) *Lecture Notes in Computer Science*, 2594, pp. 162–174 (2003) http://dx.doi.org/10.1007/3-540-36469-2_13
- [40] W. Schwabhäuser, W. Szmielew, and A. Tarski: *Metamathematische Methoden in der Geometrie*, Springer-Verlag (1983)
- [41] B. Shminke: Routh's, Menelaus' and Generalized Ceva's Theorems, *Formalized Mathematics*, 20(2), pp. 157–159 (2012) <http://dx.doi.org/10.2478/v10037-012-0018-9>
- [42] L.W. Szczerba: Tarski and geometry, *Journal of Symbolic Logic*, 51, pp. 907–912 (1986)
- [43] A. Tarski: What is elementary geometry?, in *Studies in Logic and the Foundations of Mathematics*, North-Holland, pp. 16–29 (1959)
- [44] A. Tarski and S. Givant: Tarski's system of geometry, *The Bulletin of Symbolic Logic*, 5(2), pp. 175–214 (1999)
- [45] W.A. Trybulec: Axioms of incidence, *Formalized Mathematics*, 1(1), pp. 205–213 (1990)
- [46] J. Urban and G. Sutcliffe: Automated reasoning and presentation support for formalizing mathematics in Mizar, *Lecture Notes in Computer Science*, 6167, Springer, 132–146 (2010) http://dx.doi.org/10.1007/978-3-642-14128-7_12
- [47] F. Wiedijk: A synthesis of the procedural and declarative styles of interactive theorem proving, *Logical Methods in Computer Science*, 8(1):30 (2012) [http://dx.doi.org/10.2168/LMCS-8\(1:30\)2012](http://dx.doi.org/10.2168/LMCS-8(1:30)2012)
- [48] F. Wiedijk: Formal proof – getting started, *Notices of the American Mathematical Society*, 55(11), 1408–1414 (2008)
- [49] T. Zurek, Modelling of a fortiori reasoning, *Expert Systems with Applications*, 39(12), 10772–10779 (2012) <http://dx.doi.org/10.1016/j.eswa.2012.02.188>