A Formalization of Metric Spaces in HOL Light

A formalization of metric spaces in HOL Light

Marco Maggesi*
Dipartimento di Matematica e Informatica "Ulisse Dini"
University of Florence, Italy

March 2017

Abstract

We present a computer formalization of metric spaces in the HOL Light theorem prover. Basic results of the theory of complete metric spaces are provided, including the Banach Fixed-Point Theorem, the Baire Category Theorem and the completeness of the space of continuous bounded functions. A decision procedure for a fragment of the elementary theory of metric spaces is also implemented. As an application, the Picard-Lindelöf theorem on the existence of the solutions of ordinary differential equations is proved by using the well-known argument which appeals to the Banach theorem.

1 Introduction

Metric spaces constitute a fundamental concept in several mathematical fields like geometry, topology, and analysis. In this paper we introduce the definition of metric space in HOL Light and we provide some classical results and applications.[1]

1.1 Background

Several important results of the theory of metric spaces have been already formalized by Harrison in HOL Light in the special case of the standard metrics on Euclidean spaces [Har05], like the Banach Fixed-Point Theorem and the Baire Category Theorem.

The main point of the present work is to set up in HOL Light a general theory of *abstract* metric spaces in which the theorems of metric geometry can be stated and exploited in their full generality, as it is required for their application in various fields, such as algebra or functional analysis.

HOL Light also provides a decision procedure, called `NORM_ARITH`, for a fragment of the theory of Euclidean vector spaces, which automates certain proofs involving reasoning with metric concepts such as the triangular law. The underlying theory has been developed by Solovay, Arthan and Harrison in [SAH12]. In the same paper, an analogous decidability result for the elementary theory of metric spaces is also given, which has been used in this work as a basis to implement a decision procedure, `METRIC_ARITH`, for metric spaces.

Part of the code presented in this paper is based on a previous work of Claudia Carapelle, who developed the definition of metric spaces in HOL Light as part of the dissertation for her master's thesis in Mathematics at the University of Florence [Car11].

1.2 Related work

Being such a fundamental tool, the theory of metric spaces appears in several formalization efforts with a variable degree of details and generality.

*The author has been supported by INdAM-GNSAGA and MIUR.

[1] Our code has been included in the HOL Light distribution and is also available from its original repository at `https://bitbucket.org/maggesi/metric/` .

The QED Manifesto [QED94] cites (in section 2, Reply to Objection 11), among several other examples, the Banach fixed point theorem and the Picard-Lindelöf as successful mechanizations of non-trivial mathematical results, witnessing that computer formalization of metric spaces dates before 1994. The Manifesto does not provide precise references, but interesting achievements can be found in the Mizar Mathematical Library [dlC91], in the IMPS Library [FT91, FGJT92, FGT93] and in the LEGO theorem prover [Jon93].

In more recent years we can find a number of other examples. Some of them are hard to compare with our work because they build on setups with deep methodological and foundational differences with respect to ours. E.g., we can cite Madsen's proof of the Banach Fixed-Point Theorem [Mad09], which has been done in Metamath, a system that does not provide proof automation, using an untyped foundation (Zermelo–Fraenkel set theory).

As mentioned above, important concepts about the geometry of complete metric space have already been formalized by Harrison in HOL Light for Euclidean spaces. Harrison's technique has been used to build an analogous library in Isabelle/HOL, which has been later extended to include a general theory of metric spaces. In this setting, Immler and Hölzl [IH12] give a proof of the Picard-Lindelöf theorem with a constructive approach, thus providing a numerical approximation method for the solution of ordinary differential equations. Their implementation choices are different from the ones taken in this work. The details are discussed in the next session.

Another work with a strong emphasis on constructivity has been carried out in Coq by Makarov and Spitters in [MS13], where they provide an intuitionistic proof of the Picard-Lindelöf theorem.

2 Metric spaces in Higher-Order Logic

2.1 Introducing a new HOL type for metric spaces

This section discusses some crucial implementation choices about the very definition of metric space in Higher-Order Logic.

In mainstream mathematics, with the language of sets, a metric space M is defined as a pair (X, d) where X is a set of *points* and d is a *metric* on X, that is, a function $d\colon X \times X \to \mathbb{R}$ such that the following holds: for any three points x, y, z of X

1. $d(x, y) \geq 0$,
2. $d(x, y) = 0$ if and only if $x = y$,
3. $d(x, y) = d(y, x)$,
4. $d(x, z) \leq d(x, y) + d(y, z)$.

Our first task is to translate the previous definition from Set Theory to Higher-Order Logic. This immediately presents us with a number of possible design options.

Possibly, the simplest choice would be to introduce a predicate `‘is_metric d‘` that expresses the fact that $d : X \times X \to \mathbb{R}$ is a metric on the type X:[2]

```
let is_metric = new_definition
  ‘is_metric(d:X×X->real) ⟺
   (∀x y. &0 <= d(x,y)) ∧
   (∀x y. d(x,y) = &0 ⟺ x = y) ∧
   (∀x y. d(x,y) = d(y,x)) ∧
   (∀x y z. d(x,z) <= d(x,y) + d(y,z))‘;;
```

The previous definition can be referred to as a formalization of *total* metric spaces, by the fact that the distance function is total, that is, the space of points coincides with the whole 'carrier' type `‘:X‘`.

This is, in essence, the approach followed in the `Multivariate_Analysis` library of Isabelle/HOL, where this method is especially convenient, since it allows one to exploit the

[2] The symbol `‘&‘` is the embedding $\mathbb{N} \to \mathbb{R}$.

mechanism of axiomatic classes provided by the system. It also has been adopted in a earlier formalization of real analysis in HOL Light[3].

The above definition is clear and simple. However, in practice, we have important examples of metric spaces where the set of points X arises naturally as a subset of some *ambient* type, instead as a type in its own. We can cite, for instance, the Poincaré disk $\Delta \subset \mathbb{R}^2$. Since HOL does not have a notion of subtype, such situation cannot be represented directly.

In most cases this problem can be circumvented by introducing, for each metric space we want to consider, a corresponding type for its set of points. This is, for instance, how the metric of bounded continuous functions are implemented by Immler and Hölzl [IH12]. One drawback of this approach is that it forces us to translate the points of the space back and forth from the original representation to new type of *points*.

More importantly, we also have significant examples of *families* of metric spaces which cannot be formalized with this approach. One paradigmatic example is the class of L^p function spaces:[4] since they form an infinite family of spaces of functions (one for every p in$[1, \infty]$) they cannot be represented as total metric spaces, since HOL lacks dependent types.[5]

If we want to model these examples in a natural way, we are forced to take into account the fact that the metric d must be considered as a *partial* function on the product type `‘:X×X‘`.

We considered a few possible alternatives to describe the domain of the metric d, that is, the set of points. One feasible choice would be to introduce some conventions to deduce the domain from the values of the distance function d. For instance, we could decide that, outside of its intended domain, the function d takes a negative value, say -1. This would be certainly a sound approach, but has the drawback that the metric axioms gets polluted by this exotic behavior and the initial definition loses its original elegance. Moreover, encoding the carrier set of the space in the metric also means that, when we are about to consider subspaces, we have to modify the metric. This lack of monotonicity would be a further source of complications.

In the end, we decided to follow the most direct path, that is, formulate a definition that involves explicitly a set of point (that is, a predicate `‘s:X->bool‘`):

```
let is_metric_space = new_definition
  ‘is_metric_space (s,d) ⟺
   (∀x y:X. x ∈ s ∧ y ∈ s ⟹ &0 <= d(x,y)) ∧
   (∀x y. x ∈ s ∧ y ∈ s ⟹ (d(x,y) = &0 ⟺ x = y)) ∧
   (∀x y. x ∈ s ∧ y ∈ s ⟹ d(x,y) = d(y,x)) ∧
   (∀x y z. x ∈ s ∧ y ∈ s ∧ z ∈ s
            ⟹ d(x,z) <= d(x,y) + d(y,z))‘;;
```

The resulting definition reflects perfectly the traditional one. Notice that for elements outside the domain s, nothing can be deduced about the behavior of the metric d.

We considered the idea of imposing an additional requirement: that the metric d assumes a fixed value outside the domain s. This would be useful to establish later a more terse extensional principle for metric spaces. But we avoided doing that because we did not have any immediate need for it and we felt that it would make the formalization of the basics of the theory unnecessarily more complicated.

The predicate `‘is_metric_space‘` is sound from the semantic point of view, but it is not yet very practical in several circumstances. The problem is that this notion of metric space is given by a pair of terms, s and d, that satisfies a predicate. Instead we would like to identify a metric space with a single HOL term as it is customary in Set Theory or in informal language.

Thus we introduce a new type `‘:(X)metric‘` whose inhabitants represent all the metric spaces with points in the type X. Then we define two projections `‘mspace‘` and `‘mdist‘`, which give respectively the support space and the metric associated to a space. Their characterizing property is the following theorem

[3]See file `Library/analysis.ml` of the HOL Light distribution. However, this code is now superseded by (and incompatible with) the much richer formalization in `Multivariate`

[4]On the other hand, this example of L^p function spaces suggests us considering also the definition of *pseudo*metric space, i.e., when reflexivity ($d(x, x) = 0$) is assumed, but the *identity of indiscernibles* ($d(x, y) = 0 \implies x = y$) is not. However, this is beyond the scope of the present work.

[5]Notice that *Harrison's trick* [Har05] of using a type parameter to emulate certain features of dependent types does not extend to the case of parameters ranging over an uncountable set.

```
⊢ ∀s d. is_metric_space(s,d)
        ⟹ mspace(metric(s,d)) = s ∧ mdist(metric(s,d)) = d
```

which guarantees that our construction has the intended meaning. From this we also get the fundamental properties of metric spaces:

```
⊢ ∀m x y. x ∈ mspace m ∧ y ∈ mspace m ⟹ &0 <= mdist m (x,y)

⊢ ∀m x y. x ∈ mspace m ∧ y ∈ mspace m
          ⟹ (mdist m (x,y) = &0 ⟺ x = y)

⊢ ∀m x y. x ∈ mspace m ∧ y ∈ mspace m
          ⟹ mdist m (x,y) = mdist m (y,x)

⊢ ∀m x y z. x ∈ mspace m ∧ y ∈ mspace m ∧ z ∈ mspace m
            ⟹ mdist m (x,z) <= mdist m (x,y) + mdist m (y,z)
```

2.2 Basic examples of metric spaces

As we mentioned before, Euclidean spaces $\mathbb{R}^N$ are already defined in HOL Light. They can be easily imported in our framework by the following definition:

```
let euclidean_metric = new_definition
  `euclidean_metric = metric((:real^N), dist)`;;
```

Here the notation `(:real^N)` denotes the trivial subset of all elements of the type `:real^N` and the term `dist` is the usual Euclidean distance, already defined in HOL Light.

Analogously, here is the definition of the Poincaré model (in dimension N):

```
let poincare_disk = new_definition
  `metric ({x:real^N | norm x < &1},
           λx y. &2 * dist(x,y) pow 2 /
                 ((1 - norm x pow 2) * (1 - norm y pow 2)))`;;
```

Another cheap, but important, example is the subset of a metric space, which is by restriction of the distance function, a metric space in its own. As stressed above, the notion of metric space adopted in this work is conceived to be convenient when we come to consider subsets. Given a metric space `m:(A)metric` and a set `s:A->bool`, we can form the subspace `submetric m s` which enjoys the properties:

```
⊢ ∀m s. mspace (submetric m s) = s ∩ mspace m

⊢ ∀m s. mdist (submetric m s) = mdist m
```

Notice that, in practical situations, `s` is expected to be a subset of the set of points of `m`. In that case, the former relation reduces to the following simpler and directed one: `mspace (submetric m s) = s`. However the definition we adopted makes sense in the general situation.

A further example of metric space arising in this work is the one of bounded metric functions, described in section 5.4.

3 Metric-induced topology

Metric spaces are one of the prominent examples of topological spaces. In this section we give a brief account of the construction in HOL of the topology induced by a metric. We also overview the proofs of the most basic topological properties of metric spaces.

3.1 The topology induced by a metric

The notion of topology, is already present in HOL Light as part of Harrison's formalization of Multivariate Real Analysis [Har05]. There is also an alternative library (also included in the HOL Light distribution) recently developed by Richter that aims at providing a readable formal development of topological spaces. We decided to refer to Harrison's library partly because we already used it in the past and we were already familiar with it, but also because it seems, at the moment, better integrated in the rest of the standard library.

That said, our formalization follows the usual practice. We first give a definition of open balls:

```
let mball = new_definition
  `mball m (x:X,r) =
     {y | x ∈ mspace m ∧ y ∈ mspace m ∧ mdist m (x,y) < r}`;;
```

notice that when the center of the ball x is outside the support of the space our definition gives the empty set, because the set specification contains the condition `` `x ∈ mspace m` ``. This ensures a sensible definition in all cases.

We also observe that our definition of ball plays nicely with subspaces:

```
⊢ ∀m s x r. x ∈ s ⟹ mball (submetric m s) (x,r) = mball m (x,r) ∩ s
```

Then we define the metric topology through the collection of its open sets in the obvious way. As a result we get a function `` `mtopology:(X)metric->(X)topology` `` together with the expected properties:

```
⊢ ∀m. topspace (mtopology m) = mspace m
```

```
⊢ ∀m u.
      open_in (mtopology m) u ⟺
      u ⊆ mspace m ∧
      (∀x. x ∈ u ⟹ ∃r. &0 < r ∧ mball m (x,r) ⊆ u)
```

As an example, we prove that the Euclidean topology of $\mathbb{R}^N$, already defined in HOL, corresponds to our definition of metric topology of the Euclidean metric:

```
⊢ mtopology euclidean_metric = euclidean:(real^N)topology
```

We also verify that the notions of metric subspace and subtopology are coherent in the obvious sense:

```
⊢ ∀m s. mtopology(submetric m s) = subtopology(mtopology m) s
```

3.2 Limits and continuity

The HOL Light library provides a very rich collection of results about limits and continuity in the special case of the Euclidean spaces $\mathbb{R}^N$. Unfortunately, these results cannot be used in our more general context of topological spaces and metric spaces. We thus provide our own formalization of the notion of limit and continuity as it is usually done in General Topology.

We follow, whenever it is possible, the style used by Harrison for the multivariate development [Har05]. In particular, our notion of limit makes use of the concept of *net*. The resulting definition is as follows:

```
let limit = new_definition
  `limit top (f:A->B) l net ⟺
   l ∈ topspace top ∧
   (∀u. open_in top u ∧ l ∈ u ⟹ eventually (\x. f x ∈ u) net)`;;
```

We can easily prove that this new definition is equivalent to the expected one when it is instantiated to the case of metric spaces. For instance, from the above definition and the following theorem

```
⊢ ∀p. eventually p sequentially ⟺ (∃N. ∀n. N <= n ⟹p n)
```

we obtain the classical formulation for the limit of sequences

```
⊢ ∀m f l.
      limit (mtopology m) f l sequentially ⟺
      l ∈ mspace m ∧
      (∀e. &0 < e
           ⟹ (∃N. ∀n. N <= n
                       ⟹ f n ∈ mspace m ∧ mdist m (f n,l) < e))
```

and that it readily generalizes the already established case of Euclidean spaces:

```
⊢ ∀f x net. limit euclidean f x net ⟺ (f --> x) net
```

Several topological notions admit interesting reformulations in the special case of metric spaces. We provide some of these characterizations. For instance, a set is closed in a metric space if and only if it is *sequentially closed*:

```
⊢ ∀m s.
      closed_in (mtopology m) s ⟺
      s ⊆ mspace m ∧
      (∀a l. (∀n. a n ∈ s) ∧ limit (mtopology m) a l sequentially
             ⟹ l ∈ s)
```

Once we have laid down the essential results about limits, we can give the classical definition of continuity between topological spaces together with its usual reformulations for the case of metric spaces. Following the mainstream practice, we say that a function f is continuous if the preimage of an open set through f is open:

```
let topcontinuous = new_definition
  `topcontinuous top top' (f:A->B) ⟺
   (∀x. x ∈ topspace top ⟹ f x ∈ topspace top') ∧
   (∀u. open_in top' u
        ⟹ open_in top {x | x ∈ topspace top ∧ f x ∈ u})`;;
```

As done before, we certify that our notion generalizes the case of Euclidean spaces. It is also useful to give the usual notion of pointwise continuity and have a collection of lemmata that relate all these notions of continuity and limits. To give just a couple of examples, the following is the result that links continuity and pointwise continuity:

```
⊢ ∀top top' f.
      topcontinuous top top' f ⟺
      (∀x. x ∈ topspace top ⟹ topcontinuous_at top top' f x)
```

and here we prove that continuous functions preserve limits:

```
⊢ ∀net top top' f g l.
      topcontinuous top top' g ∧ limit top f l net
      ⟹ limit top' (g o f) (g l) net
```

3.3 Compactness

Another fundamental topological notion which plays an important role in the study of metric space is compactness. As for the case of limits and continuity discussed in the previous section, the notion of compact set is defined in the standard library of HOL Light only in the case of Euclidean spaces.

We introduced the corresponding definition in the case of general topological spaces with the usual approach: a set is compact if every open cover admits a finite subcover:

```
let compact_in = new_definition
  `∀top s:X->bool.
     s compact_in top ⟺
     s ⊆ topspace top ∧
     (∀U. (∀u. u ∈ U ⟹ open_in top u) ∧ s ⊆ UNIONS U
          ⟹ (∃V. FINITE V ∧ V ⊆ U ∧ s ⊆ UNIONS V))`;;
```

Notice that it is more usual to define the notion of compactness for the entire topological space. Here we prefer to give the equivalent notion for the subsets of the space, since it seems generally more convenient in practice.

We prove some basic results about compact sets. For instance, the following theorems state that: (i) every closed subset of a compact set is compact; (ii) a compact set in a metric space is bounded; (iii) continuous functions preserve compactness; (iv) our definition of compactness coincides with the one already given for the Euclidean case:

```
⊢ ∀top k c. k compact_in top ∧ c ⊆ k ∧ closed_in top c
            ⟹ c compact_in top

⊢ ∀m s. s compact_in (mtopology m) ⟹ mbounded m s

⊢ ∀top top' f s.
     s compact_in top ∧ topcontinuous top top' f
     ⟹ (IMAGE f s) compact_in top'

⊢ ∀s. s compact_in euclidean ⟺ compact s
```

4 Implementation of a decision procedure

Simple statements in the theory of metric spaces are often geometrically intuitive and, in such a case, they are sometimes asserted without proof or justified with a drawing. These informal arguments have to be translated into rigorous proofs that can be accepted by the system, a task which is frequently time consuming and mathematically uninteresting.

For this reason, following Solovay, Arthan and Harrison [SAH12], we implemented a HOL rule, called `METRIC_ARITH`, that can prove automatically certain simple theorems in the theory of elementary metric spaces.

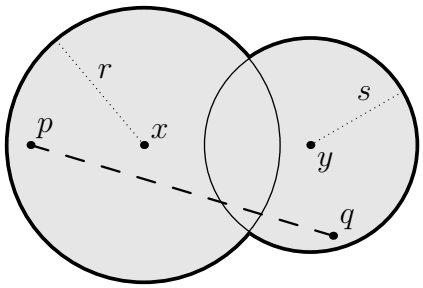
Our procedure is entirely implemented in ML; no external tool is required. It takes as input a HOL term, which is the statement to be proved, and either returns the corresponding HOL theorem or fails. We also provide an associated tactic, `METRIC_ARITH_TAC`, which uses the procedure to solve the current goal in an interactive proof session.

In what follows, we consider the language of first-order logic with atoms of one of the following three forms:

- $x \in M$ (the point x belongs to the given metric space M);
- $x = y$ (x and y are the same point);
- arithmetic equalities and inequalities involving the distance between points (as $d(x, y) < d(x, z) + 3$).

4.1 Examples of use of `METRIC_ARITH`

Consider the following proposition: given B_1, B_2 two intersecting open balls of radius r, s respectively, the diameter of their union $B_1 \cup B_2$ is less than $2(r + s)$.



The above statement, formalized in our framework, is:

```
# g `∀m x y:A r s.
       ¬(DISJOINT (mball m (x,r)) (mball m (y,s)))
       ⟹ ∀p q. p ∈ mball m (x,r) ∪ mball m (y,s) ∧
                q ∈ mball m (x,r) ∪ mball m (y,s)
                ⟹ mdist m (p,q) < &2 * (r + s)`;;
```

Let us show how we can prove this goal using our decision procedure. This requires just two steps in this case. First, we use rewriting to eliminate set-theoretical operations

```
# e (REWRITE_TAC[DISJOINT; FORALL_IN_UNION; IMP_CONJ;
       RIGHT_FORALL_IMP_THM; EXTENSION; IN_INTER;
       NOT_IN_EMPTY; IN_MBALL]);;
```

This reduces our goal to the following sentence in the language of elementary metric spaces:

```
val it : goalstack = 1 subgoal (1 total)

`∀m x y r s.
   ~(∀x'. ~((x ∈ mspace m ∧ x' ∈ mspace m ∧ mdist m (x,x') < r) ∧
            y ∈ mspace m ∧
            x' ∈ mspace m ∧
            mdist m (y,x') < s))
   ⟹ (∀p. x ∈ mspace m ∧ p ∈ mspace m ∧ mdist m (x,p) < r
           ⟹ (∀q. x ∈ mspace m ∧ q ∈ mspace m ∧ mdist m (x,q) < r
                  ⟹ mdist m (p,q) < &2 * (r + s)) ∧
              (∀q. y ∈ mspace m ∧ q ∈ mspace m ∧ mdist m (y,q) < s
                  ⟹ mdist m (p,q) < &2 * (r + s))) ∧
       (∀p. y ∈ mspace m ∧ p ∈ mspace m ∧ mdist m (y,p) < s
           ⟹ (∀q. x ∈ mspace m ∧ q ∈ mspace m ∧ mdist m (x,q) < r
                  ⟹ mdist m (p,q) < &2 * (r + s)) ∧
              (∀q. y ∈ mspace m ∧ q ∈ mspace m ∧ mdist m (y,q) < s
                  ⟹ mdist m (p,q) < &2 * (r + s)))`
```

Then, we can use our tactic to automatically solve the goal:

```
# e METRIC_ARITH_TAC;;
# top_thm();;
```

Another interesting example is the *reverse triangle inequality*

$$|d(x,y) - d(y,z)| \leq d(x,z)$$

which `METRIC_ARITH` can prove directly:

```
# METRIC_ARITH
    `∀m x y z:A.
       x ∈ mspace m ∧ y ∈ mspace m ∧ z ∈ mspace m
       ⟹ abs (mdist m (x,y) - mdist m (y,z)) <= mdist m (x,z)`;;
val it : thm =
  |- ∀m x y z.
       x ∈ mspace m ∧ y ∈ mspace m ∧ z ∈ mspace m
       ⟹ abs (mdist m (x,y) - mdist m (y,z)) <= mdist m (x,z)
```

4.2 How `METRIC_ARITH` works

We now describe the principle of operation of our decision procedure `METRIC_ARITH`. More details and further discussions about decidability and undecidability results for the theory of metric spaces can be found in Section 4 of [SAH12].

Definition 1. A sentence (of first-order logic) in prenex normal form is said to be ∀∃ (or AE) if no universal quantifier occurs in the scope of an existential one, i.e., it is of the form:

$$\forall x_1 \dots x_m . \exists y_1 \dots y_m . \psi$$

for some $m \geq 0$ and $n \geq 0$. The set of $\exists\forall$ sentences is defined analogously exchanging '$\forall$' with '$\exists$'.

The Bernays-Schönfinkel theorem asserts that the set of valid $\forall\exists$ sentences without function symbols is decidable. In fact, if ϕ is one such sentence, then its Skolemization has no function symbols except nullary ones (i.e., constants), thus the set of Herbrand's interpretation of ϕ is finite. Hence the satisfiability of $\exists\forall$ sentences is checkable algorithmically. By taking the negation, the set of valid $\forall\exists$ sentences is also decidable.

In the case of metric spaces, we can consider a more general fragment of the language that can contain the distance d and the arithmetic operations as function symbols and is slightly more flexible on the order of quantifiers.

Definition 2. We say a sentence of the language of metric spaces is $\forall\exists_p$ if it is prenex and no universal quantifier over points is in the scope of an existential quantifier (of any sort).

Again, the set of $\exists\forall_p$ sentences is defined analogously exchanging '$\forall$' with '$\exists$'.

Solovay, Arthan and Harrison prove in [SAH12] (see Theorem 8) that the set of valid $\forall\exists_p$ sentences is decidable. Our procedure follows their proof. However, for efficiency reasons, no non-trivial nonlinear reasoning is performed, which means essentially that our procedure will succeed only if the goal involves, after suitable normalizations, only linear equalities and inequalities (see Step 4 and Remark 3 below).

Then let ϕ be a $\forall\exists_p$ sentence in the language of metric spaces. Since universal quantifiers commute up to logical equivalence, any $\forall\exists_p$ sentence ϕ can be assumed, without loss of generality, to be formed of an initial block of $n \geq 0$ universal quantifiers over points followed by a block comprising existential quantifiers over points and scalar quantifiers of either kind. We write this as follows:

$$\phi \equiv \forall x_1 \ldots x_n.\, \exists\, \bar{y}/\mathsf{Q}\bar{z}.\, \psi \tag{1}$$

where ψ is a quantifier-free, the variables x_i, y_j range over points, the variables z_k range over reals and the set of free variables of ψ is contained in

$$\{x_1, \ldots, x_n, y_1, \ldots, y_k, z_1, \ldots, z_l\}.$$

We are now ready to outline the essential steps of the `METRIC_ARITH` procedure. Let ϕ be a $\forall\exists_p$ sentence as before, M`:(`α`)metric` a metric space in HOL and ϕ_M be the HOL term which is the interpretation of ϕ on M. Our procedure will try to prove ϕ_M; however, as it will become clear shortly, it fails unless ϕ is logically valid (i.e., valid on any nonempty metric space).

Step 1 We start by replacing every subformula of ϕ of the form $\exists y.\, \rho$ with $\rho[x_1/y] \vee \cdots \vee \rho[x_n/y]$. We obtain a formula $\phi' \equiv \forall x_1 \ldots x_n \mathsf{Q}\bar{z}.\, \psi'$ with no existential quantifiers on points. The system can easily prove the theorem $\vdash \phi'_M \implies \phi_M$ and then proceed to seek a proof of ϕ'_M.

Notice, that the converse implication is not true in general. See Remark 1.

Step 2 The procedure attempts to prove a theorem of the form

$$\vdash (x_1 \in M \wedge \cdots \wedge x_n \in M \implies \mathsf{Q}\bar{z}.\, \psi''_M) \implies \mathsf{Q}\bar{z}.\, \psi'_M$$

where ψ'' is obtained by replacing the atoms $x_i \in M$ with true in ψ'. The generation of this theorem may fail, see the example discussed in Remark 2.

Step 3 Now consider the finite metric space $M' = \{x_1, \ldots, x_n\}$ and the function $f : M' \longrightarrow \mathbb{R}^n$ defined by $f(p) = (d(p, x_1), \ldots, d(p, x_n))$. It is easy to prove that f is an isometry between M' and $(\mathbb{R}^n, d_\infty)$, where d_∞ is the metric $d_\infty(u, v) = \max\{|u_i - v_i| : 1 \leq i \leq n\}$. This is formalized by the following theorem:

```
⊢ ∀m s x y.
     mbounded m s ∧ x ∈ s ∧ y ∈ s
     ⟹ mdist m (x,y) =
         sup (IMAGE (λa. abs(mdist m (x,a) - mdist m (a,y))) s)
```

Then, we can prove our formula on the model $(\mathbb{R}^n, d_\infty)$ instead on M. More precisely, we introduce the abbreviations $x_{ij} = d(x_i, x_j) = d(x_j, x_i)$ for all $i \leq j$ and we replace each subterm of the form $d(x_s, x_t)$ with $\max\{|x_{s1} - x_{t1}|, \ldots, |x_{sn} - x_{tn}|\}$ and each term of the form $x_s = x_t$ with $x_{s1} = x_{t1} \wedge \cdots \wedge x_{sn} = x_{tn}$.

Step 4 At the end of this process, we have a sentence with no point variables, and we use the `REAL_ARITH` procedure to terminate the proof. This may fail if the final goal is not a theorem of the language of linear real arithmetic (see Remark 3).

Remark 1. The reduction from ϕ to ϕ' in Step 1, is motivated by a semantic argument. In fact, ϕ it is logically valid if and only if ϕ' is. The proof follows from the basic observation that $M' = \{x_1, \ldots, x_n\}$ is a space with the metric inherited from M. Thus, if ϕ is logically valid, it must be valid on the model M'. (See Corollary 7 in Section 4.1 in [SAH12] for a detailed proof.)

Remark 2. The goal

```
‘∀x:real^N. mdist euclidean_metric (x,x) = &0‘
```

is provable, as it can be immediately reduced to the equivalent goal

```
‘∀x:real^N. dist(x,x) = &0‘
```

However, our procedure is not able to prove it since it cannot find that x belongs to the metric space.

Indeed, the formula `‘∀x:A. mdist m (x,x) = &0‘` is not valid, since, as we already stressed, `‘mspace m‘` may be a proper subset of the type `‘:A‘` and nothing can be deduced about a point outside the support of the metric space.

To make the goal suitable for our procedure, the membership condition must be given explicitly, even if it is trivial in this case:

```
‘∀x:real^N. x ∈ mspace euclidean_metric
            ⟹ mdist euclidean_metric (x,x) = &0‘
```

Remark 3. The `REAL_ARITH` procedure used in the final Step 4, is basically a solver for linear analysis based on the Fourier-Motzkin elimination. In principle, we could use a complete decision procedure for the theory of real closed fields, like Tarski's quantifier elimination algorithm. This would also provide us with a constructive proof of the decidability of $\forall\exists_p$ sentences in the language of metric spaces, as done in [SAH12], Theorem 8. However, in practice, it is preferable to employ a much more efficient and faster, although incomplete, procedure.

5 Sequences and complete metric spaces

In most practical situations, several topological and metric phenomena can be studied and understood through the behavior of *sequences*. From a technical point of view, a sequence is nothing more than a function over the natural numbers and this is how they are introduced and used in HOL Light. In informal reasoning, for a sequence s we will write s_n instead of $s(n)$ to denote the *n-th element* of the sequence as it is customary in traditional mathematical texts.

5.1 Cauchy sequences and complete metric spaces

A crucial notion in metric space is the one of *Cauchy sequence*, which means that the distance between its elements become arbitrary small after a certain index. Here is our formal definition:

```
let cauchy_in = new_definition
  ‘∀m:A metric s:num->A.
     cauchy_in m s ⟺
     (∀n. s n ∈ mspace m) ∧
     (∀e. &0 < e
          ⟹ (∃N. ∀n n'. N <= n ∧ N <= n'
                         ⟹ mdist m (s n,s n') < e))‘;;
```

It is an easy but fundamental fact that every convergent sequence is a Cauchy sequence. A metric space is said to be *complete* if the converse is also true:

```
let mcomplete = new_definition
  `∀m:A metric.
     mcomplete m ⟺
     (∀s. cauchy_in m s
          ⟹ ∃x. limit (mtopology m) s x sequentially)`;;
```

For Euclidean space, completeness is very simple: a subset of $\mathbb{R}^N$ is complete if and only if it is closed. This has already proved in the standard library:

```
⊢ ∀s. complete s ⟺ closed s
```

In the general case (e.g. in infinite-dimensional Banach spaces and Hilbert spaces), completeness is a much stronger property than being closed. We prove some basic results. For instance, a closed subset of a complete metric space is complete:

```
⊢ ∀m s. closed_in (mtopology m) s ∧ mcomplete m
        ⟹ mcomplete (submetric m s)
```

5.2 The Banach fixed-point theorem

We are now ready to state and prove the Banach fixed-point theorem (also known as the Contraction Mapping Principle), a significant result of the theory of complete metric spaces, which has various noteworthy applications. One such application to the theory of ordinary differential equations is presented later in this paper, in section 6.

Let us start by recalling the necessary definitions and the statement of the theorem. A map between metric spaces $f : M \to N$ is said to be *Lipschitzian* if there exists a real constant k such that

$$d(f(x), f(y)) \leq kd(x, y)$$

for all x, y in M. When M and N are the same space and $k < 1$ we say that f is a *contraction*.

Theorem 1 (Banach)**.** *Every contraction $f : M \to M$ on a non empty, complete metric space M has an unique fixed-point.*

The idea of the proof is as follows. Take an arbitrary point x_0 of M and consider the iterated sequence $x_n = f^n(x_0)$ (i.e., the point obtained by applying n times f to x_0). Such sequence is Cauchy and its limit is the fixed-point of f.

We report the formal statement of the theorem:

```
⊢ ∀m f k.
     ¬(mspace m = ∅) ∧
     mcomplete m ∧
     (∀x. x ∈ mspace m ⟹ f x ∈ mspace m) ∧
     k < &1 ∧
     (∀x y. x ∈ mspace m ∧ y ∈ mspace m
            ⟹ mdist m (f x, f y) <= k * mdist m (x,y))
     ⟹ (∃!x. x ∈ mspace m ∧ f x = x)
```

5.3 The Baire Category Theorem

As a second example, we illustrate the statement of the Baire Category Theorem:

```
⊢ ∀m g. mcomplete m ∧ COUNTABLE g ∧
          (∀t. t ∈ g ⟹ open_in (mtopology m) t ∧
          mtopology m closure_of t = mspace m)
          ⟹ mtopology m closure_of INTERS g = mspace m
```

In the above statement, we have a complete metric space `` `m` `` (whose associated set of points is `` `mspace m` `` and whose associated topology is `` `mtopology m` ``) and a countable family of dense open sets `` `g` ``. The thesis is that the intersection of the family `` `g` `` is dense in `` `m` ``.

5.4 The complete metric space of continuous functions

As we outlined in the introduction, our motivation for moving out of the concrete setting of Euclidean metric and considering the more abstract notion of metric space is to give a suitable framework for certain important examples of 'spaces', for which a profitable notion of distance can be given. In this section we define a metric structure on the set of bounded functions and the set of continuous and bounded functions.

We start by considering functions f that have, over a selected domain S, a bounded codomain in some given metric space M. This is expressed in HOL by the condition `‘mbounded m (IMAGE f s)‘`.

Under this condition, we can consider the *sup-metric* (also called *uniform metric* or L^∞-metric)

$$d(f,g) = \sup_{x \in S} d(f(x), g(x)).$$

One problem with this metric comes from the fact that functions in HOL are total. In fact, the above metric involves only the values that the functions take on the set S. However, to preserve the property of *indiscernibility*, namely that if the distance between two functions is zero, they are the same function, we are forced to set up a mechanism for 'truncating' a function to its 'domain of definition'. The trick is well-known by HOL programmers and has been used by other authors for different purposes, but, to our knowledge, never formalized in HOL Light specifically.

We start by using the Hilbert ε operator to find, for any type `‘:A‘`, a distinguished element that we conventionally call `‘UNDEFINED‘`:

```
let UNDEFINED = new_definition ‘UNDEFINED = (εx:A. F)‘;;
```

Note that nothing non trivial can be proved for this special element. We then use this special element for indicating when a function is not defined. In particular, we say that a function is extensional on a given set if it is 'undefined' elsewhere:

```
let EXTENSIONAL = new_definition
  ‘EXTENSIONAL s = {f:A->B | ∀x. ¬(x ∈ s) ⟹ f x = UNDEFINED}‘;;
```

We stress that it is not possible to recover s from f, because it is perfectly possible that `f x = UNDEFINED` for x in s.[6]

It is also easy to 'restrict' a given function over a certain set:

```
let RESTRICTION = new_definition
  ‘RESTRICTION s (f:A->B) x = if x ∈ s then f x else UNDEFINED‘;;
```

The essential point of this construction is to have an appropriate extensional principle for such *restricted* functions:

```
⊢ ∀s f g. f ∈ EXTENSIONAL s ∧ g ∈ EXTENSIONAL s ∧
          (∀x. x ∈ s ⟹ f x = g x)
          ⟹ f = g
```

With these definitions in place we can give an appropriate definition of the metric space `‘funspace s m‘` of bounded functions from the set `‘s‘` to the metric space `‘m‘`. We thus get the following characterizing theorem:

```
⊢ ∀s m. mspace (funspace s m) =
          {f | (∀x. x ∈ s ⟹ f x ∈ mspace m) ∧ f ∈ EXTENSIONAL s ∧
               mbounded m (IMAGE f s)} ∧
        (∀f g. mdist (funspace s m) (f,g) =
                 if s = ∅ then &0 else
                 sup {mdist m (f x,g x) | x | x ∈ s})
```

We prove some basic but fundamental properties of this construction. For instance, the fact that this space inherits the completeness from the codomain:

[6] One different approach would be to carry s around with f by working systematically with pairs (s, f), as one would do to implement a good model of the category of sets and functions. But this would be needlessly complicated for our purposes here.

```
⊢ ∀s m. mcomplete m ⟹ mcomplete (funspace s m)
```

Several other important examples of metric spaces can be obtained as subsets of those mentioned above. In this paper we consider $C(X, M)$, the metric subspace of bounded and continuous functions from a topological space X to a metric space M. Its definition it is now straightforward:

```
let cfunspace = new_definition
  `cfunspace top m =
   submetric (funspace (topspace top) m)
     {f:A->B | topcontinuous top (mtopology m) f}`;;
```

As for bounded functions, we get an important criterion for completeness:

```
⊢ ∀top m. topspace top compact_in top ∧ mcomplete m
          ⟹ mcomplete (cfunspace top m)
```

6 The Picard-Lindelöf theorem

To conclude our formalization we give one classical application of the Banach fixed-point theorem in analysis: the Picard-Lindelöf theorem (also known as the Cauchy–Lipschitz theorem) about the existence of solutions of ordinary differential equations. Beside being an interesting achievement in its own, we present this further development as a testbed for checking the applicability of our constructions.

An *initial value problem* (or *Cauchy problem*) is a differential equation together with an initial condition:

$$\begin{cases} u'(t) = f(t, (u(t))) \\ u(t_0) = u_0 \end{cases} \tag{2}$$

A *solution* to an initial value problem is a function u satisfying the above two conditions.

Theorem 2. *Let f be a function which is continuous from an open set Ω of $\mathbb{R} \times \mathbb{R}^N$ to $\mathbb{R}^N$ and assume that $f(t, v)$ is Lipschitz on v and bounded, that is, that there are two real constants B and c such that*

$$\|f(s, v)\| \leq B$$
$$\|f(s, v) - f(s, w)\| \leq c\|v - w\|$$

for all points (s, v) and (s, w) of Ω. Then consider two positive numbers r_0, r_1 such that $r_0 < \min(r_1/B, 1/c)$ and $[t_0, t_0 + r_0] \times \overline{B(u_0, r_1)}$ is a subset of Ω. Then there exists an unique solution $u : [t_0, t_0 + r_0] \to \mathbb{R}^N$ of the initial value problem (2).

We used the notation $\overline{B(u_0, r_1)}$ for the closed ball of centre u_0 and radius r_1.

We give a very brief sketch of the proof to make clear the link with theory of complete metric spaces. A more detailed account can be found in the classical references on ordinary differential equations (such as [CL55, Sim78]) and in several undergraduate textbooks on calculus.

We consider the metric space $X = C\big([t_0, t_0 + r_0], \overline{B(u_0, r_1)}\big)$ and the operator H which, at every function $v \in X$, associates the function $y = H(v)$ defined by

$$y(t) = u_0 + \int_{t_0}^{t_0 + r} f(t, v(t))\, \mathrm{d}t.$$

The important observation, which follows from the fundamental theorem of calculus, is that the solutions of the Cauchy problem are precisely the fixed points of the operator H. We want to invoke the Banach fixed-point theorem to conclude the proof. The remaining verifications are not difficult.

Here is the formal statement we obtained in the HOL Light theorem prover corresponding to the above Theorem 2:

```
⊢ ∀s f t0 u0 r0 r1 B c.
    open s ∧ f continuous_on s ∧
    &0 < r0 ∧ &0 < r1 ∧ B * r0 < r1 ∧ c * r0 < &1 ∧
    interval[t0,t0 + lift r0] PCROSS cball(u0,r1) ⊆ s ∧
    (∀x. x ∈ s ⟹ norm(f x) <= B) ∧
    (∀t v w. t ∈ interval[t0,t0 + lift r0] ∧
             v ∈ cball(u0,r1) ∧ w ∈ cball(u0,r1)
             ⟹ norm(f(pastecart t v) - f(pastecart t w)) <=
                 c * norm(v - w))
    ⟹ ∃u. u t0 = u0 ∧
           (∀t. t ∈ interval[t0,t0 + lift r0]
                ⟹ (u has_vector_derivative f(pastecart t (u t)))
                     (at t within interval[t0,t0 + lift r0])) ∧
           (∀v. (∀t. t ∈ interval[t0,t0 + lift r0]
                     ⟹ pastecart t (v t) ∈ s) ∧
                v t0 = u0 ∧
                (∀t. t ∈ interval[t0,t0 + lift r0]
                     ⟹ (v has_vector_derivative
                          f(pastecart t (v t)))
                          (at t within interval[t0,t0 + lift r0]))
                ⟹ (∀t. t ∈ interval[t0,t0 + lift r0]
                       ⟹ v t = u t))
```

7 Conclusions

Metric spaces are an indispensable tool in modern mathematics. We introduced a definition of metric space in Higher-Order Logic, which allows us to state and prove theorems about metric geometry in their full generality. We implemented a simple decision procedure, we proved some notable results about complete metric spaces and we gave some basic applications to functional analysis and ordinary differential equations. Our hope is that the present work can lay a foundation for the theory of metric spaces in the HOL Light theorem prover.

References

[Car11] Claudia Carapelle. La formalizzazione degli spazi metrici e teorema di punto fisso nella logica di ordine superiore. Master's thesis, Corso di Laurea Specialistica in Matematica, Università degli Studi di Firenze, 2011.

[CL55] Earl A. Coddington and Norman Levinson. *Theory of ordinary differential equations*. International series in pure and applied mathematics. McGraw-Hill, 1955.

[dlC91] Alicia de la Cruz. Totally bounded metric spaces. *Journal of Formalized Mathematics*, 2(4), September-October 1991.

[FGJT92] William M. Farmer, Joshua D. Guttman, and F. Javier Thayer. *Little theories*, pages 567–581. Springer Berlin Heidelberg, Berlin, Heidelberg, 1992.

[FGT93] William M. Farmer, Joshua D. Guttman, and F. Javier Thayer. Imps: An interactive mathematical proof system. *Journal of Automated Reasoning*, 11(2):213–248, 1993.

[FT91] William M Farmer and F Javier Thayer. Two computer-supported proofs in metric space topology. *Notices of the American Mathematical Society*, 38(9):1133–1138, 1991.

[Har05] John Harrison. A HOL theory of Euclidean space. In Joe Hurd and Tom Melham, editors, *Theorem Proving in Higher Order Logics, 18th International Conference, TPHOLs 2005*, volume 3603 of *Lecture Notes in Computer Science*, pages 114–129, Oxford, UK, August 2005. Springer-Verlag.

[IH12] Fabian Immler and Johannes Hölzl. Numerical analysis of ordinary differential equations in Isabelle/HOL. In Lennart Beringer and Amy Felty, editors, *Interactive Theorem Proving*, volume 7406 of *Lecture Notes in Computer Science*, pages 377–392. Springer Berlin Heidelberg, 2012.

[Jon93] Claire Jones. Completing the rationals and metric spaces in LEGO. In *Papers Presented at the Second Annual Workshop on Logical Environments*, pages 297–316, New York, NY, USA, 1993. Cambridge University Press.

[Mad09] Jeff Madsen. Proof of Banach Fixed-Point Theorem in Metamath. From the metamath website: `http://us.metamath.org/mpegif/bfp.html`, September 2009.

[MS13] Evgeny Makarov and Bas Spitters. The Picard algorithm for ordinary differential equations in Coq. In Sandrine Blazy, Christine Paulin-Mohring, and David Pichardie, editors, *Interactive Theorem Proving*, volume 7998 of *Lecture Notes in Computer Science*, pages 463–468. Springer Berlin Heidelberg, 2013.

[QED94] The QED Manifesto. In *Proceedings of the 12th International Conference on Automated Deduction*, CADE-12, pages 238–251, London, UK, UK, 1994. Springer-Verlag.

[SAH12] Robert M. Solovay, R.D. Arthan, and John Harrison. Some new results on decidability for elementary algebra and geometry. *Annals of Pure and Applied Logic*, 163(12):1765 – 1802, 2012.

[Sim78] George F. Simmons. *Differential Equations: With Applications and Historical Notes*. Tata McGraw-Hill Publishing Company, 1978.