



**HAL**  
open science

# Hierarchy Builder: algebraic hierarchies made easy in Coq with Elpi

Cyril Cohen, Kazuhiko Sakaguchi, Enrico Tassi

► **To cite this version:**

Cyril Cohen, Kazuhiko Sakaguchi, Enrico Tassi. Hierarchy Builder: algebraic hierarchies made easy in Coq with Elpi. 2020. hal-02478907v3

**HAL Id: hal-02478907**

**<https://inria.hal.science/hal-02478907v3>**

Preprint submitted on 11 Mar 2020 (v3), last revised 23 Sep 2022 (v6)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Hierarchy Builder: algebraic hierarchies made easy in Coq with Elpi

Cyril Cohen

Inria, Université Côte d’Azur, France

Cyril.Cohen@inria.fr

Kazuhiko Sakaguchi

University of Tsukuba, Japan

sakaguchi@logic.cs.tsukuba.ac.jp

Enrico Tassi

Inria, Université Côte d’Azur, France

Enrico.Tassi@inria.fr

## Abstract

It is nowadays customary to organize libraries of machine checked proofs around hierarchies of algebraic structures [2, 6, 8, 16, 18, 23, 27]. One influential example is the Mathematical Components library on top of which the long and intricate proof of the Odd Order Theorem could be fully formalized [14].

Still, building algebraic hierarchies in a proof assistant such as Coq [9] requires a lot of manual labor and often a deep expertise in the internals of the prover [13, 17]. Moreover, according to our experience [26], making a hierarchy evolve without causing breakage in client code is equally tricky: even a simple refactoring such as splitting a structure into two simpler ones is hard to get right.

In this paper we describe *HB*, a high level language to *build* hierarchies of algebraic structures and to make these hierarchies *evolve* without breaking user code. The key concepts are the ones of *factory*, *builder* and *abbreviation* that let the hierarchy developer describe an actual interface for their library. Behind that interface the developer can provide appropriate code to ensure backward compatibility. We implement the *HB* language in the `hierarchy-builder` addon for the Coq system using the Elpi [11, 28] extension language.

**2012 ACM Subject Classification** General and reference → General literature; General and reference

**Keywords and phrases** Algebraic Hierarchy, Packed Classes, Coq, Elpi, Metaprogramming, λProlog

**Digital Object Identifier** 10.4230/LIPIcs.CVIT.2016.23

## 1 Introduction

Modern libraries of machine checked proofs are organized around hierarchies of algebraic structures [2, 6, 8, 16, 18, 23, 27]. For example the Mathematical Components library for the Coq system [9] provides a very rich, ever growing, hierarchy of structures such as group, ring, module, algebra, field, partial order, order, lattice. . . The hierarchy does not only serve the purpose of organizing knowledge, but also to make it easy to exploit it. Indeed the interactive prover can take advantage of the structure of the library and the relation between its concepts to infer part of information usually left implicit by the user, a capability that turned out to be key to tame the complexity and size of the formal proof of the Odd Order Theorem [14].

The hierarchy of the Mathematical Components library is implemented following the discipline of Packed Classes initially introduced in [13] and later also adopted in [2] and [6]. We call Packed Classes a discipline, and not a language, because, in spite of its many virtues, it is unwieldy to use. In particular it leaks to the user many of the technical details of the Coq system. As a result, one needs to be a Coq expert in order to build or modify a hierarchy, and even experts make mistakes as shown in [26]. Another inconvenience of the



© Cyril Cohen and Kazuhiko Sakaguchi and Enrico Tassi;  
licensed under Creative Commons License CC-BY

42nd Conference on Very Important Topics (CVIT 2016).

Editors: John Q. Open and Joan R. Access; Article No. 23; pp. 23:1–23:18

Leibniz International Proceedings in Informatics



LIPICs Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

## 23:2 Hierarchy Builder

46 Packed Classes discipline is that even simple changes to the hierarchy, such as splitting a  
47 structure into two simpler ones, break user code.

48 In this paper we describe  $\mathcal{HB}$ , a high level language to *build* hierarchies of algebraic  
49 structures and to make these hierarchies *evolve* without breaking user code. The key concepts  
50 are the ones of *factory* and *abbreviation* that let the hierarchy developer describe an actual  
51 interface for their library. Behind that interface the developer can provide appropriate code  
52 to ensure backward compatibility. We implement the  $\mathcal{HB}$  language by compiling it to a  
53 variant of the Packed Classes discipline, that we call *flat*, in the `hierarchy-builder` addon  
54 for Coq. We write this addon using the Elpi [11, 28] extension language.

55

56 To sum up, the main contributions of the paper are:

- 57 ■ the design of the  $\mathcal{HB}$  language,
- 58 ■ the compilation of  $\mathcal{HB}$  to the (flat) Packed Classes discipline, and
- 59 ■ the implementation of  $\mathcal{HB}$  in the `hierarchy-builder` addon for Coq.

60 The paper is organized as follows. Via an example we introduce  $\mathcal{HB}$  and its key ideas. We  
61 then describe the discipline of Packed Classes and we show how  $\mathcal{HB}$  can be compiled to it.  
62 We then discuss the implementation of the Coq addon via the Elpi extension language and  
63 we position  $\mathcal{HB}$  in the literature.

### 64 **2** $\mathcal{HB}$ by examples: building and evolving a hierarchy

65 The first version of our hierarchy (that we name V1) features only two structures: `Monoid`  
66 and `Ring`.

```
1  HB.mixin Record Monoid_of_Type A := {
2    zero : A;
3    add : A -> A -> A;
4    addrA : associative add;           (* `add` is associative.           *)
5    addOr : left_id zero add;         (* `zero` is the left and right neutral *)
6    addr0 : right_id zero add;        (* element with respect to `add`.     *)
7  }.
8  HB.structure Definition Monoid := { A of Monoid_of_Type A }.
9
10 HB.mixin Record Ring_of_Monoid A of Monoid A := {
11   one : A;
12   opp : A -> A;
13   mul : A -> A -> A;
14   addNr : left_inverse zero opp add; (* `opp x` is the left and right additive *)
15   addrN : right_inverse zero opp add; (* inverse of `x`.                       *)
16   mulrA : associative mul;           (* `mul` is associative.                 *)
17   mul1r : left_id one mul;           (* `one` is the left and right neutral  *)
18   mulr1 : right_id one mul;         (* element with respect to `mul`.       *)
19   mulrDl : left_distributive mul add; (* `mul` is left and right distributive *)
20   mulrDr : right_distributive mul add; (* over `add`.                          *)
21 }.
22 HB.structure Definition Ring := { A of Monoid A & Ring_of_Monoid A }.
```

67 In order to build a structure we need to declare some factories and later assemble them.  
68 One kind of factory supported by  $\mathcal{HB}$ , the simplest one, is called *mixin* and is embodied by  
69 a record, by convention named `axioms`, that gathers operations and properties.

70 Mixins are declared via the `HB.mixin` command that takes a record declaration with a  
71 type variable name and a possibly empty list of factories for that type. The code between  
72 lines 1 and 7 declares a mixin that can turn a naked type `A` into a monoid, hence we chose  
73 the name to be `Monoid_of_Type`.

The `HB.structure` command takes in input the definition `S` for a type `A` together with a list of factories for `A`. It registers a structure `S` in the hierarchy placing any definition specific to that structure inside a Coq module named `S`. Line 8 hence forges the `Monoid` structure.

Line 10 declares a second mixin collecting the operations and properties that are needed in order to enrich a monoid to a ring, hence the name `Ring_of_Monoid`. Indeed this time the type variable `A` is followed by `Monoid` that enriches `A` with the operations and properties of monoids. As a consequence `add` and `zero` can be used to express the new properties.

The last line declares the `Ring` structure to hold all the axioms declared so far. We can now inspect the contents of the hierarchy and then proceed to build a theory about abstract rings, register examples (instances) of ring structures and finally use the abstract theory on these examples.

```

1 Print Monoid.type. (* Monoid.type := { sort : Type; ... } *)
2 Check @add. (* add : forall M : Monoid.type, M -> M -> M *)
3 Check @addNr. (* addNr : forall R : Ring.type, left_inverse zero opp add *)
4
5 Lemma addrC : forall R : Ring.type, commutative (@add R).
6 Proof. (* Proof by Hankel 1867, omitted for brevity *) Qed.
7
8 Definition Z_Monoid_axioms : Monoid_of_Type Z :=
9   Monoid_of_Type.Build Z 0 Z.add Z.add_assoc Z.add_0_l Z.add_0_r.
10
11 HB.instance Z Z_Monoid_axioms.
12
13 Definition Z_Ring_axioms : Ring_of_Monoid Z _ :=
14   Ring_of_Monoid.Build Z 1 Z.opp Z.mul
15     Z.add_opp_diag_l Z.add_opp_diag_r Z.mul_assoc Z.mul_1_l Z.mul_1_r
16     Z.mul_add_distr_r Z.mul_add_distr_l.
17
18 HB.instance Z Z_Ring_axioms.
19
20 Lemma exercise (m n : Z) : (n + m) - n * 1 = m.
21 Proof. by rewrite mulr1 (addrC n) -(addrA m) addrN addr0. Qed.

```

We can print the type for monoids as forged by `HB` (line 1). It packs a carrier, called `sort`, and the collection of operation and properties that we omit for brevity. We can also look at the type of two constants synthesized by `HB` out of the hierarchy declaration. Remark that while the names of the constants come from the names of mixin fields, their types differ. In particular they are quantified over a `Monoid.type` or `Ring.type`, and not a simple type `A` as in the mixins. Moreover we evince that the `sort` projection is declared as an implicit coercion [24] and is automatically inserted in order to make `M -> M -> M` a meaningful type for binary operations on the carrier of `M`. Last, we see that properties are quantified on (hence apply to) the structure they belong to but use, in their statements, operations belonging to simpler structures. For example `addNr` is a property of a ring but its statement mentions `add`, the operation of the underlying monoid.

We then follow the proof of Hankel [5] to show that ring axioms imply the commutativity of the underlying monoid (line 5). This simple example shows we can populate the theory of abstract rings with new results.

We use the `Monoid_of_Type.Build abbreviation` (line 8) in order to build an instance of the `Monoid` structure for binary integers `Z`. We then register that monoid instance as the canonical one on `Z` (line 11) via the command `hb.canonical`. We can similarly declare that `Z` forms a ring by using the `Ring_of_Monoid.Build abbreviation` (lines 13 and 18). Note that the `Ring_of_Monoid.Build abbreviation` is not a plain record constructor for `Ring_of_Monoid`, since that would require more arguments, namely the monoid ones (see the `_` at line 13).

## 23:4 Hierarchy Builder

105 The abbreviation synthesized by *HB* infers them automatically (as in [17, Section 7]) thanks  
106 to the `hb.canonical` declaration given just above.

107 From now on the axioms as well as the abstract theory of rings apply to integers, as shown  
108 in lemma `exercise`. The details of the proof do not matter here, what is worth pointing out  
109 is that in a single statement we mix monoid (e.g. `+`) and ring (e.g. `-`) operations and in the  
110 proof we use monoid axioms (e.g. `addrA`), ring axioms (e.g. `addrN`) and ring lemmas (e.g.  
111 `addrC`), all seamlessly.

### 112 2.1 Evolution of the hierarchy

113 We proceed by accommodating the intermediate structure of Abelian groups.

```
1  HB.mixin Record Monoid_of_Type A := { ... (* unchanged *) ... }.
2  HB.structure Definition Monoid := { A of Monoid_of_Type A }.
3
4  HB.mixin Record AbelianGroup_of_Monoid A of Monoid A := {
5    opp : A -> A;
6    addrC : commutative (add : A -> A -> A);
7    addrNr : left_inverse zero opp add;
8  }.
9  HB.structure Definition AbelianGroup := { A of Monoid A & AbelianGroup_of_Monoid A }.
10
11 HB.mixin Record Ring_of_AbelianGroup A of AbelianGroup A := {
12   one : A;
13   mul : A -> A -> A;
14   mulrA : associative mul;
15   mulr1 : left_id one mul;           mulr1 : right_id one mul;
16   mulrDl : left_distributive mul add; mulrDr : right_distributive mul add;
17 }.
18 HB.structure Definition Ring := { A of AbelianGroup A & Ring_of_AbelianGroup A }.
19
20 Lemma addrN {R : AbelianGroup.type} : right_inverse zero opp add.
21 Proof. by move=>x; rewrite addrC addrNr. Qed.
```

114 Some operations and properties were moved from the old mixin for rings into a newborn  
115 mixin `AbelianGroup_of_Monoid` that gathers the axioms needed to turn a monoid into an  
116 Abelian group. Consequently the mixin for rings is now called `Ring_of_AbelianGroup` (instead  
117 of `Ring_of_Monoid`) since it expects the type `A` to be already an Abelian group and hence  
118 gathers fewer axioms.

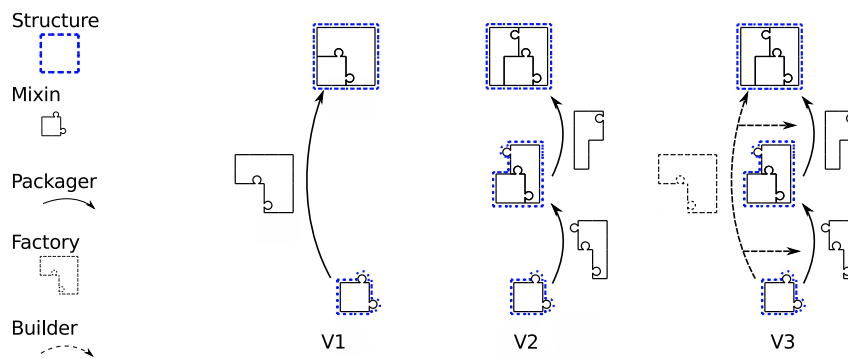
119 While operations moved from one structure to another, some properties undergo a deep  
120 change in their status. The lemma `addrC` part of the abstract theory of rings is now an axiom  
121 of Abelian groups, while `addrN` is no more an axiom of rings, but rather a theorem of the  
122 abstract theory of Abelian groups.

123 With this new version of the hierarchy, that we name `V2`, code written for version `V1`  
124 breaks. For example the declaration of the canonical ring over the integers fails, if only  
125 because we do not have a `Ring_of_Monoid.Build` abbreviation anymore.

126 Our objective is to obtain a version of the hierarchy, that we name `V3`, that does not  
127 only feature Abelian groups but that is also backward compatible with `V1`.

### 128 2.2 The missing puzzle piece

129 The key to make a hierarchy evolve without breaking user code is the full fledged notion  
130 of *factory* (the mixins seen so far are degenerate, trivial, factories). Factories, like mixins,  
131 are packages for operations and properties but are not directly used in the definition of  
132 structures. Instead a factory is equipped with *builders*: user provided pieces of code that  
133 extract from the factory the contents of mixins, so that existing abbreviations can be used.



■ **Figure 1** The evolution of the hierarchy. V3 is backward compatible with V1, while V2 is not.

134 As depicted in figure 1 we change again the hierarchy by declaring a `Ring_of_Monoid` factory,  
 135 that, from the user point of view, will look indistinguishable from the old `Ring_of_Monoid`  
 136 mixin and hence grant backward compatibility between version V3 and version V1.

```

1  HB.factory Record Ring_of_Monoid A of Monoid A := {
2  ... (* unchanged, see page 2 line 13 *) ...
3  }.
4  HB.builders Context A (f : Ring_of_Monoid A).
5
6  Lemma addrC : commutative add. Proof. (* The same proof as before *) Qed.
7
8  Definition to_AbelianGroup_of_Monoid :=
9  AbelianGroup_of_Monoid.Build A (opp f) addrC (addNr f). (* addrN unused *)
10 HB.instance A to_AbelianGroup_of_Monoid.
11
12 Definition to_Ring_of_AbelianGroup :=
13 Ring_of_AbelianGroup.Build A (one f) (mul f)
14 (mulrA f) (mul1r f) (mulr1 f) (mulrDl f) (mulrDr f).
15 HB.instance A to_Ring_of_AbelianGroup.
16
17 HB.end.
```

137 The record `Ring_of_Monoid` is the same we declared as a mixin in version V1. In order to  
 138 make a factory out of it we equip it with two definitions that embody the *builders*. The first  
 139 is `to_AbelianGroup_of_Monoid` and explains how to build an `AbelianGroup` structure out of the  
 140 factory axioms (named `f`, line 4). This construction is also registered as canonical for `A`, so  
 141 that the next construction `to_Ring_of_AbelianGroup` can call the `Ring_of_AbelianGroup.Build`  
 142 abbreviation that requires `A` to be an `AbelianGroup`. It is worth pointing out that the proof  
 143 of `addrC` we had in V1 is now required in order to write the builder for Abelian groups, while  
 144 the `addrN` field is not used (the same statement is already part of the theory of Abelian  
 145 groups, see line 20 of the previous code snippet).

146 Thanks to this factory we can now declare `Z` to be an instance of a ring using the  
 147 `Ring_of_Monoid.Build` abbreviation. The associated builders generate, behind the scenes,  
 148 instances of the `Ring_of_AbelianGroup` and `AbelianGroup_of_Monoid` mixins that in turn are  
 149 used to build instances of the `AbelianGroup` and `Ring` structures. Indeed, when used in the  
 150 context of the hierarchy version V3, the command `HB.instance Z Z_Ring_axioms` makes `Z` an  
 151 instance of *both* structures, and not just the `Ring` one as in version V1. Thanks to that the  
 152 proof of `example` can use the theory of both structures, for example `addrC` holds on Abelian  
 153 groups, `addrA` holds on monoids, while `addr1` holds on rings. As a result the very same proof  
 154 works on both version V1 and V3.

## 23:6 Hierarchy Builder

155 Last, it is worth pointing out that the new factory makes the following two lines equivalent  
156 (the former declares rings in V1) since they both describe the same set of mixins.

```
1 HB.structure Definition Ring := { A of Monoid A & Ring_of_Monoid A }.  
2 HB.structure Definition Ring := { A of AbelianGroup A & Ring_of_AbelianGroup A }.
```

157 This is another example of client code that would not break: the client of the hierarchy  
158 is allowed to declare new structures on top of existing ones.

### 159 2.3 $\mathcal{HB}$ in a nutshell

160 By using  $\mathcal{HB}$  the hierarchy designer has the following freedoms and advantages.

- 161 ■ Operations and properties (axioms) are made available to the user as soon as they are  
162 used in a structure. The hierarchy designer is free to move them from one to another and  
163 replace an axiom by a lemma and vice versa.
- 164 ■ Structures cannot disappear but the way they are built may change. The hierarchy  
165 designer is free to split structures into smaller ones in order to better factor and reuse  
166 parts of the hierarchy and the library that follows it.
- 167 ■ Mixins cannot disappear but can change considerably in their implementation. A mixin  
168 can become a full fledged factory equipped with builders to ensure backward compatibility.
- 169 ■  $\mathcal{HB}$  high level commands compile to the discipline of Packed Classes (Coq modules, records,  
170 coercions, implicit arguments, canonical structure instances, notations, abbreviations).  
171 This process lifts a considerable burden from the shoulders of the hierarchy designer and  
172 the final user who are no longer required to master the details of Packed Classes.

## 173 3 $\mathcal{HB}$ language

174 The Coq terms handled by  $\mathcal{HB}$  are subdivided in five categories, the mixins  $\mathcal{M}$ , the factories  $\mathcal{F}$ ,  
175 builders  $\mathcal{B}$ , the classes  $\mathcal{C}$  and the structures  $\mathcal{S}$ . Mixins, factories and instances are tagged  
176 by the user, through the commands `HB.mixin`, `HB.factory` and `HB.instance`, and the user  
177 may rely on their implementation. However structures and classes are generated with the  
178 command `HB.structure` and builders are generated when using `HB.instance` while declaring  
179 a factory, and the user may only refer to structure types, but should never rely on their  
180 implementation, neither should they rely on explicit builders.

### 181 3.1 $\mathcal{HB}$ mixins, factories and instances

182 ► **Definition 1** ( $\mathcal{M}$ , mixins). A mixin  $m \in \mathcal{M}$  is a Coq record with one or more parameters.  
183 The first parameter must be a  $(T : \text{Type})$ , while the other parameters are mixins applied to  $T$   
184 and possibly  $n$  other distinct mixins. I.e.

```
1 Record m (T : Type)  $\overline{(p : m T p_\sigma)^n}$  : Type := { .. }.
```

185 where  $\overline{(p : m T p_\sigma)^n} = \overline{(p_i : m_i T p_{\sigma(i,1)} \dots p_{\sigma(i,n_i)})_{i \in \{1, \dots, n\}}}$ . For all  $i \in \{1, \dots, n\}$  we  
186 have  $m_i \in \mathcal{M}$  and its arguments consist in  $n_i$  of the previously quantified mixin parameters  
187  $p_k$ , i.e. for all  $k \in \{1, \dots, n_i\}$ , we have  $\sigma(i, k) \in \{1, \dots, i - 1\}$ .

188 ► **Definition 2** ( $\text{dep}$ , mixin dependencies). Given  $m \in \mathcal{M}$ , we define  $\text{dep}(m) \in \mathcal{P}(\mathcal{M})$  as the  
189 set of all mixins that occur as parameters of  $m$ , i.e.  $\text{dep}(m) = \{m_1, \dots, m_n\}$ .

190 ► **Remark 3.** For all  $i \in \{1, \dots, n\}$ , we have  $\text{dep}(m_i) = \{m_{\sigma(i,1)}, \dots, m_{\sigma(i,n_i)}\}$ .

191 ► **Definition 4.** If  $f : \mathcal{X} \rightarrow \wp(\mathcal{Y})$ , then  $f^* : \wp(\mathcal{X}) \rightarrow \wp(\mathcal{Y})$  is defined as  $f^*(X) = \bigcup_{x \in X} f(x)$ .

192 ► **Proposition 5.**  $dep$  is transitively closed since records are well typed in the empty context  
 193 and describes a DAG since Coq does not admit circular definitions:  $dep^*(dep^*(M)) \subseteq$   
 194  $dep^*(M)$

195 ► **Definition 6** (factories  $\mathcal{F}$ , builders  $\mathcal{B}$  and from). A builder  $\mu \in \mathcal{B}$  is any function which para-  
 196 meters are the carrier type ( $T$ : **Type**), an arbitrary number of distinct mixins  $\{m_1, \dots, m_n\}$   
 197 and an additional argument  $f$ , which we call a factory, and which depends on the mixins. Its  
 198 return type is a mixin  $m_{n+1} \in \mathcal{M}$  i.e.

1 **Definition**  $f (T : \mathbf{Type}) \overline{(p : m \ T \ p_\sigma)}^n : \mathbf{Type} := \dots$

2 **Definition**  $\mu (T : \mathbf{Type}) (p : m \ T \ p_\sigma)^n : f \ T \ p_1 \dots p_n \rightarrow m_{n+1} \ T \ p_{\sigma(n+1,1)} \dots p_{\sigma(n+1,n_i)} := \dots$

199 We define  $from(f, m_{n+1}) = \mu$  to be the (unique) builder for  $m_{n+1}$  from the factory  $f$ .

200 ► **Definition 7** (provides, requires). We define the following functions

201 ■  $requires(f) = \{m_1, \dots, m_n\} \in \wp(\mathcal{M})$  is the set of mixins a factory  $f \in \mathcal{F}$  depends on. As  
 202 a consequence all the builders of a given factory have the same set of dependencies.

203 ■  $provides(f) = \{m \mid from(f, m) \text{ is defined}\} \in \wp(\mathcal{M})$  the set of mixins that a factory  $f \in \mathcal{F}$   
 204 provides, through its builders.

205 Moreover the following properties hold for all  $f \in \mathcal{F}$ :

206 1.  $dep^*(requires(f)) \subseteq requires(f)$

207 2.  $requires(f) \cap provides(f) = \emptyset$

208 3.  $\forall m \in provides(f)$ ,  $from(f, m)$  is defined

209 Mixins are declared by the user as the fundamental building blocks of a hierarchy. As  
 210 the next proposition shows they shall not be distinguished from regular factories, since they  
 211 are a special cases.

212 ► **Proposition 8** ( $\mathcal{M} \subseteq \mathcal{F}$ ). For all  $m \in \mathcal{M}$  we have  $requires(m) = dep(m)$ ,  $provides(m) =$   
 213  $\{m\}$  and  $from(m, m) = \left( \mathbf{fun} \ T \ \overline{(p : m \ T \ p_\sigma)}^n \ (x : m \ T \ p_1 \dots p_n) \Rightarrow x \right) \in \mathcal{B}$ .

214 ► **Coq command** to declare a new mixin:

1 **HB.mixin Record**  $M \ T$  **of**  $f_1 \dots f_n := \{ \dots \}$ .

215 Creates a module  $M$  with a record  $M$ , which depends on the mixins  $requires(M) = provides^*(\{f_1, \dots, f_n\})$ ,  
 216 and registers this new record both as a mixin and a factory. It also creates a abbreviation  
 217  $M.Build$ .

218 ► **Coq command** to declare an instance: **HB.instance**  $X \ b_1 \dots b_k$ , synthesizes terms  
 219 corresponding to all the mixins that can be built from the  $b_i$ . Indeed if  $b_i : f_i \ T \dots$ , then this  
 220 command creates elements of type  $provides^*(\{f_1, \dots, f_k\})$ . This command also generates  
 221 unification hints as described in Section 4.

222 ► **Coq command** to declare a new factory and generate new builders:

1 **HB.factory Record**  $M \ T$  **of**  $f_1 \dots f_n := \{ \dots \}$ .

2 **HB.builders Context**  $T (a : M \ A)$ .

3

4 **Definition**  $b_{n+1} : f_{n+1} \ T \dots := \dots$

5 **HB.instance**  $T \ b_{n+1}$ .

6

7 **Definition**  $b_{n+k} : f_{n+k} \ T \dots := \dots$

8 **HB.instance**  $T \ b_{n+k}$ .

9

**HB.end**.



223 Creates a module  $M$  with a record  $M$ , which depends on the mixins  $requires(M) = provides^*(\{f_1, \dots, f_n\})$   
 224 and registers this new record as a factory. We then use the factory instances  $b_{n+1}, \dots, b_{n+k}$   
 225 provided by the user and derive builders from them, so that  $provides(M) = provides^*(\{f_{n+1}, \dots, f_{n+k}\})$ .  
 226 It is thus necessary that  $requires^*(\{f_{n+1}, \dots, f_{n+k}\}) \subseteq provides^*(\{f_1, \dots, f_n\})$ .

227 Note that the  $b_i$  are not builders since their return types are not necessarily mixins, but  
 228 could be a factory. However, since all factories provide mixin through builders, we can obtain  
 229 builders out of each  $b_i$  by function composition.

### 230 3.2 $\mathcal{HB}$ classes and structures

231 ► **Definition 9** ( $\mathcal{C}$ , class). A class  $c \in \mathcal{C}$  is a Coq record with one parameter ( $T : \mathbf{Type}$ ). The  
 232 type of each field is a mixin in  $\mathcal{M}$  applied to  $T$  and, if needed, any number of other fields:

1 **Record**  $c$  ( $T : \mathbf{Type}$ ) := {  $p_1 : m_1 T$ ; ... ;  $p_i : m_i T$   $p_{\sigma(i,1)} \dots p_{\sigma(i,n_i)}$ ; ... }.

233 ► **Definition 10** ( $def$ , class definition). We call  $def(c) \in \mathcal{O}(\mathcal{M})$  the set of mixins mentioned  
 234 in the fields of the class, i.e.  $\{m_1, \dots, m_n\}$ . Given that class records are well typed in the  
 235 empty context the set of mixin records is closed transitively. The implementation enforces  
 236 that no two class records contains the same set of mixins (disregarding the order of the fields),  
 237 i.e.  $def$  is injective.

238 ► **Programming invariant for  $def$**  For all  $f \in \mathcal{F}$ :

- 239 1.  $\exists c \in \mathcal{C}, def(c) = requires^*(f) \cup provides^*(f)$ ,
- 240 2.  $\exists C \subseteq \mathcal{C}, def^*(C) = requires^*(f)$ .

241 Classes could also be seen as factories.

242 ► **Proposition 11** ( $\mathcal{C} \subseteq \mathcal{F}$ ). for all  $m \in def(c)$  we have  $from(c, m_i) = p_i$ , and it follows that  
 243  $requires(c) = \emptyset$ ,  $provides(c) = def(c)$ .

244 However since classes are generated, their implementation may change, thus users should  
 245 not rely on constructors of classes. Hence the only way users may refer to a class is as an  
 246 argument of `HB.mixin`, `HB.factory` or `HB.structure`.

247 ► **Definition 12** ( $\mathcal{S}$  Structure). A structure  $s \in \mathcal{S}$  is a dependent pair: a Coq record where  
 248 the value of the first field occurs in the type of the second. The first field is ( $sort : \mathbf{Type}$ )  
 249 and the second field is ( $class : c$   $sort$ ) for some  $c \in \mathcal{C}$ . As a consequence structures are in  
 250 bijection with classes.

251 ► **Coq command** to declare a class and structure: `HB.structure Definition M := { A of`  
 252  `$f_1 \dots f_n$  }`. crafts a class  $M.class\_of \in \mathcal{C}$  where  $def(c) = provides^*(\{f_1, \dots, f_n\})$  and the  
 253 corresponding structure  $M.type \in \mathcal{S}$ , together with unification hints as described in Section 4.

254 ► **Definition 13** ( $\leq \in \mathcal{C} \times \mathcal{C}$ , subclass).  $c_1 \leq c_2$  iff  $def(c_2) \subseteq def(c_1)$

### 255 3.3 Automatic inference of mixins

256 Since mixins may change but factories stay the same,  $\mathcal{HB}$  commands must never rely on a  
 257 particular set of mixins as arguments, and factory arguments must never be given explicitly  
 258 by the user. As described in Sections 3.1 and 3.2,  $\mathcal{HB}$  commands take a list of factories as  
 259 arguments, which they expand into lists of mixins behind the scene. However factory types  
 260 and constructors have mixin arguments that must be inferred automatically when used. To  
 261 this end, the commands `HB.mixin` and `HB.factory` generate abbreviations for the user to

262 replace uses of constructors of factories. These commands first create a record  $f_{aux}$  with a  
 263 constructor  $F_{aux}$  and then create abbreviations  $f$  and  $F$  that automatically fill the mixin  
 264 arguments of  $f_{aux}$  and  $F_{aux}$  respectively. See [17, Section 7] for a detailed description of how  
 265 to implement these abbreviations in Coq.

```

1 Record  $f_{aux}$  T  $\overline{(p : m T p_\sigma)^n} := F_{aux} \{ \dots \}$ .
2 Notation  $f$  T := ( $f_{aux}$  T ... (*  $p_1 \dots p_n$  inferred when T is known *)).
3 Notation  $F$  T  $x_1 \dots x_k := (F_{aux}$  T ... (*  $p_1 \dots p_n$  inferred when T is known *)  $x_1 \dots x_k$ ).
4 Definition  $b$  :  $f$  T :=  $F$  T  $x_1 \dots x_n$ 

```

## 266 4 The target language: Coq with Packed, flat, Classes

267 The language of packed classes were introduced in [13] and used directly to describe the  
 268 algebraic hierarchy of the Mathematical Components library. It is based on a disciplined use  
 269 of records and projections and on the possibility of extending the elaborator of Coq via the  
 270 declaration of Canonical Structures instances. In this section we describe the *flat* variant of  
 271 Packed Classes, the target language of  $\mathcal{HB}$ .

### 272 4.1 Describing structures with records and projections

273 We describe mathematical structures with three kinds of dependent records: mixins, classes,  
 274 and structures. As shown in section 2, a mixin gathers operators and axioms newly introduced  
 275 by a structure. As in [13, section 2.4][26, section 2], a class record is parametrized by the  
 276 carrier type (T : **Type**) and gathers all the operators and axioms of a structure by assembling  
 277 mixins, and a **Structure** type (a record) bundles a carrier and its class instance, as follows.

```

1 Module Monoid.
2 Record axioms (T : Type) : Type :=
3   Class { Monoid_of_Type_mixin : Monoid_of_Type T; }.
4 Structure type : Type := Pack { sort : Type; class : axioms sort; }.
5 End Monoid.

```

278 The **Monoid** module plays the role of a name space and forces us to write qualified names such  
 279 as **Monoid.type**; as a consequence we can reuse the same unqualified names for other structures,  
 280 i.e., class and structure record can always be named as **axioms** and **type** respectively.

281 Mixins and classes are internal definitions to structures; in contrast, **Monoid.type** is part  
 282 of the interface of the monoid structure. As seen in section 2, we declare **Monoid.sort** as  
 283 an implicit coercion and lift monoid operators and axioms from projections for the monoid  
 284 mixin to definitions and lemmas for **Monoid.type** as follows.

```

1 Coercion Monoid.sort : Monoid.type >-> Sortclass.
2
3 Definition zero {T : Monoid.type} : T :=
4   Monoid_of_Type.zero T (Monoid.Monoid_of_Type_mixin T (Monoid.class T)).
5 Definition add {T : Monoid.type} : T -> T -> T :=
6   Monoid_of_Type.add T (Monoid.Monoid_of_Type_mixin T (Monoid.class T)).
7 Lemma addrA {T : Monoid.type} : associative (@add T).
8 (* Two monoid axioms `addOr` and `addr0` are omitted. *)

```

285 Next we define Abelian groups. Since the monoid structure is the bottom of this hierarchy  
 286 its class record **Monoid** consists just one mixin. In contrast the class record of Abelian groups  
 287 consists of two mixins where the second one depends on the former (since Abelian groups  
 288 inherit from monoids).

```

1 Module AbelianGroup.
2 Record axioms (T : Type) : Type := Class {

```

## 23:10 Hierarchy Builder

```

3   Monoid_of_Type_mixin : Monoid_of_Type T;
4   AbelianGroup_of_Monoid_mixin : AbelianGroup_of_Monoid T Monoid_of_Type_mixin; }.
5   Structure type : Type := Pack { sort : Type; class : axioms sort }.
6   End AbelianGroup.

```

As in the monoid structure, we declare `AbelianGroup.sort` as an implicit coercion and lift additive inverse `opp` and Abelian group axioms. Since Abelian groups inherit from monoids, we also declare an implicit coercion from Abelian groups to monoids. A coercion between structures can be defined in two steps: first a coercion between the class record of the superclass to the class record of the subclass (line 2); and a coercion between structure records (line 4) relying on the first one.

```

1   Coercion AbelianGroup.sort : AbelianGroup.type >-> Sortclass.
2   Coercion AbelianGroup_class_to_Monoid_class (T : Type) (c : AbelianGroup T) :
3     Monoid T := Monoid.Class T (AbelianGroup.Monoid_of_Type_mixin T c).
4   Coercion AbelianGroup_to_Monoid (T : AbelianGroup.type) : Monoid.type :=
5     Monoid.Pack (AbelianGroup.sort T) (AbelianGroup.class T).
6
7   Definition opp {T : AbelianGroup.type} : T -> T := ... .
8   (* Two Abelian group axioms `addrC` and `addNr` are omitted. *)

```

Generally, a coercion from a structure A to another structure B can (and should) have the following form, thanks to the corresponding coercion between classes `A_class_to_B_class`.

```

1   Coercion A_to_B (T : A.type) : B.type :=
2     B.Pack (A.sort T) ((* A_class_to_B_class _ *) (A.class T)).

```

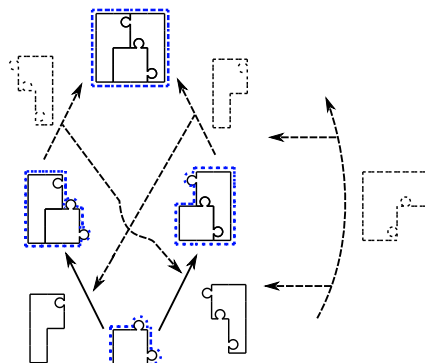
## 4.2 Multiple inheritance

In order to introduce multiple inheritance, we extend the hierarchy described in section 2.2 with the structure of semirings as depicted in figure 2. Semirings introduce the binary multiplication operator `mul` and its identity element `one`.

```

1   HB.mixin Record SemiRing_of_Monoid A
2     of Monoid A := {
3     one : A;
4     mul : A -> A -> A;
5     (* 7 axioms are omitted. *)
6   }.
7
8   HB.factory Record Ring_of_AbelianGroup A
9     of AbelianGroup A := {
10    (* 2 operators and 5 axioms are omitted. *)
11  }.
12  HB.builders Context A (a : Ring_of_AbelianGroup A).
13  ..
14  HB.end.

```



■ **Figure 2** V4 introduces multiple inheritance. For brevity we omit the factories/builders for the upper arrows.

Since semirings and Abelian groups do not inherit from each other, the definition of semirings and Abelian groups are quite similar:

```

1   Module SemiRing.
2   Record axioms (T : Type) : Type := Class {
3     Monoid_of_Type_mixin : Monoid_of_Type T;
4     SemiRing_of_Monoid_mixin : SemiRing_of_Monoid T Monoid_of_Type_mixin; }.
5   Structure type : Type := Pack { sort : Type; class : axioms sort; }.
6   End SemiRing.

```

We define implicit coercions from the semiring structure to the carrier and the monoid structure, and we lift semiring operators and axioms as follows:

```

1  Coercion SemiRing.sort : SemiRing.type >-> Sortclass.
2  Coercion SemiRing_to_Monoid : SemiRing.type >-> Monoid.type.
3
4  Definition one {T : SemiRing.type} : T := ... .
5  Definition mul {T : SemiRing.type} : T -> T -> T := ... .
6  (* 7 semiring axioms are omitted. *)

```

307 The class record is defined by gathering the monoid, Abelian group, and semiring mixins.  
308 Since the rings inherit from monoids, semirings, and Abelian groups, we define implicit  
309 coercions from the ring structure to those three structures.

```

1  Module Ring.
2  Record axioms (T : Type) : Type := Class {
3    Monoid_of_Type_mixin : Monoid_of_Type T;
4    SemiRing_of_Monoid_mixin : SemiRing_of_Monoid T Monoid_of_Type_mixin;
5    AbelianGroup_of_Monoid_mixin : AbelianGroup_of_Monoid T Monoid_of_Type_mixin; }.
6  Structure type : Type := Pack { sort : Type; class : axioms sort; }.
7  End Ring.
8
9  Coercion Ring.sort : Ring.type >-> Sortclass.
10 Coercion Ring_to_Monoid : Ring.type >-> Monoid.type.
11 Coercion Ring_to_SemiRing : Ring.type >-> SemiRing.type.
12 Coercion Ring_to_AbelianGroup : Ring.type >-> AbelianGroup.type.

```

### 310 4.3 Linking structures and instances via Canonical Structures

311 We recall here how to use the Canonical Structures [17, 25] mechanism, which lets the user  
312 extend the elaborator of Coq, in order to handle inheritance and inference of structure [1,  
313 26, 13]. This software component takes in input a term as written by the user and has to  
314 infer all the missing information that is necessary in order to make the term well typed.

315 A first example of elaboration that requires canonical structures is  $0 + 1$ . After removing  
316 all syntactic facilities, the underlying Coq term is `(add _ (zero _)) (one _)` where `_` stands  
317 for an implicit piece of information to be inferred and the constants `add` and `zero` come from  
318 the monoid structure while `one` comes from semirings. When the term is type checked each `_`  
319 is replaced by a unification variable written `?v`. Respectively, the head and the argument of  
320 the top application can be typed as follows:

```

1  add ?M (zero ?M) : Monoid.sort ?M -> Monoid.sort ?M
2  one ?SR : SemiRing.sort ?SR

```

321 where `?M : Monoid.type` and `?SR : SemiRing.type`. In order to type check the applica-  
322 tion Coq has to unify `Monoid.sort ?M` with `SemiRing.sort ?SR`, which is not trivial: it  
323 amounts at finding a structure that is both a monoid and a semiring, possibly the smal-  
324 lest one [26, Sect. 4]. This piece of information can be inferred from the hierarchy and  
325 its inheritance relation (Definition 13) and we can tell Coq to exploit it by declaring  
326 `SemiRing_to_Monoid : SemiRing.type -> Monoid.type` as canonical. With that hint Coq will  
327 pick `?M` to be `SemiRing_to_Monoid ?SR` as the canonical solution for this unification problem.  
328 In general, all the coercions between structures must be declared as canonical.

329 Another example of elaboration problem is  $-1$ , which hides the term `opp _ (one _)`. Here  
330 `opp` and `one` are respectively from Abelian groups and semirings, which do not inherit each  
331 other but whose smallest common substructure is rings; thus we have to extend the unifier  
332 to solve a unification problem `AbelianGroup.sort ?AbG = SemiRing.sort ?SR` by instantiating  
333 `?AbG` and `?SR` with `Ring_to_AbelianGroup ?R` and `Ring_to_SemiRing ?R` respectively where `?R` is  
334 a fresh unification variable of type `Ring.type`. This hint can be given as follows:

```

1  Canonical AbelianGroup_to_SemiRing (T : Ring.type) :=
2    SemiRing.Pack (AbelianGroup.sort (Ring_to_AbelianGroup T)) (Ring.class T).

```

## 23:12 Hierarchy Builder

335 Similarly, one can apply an algebraic theory to an instance (an example) of that structure,  
336 e.g.,  $2 \times 3$  where 2 and 3 have type  $\mathbb{Z}$ . The same mechanism of canonical structures let us  
337 extend the unifier to solve `SemiRing.type ? =  $\mathbb{Z}$` .

### 338 5 The implementation in Coq-Elpi

339 The implementation is based on the Elpi extension language for Coq. In this section we  
340 introduce the features of the programming language that came in handy in the development  
341 of *HB* and comment a few code snippets.

342 Coq-Elpi [28] is a Coq plugin embedding Elpi and providing an extensive, high level,  
343 API to access and script the Coq system at the vernacular level. This API lives in the  
344 `coq` namespace and lets one easily declare records, coercions, canonical structures, modules,  
345 implicit arguments, etc. The most basic Coq data type exposed to Elpi is the one of references  
346 to global declarations:

```
1 kind gref type. % The data type of references to global terms
2 type indt inductive -> gref. % eg: Coq.Init.Datatypes.nat
3 type indc constructor -> gref. % eg: Coq.Init.Datatypes.O
4 type const constant -> gref. % eg: Coq.Init.Peano.plus
```

347 The arguments of the three constructors are opaque to Elpi, that can only use values of  
348 these types via dedicated APIs. For example the API for declaring an inductive type will  
349 generate a value of type `inductive` that is printed as, for example, `<nat>`.

350 Elpi [11] is a dialect of  $\lambda$ Prolog [19], an higher order logic programming language that  
351 makes it easy to manipulate abstract syntax tree with binders. Coq-Elpi takes full advantage  
352 of this capability by representing Coq terms in Higher Order Abstract Syntax [20] style,  
353 reusing the binder of the programming language in order to represent Coq's ones. Here is an  
354 excerpt of the data type of Coq terms:

```
1 kind term type. % The data type of Coq terms
2 type global gref -> term. % eg: nat, 0, S, plus, ...
3 type fun term -> (term -> term) -> term. % eg: fun x : t => b(x)
4 type app list term -> term. % eg: app [hd/args]
5 ... % all other term constructors are omitted for brevity
```

355 Note that the `fun` constructor holds a  $\lambda$ Prolog function. In this syntax the Coq term  
356 `(fun x : nat => x x)` becomes `(fun (global (indt <nat>)) x\ app[x, x])` where `x\` binds  
357 `x` in the body of the function. Substitution of a bound variable for a term can be computed  
358 by applying a term (of function type) to an argument.

359 Data types with binders are also used as input to high level APIs that build terms behind  
360 the scenes. For example a Coq record is just an inductive type and the API to declare one  
361 must allow the type of a field to depend on the fields that comes before it. Note that the  
362 field constructor takes a coercion flag, the name of the field, its type and binds a term in  
363 the remaining record declaration.

```
1 kind indt-decl type. % The type of an inductive type declaration
2 type record string -> term -> string -> record-decl -> indt-decl.
3 type field bool -> string -> term -> (term -> record-decl) -> record-decl.
4 type end-record record-decl.
5 ... % constructors for non-record inductive types are omitted for brevity
6
7 external pred coq.env.add-indt i:indt-decl, o:inductive.
8 external pred coq.CS.canonical-projections i:inductive, o:list (option constant).
```

364 The `pred` keyword documents types and modes (input or output) of the arguments of  
 365 a predicate, while `external` signals that the predicate is a builtin (in other words it is  
 366 implemented in OCaml rather than  $\lambda$ Prolog).

367 We comment these two builtin predicates and the `indt-decl` type while looking at the  
 368 code of `declare-structure` that is in charge of scripting the following Coq code:

```
1 Structure type : Type := Pack { sort : Type; class : axioms sort }.
```

369 The following Elpi code builds the declaration, type checks it, adds it to the Coq  
 370 environment and finally returns the projections for the sort and the class fields.

```
1 pred declare-structure i:grep, o:term, o:term, o:term.
2
3 declare-structure ClassName Structure SortProjection ClassProjection :- std.do! [
4   StructureDeclaration =
5     record "type" {{ Type }} "Pack" (
6       field tt "sort" {{ Type }} s\
7         field ff "class" (app [global ClassName, s]) _\
8     end-record),
9   coq.typecheck-indt-decl StructureDeclaration,
10  coq.env.add-indt StructureDeclaration StructureName,
11  coq.CS.canonical-projections StructureName [some SortP, some ClassP],
12  Structure = global (indt StructureName),
13  SortProjection = global (const SortP),
14  ClassProjection = global (const ClassP),
15 ].
```

371 Note that the binder `s\` at line 6 lets one mention the first field in the type of the second.  
 372 The syntax `{{ Type }}` is a quotation: it lets one use the syntax of Coq to write an Elpi  
 373 expression of type `term`. The API `coq.CS.canonical-projections` lets us find the projections  
 374 automatically generated by Coq for a given record. The last detail worth mentioning is that  
 375 this program makes no use of backtraking: the `std.do!` combinator signals that.

376 In the simple case of the structure record, the number of fields, and hence the number of  
 377 binders, is fixed. On the contrary the class record has one field per mixin and each of them  
 378 can depend on the previous ones. In order to synthesize terms with binders in an inductive  
 379 fashion (using a recursive predicate)  $\lambda$ Prolog lets one postulate fresh nominal constants using  
 380 the `pi` operator and attach to the nominal some knowledge in the form of a clause via the `=>`  
 381 operator. This process is called binder mobility: the binder is moved from the data (that we  
 382 are building) to the program (the context of the current computation). This feature is key  
 383 to the following code that synthesizes the declaration of the fields of the class record.

```
1 pred synthesize-fields.field-for-mixin i:mixinname, o:term.
2 pred synthesize-fields i:list mixinname, i:term, o:record-decl.
3
4 synthesize-fields [] _ Decl :- Decl = end-record.
5 synthesize-fields [M|ML] T Decl :- std.do! [
6   get-mixin-modname M ModName, Name is ModName ^ "_mixin",
7   dep1 M Deps,
8   std.map Deps synthesize-fields.field-for-mixin Args,
9   Type = app[ global M, T | Args ],
10  Decl = (field ff Name Type f\ Fields f),
11  pi m\
12    synthesize-fields.field-for-mixin M m =>
13    synthesize-fields ML T (Fields m)
14 ].
```

384 The first predicate `synthesize-fields.field-for-mixin` is used to link a mixin to a  
 385 nominal that corresponds to the record field for that mixin. It has no clauses in the base  
 386 program but some clauses are added dynamically by `synthesize-fields`.

## 23:14 Hierarchy Builder

387 The second predicate proceeds by recursion on the (topologically sorted) list of mixins,  
388 and terminates when the list is empty. If the list contains a mixin  $M$  then it crafts a `Name` for  
389 it (line 6), fetches its dependencies (line 7) and finds the (previously declared) record fields  
390 holding these mixins (line 8). The `(std.map L1 P L2)` predicate relates the two lists  $L1$  and  
391  $L2$  point wise using the predicate  $P$ .

392 Line 9 builds the type of the field: the mixin name applied to the type (sort) and all its  
393 dependencies. Note that the `Fields` variable, representing the declaration of the next fields,  
394 is under the binder for  $f$  (the current field). In order to make the recursive call under that  
395 binder (line 13) and recursively process  $ML$  we postulate a nominal  $m$  (line 11) that is a term  
396 satisfying any future dependency on the current mixin (line 12) and we replace  $f$  by  $m$  in  
397 `Fields` by writing `(Fields m)`.

### 398 **6** Related work

399 The most closely related work is the one about Packed Classes [13] on which we build. The  
400 main differences are that  $\mathcal{HB}$  is a higher level language that is compiled down to (flat)  
401 Packed Classes. The systematic use of factories makes the user interface of a hierarchy stable  
402 under the insertion of structures, a property that Packed Classes lacks. Finally many of the  
403 intricacies of Packed Classes are hidden to the user by the compilation step, in particular  
404 creating all the necessary coercions and canonical structure declarations, especially in the case  
405 of diamonds or when merging several libraries, which used to cause the need for *a posteriori*  
406 validation of a hierarchy design [26]. It also opens the way to automatically detect and solve  
407 problems tied to non judgmental commutative diagrams when forgetting structure [1].

408 In [7] Carette and O'Connor describe the language of Theory Presentation Combinators  
409 that can be used to describe a hierarchy of algebraic structures. They focus on the categorical  
410 semantics of the language that is built upon the category of context. They do not describe  
411 any actual compilation to the language of a mainstream interactive prover, indeed they claim  
412 their language to be mostly type theory independent. We know they considered targeting  
413 type theory and the language of unification hints [4] (a super set of the one of Canonical  
414 Structures), but we could not find any written trace of that. Language wise they provides  
415 keywords such as `combine` and `over` to share (reuse) a property when defining a new structure.  
416 For example in order to avoid restating the commutativity property when defining Abelian  
417 groups they combine a commutative monoid and a group forcing the subjacent monoid  
418 to coincide: `CommutativeGroup := combine CommutativeMonoid , Group over Monoid`.  
419 In our language  $\mathcal{HB}$  the same role is played by *mixins*. A mixin lets one write once and for  
420 all a property and abstract it over types and operations so that it can be reused in all the  
421 structures that need it. One operation  $\mathcal{HB}$  allows for but that does not seem to be possible  
422 in the setting of Presentation Combinators is the one of replacing an axioms with a lemma  
423 and vice versa. As shown in subsection 2.1  $\mathcal{HB}$  supports that.

424 The MMT system [22] provides a framework to describe formal languages in a logical  
425 framework, providing good support for binders and notations. It also provides an expressive  
426 module system to organize theories and express relations among them in the form of functors.  
427 At the time of writing it provides limited support for elaborating user input taking systematic  
428 advantage of the contents of the theories. The elaborator can be extended by the user writing  
429 Scala code, and in principle use the contents of the libraries to make sense of an incomplete  
430 expressions, but no higher level language or mechanism is provided.

431 The library of the Mizar system features a hierarchy of algebraic structures [15]. In spite  
432 of lacking dependent types, Mizar provides the concept of attributed types and adjectives

433 that can be used to describe the signature of structures as one would do with a dependently  
434 typed record and their properties as one would define a conjunctive predicate. The Mizar  
435 language also provides the notion of cluster that is used to link structures: by showing that  
436 property  $P$  implies property  $Q$  one can inform automation that structures characterized by  
437  $P$  are instances of structures characterized by  $Q$ . The foundational theory of Mizar features  
438 an extensional notion of equality that makes it easy to share the signature or the properties  
439 of structures by just requiring a proof of their equivalence that is in turn used by automation  
440 to treat equivalent structures as equal.

441 The concepts of factory and builder presented here is akin to the `AbstractFactory` and  
442 `Builder` pattern from "the Gang Of Four" design patterns [12] in the sense that factories are  
443 used to build an arbitrary number of objects (here mixin instances).

## 444 **7** Conclusion

445 In this paper we design and implement  $\mathcal{HB}$ , a high level language to describe hierarchies of  
446 algebraic structures. The implementation of  $\mathcal{HB}$  is based on the Elpi extension language for  
447 the Coq system and is available at <https://github.com/math-comp/hierarchy-builder>.  
448 The implementation amounts to approximately a thousand lines of (commented) Elpi code  
449 and tree hundred lines of Coq vernacular. It took less than one month to implement  $\mathcal{HB}$   
450 while it took several years of attempts and fruitful discussions with the other developers of  
451 the Mathematical Components library to design it, as well as Coq users and developers and  
452 members of Dagstuhl seminar 18341. In particular, discussions with Georges Gonthier, Assia  
453 Mahboubi and Florent Hivert were instrumental in coining the concepts formalized here.

454 The  $\mathcal{HB}$  language is only loosely tied to Coq or even Type Theory. We believe it could  
455 be adopted with no major change to other tools. Indeed the properties and invariants that  
456 link factories, mixins and classes are key to rule out meaningless or ambiguous sentences  
457 and are not specific to our logic setting. Moreover most logics feature packing construction  
458 similar to records.

459 The compilation scheme we present in section 4 is tied to dependent records, that are  
460 available in Coq but also in other provers based on Type Theory such a Matita [3] and  
461 Lean [10]. We chose the flat variant of Packed Classes as the target for  $\mathcal{HB}$  because the  
462 Mathematical Components library uses Packed Classes as well: As of today Packed Classes  
463 offer the best compromise between flexibility and performance [17, Section 8] in Coq. Of  
464 course one could imagine a future were other approaches such as telescopes [21] or unbundled  
465 classes [27] would offer equal or better performances in Coq, or in another system. It is a  
466 virtue of our work to provide a user language that is separate from the implementation one.  
467 We believe it would take a minor coding effort to retarget  $\mathcal{HB}$  to another bundling approach.

468 We leave to future work extending  $\mathcal{HB}$  to support structures parametrized by structures  
469 such as the one of module over a ring. We also leave to future work the automatic synthesis  
470 of the notion of morphism between structures of the hierarchy.

## 471 ——— References ———

- 472 1 Reynald Affeldt, Cyril Cohen, Marie Kerjean, Assia Mahboubi, Damien Rouhling, and  
473 Kazuhiko Sakaguchi. Formalizing functional analysis structures in dependent type theory.  
474 working paper or preprint, January 2020. URL: <https://hal.inria.fr/hal-02463336>.
- 475 2 Reynald Affeldt, David Nowak, and Takafumi Saikawa. A hierarchy of monadic effects for  
476 program verification using equational reasoning. In *Mathematics of Program Construction -*  
477 *13th International Conference, MPC 2019, Porto, Portugal, October 7-9, 2019, Proceedings*,



- 478 pages 226–254, 2019. URL: [https://doi.org/10.1007/978-3-030-33636-3\\_9](https://doi.org/10.1007/978-3-030-33636-3_9), doi:10.1007/  
479 978-3-030-33636-3\_9.
- 480 **3** Andrea Asperti, Wilmer Ricciotti, Claudio Sacerdoti Coen, and Enrico Tassi. The mat-  
481 ita interactive theorem prover. In *Automated Deduction - CADE-23 - 23rd Interna-*  
482 *tional Conference on Automated Deduction, Wroclaw, Poland, July 31 - August 5, 2011.*  
483 *Proceedings*, pages 64–69, 2011. URL: [https://doi.org/10.1007/978-3-642-22438-6\\_7](https://doi.org/10.1007/978-3-642-22438-6_7),  
484 doi:10.1007/978-3-642-22438-6\_7.
- 485 **4** Andrea Asperti, Wilmer Ricciotti, Claudio Sacerdoti Coen, and Enrico Tassi. Hints in  
486 unification. In Stefan Berghofer, Tobias Nipkow, Christian Urban, and Makarius Wenzel,  
487 editors, *Theorem Proving in Higher Order Logics*, pages 84–98, Berlin, Heidelberg, 2009.  
488 Springer Berlin Heidelberg.
- 489 **5** Gerhard Betsch. On the beginnings and development of near-ring theory. In *Near-rings and*  
490 *near-fields. Proceedings of the conference held in Fredericton, New Brunswick, July 18-24,*  
491 *1993*, pages 1–11. Mathematics and its Applications, 336. Kluwer Academic Publishers Group,  
492 Dordrecht, 1995.
- 493 **6** Sylvie Boldo, Catherine Lelay, and Guillaume Melquiond. Coquelicot: A user-friendly library  
494 of real analysis for Coq. *Mathematics in Computer Science*, 9(1):41–62, 2015. URL: <https://doi.org/10.1007/s11786-014-0181-1>,  
495 doi:10.1007/s11786-014-0181-1.
- 496 **7** Jacques Carette and Russell O’Connor. Theory presentation combinators. In Johan Jeuring,  
497 John A. Campbell, Jacques Carette, Gabriel Dos Reis, Petr Sojka, Makarius Wenzel, and  
498 Volker Sorge, editors, *Intelligent Computer Mathematics*, pages 202–215, Berlin, Heidelberg,  
499 2012. Springer Berlin Heidelberg.
- 500 **8** Cyril Cohen. *Formalized algebraic numbers: construction and first-order theory*. PhD  
501 thesis, Ecole Polytechnique X, 2012. URL: [https://pastel.archives-ouvertes.fr/  
502 pastel-00780446](https://pastel.archives-ouvertes.fr/pastel-00780446).
- 503 **9** The Coq Development Team. *The Coq Proof Assistant Reference Manual, version 8.11*,  
504 October 2020. URL: <http://coq.inria.fr>.
- 505 **10** Leonardo Mendonça de Moura, Soonho Kong, Jeremy Avigad, Floris van Doorn, and Jakob von  
506 Raumer. The lean theorem prover (system description). In *Automated Deduction - CADE-25*  
507 *- 25th International Conference on Automated Deduction, Berlin, Germany, August 1-7, 2015,*  
508 *Proceedings*, pages 378–388, 2015. URL: [https://doi.org/10.1007/978-3-319-21401-6\\_26](https://doi.org/10.1007/978-3-319-21401-6_26),  
509 doi:10.1007/978-3-319-21401-6\_26.
- 510 **11** Cvetan Dunchev, Ferruccio Guidi, Claudio Sacerdoti Coen, and Enrico Tassi. ELPI: fast,  
511 embeddable, λProlog interpreter. In *Logic for Programming, Artificial Intelligence, and*  
512 *Reasoning - 20th International Conference, LPAR-20 2015, Suva, Fiji, November 24-28, 2015,*  
513 *Proceedings*, pages 460–468, 2015. URL: [https://doi.org/10.1007/978-3-662-48899-7\\_32](https://doi.org/10.1007/978-3-662-48899-7_32),  
514 doi:10.1007/978-3-662-48899-7\_32.
- 515 **12** Erich Gamma. *Design patterns: elements of reusable object-oriented software*. Pearson  
516 Education India, 1995.
- 517 **13** François Garillot, Georges Gonthier, Assia Mahboubi, and Laurence Rideau. Packaging  
518 mathematical structures. In *Theorem Proving in Higher Order Logics, 22nd International*  
519 *Conference, TPHOLs 2009, Munich, Germany, August 17-20, 2009. Proceedings*, pages  
520 327–342, 2009. URL: [https://doi.org/10.1007/978-3-642-03359-9\\_23](https://doi.org/10.1007/978-3-642-03359-9_23), doi:10.1007/  
521 978-3-642-03359-9\_23.
- 522 **14** Georges Gonthier, Andrea Asperti, Jeremy Avigad, Yves Bertot, Cyril Cohen, François Garillot,  
523 Stéphane Le Roux, Assia Mahboubi, Russell O’Connor, Sidi Ould Biha, Ioana Pasca, Laurence  
524 Rideau, Alexey Solovyev, Enrico Tassi, and Laurent Théry. A machine-checked proof of the  
525 odd order theorem. In *Interactive Theorem Proving - 4th International Conference, ITP*  
526 *2013, Rennes, France, July 22-26, 2013. Proceedings*, pages 163–179, 2013. URL: [https://doi.org/10.1007/978-3-642-39634-2\\_14](https://doi.org/10.1007/978-3-642-39634-2_14), doi:10.1007/978-3-642-39634-2\_14.  
527

- 528 **15** Adam Grabowski, Artur Kornilowicz, and Christoph Schwarzweiler. On algebraic hierarchies  
529 in mathematical repository of Mizar. In *2016 Federated Conference on Computer Science and*  
530 *Information Systems (FedCSIS)*, pages 363–371, Sep. 2016.
- 531 **16** Johannes Hölzl, Fabian Immler, and Brian Huffman. Type classes and filters for math-  
532 ematical analysis in Isabelle/HOL. In *Interactive Theorem Proving - 4th International*  
533 *Conference, ITP 2013, Rennes, France, July 22-26, 2013. Proceedings*, pages 279–294.  
534 Springer, 2013. URL: [https://doi.org/10.1007/978-3-642-39634-2\\_21](https://doi.org/10.1007/978-3-642-39634-2_21), doi:10.1007/  
535 978-3-642-39634-2\_21.
- 536 **17** Assia Mahboubi and Enrico Tassi. Canonical structures for the working coq user. In *Interactive*  
537 *Theorem Proving - 4th International Conference, ITP 2013, Rennes, France, July 22-26, 2013.*  
538 *Proceedings*, pages 19–34, 2013. URL: [https://doi.org/10.1007/978-3-642-39634-2\\_5](https://doi.org/10.1007/978-3-642-39634-2_5),  
539 doi:10.1007/978-3-642-39634-2\_5.
- 540 **18** The mathlib Community. The Lean mathematical library. In *Proceedings of the 9th ACM*  
541 *SIGPLAN International Conference on Certified Programs and Proofs, CPP 2020*, pages  
542 367–381, New York, NY, USA, 2020. Association for Computing Machinery. URL: <https://doi.org/10.1145/3372885.3373824>, doi:10.1145/3372885.3373824.
- 544 **19** Dale Miller and Gopalan Nadathur. *Programming with Higher-Order Logic*. Cambridge  
545 University Press, New York, NY, USA, 1st edition, 2012.
- 546 **20** Frank Pfenning and Conal Elliott. Higher-order abstract syntax. In *Proceedings of the ACM*  
547 *SIGPLAN 1988 Conference on Programming Language Design and Implementation, PLDI*  
548 *'88*, page 199–208, New York, NY, USA, 1988. Association for Computing Machinery. URL:  
549 <https://doi.org/10.1145/53990.54010>, doi:10.1145/53990.54010.
- 550 **21** Robert Pollack. Dependently typed records in type theory. *Formal Aspects of Computing*,  
551 13:386–402, 2002. doi:10.1007/s001650200018.
- 552 **22** Florian Rabe and Michael Kohlhase. A scalable module system. *Information and Com-*  
553 *putation*, 230:1–54, 2013. URL: [http://www.sciencedirect.com/science/article/pii/](http://www.sciencedirect.com/science/article/pii/S0890540113000631)  
554 [S0890540113000631](http://www.sciencedirect.com/science/article/pii/S0890540113000631), doi:<https://doi.org/10.1016/j.ic.2013.06.001>.
- 555 **23** Damien Rouhling. *Formalisation Tools for Classical Analysis – A Case Study in Control Theory*.  
556 PhD thesis, Université Côte d’Azur, 2019. URL: <https://hal.inria.fr/tel-02333396>.
- 557 **24** Amokrane Saïbi. Typing algorithm in type theory with inheritance. In *24th ACM SIGPLAN-*  
558 *SIGACT Symposium on Principles of Programming Languages (POPL 1997), Paris, France,*  
559 *15–17 January 1997*, pages 292–301. ACM, 1997.
- 560 **25** Amokrane Saïbi. *Outils Génériques de Modélisation et de Démonstration pour la Formalisation*  
561 *des Mathématiques en Théorie des Types. Application à la Théorie des Catégories. (Formal-*  
562 *ization of Mathematics in Type Theory. Generic tools of Modelisation and Demonstration.*  
563 *Application to Category Theory)*. PhD thesis, Pierre and Marie Curie University, Paris, France,  
564 1999. URL: <https://tel.archives-ouvertes.fr/tel-00523810>.
- 565 **26** Kazuhiko Sakaguchi. Validating mathematical structures, 2020. [arXiv:2002.00620](https://arxiv.org/abs/2002.00620).
- 566 **27** Bas Spitters and Eelis van der Weegen. Type classes for mathematics in type theory. *Mathemat-*  
567 *ical Structures in Computer Science*, 21(4):795–825, 2011. doi:10.1017/S0960129511000119.
- 568 **28** Enrico Tassi. Coq-Elpi, Coq plugin embedding Elpi. <https://github.com/LPCIC/coq-elpi>,  
569 2020.

570 **A** Coq reference

```

1 Section OperationProperties.
2 Variable T : Type.
3 Variable e : T.
4 Variable inv : T -> T.
5 Variable op : T -> T -> T.
6 Variable add : T -> T -> T.
7
8 Definition left_id := forall x, op e x = x.
9 Definition right_id := forall x, op x e = x.
10
11 Definition left_inverse := forall x, op (inv x) x = e.
12
13 Definition commutative := forall x y, op x y = op y x.
14 Definition associative := forall x y z, op x (op y z) = op (op x y) z.
15
16 Definition left_distributive := forall x y z, op (add x y) z = add (op x z) (op y z).
17 Definition right_distributive := forall x y z, op x (add y z) = add (op x y) (op x z).

```

571 **B** Proof of addrC

```

1 Lemma addrC {R : Ring.type} : commutative (@add R).
2 Proof.
3 have innerC (a b : R) : (a + b) + (a + b) = (a + a) + (b + b).
4   by rewrite -[a+b]mul1r -mulrD1 mulrDr !mulrD1 !mul1r.
5 have addKl (a b c : R) : a + b = a + c -> b = c.
6   apply: can_inj (add a) (add (-a)) _ _ _ .
7   by move=> x; rewrite addrA addNr addOr.
8 have addKr (a b c : R) : b + a = c + a -> b = c.
9   apply: can_inj (add ^~ a) (add ^~ (-a)) _ _ _ .
10  by move=> x; rewrite /= -addrA addrN addrO.
11 move=> x y; apply: addKl (x) _ _ _; apply: addKr (y) _ _ _ .
12 by rewrite -!addrA [in RHS]addrA innerC !addrA.
13 Qed.

```