

Elaborating dependent (co)pattern matching

JESPER COCKX and ANDREAS ABEL, Department of Computer Science and Engineering, Chalmers and Gothenburg University, Sweden

In a dependently typed language, we can guarantee correctness of our programs by providing formal proofs. To check them, the typechecker elaborates these programs and proofs into a low level core language. However, this core language is by nature hard to understand by mere humans, so how can we know we proved the right thing? This question occurs in particular for dependent copattern matching, a powerful language construct for writing programs and proofs by dependent case analysis and mixed induction/coinduction. A definition by copattern matching consists of a list of *clauses* that are elaborated to a *case tree*, which can be further translated to primitive *eliminators*. In previous work this second step has received a lot of attention, but the first step has been mostly ignored so far.

We present an algorithm elaborating definitions by dependent copattern matching to a core language with inductive datatypes, coinductive record types, an identity type, and constants defined by well-typed case trees. To ensure correctness, we prove that elaboration preserves the first-match semantics of the user clauses. Based on this theoretical work, we reimplement the algorithm used by Agda to check left-hand sides of definitions by pattern matching. The new implementation is at the same time more general and less complex, and fixes a number of bugs and usability issues with the old version. Thus we take another step towards the formally verified implementation of a practical dependently typed language.

1 INTRODUCTION

Dependently typed functional languages such as Agda [2017], Coq [INRIA 2017], Idris [2013], and Lean [de Moura et al. 2015] combine programming and proving into one language, so they should be at the same time expressive enough to be useful and simple enough to be sound. These apparently contradictory requirements are addressed by having two languages: a high-level surface language that focuses on expressivity and a small core language that focuses on simplicity. The main role of the typechecker is then to *elaborate* the high-level surface language into the low-level core.

Since the difference between the surface and core languages can be quite large, the elaboration process can be, well, elaborate. If there is an error in the elaboration process, our program or proof may still be accepted by the system but its meaning is not what was intended [Pollack 1998]. In particular, the statement of a theorem may depend on the correct behaviour of some defined function, so if something went wrong in the elaboration of these definitions, the theorem statement may not be what it seems. As an extreme example, we may think we have proven an interesting theorem when in fact, we have only proven something trivial. This may be detected in a later phase when trying to use this proof, or it may not be detected at all. Unfortunately, there is no bulletproof way to avoid such problems: each part of the elaboration process has to be verified independently to make sure it produces something sensible.

One important part of the elaboration process is the elaboration of definitions by dependent pattern matching [Coquand 1992]. Dependent pattern matching provides a convenient high-level interface to the low-level constructions of case splitting, structural induction, and specialization by unification. The elaboration of dependent pattern matching goes in two steps: first the list of

¹This paper is best viewed in `co`lor.

Authors' address: Jesper Cockx, cockx@chalmers.se; Andreas Abel, andreas.abel@gu.se, Department of Computer Science and Engineering, Chalmers and Gothenburg University, Rännvägen 6b, Gothenburg, 41296, Sweden.

2018. 2475-1421/2018/1-ART1 \$15.00
<https://doi.org/0000001.0000001>

clauses given by the user is translated to a case tree, and then the case tree is further translated to a term that only uses the primitive datatype eliminators.² The second step has been studied in detail and is known to preserve the semantics of the case tree precisely [Cockx 2017; Goguen et al. 2006]. In contrast, the first step has received much less attention.

The goal of this paper is to formally describe an elaboration process of definitions by dependent pattern matching to a well-typed case tree for a realistic dependently typed language. Compared to the elaboration processes described by Norell [2007] and Sozeau [2010], we make the following improvements:

- We include both pattern and copattern matching.
- We are more flexible in the placement of forced patterns.
- We prove that the translation preserves the first-match semantics of the user clauses.

We discuss each of these improvements in more detail below.

Copatterns. Copatterns provide a convenient way to define and reason about infinite structures such as streams [Abel et al. 2013]. They can be nested and mixed with regular patterns. Elaboration of definitions by copattern matching has been studied for simply typed languages by Setzer et al. [2014], but so far the combination of copatterns with general dependent types has not been studied in detail, even though it has already been implemented in Agda.

One complication when dealing with copatterns in a dependently typed language is that the type of a projection can depend on the values of the previous projections. For example, define the coinductive type `CoNat` of possibly infinite natural numbers by the two projections `iszero : Bool` and `pred : iszero \equiv_{Bool} false \rightarrow CoNat`. We use copatterns to define the co-natural number `cozero`:

$$\begin{aligned} \text{cozero} &: \text{CoNat} \\ \text{cozero} \text{ .iszero} &= \text{true} \\ \text{cozero} \text{ .pred} &= \emptyset \end{aligned} \tag{1}$$

Here the new constant `cozero` is being defined with the field `iszero` equal to `true` (and no value for `pred`).

To refute the proof of `cozero .iszero \equiv_{Bool} false` with an absurd pattern \emptyset , the typechecker needs to know already that `cozero .iszero = true`, so it needs to check the clauses in the right order.

This example also shows that with mixed pattern/copattern matching, some clauses can have more arguments than others, so the typechecker has to deal with *variable arity*. This means that we need to consider introducing a new argument as an explicit node in the constructed case tree.

Flexible placement of forced patterns. When giving a definition by dependent pattern matching that involves forced patterns (also called presupposed terms [Brady et al. 2003] or inaccessible patterns [Norell 2007] or, in Agda, dot patterns), there are often multiple positions where to place them. For example, in the proof of symmetry of equality

$$\begin{aligned} \text{sym} &: (x \ y : A) \rightarrow x \equiv_A y \rightarrow y \equiv_A x \\ \text{sym } x \ [x] \ \text{refl} &= \text{refl} \end{aligned} \tag{2}$$

it should not matter if we instead write `sym [x] x refl = refl`. In fact, we even allow the apparently non-linear definition `sym x x refl = refl`.

Our elaboration algorithm addresses this by treating forced patterns as *laziness annotations*: they guarantee that the function will not match against a certain argument. This allows the user to be free in the placement of the forced patterns. For example, it is always allowed to write `zero` instead of `[zero]`, or `x` instead of `[x]`.

²In Agda, case trees are part of the core language so the second step is skipped in practice, but it is still important to know that it could be done in theory.

With our elaboration algorithm, it is easy to extend the pattern syntax with *forced constructor patterns* such as `[suc] n` (Brady et al. [2003]’s presupposed-constructor patterns). These allow the user to annotate that the function should not match on the argument but still bind some of the arguments of the constructor.

Preservation of first-match semantics. Like Augustsson [1985] and Norell [2007], we allow the clauses of a definition by pattern matching to overlap and use the first-match semantics in the construction of the case tree. For example, when constructing a case tree from the definition

$$\begin{aligned}
 \max & : \mathbb{N} \rightarrow \mathbb{N} \rightarrow \mathbb{N} \\
 \max \text{ zero } y & = y \\
 \max x \text{ zero } & = x \\
 \max (\text{suc } x) (\text{suc } y) & = \text{suc } (\max x y)
 \end{aligned} \tag{3}$$

we do not get `max x zero = x` but only `max (suc x') zero = suc x'`. This makes a difference for dependent type checking where we evaluate *open terms* with free variables like x . In this paper we provide a proof that the translation from a list of clauses to a case tree preserves the first-match semantics of the clauses. More precisely, we prove that if the arguments given to a function match a clause and all previous clauses produce a mismatch,³ then the case tree produced by elaborating the clauses also computes for the given arguments and the result is the same as the one given by the clause.

Contributions.

- We present a dependently typed core language with inductive datatypes, coinductive record types and an identity type. The language is focused [Andreoli 1992; Krishnaswami 2009; Zeilberger 2008]: terms of our language correspond to the non-invertible rules to introduce and eliminate these types, while the invertible rules constitute case trees.
- We are the first to present a coverage checking algorithm for fully dependent copatterns. Our algorithm desugars deep copattern matching to well-typed case trees in our core language.
- We prove correctness: if the desugaring succeeds, then the behaviour of the case tree corresponds precisely to the first-match semantics of the given clauses.
- We have implemented a new version of the algorithm used by Agda for checking the left-hand sides of a definition by dependent (co)pattern matching, which has been released as part of Agda 2.5.4.⁴ At the time of writing the effort to remodel the elaboration to a case tree according to the theory presented in this paper is still ongoing, but our work so far has already uncovered and fixed multiple issues in the old implementation [Agda issue 2017a,b,c,d, 2018a,b]. Our algorithm could also be used by other implementations of dependent pattern matching such as the Equations package for Coq [Sozeau 2010], Idris [2013], and Lean [de Moura et al. 2015].

This paper was born out of a practical need that arose while reimplementing the elaboration algorithm for Agda: it was not clear to us what exactly we wanted to implement, and we did not find sufficiently precise answers in the existing literature. Our main goal in this paper is therefore to give a precise description of the language, the elaboration algorithm, and the high-level properties we expect them to have. This also means we do not focus on fully developing the metatheory of the language or giving detailed proofs for all the basic properties one would expect.

³ Note that, in the example, the open term `max x zero` does not produce a mismatch with the first clause since it could match if variable x was replaced by `zero`. In the first-match semantics, evaluation of `max x zero` is stuck.

⁴ Agda 2.5.4 released on 2018/06/02, changelog: <https://hackage.haskell.org/package/Agda-2.5.4/changelog>.

We start by introducing definitions by dependent (co)pattern matching and our elaboration algorithm to a case tree by a number of examples in Sect. 2. We then describe our core language in Sect. 3: the syntax, the rules for typing and equality, and the evaluation rules. In Sect. 4 we give the syntax and rules for case trees, and prove that a function defined by a well-typed case tree satisfies type preservation and coverage. Finally, in Sect. 5 we describe the rules for elaborating a definition by dependent (co)pattern matching to a well-typed case tree, and prove that this translation preserves the computational meaning of the given clauses. Sect. 6 discusses related work, and Sect. 7 concludes.

2 ELABORATING DEPENDENT (CO)PATTERN MATCHING BY EXAMPLE

Before we move on to the general description of our core language and the elaboration process, we give some examples of definitions by (co)pattern matching and how our algorithm elaborates them to a case tree. The elaboration works on a configuration $\Gamma \vdash P \mid u : C$ consisting of:

- A context Γ , i.e. a list of variables annotated with types. Initially Γ is the empty context ϵ .
- The current target type C . This type may depend on variables bound in Γ . Initially C is the type of the function being defined.
- A representation of the left-hand side u . In the end u should have type C in context Γ . Initially u is the function being defined itself.
- A list of partially deconstructed user clauses P . Initially these are the clauses as written by the user.

These four pieces of data together describe the current state of elaborating the definition.

The elaboration algorithm transforms this state step by step until the user clauses are deconstructed completely. In the examples below, we annotate each step with a label such as SPLITCON or INTRO, linking it to the general rules given in Sect. 5.

Example 1. Let us define a function $\max : \mathbb{N} \rightarrow \mathbb{N} \rightarrow \mathbb{N}$ by pattern matching as in the introduction (3). The initial configuration is $\vdash P_0 \mid \max : \mathbb{N} \rightarrow \mathbb{N} \rightarrow \mathbb{N}$ where

$$P_0 = \begin{cases} \text{zero} & j & \hookrightarrow j \\ i & \text{zero} & \hookrightarrow i \\ (\text{suc } k) & (\text{suc } l) & \hookrightarrow \text{suc } (\max k l) \end{cases} \quad (4)$$

The first operation we need is to introduce a new variable m (rule INTRO). It transforms the initial problem into $(m : \mathbb{N}) \vdash P_1 \mid \max m : \mathbb{N} \rightarrow \mathbb{N}$ where

$$P_1 = \begin{cases} [m / ? \text{zero}] & j & \hookrightarrow j \\ [m / ? i] & \text{zero} & \hookrightarrow i \\ [m / ? \text{suc } k] & (\text{suc } l) & \hookrightarrow \text{suc } (\max k l) \end{cases} \quad (5)$$

This operation strips the first user pattern from each clause and replaces it by a constraint $m / ? p$ that it should be equal to the newly introduced variable m . We write these constraints between brackets in front of each individual clause.

The next operation we need is to perform a case analysis on the variable m (rule SPLITCON).⁵ This transforms the problem into two subproblems $\vdash P_2 \mid \max \text{zero} : \mathbb{N} \rightarrow \mathbb{N}$ and $(p : \mathbb{N}) \vdash P_3 \mid \max (\text{suc } p) : \mathbb{N} \rightarrow \mathbb{N}$ where

$$P_2 = \begin{cases} [\text{zero} / ? \text{zero}] & j & \hookrightarrow j \\ [\text{zero} / ? i] & \text{zero} & \hookrightarrow i \\ [\text{zero} / ? \text{suc } k] & (\text{suc } l) & \hookrightarrow \text{suc } (\max k l) \end{cases} \quad (6)$$

⁵At this point we could also introduce the variable for the second argument of \max , the elaboration algorithm is free to choose either option.

$$P_3 = \begin{cases} [\text{succ } p / ? \text{ zero}] & j \quad \hookrightarrow j \\ [\text{succ } p / ? i] & \text{zero} \quad \hookrightarrow i \\ [\text{succ } p / ? \text{succ } k] & (\text{succ } l) \quad \hookrightarrow \text{succ } (\text{max } k l) \end{cases} \quad (7)$$

We simplify the constraints, removing those clauses with absurd constraints:

$$P_2 = \begin{cases} j \quad \hookrightarrow j \\ [\text{zero} / ? i] & \text{zero} \quad \hookrightarrow i \end{cases} \quad P_3 = \begin{cases} [\text{succ } p / ? i] & \text{zero} \quad \hookrightarrow i \\ [p / ? k] & (\text{succ } l) \quad \hookrightarrow \text{succ } (\text{max } k l) \end{cases} \quad (8)$$

We continue applying these operations `INTRO` and `SPLITCON` (introducing a new variable and case analysis on a variable) until the first clause has no more user patterns and no more constraints where the left-hand side is a constructor. For example, for P_2 we get after one more introduction step $(n : \mathbb{N}) \vdash P_4 \mid \text{max zero } n : \mathbb{N}$ where

$$P_4 = \begin{cases} [n / ? j] \quad \hookrightarrow j \\ [\text{zero} / ? i, n / ? \text{zero}] \quad \hookrightarrow i \end{cases} \quad (9)$$

We solve the remaining constraint in the first clause by instantiating $j := n$. This means we are done and we have $\text{max zero } n = j[n/j] = n$ (rule `DONE`). Similarly, elaborating $(p : \mathbb{N}) \vdash P_3 \mid \text{max } (\text{succ } p) : \mathbb{N} \rightarrow \mathbb{N}$ (with rules `INTRO`, `SPLITCON`, and `DONE`) gives us $\text{max } (\text{succ } p) \text{ zero} = \text{succ } p$ and $\text{max } (\text{succ } p) (\text{succ } q) = \text{succ } (\text{max } p q)$.

We record the operations used when elaborating the clauses in a *case tree*. Our syntax for case trees is close to the normal term syntax in other languages: λx . for introducing a new variable and $\text{case}_x\{\}$ for a case split. For `max`, we get the following case tree:

$$\lambda m. \text{case}_m \left\{ \begin{array}{l} \text{zero} \mapsto \lambda n. n \\ \text{succ } p \mapsto \lambda n. \text{case}_n \left\{ \begin{array}{l} \text{zero} \mapsto \text{succ } p \\ \text{succ } q \mapsto \text{succ } (\text{max } p q) \end{array} \right\} \end{array} \right\} \quad (10)$$

Example 2. Next we take a look at how to elaborate definitions using copatterns. For the `cozero` example (1), we have the initial configuration $\vdash P_0 \mid \text{cozero} : \text{CoNat}$ where:

$$P_0 = \begin{cases} \text{.iszero} \quad \hookrightarrow \text{true} \\ \text{.pred } \emptyset \quad \hookrightarrow \text{impossible} \end{cases} \quad (11)$$

Here we need a new operation to split on the result type `CoNat` (rule `COSPLIT`). This produces two subproblems $\vdash P_1 \mid \text{cozero .iszero}$ and $\vdash P_2 \mid \text{cozero .pred} : \text{cozero .iszero} \equiv_{\text{Bool}} \text{false} \rightarrow \text{CoNat}$ where

$$P_1 = \{ \hookrightarrow \text{true} \} \quad P_2 = \{ \emptyset \hookrightarrow \text{impossible} \} \quad (12)$$

The first problem is solved immediately with `cozero .iszero = true` (rule `DONE`). In the second problem we introduce the variable $x : \text{cozero .iszero} \equiv_{\text{Bool}} \text{false}$ (rule `INTRO`) and note that `cozero .iszero = true` from the previous branch, hence $x : \text{true} \equiv_{\text{Bool}} \text{false}$. Since $\text{true} \equiv_{\text{Bool}} \text{false}$ is an empty type (technically, since unification of `true` with `false` results in a conflict), we can perform a case split on x with zero cases (rule `SPLITEMPTY`), solving the problem.

In the resulting case tree, the syntax for a split on the result type is `record\{\}`, with one subtree for each field of the record type:

$$\text{record} \left\{ \begin{array}{l} \text{iszero} \mapsto \text{true} \\ \text{pred} \mapsto \lambda x. \text{case}_x\{\} \end{array} \right\} \quad (13)$$

For the next examples, we omit the details of the elaboration process and only show the definition by pattern matching and the resulting case tree.

246 **Example 3.** Consider the type `CStream` of C streams: potentially infinite streams of numbers that
 247 end on a zero. We define this as a record where the `tail` field has two extra arguments enforcing
 248 that we can only take the tail if the head is `suc m` for some `m`.

```
249
250     record self : CStream : Set where
251       head : ℕ
252       tail  : (m : ℕ) → self .head ≡ℕ suc m → CStream
```

(14)

253 Here, the name `self` is bound to the current record instance, allowing later projections to depend
 254 on prior projections.

255 Now consider the function `countdown` that creates a C stream counting down from a given
 256 number `n`:

```
257     countdown : ℕ → CStream
258     countdown n     .head      = n
259     countdown zero .tail m 0
260     countdown (suc m) .tail m refl = countdown m
```

(15)

262 Our elaboration algorithm applies rules `INTRO`, `COSPLIT`, `SPLITCON`, `SPLITEMPTY`, `SPLITEQ`, and
 263 `DONE` in sequence to translate this definition to the following case tree:

$$264 \quad \lambda n. \text{record} \left\{ \begin{array}{l} \text{head} \mapsto n \\ \text{tail} \mapsto \lambda m, p. \text{case}_n \left\{ \begin{array}{l} \text{zero} \mapsto \text{case}_p \{ \} \\ \text{suc } n' \mapsto \text{case}_p \{ \text{refl} \mapsto \mathbb{1}^m (\text{countdown } m) \} \end{array} \right\} \end{array} \right\} \quad (16)$$

268 Note the extra annotation $\mathbb{1}_m$ after the case split on $p : \text{suc } m \equiv_{\mathbb{N}} \text{suc } n'$. This is a substitution
 269 (in this case the identity substitution on $(m : \mathbb{N})$) necessary for the evaluation rules of the case
 270 tree when matching on `refl`. It reflects the fact that n' went out of scope after the case split on
 271 `refl : suc n' ≡ℕ suc m` (since unification instantiated it with m) so only the variable m can still be
 272 used after this point.

274 **Example 4.** This example is based on issue #2896 on the Agda bug tracker [Agda issue 2018b].
 275 The problem was that Agda's old elaboration algorithm threw away a part of the pattern written
 276 by the user. This meant the definition could be elaborated to a different case tree from the one
 277 intended by the user.

278 The (simplified) example consists of the following datatype `D` and function `foo`:

```
279     data D (m : ℕ) : Set where
280       c : (n : ℕ) (p : n ≡ℕ m) → D m
281       foo : (m : ℕ) → D (suc m) → ℕ
282       foo m (c (suc n) refl) = m + n
```

(17)

283 The old algorithm would ignore the pattern `suc n` in the definition of `foo` because it corresponds
 284 to a forced pattern after the case split on `refl`. Our elaboration instead produces the following case
 285 tree (using rules `INTRO`, `SPLITCON`, `SPLITEQ`, and `DONE`):

$$286 \quad \lambda m, x. \text{case}_x \left\{ c \ n \ p \mapsto \text{case}_p \left\{ \text{refl} \mapsto \mathbb{1}^m (m + m) \right\} \right\} \quad (18)$$

288 Even though this case tree does not match on the `suc` constructor, it implements the same compu-
 289 tational behaviour as the clause in the definition of `foo` because the first argument of `c` is forced to
 290 be `suc m` by the typing rules.

291 This example also shows another feature supported by our elaboration algorithm, namely that
 292 two different variables m and n in the user syntax may correspond to the same variable m in the
 293 core syntax. In effect, n is treated as a let-bound variable with value m .

294

Example 5. This example is based on issue #2964 on the Agda bug tracker [Agda issue 2018a]. The problem was that Agda was using a too liberal version of the first-match semantics that was not preserved by the translation to a case tree. The problem occurred for the following definition:

$$\begin{aligned} f &: (A : \text{Set}) \rightarrow A \rightarrow \text{Bool} \rightarrow (A \equiv_{\text{Set}} \text{Bool}) \rightarrow \text{Bool} \\ f \text{ [Bool] true true refl} &= \text{true} \\ f \text{ _ _ _ _} &= \text{false} \end{aligned} \quad (19)$$

This function is elaborated (both by Agda's old algorithm and by ours) to the following case tree (using rules INTRO, SPLITCON, SPLITEQ, and DONE):

$$\lambda A, x, y, p. \text{case}_y \left\{ \begin{array}{l} \text{true} \mapsto \text{case}_p \left\{ \begin{array}{l} \text{refl} \mapsto \mathbb{1}_{x,y} \text{case}_x \left\{ \begin{array}{l} \text{true} \mapsto \text{true} \\ \text{false} \mapsto \text{false} \end{array} \right\} \\ \text{false} \mapsto \text{false} \end{array} \right\} \\ \text{false} \mapsto \text{false} \end{array} \right\} \quad (20)$$

According to the (liberal) first-match semantics, we should have $f \text{ Bool false } y \text{ } p = \text{false}$ for any $y : \text{Bool}$ and $p : \text{Bool} \equiv_{\text{Set}} \text{Bool}$, but this is not true for the case tree since evaluation gets stuck on the variable y . Another possibility is to start the case tree by a split on p (after introducing all the variables), but this case tree still gets stuck on the variable p . In fact, there is no well-typed case tree that implements the first-match semantics of these clauses since we cannot perform a case split on $x : A$ before splitting on p .

One radical solution for this problem would be to only allow case trees where the case splits are performed in order from left to right. However, this would mean the typechecker must reject many definitions such as f in this example, because the type of x is not known to be a datatype until the case split on $A \equiv_{\text{Set}} \text{Bool}$. Instead we choose to keep the elaboration as it is and restrict the first-match semantics of clauses. In the example of f , this change means that we can only go to the second clause once all three arguments x , y and p are constructors, and at least one of them produces a mismatch.

3 CORE LANGUAGE

In this section we introduce a basic type theory for studying definitions by dependent (co)pattern matching. It has support for dependent function types, an infinite hierarchy of predicative universes, equality types, inductive datatypes and coinductive records.

To keep the work in this paper as simple as possible, we leave out many features commonly included in dependently typed languages, such as lambda expressions and inductive families of datatypes (other than the equality type). These features can nevertheless be encoded in our language, see Sect. 3.5 for details.

Note also that we do not include any rules for η -equality, neither for lambda expressions (which do not exist) nor for records (which can be coinductive hence do not satisfy η). Sect. 3.5 discusses how our language could be extended with η -rules.

3.1 Syntax of the core type theory

Expressions of our type theory are almost identical to Agda's internal term language. All function applications are in spine-normal form, so the head symbol of an application is exposed, be it variable x , data D or record type R , or defined function f . We generalize applications to eliminations e by including projections $\cdot\pi$ in spines \vec{e} . Any expression is in weak head normal form but $f \vec{e}$, which is computed via pattern matching (see Sect. 3.4).

344			
345	$A, B, u, v ::= w$		weak head normal form
346	$f \bar{e}$		defined function applied to eliminations
347	$W, w ::= (x : A) \rightarrow B$		dependent function type
348	Set_ℓ		universe ℓ
349	$D \bar{u}$		datatype fully applied to parameters
350	$R \bar{u}$		record type fully applied to parameters
351	$u \equiv_A v$		equality type
352	$x \bar{e}$		variable applied to eliminations
353	$c \bar{u}$		constructor fully applied to arguments
354	refl		proof of reflexivity

355 Any expression but $c \bar{u}$ or refl can be a type; the first five weak head normal forms are definitely
 356 types. Any type has in turn a type, specifically some universe Set_ℓ . Syntax is colored according
 357 to the Agda conventions: primitives and defined symbols are blue, constructors are green, and
 358 projections are pink.

360			
361	$e ::= u$	application	(22)
362	$\cdot \pi$	projection	

363 Binary application $\boxed{u e}$ is defined as a partial function on the syntax: for variables and functions
 364 it is defined by $(x \bar{e}) e = x (\bar{e}, e)$ and $(f \bar{e}) e = f (\bar{e}, e)$ respectively, otherwise it is undefined.

365 Patterns are generated from variables and constructors. In addition, we have *forced* and *absurd*
 366 patterns. Since we are matching spines, we also consider projections as patterns, or more precisely,
 367 as *copatterns*.

368			
369	$p ::= x$	variable pattern	
370	refl	pattern for reflexivity proof	
371	$c \bar{p}$	constructor pattern	
372	$[c] \bar{p}$	forced constructor pattern	
373	$[u]$	forced argument	(23)
374	\emptyset	absurd pattern	

375			
376	$q ::= p$	application copattern	
377	$\cdot \pi$	projection copattern	

378 Forced patterns [Brady et al. 2003] appear with dependent types; they are either entirely forced
 379 arguments $[u]$, which are Agda's *dot patterns*, or only the constructor is forced $[c] \bar{p}$. An argument
 380 can be forced by a match against refl somewhere in the surrounding (co)pattern. However, some-
 381 times we want to bind variables in a forced argument; in this case, we revert to forced constructors.
 382 Absurd patterns⁶ are used to indicate that the type at this place is empty, i.e. no constructor can
 383 possibly match. They are also used to indicate an empty copattern split, i.e. a copattern split on a
 384 record type with no projections. This allows us in particular to define the unique element tt of the
 385 unit record, which has no projections at all, by the clause $\text{tt } \emptyset = \text{impossible}$.

386 The *pattern variables* $PV(\bar{q})$ is the list of variables in \bar{q} that appear outside forcing brackets $[\cdot]$. By
 387 removing the forcing brackets, patterns p embed into terms $[p]$, and copatterns q into eliminations
 388 $[q]$, except for the absurd pattern \emptyset .

389			
390	$[x] = x$	$[c \bar{p}] = c [p]$	$[[u]] = u$
391	$[\text{refl}] = \text{refl}$	$[[c] \bar{p}] = c [p]$	$[\cdot \pi] = \cdot \pi$

392 ⁶Absurd patterns are written $()$ in Agda syntax.

Constructors take a list of arguments whose types can depend on all previous arguments. The constructor parameters are given as a list $x_1:A_1, \dots, x_n:A_n$ with pairwise distinct x_i where A_i can depend on x_1, \dots, x_{i-1} . This list can be conceived as a *cons*-list, then it is called a *telescope*, or as a *snoc*-list, then we call it a *context*.

$$\begin{array}{ll} \Gamma ::= \epsilon & \text{empty context} \\ | \Gamma(x:A) & \text{context extension} \end{array} \quad \begin{array}{ll} \Delta ::= \epsilon & \text{empty telescope} \\ | (x:A)\Delta & \text{non-empty telescope} \end{array} \quad (25)$$

Context and telescopes can be regarded as finite maps from variables to types, and we require $x \notin \text{dom}(\Gamma)$ and $x \notin \text{dom}(\Delta)$ in the above grammars. We implicitly convert between contexts and telescopes, but there are still some conceptual differences. Contexts are always *closed*, i.e. its types only refer to variables bound prior in the same context. In contrast, we allow *open* telescopes whose types can also refer to some surrounding context. Telescopes can be naturally thought of as *context extensions*, and if Γ is a context and Δ a telescope in context Γ where $\text{dom}(\Gamma)$ and $\text{dom}(\Delta)$ are disjoint, then $\boxed{\Gamma\Delta}$ defined by $\Gamma\epsilon = \Gamma$ and $\Gamma((x:A)\Delta) = (\Gamma(x:A))\Delta$ is a new valid context. We embed telescopes in the syntax of declarations, but contexts are used in typing rules exclusively.

Given a telescope Δ , let $\hat{\Delta}$ be Δ without the types, i.e. the variables of Δ in order. Further, we define $\boxed{\Delta \rightarrow C}$ as the iterated dependent function type via $\epsilon \rightarrow C = C$ and $(x:A)\Delta \rightarrow C = (x:A) \rightarrow (\Delta \rightarrow C)$.

A development in our core type theory is a list of declarations, of which there are three kinds: data type, record type, and function declarations. The input to the type checker is a list of unchecked declarations decl^\ominus , and the output a list of checked declarations decl^\oplus , called a *signature* Σ .

$$\begin{array}{ll} s & ::= \ominus & \text{status: unchecked} \\ & | \oplus & \text{status: checked} \\ \text{decl}^\ominus & ::= \text{data } D \Delta : \text{Set}_\ell \text{ where } \overline{\text{con}} & \text{datatype declaration} \\ & | \text{record } \text{self} : R \Delta : \text{Set}_\ell \text{ where } \overline{\text{field}} & \text{record declaration} \\ & | \text{definition } f : A \text{ where } \overline{\text{cls}}^s & \text{function declaration} \\ \text{con} & ::= c \Delta & \text{constructor declaration} \\ \text{field} & ::= \pi : A & \text{field declaration} \\ \text{cls}^\ominus & ::= \bar{q} \hookrightarrow \text{rhs} & \text{unchecked clause} \\ \text{cls}^\oplus & ::= \Delta \vdash \bar{q} \hookrightarrow u : B & \text{checked clause} \\ \text{rhs} & ::= u & \text{clause body: expression} \\ & | \text{impossible} & \text{empty body for absurd pattern} \\ \Sigma & ::= \overline{\text{decl}^\oplus} & \text{signature} \end{array} \quad (26)$$

A *data type* D can be parameterized by telescope Δ and inhabits one of the universes Set_ℓ . Each of its constructors c_i (although there might be none) takes a telescope Δ_i of arguments that can refer to the parameters in Δ . The full type of c_i could be $\Delta\Delta_i \rightarrow D \hat{\Delta}$, but we never apply constructors to the data parameters explicitly.

A *record type* R can be thought of as a single constructor data type; its fields $\pi_1:A_1, \dots, \pi_n:A_n$ would be the constructor arguments. The field list behaves similar to a telescope, the type of each field can depend on the value of the previous fields. However, these values are referred to via *self*. π_i

442 where variable *self* is a placeholder for the value of the whole record.⁷ The full type of projection
 443 π_i could be $\Delta(\text{self} : R \hat{\Delta}) \rightarrow A_i$, but like for constructors, we do not apply a projection explicitly
 444 to the record parameters.

445 Even though we do not spell out the conditions for ensuring totality in this paper, like *positiv-*
 446 *ity*, *termination*, and *productivity* checking, data types, when recursive, should be thought of as
 447 inductive types, and record types, when recursive, as coinductive types [Abel et al. 2013]. Thus,
 448 there is no dedicated constructor for records; instead, concrete records are defined by what their
 449 projections compute.

450 Such definitions are subsumed under the last alternative dubbed *function declaration*. More pre-
 451 cisely, these are *definitions by copattern matching* which include record definitions. Each clause
 452 defining the constant $f : A$ consists of a list of copatterns \bar{q} and right hand side *rhs*. The copatterns
 453 eliminate type A into the type of the *rhs* which is either a term u or the special keyword **impossible**,
 454 in case one of the copatterns q_i contains an absurd pattern \emptyset . The intended semantics is that if an
 455 application $f \bar{e}$ matches a left hand side $f \bar{q}$ with substitution σ , then $f \bar{e}$ reduces to *rhs* under σ . For
 456 efficient computation of matching, we require *linearity of pattern variables* for checked clauses:
 457 each variable in \bar{q} occurs only once in a non-forced position.

458 While checking declarations, the typechecker builds up a signature Σ of already checked (parts
 459 of) declarations. Checked clauses are the elaboration (sections 2 and 5) of the corresponding
 460 unchecked clauses: they are non-overlapping and supplemented by a telescope Δ holding the types
 461 of the pattern variables and the type B of left and right hand side. Further, checked clauses do not
 462 contain absurd patterns.

463 In the signature, the last entry might be incomplete, e.g. a data type missing some constructors, a
 464 record type missing some fields, or a function missing some clauses. During checking a declaration,
 465 we might add already checked parts of the declaration, dubbed *snippets*, to the signature.

467	$Z ::=$	data $D \Delta : \text{Set}_\ell$	data type signature	
468		constructor $c \Delta_c : D \Delta$	constructor signature	
469		record $R \Delta : \text{Set}_\ell$	record type signature	
470		projection $\text{self} : R \Delta \vdash .\pi : A$	projection signature	(27)
471		definition $f : A$	function signature	
472		clause $\Delta \vdash f \bar{q} \leftrightarrow v : B$	function clause	

473
 474 Adding a snippet Z to a signature Σ , written $\boxed{\Sigma, Z}$ is always defined if Z is a data or record
 475 type or function signature; in this case, the corresponding declaration is appended to Σ . Adding
 476 a constructor signature **constructor** $c \Delta_c : D \Delta$ is only defined if the *last* declaration in Σ is
 477 (**data** $D \Delta : \text{Set}_\ell$ **where** $\overline{c\bar{o}n}$) and c is not part of $\overline{c\bar{o}n}$ yet. Analogous conditions apply when
 478 adding projection snippets. Function clauses can be added if the last declaration of Σ is a function
 479 declaration with the same name. We trust the formal definition of Σ, Z to the imagination of the
 480 reader. The conditions ensure that we do not add new constructors to a data type that is already
 481 complete or new fields to a completed record declaration. Such additions could destroy coverage
 482 for functions that have already been checked. Late addition of function clauses would not pose a
 483 problem, but that feature would be obsolete for our type theory anyway.

484 *Membership of a snippet* is written $\boxed{Z \in \Sigma}$ and a decidable property with the obvious definition.
 485 These operations on the signature will be used in the inference rules of our type theory. Since we
 486 only refer to a constructor c in conjunction with its data type D , constructors can be overloaded,
 487 and likewise projections.

488
 489 ⁷*self* is the analog of Java's *this*, but like in Scala's *trait*, the name can be chosen.

3.2 Typing and equality

Our type theory employs the following basic typing and equality judgments, which are relative to a signature Σ .

$\Sigma \vdash \Gamma$	context Γ is well-formed
$\Sigma; \Gamma \vdash_{\ell} \Delta$	in context Γ , telescope Δ is well-formed and ℓ -bounded
$\Sigma; \Gamma \vdash u : A$	in context Γ , term u has type A
$\Sigma; \Gamma \vdash \bar{u} : \Delta$	in context Γ , term list \bar{u} instantiates telescope Δ
$\Sigma; \Gamma \mid u : A \vdash \bar{e} : B$	in context Γ , head u of type A is eliminated via \bar{e} to type B
$\Sigma; \Gamma \vdash u = v : A$	in context Γ , terms u and v are equal of type A
$\Sigma; \Gamma \vdash \bar{u} = \bar{v} : \Delta$	in context Γ , term lists \bar{u} and \bar{v} are equal instantiations of Δ
$\Sigma; \Gamma \mid u : A \vdash \bar{e} = \bar{e}' : B$	\bar{e} and \bar{e}' are equal eliminations of head $u : A$ to type B in Γ

In all these judgements, the signature Σ is fixed, thus we usually omit it, e.g. in the inferences rules.

We further define some shorthands for type-level judgements when we do not care about the universe level ℓ :

$\Sigma; \Gamma \vdash \Delta$	$\iff \exists \ell. \Sigma; \Gamma \vdash_{\ell} \Delta$	well-formed telescope
$\Sigma; \Gamma \vdash A$	$\iff \exists \ell. \Sigma; \Gamma \vdash A : \mathbf{Set}_{\ell}$	well-formed type
$\Sigma; \Gamma \vdash A = B$	$\iff \exists \ell. \Sigma; \Gamma \vdash A = B : \mathbf{Set}_{\ell}$	equal types

$\boxed{\Gamma \vdash u : A}$ Entails $\vdash \Gamma$ and $\Gamma \vdash A$.

Types.

$\vdash \Gamma$	$\Gamma \vdash A : \mathbf{Set}_{\ell}$	$\Gamma(x : A) \vdash B : \mathbf{Set}_{\ell'}$
$\Gamma \vdash \mathbf{Set}_{\ell} : \mathbf{Set}_{\ell+1}$	$\Gamma \vdash (x : A) \rightarrow B : \mathbf{Set}_{\max(\ell, \ell')}$	
$\mathbf{data} \ D \ \Delta : \mathbf{Set}_{\ell} \in \Sigma$	$\Gamma \vdash \bar{u} : \Delta$	$\mathbf{record} \ R \ \Delta : \mathbf{Set}_{\ell} \in \Sigma$
$\Gamma \vdash D \ \bar{u} : \mathbf{Set}_{\ell}$		$\Gamma \vdash R \ \bar{u} : \mathbf{Set}_{\ell}$
$\Gamma \vdash A : \mathbf{Set}_{\ell}$		
$\Gamma \vdash u : A$		
$\Gamma \vdash v : A$		
$\Gamma \vdash u \equiv_A v : \mathbf{Set}_{\ell}$		

Heads ($h ::= x \in \mid f \in$) and applications $h \bar{e}$.

$\vdash \Gamma$	$x : A \in \Gamma$	$\vdash \Gamma$	$\mathbf{definition} \ f : A \in \Sigma$	$\Gamma \vdash h : A$	$\Gamma \mid h : A \vdash \bar{e} : C$
$\Gamma \vdash x \in : A$		$\Gamma \vdash f \in : A$		$\Gamma \vdash h \bar{e} : C$	

Values.

$\mathbf{constructor} \ c \ \Delta_c : D \ \Delta \in \Sigma$	$\Gamma \vdash \bar{u} : \Delta$	$\Gamma \vdash \bar{v} : \Delta_c[\bar{u} / \Delta]$	$\Gamma \vdash u : A$
$\Gamma \vdash c \ \bar{v} : D \ \bar{u}$			$\Gamma \vdash \mathbf{refl} : u \equiv_A u$

Conversion.

$\Gamma \vdash u : A$	$\Gamma \vdash A = B$
$\Gamma \vdash u : B$	

Fig. 1. Typing rules for expressions.

$$\begin{array}{c}
540 \quad \boxed{\Gamma \mid u : A \vdash \bar{e} : C} \quad \text{If } \Gamma \vdash u : A \text{ then } \Gamma \vdash C. \\
541 \\
542 \quad \frac{}{\Gamma \mid u : A \vdash \epsilon : A} \quad \frac{\Gamma \vdash v : A \quad \Gamma \mid u \ v : B[v/x] \vdash \bar{e} : C}{\Gamma \mid u : (x : A) \rightarrow B \vdash v \ \bar{e} : C} \\
543 \\
544 \\
545 \quad \frac{\text{projection } \text{self} : \mathbf{R} \ \Delta \vdash .\pi : A \in \Sigma \quad \Gamma \mid u .\pi : A[\bar{v}/\Delta, u/\text{self}] \vdash \bar{e} : C}{\Gamma \mid u : \mathbf{R} \ \bar{v} \vdash .\pi \ \bar{e} : C} \\
546 \\
547 \\
548 \quad \frac{\Gamma \vdash A = A' \quad \Gamma \mid u : A' \vdash \bar{e} : C}{\Gamma \mid u : A \vdash \bar{e} : C} \\
549 \\
550 \\
551 \\
552 \\
553 \\
554
\end{array}$$

Fig. 2. The typing rules for eliminations.

$$\begin{array}{c}
555 \quad \boxed{\Gamma \vdash_{\ell} \Delta} \quad \text{Entails } \vdash \Gamma. \\
556 \\
557 \quad \frac{\vdash \Gamma}{\Gamma \vdash_{\ell} \epsilon} \quad \frac{\Gamma \vdash A : \mathbf{Set}_{\ell'} \quad \Gamma(x : A) \vdash_{\ell} \Delta}{\Gamma \vdash_{\ell} (x : A)\Delta} \quad \ell' \leq \ell \\
558 \\
559 \quad \boxed{\Gamma \vdash \bar{u} : \Delta} \quad \text{Precondition: } \Gamma \vdash \Delta. \\
560 \\
561 \quad \frac{}{\Gamma \vdash \epsilon : \epsilon} \quad \frac{\Gamma \vdash u : A \quad \Gamma \vdash \bar{u} : \Delta[u/x]}{\Gamma \vdash u \ \bar{u} : (x : A)\Delta} \\
562 \\
563 \\
564 \\
565 \\
566 \\
567
\end{array}$$

Fig. 3. The typing rules for telescopes and lists of terms.

567 In the inference rules, we make use of substitutions. *Substitutions* σ, τ, ν are partial maps from
568 variable names to terms with a finite domain. If $\text{dom}(\sigma)$ and $\text{dom}(\tau)$ are disjoint, then $\boxed{\sigma \uplus \tau}$
569 denotes the union of these maps. We write the substitution that maps the variables x_1, \dots, x_n to
570 the terms v_1, \dots, v_n (and is undefined for all other variables) by $[v_1/x_1; \dots; v_n/x_n]$. In partic-
571 ular, the empty substitution $[\]$ is undefined for all variables. If $\Delta = (x_1 : A_1) \dots (x_n : A_n)$ is a
572 telescope and $\bar{v} = v_1, \dots, v_n$ is a list of terms, we may write $[\bar{v}/\Delta]$ for the substitution $[\bar{v}/\hat{\Delta}]$,
573 i.e. $[v_1/x_1; \dots; v_n/x_n]$. In particular, the identity substitution $\mathbb{1}_{\Gamma} = [\hat{\Gamma}/\Gamma]$ maps all variables in
574 Γ to themselves. We also use the identity substitution as a weakening substitution, allowing us to
575 forget about all variables that are not in Γ . If $x \in \text{dom}(\sigma)$, then $\boxed{\sigma \setminus x}$ is defined by removing x
576 from the domain of σ .

577 Application of a substitution σ to a term u is written as $\boxed{u\sigma}$ and is defined as usual by replacing
578 all (free) variables in u by their values given by σ , avoiding variable capture via suitable renaming
579 of bound variables. Like function application, this is a partial operation on the syntax; for instance,
580 $(x.\pi)[c/x]$ is undefined as constructors cannot be the head of an elimination. Thus, when a substi-
581 tution appears in an inference rule, its definedness is an implicit premise of the rule. Also, such
582 pathological cases are ruled out by typing. Well-typed substitutions can always be applied to well-
583 typed terms (established in Lemma 8). Substitution composition $\boxed{\sigma; \tau}$ shall map the variable x to
584 the term $(x\sigma)\tau$. Note the difference between $\sigma; \tau$ and $\sigma \uplus \tau$: the former applies first σ and then τ
585 in sequence, while the latter applies σ and τ in parallel to disjoint parts of the context. Application
586 of a substitution to a pattern $\boxed{p\sigma}$ is defined as $[p]\sigma$.

In addition to substitutions on terms, we also make use of substitutions on patterns called *pattern substitutions*. A pattern substitution ρ assigns to each variable a pattern. We reuse the same syntax for pattern substitutions as for normal substitutions. Any pattern substitution ρ can be used as a normal substitution $[\rho]$ defined by $x[\rho] = [x\rho]$.

The rules for the typing judgement $\boxed{\Gamma \vdash t : A}$ are listed in Fig. 1. The type formation rules introduce an infinite hierarchy of predicative universes Set_ℓ without cumulativity. The formation rules for data and record types make use of the judgment $\Gamma \vdash \bar{u} : \Delta$ to type argument lists, same for the constructor rule, which introduces a data type. Further, *refl* introduces the equality type. All expressions involved in these rules are fully applied, but this changes when we come to the elimination rules. The types of heads, i.e. variables x or defined constants f are found in the context or signature.

The rules for applying heads u to spines \bar{e} , judgement $\boxed{\Gamma \mid u : A \vdash \bar{e} : C}$, are presented in Fig. 2. For checking arguments, the type of the head is sufficient, and it needs to be a function type. To check projections, we need also the value u of the head that replaces *self* in the type of the projection. We may need to convert the type of the head to a function or record type to apply these rules, hence, we supply a suitable conversion rule. The result type C of this judgement need not be converted here, it can be converted in the typing judgement for expressions.

Remark 6 (Focused syntax). The reader may have observed that our expressions cover only the *non-invertible* rules in the sense of focusing [Andreoli 1992], given that we consider data types as multiplicative disjunctions and record types as additive conjunctions: Terms introduce data and eliminate records and functions. The *invertible* rules, i.e. elimination for data and equality and introduction for function space and records are covered by pattern matching (Sect. 3.4) and, equivalently, case trees (Sect. 4). This matches our intuition that all the information/choice resides with the non-invertible rules, the terms, while the choice-free pattern matching corresponding to the invertible rules only sets the stage for the decisions taken in the terms.

Fig. 3 defines judgement $\boxed{\Gamma \vdash_\ell \Delta}$ for telescope formation. The level ℓ is an upper bound for the universe levels of the types that comprise the telescope. In particular, if we consider a telescope as a nested Σ -type, then ℓ is an upper bound for the universe that hosts this type. This is important when checking that the level of a data type is sufficiently high for the level of data it contains (Fig. 4).

Using the notation $(x_1, \dots, x_n)\sigma = (x_1\sigma, \dots, x_n\sigma)$, substitution typing can be reduced to typing of lists of terms: Suppose $\vdash \Gamma$ and $\vdash \Delta$. We write $\boxed{\Gamma \vdash \sigma : \Delta}$ for $\text{dom}(\sigma) = \Delta$ and $\Gamma \vdash \hat{\Delta}\sigma : \Delta$. Likewise, we write $\boxed{\Gamma \vdash \sigma = \sigma' : \Delta}$ for $\Gamma \vdash \hat{\Delta}\sigma = \hat{\Delta}\sigma' : \Delta$.

Definitional equality $\boxed{\Gamma \vdash u = u' : A}$ is induced by rewriting function applications according to the function clauses. It is the least typed congruence over the axiom:

$$\frac{\text{clause } \Delta \vdash f \bar{q} \hookrightarrow v : B \in \Sigma \quad \Gamma \vdash \sigma : \Delta}{\Gamma \vdash f \bar{q}\sigma = v\sigma : B\sigma}$$

If $f \bar{q} \hookrightarrow v$ is a defining clause of function f , then each instance arising from a well-typed substitution σ is a valid equation. The full list of congruence and equivalence rules is given in Fig. 18 in Appendix A, together with congruence rules for applications (Fig. 19) and lists of terms (Fig. 20). As usual in dependent type theory, definitional equality on types $\Gamma \vdash A = B : \text{Set}_\ell$ is used for type conversion.

Lemma 7. *If $\Gamma \vdash \sigma : \Delta_1(x : A)\Delta_2$ then also $\Gamma \vdash \sigma : \Delta_1(\Delta_2[x\sigma / x])$.*

Lemma 8 (Substitution). *Suppose $\Gamma' \vdash \sigma : \Gamma$. Then the following hold:*

$$\begin{array}{c}
638 \quad \boxed{\Sigma \vdash Z} \quad \text{Snipped } Z \text{ is well-formed in signature } \Sigma. \\
639 \\
640 \quad \frac{\Sigma \vdash \Delta}{\Sigma \vdash \text{data } D \Delta : \text{Set}_\ell} \quad \frac{\text{data } D \Delta : \text{Set}_\ell \in \Sigma \quad \Sigma; \Delta \vdash_\ell \Delta_c}{\Sigma \vdash \text{constructor } c \Delta_c : D \Delta} \\
641 \\
642 \\
643 \quad \frac{\Sigma \vdash \Delta}{\Sigma \vdash \text{record } R \Delta : \text{Set}_\ell} \quad \frac{\text{record } R \Delta : \text{Set}_\ell \in \Sigma \quad \Sigma; \Delta(x : R \hat{\Delta}) \vdash A : \text{Set}_{\ell'}}{\Sigma \vdash \text{projection } x : R \Delta \vdash \cdot \pi : A} \quad \ell' \leq \ell \\
644 \\
645 \\
646 \quad \frac{\Sigma \vdash A}{\Sigma \vdash \text{definition } f : A} \quad \frac{\text{definition } f : A \in \Sigma \quad \Sigma \vdash \Delta \quad \Delta \mid f : A \vdash [\bar{q}] : B \quad \Delta \vdash v : B}{\Sigma \vdash \text{clause } \Delta \vdash f \bar{q} \leftrightarrow v : B} \\
647 \\
648 \\
649 \quad \boxed{\Sigma_0 \subseteq \Sigma} \quad \text{Signature } \Sigma \text{ is a valid extension of } \Sigma_0. \\
650 \\
651 \quad \frac{}{\Sigma_0 \subseteq \Sigma_0} \quad \frac{\Sigma_0 \subseteq \Sigma \quad \Sigma \vdash Z \quad \Sigma, Z \text{ defined}}{\Sigma_0 \subseteq \Sigma, Z} \\
652 \\
653 \\
654 \\
655 \\
656 \\
657 \\
658 \\
659 \\
660 \\
661 \\
662 \\
663 \\
664 \\
665 \\
666 \\
667 \\
668 \\
669 \\
670 \\
671 \\
672 \\
673 \\
674 \\
675 \\
676 \\
677 \\
678 \\
679 \\
680 \\
681 \\
682 \\
683 \\
684 \\
685 \\
686
\end{array}$$

Fig. 4. Rules for well-formed signature snippets and extension.

- If $\Gamma \vdash u : A$ then $\Gamma' \vdash u\sigma : A\sigma$.
- If $\Gamma \mid u : A \vdash \bar{e} : B$ then $\Gamma' \mid u\sigma : A\sigma \vdash \bar{e}\sigma : B\sigma$.
- If $\Gamma \vdash_\ell \Delta$ then $\Gamma' \vdash_\ell \Delta\sigma$.
- If $\Gamma \vdash \bar{u} : \Delta$ then $\Gamma' \vdash \bar{u}\sigma : \Delta\sigma$.
- If $\Gamma \vdash u = v : A$ then $\Gamma' \vdash u\sigma = v\sigma : A\sigma$.
- If $\Gamma \mid u : A \vdash \bar{e}_1 = \bar{e}_2 : C$ then $\Gamma' \mid u\sigma : A\sigma \vdash \bar{e}_1\sigma = \bar{e}_2\sigma : C\sigma$.
- If $\Gamma \vdash \bar{u}_1 = \bar{u}_2 : \Delta$ then $\Gamma' \vdash \bar{u}_1\sigma = \bar{u}_2\sigma : \Delta$.

PROOF. By mutual induction on the derivation of the given judgement. The interesting case is when u is a variable application $x \bar{e}$. Suppose that $x : A \in \Gamma$ and $\Gamma \mid x : A \vdash \bar{e} : B$, then $\Gamma' \vdash x\sigma : A\sigma$. We also know from the induction hypothesis that $\Gamma' \mid x\sigma : A\sigma \vdash \bar{e}\sigma : B\sigma$, so we have $\Gamma' \vdash x\sigma \bar{e}\sigma : B\sigma$, as we had to prove. \square

Property 9. If $\Gamma \vdash u : A$ and $\Gamma \mid u : A \vdash \bar{e} : B$ then $u \bar{e}$ is well-defined and $\Gamma \vdash u \bar{e} : B$.

3.3 Signature well-formedness

A signature Σ extends Σ_0 if we can go from Σ_0 to Σ by adding valid snippets Z , i.e. new datatypes, record types, and defined constants, but new constructors/projections/clauses only for not yet completed definitions in Σ . A signature Σ is well-formed if it is a valid extension of the empty signature ϵ . Formally, we define signature extension $\boxed{\Sigma_0 \subseteq \Sigma}$ via snippet typing $\boxed{\Sigma \vdash Z}$ by the rules in Fig. 4, and signature well-formedness $\boxed{\vdash \Sigma}$ as $\epsilon \subseteq \Sigma$. Recall that the rules for extending the signature with a constructor (resp. projection or clause) can only be used when the corresponding data type (resp. record type or definition) is the last thing in the signature, by definition of extending the signature with a snippet Σ, Z . When adding a constructor or projection, it is ensured that the *stored data* is not too big in terms of universe level ℓ ; this preserves predicativity. However, the *parameters* Δ of a data or record type of level ℓ can be *big*, they may exceed ℓ .

All typing and equality judgements are monotone in the signature, thus, remain valid under signature extensions.

$$\boxed{\Sigma \vdash u \searrow w} \quad (\Sigma \text{ fixed, dropped from rules.})$$

$$\frac{}{w \searrow w} \quad \text{clause } \frac{\Delta \vdash f \bar{q} \hookrightarrow v : A \in \Sigma \quad [\bar{e} / \bar{q}] \searrow \sigma \quad v \sigma \searrow w}{f \bar{e} \searrow w}$$

Fig. 5. Rules for weak-head normalization.

$$\boxed{\Sigma \vdash [v / \bar{p}] \searrow \sigma_{\perp}} \quad (\Sigma \text{ fixed, dropped from rules.})$$

$$\frac{}{[v / x] \searrow [v / x]} \quad \frac{}{[v / [u]] \searrow []} \quad \frac{v \searrow \text{refl}}{[v / \text{refl}] \searrow []}$$

$$\frac{v \searrow c \bar{u} \quad [\bar{u} / \bar{p}] \searrow \sigma_{\perp}}{[v / c \bar{p}] \searrow \sigma_{\perp}} \quad \frac{v \searrow c_2 \bar{u} \quad [\bar{u} / \bar{p}] \searrow \sigma_{\perp}}{[v / [c_1] \bar{p}] \searrow \sigma_{\perp}} \quad \frac{v \searrow c_2 \bar{u} \quad c_1 \neq c_2}{[v / c_1 \bar{p}] \searrow \perp}$$

$$\boxed{\Sigma \vdash [e / \bar{q}] \searrow \sigma_{\perp}}$$

$$\frac{}{[\pi / \pi] \searrow []} \quad \frac{\pi_1 \neq \pi_2}{[\pi_2 / \pi_1] \searrow \perp}$$

$$\boxed{\Sigma \vdash [\bar{e} / \bar{q}] \searrow \sigma_{\perp}}$$

$$\frac{}{[e / \epsilon] \searrow []} \quad \frac{[e / \bar{q}] \searrow \sigma_{\perp} \quad [\bar{e} / \bar{q}] \searrow \tau_{\perp}}{[e \bar{e} / \bar{q} \bar{q}] \searrow \sigma_{\perp} \uplus \tau_{\perp}}$$

Fig. 6. Rules for the pattern matching and mismatching algorithm.

Lemma 10 (Signature extension preserves inferences). *If $\Sigma; \Gamma \vdash u : A$ and $\Sigma \subseteq \Sigma'$ then also $\Sigma'; \Gamma \vdash u : A$ (and likewise for other judgements).*

Remark 11 (Coverage). The rules for extending a signature with a function definition given by a list of clauses are not strong enough to guarantee the usual properties of a language such as type preservation and progress. For example, we could define a function with no clauses at all (violating progress), or we could add a clause where all patterns are forced patterns (violating type preservation). We prove type preservation and progress only for functions that correspond to a well-typed case tree as defined in Sect. 4.

3.4 Pattern matching and evaluation rules

Evaluation to weak-head normal form $\boxed{\Sigma \vdash u \searrow w}$ is defined inductively in Fig. 5. Since our language does not contain syntax for lambda abstraction, there is no rule for β -reduction. Almost all terms are their own weak-head normal form; the only exception are applications $f \bar{e}$.

Evaluation is mutually defined with matching against (co)patterns $\boxed{\Sigma \vdash [\bar{e} / \bar{q}] \searrow \sigma_{\perp}}$ (Fig. 6). Herein, σ_{\perp} is either a substitution σ with $\text{dom}(\sigma) = \text{PV}(\bar{q})$ or the error value \perp for mismatch. Join of lifted substitutions $\sigma_{\perp} \uplus \tau_{\perp}$ is \perp if one of the operands is \perp , otherwise the join $\sigma \uplus \tau$.

A pattern variable x matches any term v , producing singleton substitution $[v / x]$. Likewise for a forced pattern $[u]$, but it does not bind any pattern variables. Projections π only match themselves, and so do constructors $c \bar{p}$, but they require evaluation $v \searrow c \bar{u}$ of the scrutinee v and subsequent successful matching $[\bar{u} / \bar{p}] \searrow \sigma$ of the arguments. For forced constructors $[c_1] \bar{p}$, the constructor

equality test is skipped, as it is ensured by typing. Constructor ($c_1 \neq c_2$) and projection ($\pi_1 \neq \pi_2$) mismatches produce \perp . We do not need to match against the absurd pattern; user clauses with absurd matches are never added to the signature. Recall that absurd patterns are not contained in clauses of the signature, thus, we need not consider them in the matching algorithm. Evaluating a function that eliminates absurdity will be stuck for lack of matching clauses.

A priori, matching can get stuck, if none of the rules apply. In particular, this happens when we try to evaluate an underapplied function or an open term, i.e. a term with free variables. For the purpose of the evaluation judgement, we would not need to track definite mismatch (\perp) separately from getting stuck. However, for the first-match semantics [Augustsson 1985] we do: There, a function should reduce with the first clause that matches while all previous clauses produce a mismatch. If matching a clause is stuck, we must not try the next one.

The first-match semantics is also the reason why either $\Sigma \vdash [e/q] \searrow \perp$ or $\Sigma \vdash [\bar{e}/\bar{q}] \searrow \perp$ alone is not sufficient to derive $\Sigma \vdash [e \bar{e}/q \bar{q}] \searrow \perp$, i.e. mismatch does not dominate stuckness, nor does it short-cut matching. Suppose a function `and` defined by the clauses `true true \hookrightarrow true` and `x y \hookrightarrow false`. If mismatch dominated stuckness, then both open terms `and false y` and `and x false` would reduce to `false`. However, there is no case tree that accomplishes this. We have to split on the first or the second variable; either way, one of the two open terms will be stuck. We cannot even decree left-to-right splitting: see Example 5 for a definition that is impossible to elaborate to a case tree using a left-to-right splitting order. Thus, we require our pattern match semantics to be faithful with *any* possible elaboration of clauses into case trees.⁸

3.5 Other language features

In comparison to dependently typed programming languages like Agda and Idris, our core language seems rather reduced. In the following, we discuss how some popular features could be translated to our core language.

Lambda abstractions and η -equality: A lambda abstraction $\lambda x. t$ in context Γ can be lifted to the top-level and encoded as auxiliary function $f \hat{\Gamma} x \hookrightarrow t$. We obtain extensionality (η) by adding the following rule to definitional equality:

$$\frac{\Gamma \vdash t_1 : (x : A) \rightarrow B \quad \Gamma \vdash t_2 : (x : A) \rightarrow B \quad \Gamma(x : A) \vdash t_1 x = t_2 x : B}{\Gamma \vdash t_1 = t_2 : (x : A) \rightarrow B} \quad x \notin \text{dom}(\Gamma)$$

Record expressions: Likewise, a record value `record`{ $\bar{\pi} = \bar{v}$ } in Γ can be turned into an auxiliary definition by copattern matching with clauses `(f $\hat{\Gamma} \pi_i \hookrightarrow v_i$)i`. We could add an η -law that considers two values of record type R definitionally equal if they are so under each projection of R . However, to maintain decidability of definitional equality, this should only be applied to non-recursive records, as recursive records model coinductive types which do not admit η .

Indexed datatypes can be defined as regular (parameterized) datatypes with extra arguments to each constructor containing equality proofs for the indices. For example, `Vec A n` can be defined as follows:

```
data Vec (A : Setℓ)(n : ℕ) : Setℓ where
  nil   : n ≡ℕ zero → Vec A n
  cons  : (m : ℕ)(x : A)(xs : Vec A m) → n ≡ℕ suc m → Vec A n
```

⁸In a sense, this is opposite to *lazy pattern matching* [Marangot 1992], which aims to find the right clause with the least amount of matching.

Indexed record types can be defined analogously to indexed datatypes. For example, $\text{Vec } A \ n$ can also be defined as a record type:

$$\begin{aligned} \text{record } \text{Vec } (A : \text{Set}_\ell)(n : \mathbb{N}) : \text{Set}_\ell \text{ where} \\ \text{head} & : (m : \mathbb{N}) \rightarrow n \equiv_{\mathbb{N}} \text{suc } m \rightarrow A \\ \text{tail} & : (m : \mathbb{N}) \rightarrow n \equiv_{\mathbb{N}} \text{suc } m \rightarrow \text{Vec } A \ m \end{aligned}$$

The ‘constructors’ `nil` and `cons` are then defined by

$$\begin{aligned} \text{nil} : \text{Vec } A \ \text{zero} & & \text{cons} : (n : \mathbb{N})(x : A)(xs : \text{Vec } A \ n) \rightarrow \text{Vec } A \ (\text{suc } n) \\ \text{nil} . \text{head } m \ \emptyset = \text{impossible} & & \text{cons } n \ x \ xs . \text{head } [n] \ \text{refl} = x \\ \text{nil} . \text{tail } m \ \emptyset = \text{impossible} & & \text{cons } n \ x \ xs . \text{tail } [n] \ \text{refl} = xs \end{aligned}$$

Mutual recursion can be simulated by nested recursion as long as we do not define checks for positivity and termination.

Wildcard patterns can be written as variable patterns with a fresh name. Note that an unused variable may stand for either a wildcard or a forced pattern. In the latter case our algorithm treats it as a let-bound variable in the right-hand side of the clause.

Record patterns would make sense for inductive records with η . Without changes to the core language, we can represent them by first turning deep matching into shallow matching, along the lines of [Setzer et al. \[2014\]](#), and then turn record matches on the left-hand side into projection applications on the right-hand side.

This concludes the presentation of our core language.

4 CASE TREES

From a user perspective it is nice to be able to define a function by a list of clauses, but for a core language this representation of functions leaves much to be desired: it is hard to see whether a set of clauses is covering all cases [[Coquand 1992](#)], and evaluating the clauses directly can be slow for deeply nested patterns [[Cardelli 1984](#)]. Recall that for type-checking dependent types, we need to decide equality of open terms which requires computing weak head normal forms efficiently.

Thus, instead of using clauses, we represent functions by a *case tree* in our core language. In this section, we give a concrete syntax for case trees and give typing and evaluation rules for them. We also prove that a function defined by a case tree enjoys good properties such as type preservation and coverage.

$$\begin{aligned} Q ::= & u && \text{branch body (splitting done)} \\ & | \lambda x. Q && \text{bringing next argument into scope as } x \\ & | \text{record}\{\pi_1 \mapsto Q_1; \dots; \pi_n \mapsto Q_n\} && \text{splitting result by possible projections } (n \geq 0) \\ & | \text{case}_x\{c_1 \hat{\Delta}_1 \mapsto Q_1; \dots; c_n \hat{\Delta}_n \mapsto Q_n\} && \text{splitting on data } x \ (n \geq 0) \\ & | \text{case}_x\{\text{refl} \mapsto^\tau Q\} && \text{matching on equality proof } x \end{aligned} \tag{28}$$

Note that empty `case` and empty `record` are allowed, to cover the empty data type and the unit type, i.e. the record without fields.

Remark 12 (Focusing). Case trees allow us to introduce functions and records, and eliminate data. In the sense of focusing, this corresponds to the invertible rules for implication, additive conjunction, and multiplicative disjunction. (See typing rules in [Fig. 7](#).)

4.1 Case tree typing

A case tree Q for a defined constant $f : A$ is well-typed in environment Σ if $\Sigma \vdash f := Q : A \rightsquigarrow \Sigma'$. In this judgement, Σ is the signature in which case tree Q for function $f : A$ is well-typed, and Σ' is

834 $\boxed{\Sigma; \Gamma \vdash \mathbf{f} \bar{q} := Q : C \rightsquigarrow \Sigma'}$ Presupposes: $\Sigma; \Gamma \vdash \mathbf{f} [\bar{q}] : C$ and $\text{dom}(\Gamma) = \text{PV}(\bar{q})$.
 835 Checks case tree Q and outputs an extension Σ' of Σ by the clauses represented by “ $\mathbf{f} \bar{q} \hookrightarrow Q$ ”.

$$\begin{array}{c}
 \frac{\Sigma; \Gamma \vdash v : C}{\Sigma; \Gamma \vdash \mathbf{f} \bar{q} := v : C \rightsquigarrow \Sigma, (\text{clause } \Gamma \vdash \mathbf{f} \bar{q} \hookrightarrow v : C)} \text{CTDONE} \\
 \\
 \frac{\Sigma; \Gamma \vdash C = (x : A) \rightarrow B : \text{Set}_\ell \quad \Sigma; \Gamma(x : A) \vdash \mathbf{f} \bar{q} x := Q : B \rightsquigarrow \Sigma'}{\Sigma; \Gamma \vdash \mathbf{f} \bar{q} := \lambda x. Q : C \rightsquigarrow \Sigma'} \text{CTINTRO} \\
 \\
 \frac{\Sigma_0; \Gamma \vdash C = \mathbf{R} \bar{v} : \text{Set}_\ell \quad \text{record } \text{self}^f : \mathbf{R} \Delta : \text{Set}_\ell \text{ where } \overline{\pi_1} : A_i \in \Sigma_0 \\ \sigma = [\bar{v} / \Delta, \mathbf{f} [\bar{q}] / \text{self}^f] \quad (\Sigma_{i-1}; \Gamma \vdash \mathbf{f} \bar{q} .\pi_i := Q_i : A_i \sigma \rightsquigarrow \Sigma_i)_{i=1 \dots n}}{\Sigma_0; \Gamma \vdash \mathbf{f} \bar{q} := \text{record}\{\pi_1 \mapsto Q_1; \dots; \pi_n \mapsto Q_n\} : C \rightsquigarrow \Sigma_n} \text{CTCOSPLIT} \\
 \\
 \frac{\Sigma_0; \Gamma_1 \vdash A = \mathbf{D} \bar{v} : \text{Set}_\ell \quad \text{data } \mathbf{D} \Delta : \text{Set}_\ell \text{ where } \overline{c_i} \Delta_i \in \Sigma_0 \\ (\Delta'_i = \Delta_i[\bar{v} / \Delta])_{i=1 \dots n} \quad (\rho_i = \mathbb{1}_{\Gamma_1} \uplus [c_i \hat{\Delta}'_i / x])_{i=1 \dots n} \\ (\rho'_i = \rho_i \uplus \mathbb{1}_{\Gamma_2})_{i=1 \dots n} \quad (\Sigma_{i-1}; \Gamma_1 \Delta'_i(\Gamma_2 \rho_i) \vdash \mathbf{f} \bar{q} \rho'_i := Q_i : C \rho'_i \rightsquigarrow \Sigma_i)_{i=1 \dots n}}{\Sigma_0; \Gamma_1(x : A) \Gamma_2 \vdash \mathbf{f} \bar{q} := \text{case}_x\{c_1 \hat{\Delta}'_1 \mapsto Q_1; \dots; c_n \hat{\Delta}'_n \mapsto Q_n\} : C \rightsquigarrow \Sigma_n} \text{CTSPLITCON} \\
 \\
 \frac{\Sigma; \Gamma_1 \vdash A = (u \equiv_B v) : \text{Set}_\ell \quad \Sigma; \Gamma_1 \vdash_x u =^? v : B \Rightarrow \text{YES}(\Gamma'_1, \rho, \tau) \\ \rho' = \rho \uplus \mathbb{1}_{\Gamma_2} \quad \tau' = \tau \uplus \mathbb{1}_{\Gamma_2} \quad \Sigma; \Gamma'_1(\Gamma_2 \rho) \vdash \mathbf{f} \bar{q} \rho' := Q : C \rho' \rightsquigarrow \Sigma'}{\Sigma; \Gamma_1(x : A) \Gamma_2 \vdash \mathbf{f} \bar{q} := \text{case}_x\{\text{refl} \mapsto \tau' Q\} : C \rightsquigarrow \Sigma'} \text{CTSPLITEQ} \\
 \\
 \frac{\Sigma; \Gamma_1 \vdash A = (u \equiv_B v) : \text{Set}_\ell \quad \Sigma; \Gamma_1 \vdash_x u =^? v : B \Rightarrow \text{NO}}{\Sigma; \Gamma_1(x : A) \Gamma_2 \vdash \mathbf{f} \bar{q} := \text{case}_x\{\} : C \rightsquigarrow \Sigma} \text{CTSPLITABSURDEQ}
 \end{array}$$

861 Fig. 7. The typing rules for case trees.

862 the *output signature* which is Σ extended with the function clauses corresponding to case tree Q .
 863 Note that the absence of a local context Γ in this proposition implies that we only use case trees
 864 for top-level definitions.⁹

865 Case tree typing is established by the generalized judgement $\boxed{\Sigma; \Gamma \vdash \mathbf{f} \bar{q} := Q : A \rightsquigarrow \Sigma'}$ (Fig. 7)
 866 that considers a case tree Q for the instance $\mathbf{f} \bar{q}$ of the function in a context Γ of the pattern variables
 867 of \bar{q} . We have the following rules for $\Sigma; \Gamma \vdash \mathbf{f} \bar{q} := Q : A \rightsquigarrow \Sigma'$:

870 **CTDONE** A leaf of a case tree consists of a right-hand side v which needs to be of the same
 871 type C of the corresponding left-hand side $\mathbf{f} \bar{q}$ and may only refer to the pattern variables Γ
 872 of \bar{q} . If this is the case, the clause $\mathbf{f} \bar{q} \hookrightarrow v$ is added to the signature.

873 **CTINTRO** If the left-hand side $\mathbf{f} \bar{q}$ is of function type $(x : A) \rightarrow B$ we can extend it by variable
 874 pattern x . The corresponding case tree is function introduction $\lambda x. Q$.

875 **CTCOSPLIT** If the left-hand side is of record type $\mathbf{R} \bar{v}$ with projections π_i , we can do *result*
 876 *splitting* and extend it by copattern π_i for all i . We have $\text{record}\{\pi_1 \mapsto Q_1; \dots; \pi_n \mapsto Q_n\}$
 877 (where $n \geq 0$) as the corresponding case tree, and we check each sub tree Q_i for left-hand
 878 side $\mathbf{f} \bar{q} .\pi_i$ in the signature Σ_{i-1} which includes the clauses for the branches $j < i$. Note that
 879

880 ⁹It would also be possible to embed case trees into our language as terms instead, as is the case in many other languages.
 881 We refrain from doing so in this paper for the sake of simplicity.

these previous clauses may be needed to check the current case, since we have dependent records (Example 2).

CTSPPLITCON If left-hand side $f \bar{q}$ contains a variable x of data type $D \bar{v}$, we can split on x and consider all alternatives c_i ; the corresponding case tree is $\text{case}_x\{c_1 \hat{\Delta}'_1 \mapsto Q_1; \dots; c_n \hat{\Delta}'_n \mapsto Q_n\}$. The branch Q_i is checked for a refined left-hand side where x has been substituted by $c_i \hat{\Delta}'_i$ in a context where x has been replaced by the new pattern variables Δ'_i . Note also the threading of signatures as in rule CTCOSPLIT .

The rules CTSPPLITEQ and CTSPPLITABSURDEQ are explained in the next section.

4.2 Unification: splitting on the identity type

To split on an equality proof $x : u \equiv_B v$ we try to unify u and v . We may either find a most general unifier (m.g.u.); then we can build a case tree $\text{case}_x\{\text{refl} \mapsto \cdot\}$ (rule CTSPPLITEQ). We may find a disunifier and build the case tree $\text{case}_x\{\}$ (rule CTSPPLITABSURDEQ). Finally, there might be neither a m.g.u. nor a disunifier, e.g. for equality $y + z \equiv_{\mathbb{N}} y' + z'$; then type-checking fails.

In fact, in our setting we need a refinement of m.g.u.s we call *strong unifiers*. We recall the definitions of a strong unifier and a disunifier from Cockx et al. [2016], here translated to the language of this paper and specialized to the case of a single equation:

Definition 13 (Strong unifier). Let Γ be a well-formed context and u and v be terms such that $\Gamma \vdash u, v : A$. A *strong unifier* (Γ', σ, τ) of u and v consists of a context Γ' and substitutions $\Gamma' \vdash \sigma : \Gamma(x : u \equiv_A v)$ and $\Gamma(x : u \equiv_A v) \vdash \tau : \Gamma'$ such that:

- (1) $\Gamma' \vdash x\sigma = \text{refl} : u\sigma \equiv_{A\sigma} v\sigma$ (this implies the definitional equality $\Gamma' \vdash u\sigma = v\sigma : A\sigma$)
- (2) $\Gamma' \vdash \tau; \sigma = \mathbb{1}_{\Gamma'} : \Gamma'$
- (3) For any context Γ_0 and substitution σ_0 such that $\Gamma_0 \vdash \sigma_0 : \Gamma(x : u \equiv_A v)$ and $\Gamma_0 \vdash x\sigma_0 = \text{refl} : u\sigma_0 \equiv_{A\sigma_0} v\sigma_0$, we have $\Gamma_0 \vdash \sigma; \tau; \sigma_0 = \sigma_0 : \Gamma(x : u \equiv_A v)$.

Definition 14 (Disunifier). Let Γ be a well-formed context and $\Gamma \vdash u, v : A$. A *disunifier* of u and v is a function $\Gamma \vdash f : (u \equiv_A v) \rightarrow \perp$ where \perp is the empty type.

Since we use the substitution σ for the construction of the left-hand side of clauses, we require unification to output not just a substitution but a *pattern substitution* ρ . The only properly matching pattern in ρ is $x\rho = \text{refl}$; all the other patterns $y\rho$ are either a forced pattern $[t]$ (if unification instantiates y with t) or the variable y itself (if unification leaves y untouched).

We thus assume we have access to a proof relevant unification algorithm specified by the following judgements:

- $\boxed{\Sigma; \Gamma \vdash_x u =^? v : A \Rightarrow \text{YES}(\Gamma', \rho, \tau)}$ ensures that $x\rho = \text{refl}$ and the triple $(\Gamma', [\rho], \tau)$ is a strong unifier. Additionally, $\Gamma' \subseteq \Gamma$, $y\tau = y$ and $y\rho = y$ for all $y \in \Gamma'$, and $y\rho$ is a forced pattern for all variables $y \in \Gamma \setminus \Gamma'$.
- $\boxed{\Sigma; \Gamma \vdash_x u =^? v : A \Rightarrow \text{NO}}$ ensures that there exists a disunifier of u and v .

Remark 15. During the unification of u with v , each step either instantiates one variable from Γ (e.g. the solution step) or leaves it untouched (e.g. the injectivity step). We thus have the invariant that the variables in Γ' form a subset of the variables in Γ . In effect, the substitution τ makes the variables instantiated by unification go ‘out of scope’ after a match on refl . This property ceases to hold in a language with η -equality for record types and unification rules for η -expanding a variable such as the ones given by Cockx et al. [2016]. In particular, τ may contain not only variables but also projections applied to those variables.

$$\begin{array}{c}
\boxed{\Sigma \vdash Q\sigma \bar{e} \longrightarrow v} \\
\hline
\Sigma \vdash v\sigma \bar{e} \longrightarrow v\sigma \bar{e} \\
\hline
\frac{\Sigma \vdash Q(\sigma \uplus [u/x]) \bar{e} \longrightarrow v}{\Sigma \vdash (\lambda x. Q)\sigma u \bar{e} \longrightarrow v} \quad \frac{\Sigma \vdash Q_i\sigma \bar{e} \longrightarrow v}{\Sigma \vdash (\text{record}\{\pi_1 \mapsto Q_1; \dots; \pi_n \mapsto Q_n\})\sigma .\pi_i \bar{e} \longrightarrow v} \\
\hline
\frac{\Sigma \vdash x\sigma \searrow c_i \bar{u} \quad \Sigma \vdash Q_i(\sigma \setminus x \uplus [\bar{u}/\hat{\Delta}_i]) \bar{e} \longrightarrow v}{\Sigma \vdash (\text{case}_x\{c_1 \hat{\Delta}_1 \mapsto Q_1; \dots; c_n \hat{\Delta}_n \mapsto Q_n\})\sigma \bar{e} \longrightarrow v} \quad \frac{\Sigma \vdash x\sigma \searrow \text{refl} \quad \Sigma \vdash Q(\tau; \sigma) \bar{e} \longrightarrow v}{\Sigma \vdash (\text{case}_x\{\text{refl} \mapsto^\tau Q\})\sigma \bar{e} \longrightarrow v}
\end{array}$$

Fig. 8. Evaluation of case trees.

4.3 Operational semantics

If a function f is defined by a case tree Q , then we can compute the application of f to eliminations \bar{e} via the judgement $\Sigma \vdash Q\sigma \bar{e} \longrightarrow v$ (Fig. 8) with $\sigma = []$. The substitution σ acts as an accumulator, collecting the values for each of the variables introduced by a λ or by the constructor arguments in a $\text{case}_x\{\dots\}$. In particular, when evaluating a case tree of the form $\text{case}_x\{\text{refl} \mapsto^\tau Q\}$, the substitution τ is used to remove any bindings in σ that correspond to forced patterns.

4.4 Properties

If a function f is defined by a well-typed case tree, then it enjoys certain good properties such as type preservation and coverage. The goal of this section is to state and prove these properties. First, we need some basic lemmata.

Lemma 16 (Well-typed case trees preserve signature well-formedness). *Let $\vdash \Sigma$ be a well-formed signature with definition $f : A$ where $\overline{\text{cls}}^\oplus$ the last declaration in Σ and let Q be a case tree such that $\Sigma; \Gamma \vdash f \bar{q} := Q : C \rightsquigarrow \Sigma'$ where $\Sigma \vdash \Gamma$ and $\Sigma; \Gamma \mid f : A \vdash [\bar{q}] : C$. Then Σ' is also well-formed.*

PROOF. By induction on $\Sigma; \Gamma \vdash f \bar{q} := Q : C \rightsquigarrow \Sigma'$. \square

The following lemma implies that once the typechecker has completed checking a definition, we can replace the clauses of that definition by the case tree. This gives us more efficient evaluation of the function and guarantees that evaluation is deterministic.

Lemma 17 (Simulation lemma). *Consider a case tree Q such that $\Sigma_0; \Gamma \vdash f \bar{q} := Q : C \rightsquigarrow \Sigma$, let σ be a substitution with domain the pattern variables of \bar{q} , and let \bar{e} be some eliminations. If $\Sigma \vdash Q\sigma \bar{e} \longrightarrow t$ then there is some pattern substitution ρ and copatterns \bar{q}' such that $\text{clause } \Delta \vdash f \bar{q}\rho \bar{q}' \hookrightarrow v : A$ is in $\Sigma \setminus \Sigma_0$ and $t = v\theta \bar{e}_2$ where $\Sigma \vdash [\bar{q}\sigma \bar{e}_1 / \bar{q}\rho \bar{q}'] \searrow \theta$ and $\bar{e} = \bar{e}_1 \bar{e}_2$.*

Conversely, any clause in $\Sigma \setminus \Sigma_0$ is of the form $\text{clause } \Delta \vdash f \bar{q}\rho \bar{q}' \hookrightarrow v : A$, and for any σ and \bar{e}_1 and \bar{e}_2 such that $\Sigma \vdash [\bar{q}\sigma \bar{e}_1 / \bar{q}\rho \bar{q}'] \searrow \theta$ we have $\Sigma \vdash Q\sigma \bar{e}_1 \bar{e}_2 \longrightarrow v\theta \bar{e}_2$.

PROOF. We start by proving the first statement by induction on Q :

- In case $Q = v$ we have $\Sigma \vdash Q\sigma \bar{e} \longrightarrow v\sigma \bar{e}$, and $\Sigma' = \Sigma$, $\text{clause } \Gamma \vdash f \bar{q} \hookrightarrow v : A$. Thus we take $\rho = \mathbb{1}_\Gamma$, $\bar{q}' = \epsilon$, $v = v$, $\bar{e}_1 = \epsilon$ and $\bar{e}_2 = \bar{e}$. We clearly have $\Sigma \vdash [\bar{q}\sigma / \bar{q}] \searrow \sigma$, hence $t = v\sigma \bar{e}$.
- In case $Q = \lambda x. Q'$ we have $\bar{e} = u \bar{e}'$ and $\Sigma \vdash Q(\sigma \uplus [u/x]) \bar{e}' \longrightarrow t$. From the induction hypothesis we know that $\text{clause } \Delta \vdash f (\bar{q} x)\rho \bar{q}' \hookrightarrow v : A \in \Sigma$ and $\Sigma \vdash [\bar{q}\sigma u \bar{e}_1 / (\bar{q} x)\rho \bar{q}'] \searrow \theta$ and $t = v\theta \bar{e}_2$. Let $\rho = \rho' \uplus [p/x]$, then we have $\text{clause } \Delta \vdash f \bar{q}\rho' p \bar{q}' \hookrightarrow v : A \in \Sigma$ and $\Sigma \vdash [\bar{q}\sigma u \bar{e}_1 / \bar{q}\rho' p \bar{q}'] \searrow \theta$, so it suffices to take ρ' as the new ρ and $p \bar{q}'$ as the new \bar{q}' .

- 981 • In case $Q = \text{case}_x\{c_1 \hat{\Delta}'_1 \mapsto Q_1; \dots; c_n \hat{\Delta}'_n \mapsto Q_n\}$ we have $\Sigma \vdash x\sigma \searrow c_i \bar{u}$ and $\Sigma \vdash Q_i(\sigma \setminus x \uplus$
 982 $[\bar{u} / \Delta_i \sigma]) \bar{e} \longrightarrow t$. From the induction hypothesis we know that **clause** $\Delta \vdash f \bar{q}\rho_i \rho \bar{q}' \hookrightarrow v : A \in \Sigma$ and $\Sigma \vdash [\bar{q}\rho_i(\sigma \setminus x \uplus [\bar{u} / \Delta_i \sigma]) \bar{e}_1 / \bar{q}\rho_i \rho \bar{q}'] \searrow \theta$ and $t = v\theta \bar{e}_2$. Moreover, $\rho_i = \mathbb{1}_{\Gamma_1} \uplus$
 983 $[c_i \hat{\Delta}'_i / x] \uplus \mathbb{1}_{\Gamma_2}$. From the definition of matching, it follows that also $\Sigma \vdash [\bar{q}\sigma \bar{e}_1 / \bar{q}\rho_i \rho \bar{q}'] \searrow \theta$.
 984 Thus we finish this case by taking $\rho_i; \rho$ as the new ρ (and keep \bar{q}' the same).
 985
- 986 • In case $Q = \text{record}\{\pi_1 \mapsto Q_1; \dots; \pi_n \mapsto Q_n\}$ we have $\bar{e} = .\pi_1 \bar{e}'$ and $\Sigma \vdash Q_i \sigma \bar{e}' \longrightarrow t$.
 987 From the induction hypothesis we know that **clause** $\Delta \vdash f \bar{q}\rho .\pi_1 \bar{q}' \hookrightarrow v : A \in \Sigma$ and
 988 $\Sigma \vdash [\bar{q}\sigma .\pi_1 \bar{e}_1 / \bar{q}\rho .\pi_1 \bar{q}'] \searrow \theta$ and $t = v\theta \bar{e}_2$. Hence it suffices to take $.\pi_1 \bar{q}'$ as the new \bar{q}'
 989 (and keep ρ the same).
- 990 • In case $Q = \text{case}_x\{\text{refl} \mapsto \tau' Q'\}$ we have $\Sigma \vdash x\sigma \searrow \text{refl}$ and $\Sigma \vdash Q'(\tau'; \sigma) \bar{e} \longrightarrow t$. From the
 991 induction hypothesis we know that **clause** $\Delta \vdash f \bar{q}\rho' \rho \hookrightarrow v : A \in \Sigma$ and $\Sigma \vdash [\bar{q}\rho' \tau' \sigma \bar{e}_1 / \bar{q}\rho' \rho \bar{q}'] \searrow$
 992 θ and $t = v\theta \bar{e}_2$. Since ρ' and τ' are produced by unification, we have that $x\rho' = \text{refl}$ and for
 993 each pattern variable y of \bar{q} other than x , either $y\rho' = \lfloor s \rfloor$ or $y\rho' = y$ and $y\tau' = y$. It then
 994 follows from the definition of matching that $\Sigma \vdash [\bar{q}\sigma \bar{e}_1 / \bar{q}\rho' \rho \bar{q}'] \searrow \theta$. Hence we take $\rho'; \rho$
 995 as the new ρ (and keep \bar{q}' the same).
- 996 • There are no evaluation rules for $Q = \text{case}_x\{\}$ so this case is impossible.

997 In the other direction, we start again by induction on Q :

- 998 • In case $Q = v$ we have the single **clause** $\Gamma \vdash f \bar{q} \hookrightarrow v : A$ which is of the right
 999 form with $\rho = \mathbb{1}_\Gamma$ and $\bar{q}' = \epsilon$. If $\Sigma \vdash [\bar{q}\sigma \bar{e}_1 / \bar{q}] \searrow \theta$, then we have $\sigma = \theta$ and $\bar{e}_1 = \epsilon$, so
 1000 $\Sigma \vdash Q\sigma \bar{e}_1 \bar{e}_2 \longrightarrow v\theta \bar{e}_2$.
- 1001 • In case $Q = \lambda x. Q'$, we get from the induction hypothesis that any **clause** in $\Sigma \setminus \Sigma_0$ is of
 1002 the form **clause** $\Delta \vdash f (\bar{q} x)\rho \bar{q}' \hookrightarrow v : A$, which is of the right form if we take $\rho' = \rho \setminus x$
 1003 as the new ρ and $\bar{q}'' = x\rho \bar{q}'$ as the new \bar{q}' . Moreover, if $\Sigma \vdash [\bar{q}\sigma \bar{e}_1 / \bar{q}\rho' \bar{q}''] \searrow \theta$ then
 1004 $\bar{e}_1 = u \bar{e}'_1$ and $\Sigma \vdash [(\bar{q} x)(\sigma \uplus [u / x]) \bar{e}'_1 / (\bar{q} x)\rho \bar{q}'] \searrow \theta$. The induction hypothesis gives us
 1005 that $\Sigma \vdash Q'(\sigma \uplus [u / x]) \bar{e}'_1 \bar{e} \longrightarrow v\theta \bar{e}_2$, hence also $\Sigma \vdash Q\sigma \bar{e}_1 \bar{e}_2 \longrightarrow v\theta \bar{e}_2$.
- 1006 • In case $Q = \text{case}_x\{c_1 \hat{\Delta}'_1 \mapsto Q_1; \dots; c_n \hat{\Delta}'_n \mapsto Q_n\}$, we get from the induction hypothesis that
 1007 any **clause** in $\Sigma \setminus \Sigma_0$ is of the form **clause** $\Delta \vdash f \bar{q}\rho_i \rho \bar{q}' \hookrightarrow v : A$ for some $\rho_i = \mathbb{1}_{\Gamma_1} \uplus$
 1008 $[c_i \hat{\Delta}'_i / x] \uplus \mathbb{1}_{\Gamma_2}$. This is of the right form if we take $\rho' = \rho_i \rho$ as the new ρ (and keep \bar{q}'
 1009 the same). Moreover, if $\Sigma \vdash [\bar{q}\sigma \bar{e}_1 / \bar{q}\rho_i \rho \bar{q}'] \searrow \theta$ then we have $\Sigma \vdash x\sigma \searrow c_i \bar{u}$ from the
 1010 definition of matching. Let $\sigma' = \sigma \setminus x \uplus [\bar{u} / \Delta_i \sigma]$, then we also have $\Sigma \vdash [\bar{q}\rho_i \sigma' \bar{e}_1 / \bar{q}\rho_i \rho \bar{q}'] \searrow$
 1011 θ . From the induction hypothesis it now follows that $\Sigma \vdash Q_i \sigma' \bar{e}_1 \bar{e}_2 \longrightarrow v\theta \bar{e}_2$, hence also
 1012 $\Sigma \vdash Q\sigma \bar{e}_1 \bar{e}_2 \longrightarrow v\theta \bar{e}_2$.
- 1013 • In case $Q = \text{record}\{\pi_1 \mapsto Q_1; \dots; \pi_n \mapsto Q_n\}$, we get from the induction hypothesis that any
 1014 **clause** in $\Sigma \setminus \Sigma_0$ is of the form **clause** $\Delta \vdash f \bar{q}\rho .\pi_1 \bar{q}' \hookrightarrow v : A$. This is of the right form if we
 1015 take $\bar{q}'' = .\pi_1 \bar{q}'$ as the new \bar{q}' (and keep ρ the same). Moreover, if $\Sigma \vdash [\bar{q}\sigma \bar{e}_1 / \bar{q}\rho .\pi_1 \bar{q}'] \searrow \theta$
 1016 then $\bar{e}_1 = .\pi_1 \bar{e}'_1$. The induction hypothesis gives us that $\Sigma \vdash Q_i \sigma \bar{e}'_1 \bar{e}_2 \longrightarrow v\theta \bar{e}_2$, hence also
 1017 $\Sigma \vdash Q\sigma \bar{e}_1 \bar{e}_2 \longrightarrow v\theta \bar{e}_2$.
- 1018 • In case $Q = \text{case}_x\{\text{refl} \mapsto \tau' Q'\}$ we get from the induction hypothesis that any **clause** in
 1019 $\Sigma \setminus \Sigma_0$ is of the form **clause** $\Delta \vdash f \bar{q}\rho' \rho \hookrightarrow v : A$ where ρ' and τ' are produced by unifica-
 1020 tion. This is of the right form if we take $\rho'' = \rho'; \rho$ as the new ρ (and keep \bar{q}' the same).
 1021 Moreover, if $\Sigma \vdash [\bar{q}\sigma \bar{e}_1 / \bar{q}\rho' \rho \bar{q}'] \searrow \theta$ then we have $\Sigma \vdash x\sigma \searrow \text{refl}$ from the definition of
 1022 matching. Let $\sigma' = \tau'; \sigma$, then we have $x\rho' \sigma' = \text{refl}$ and for all other pattern variables
 1023 y of \bar{q} , either $y\rho'$ is a forced pattern or $y\rho' = y$ and $y\sigma' = y\sigma$. By matching, it follows
 1024 that also $\Sigma \vdash [\bar{q}\rho' \sigma' \bar{e}_1 / \bar{q}\rho' \rho \bar{q}'] \searrow \theta$. From the induction hypothesis it now follows that
 1025 $\Sigma \vdash Q' \sigma' \bar{e}_1 \bar{e}_2 \longrightarrow v\theta \bar{e}_2$, hence also $\Sigma \vdash Q\sigma \bar{e}_1 \bar{e}_2 \longrightarrow v\theta \bar{e}_2$.
- 1026 • In case $Q = \text{case}_x\{\}$ we have $\Sigma = \Sigma_0$ so there are no new clauses to worry about.

□

Before adding a clause $f \bar{q} \hookrightarrow v$ to the signature, we have to make sure that the copatterns \bar{q} only use forced patterns in places where it is justified: otherwise we might have $\Sigma \vdash [\bar{e} / \bar{q}] \searrow \sigma$ but $[\bar{q}] \sigma \neq \bar{e}$. This is captured in the notion of a *respectful pattern* [Goguen et al. 2006]. Here we generalize their definition to the case where we do not yet know that all reductions in the signature are necessarily type-preserving.

Definition 18. A signature Σ is *respectful* for $\Sigma \vdash u \searrow w$ if $\Sigma; \Gamma \vdash u : A$ implies $\Sigma; \Gamma \vdash u = w : A$. A signature Σ is respectful if it is respectful for all derivations of $\Sigma \vdash u \searrow w$.

In particular, this means $\Sigma; \Gamma \vdash w : A$, so evaluation with signature Σ is type preserving. It is immediately clear that the empty signature is respectful, since it does not contain any clauses.

Definition 19 (Respectful copatterns). Let \bar{q} be a list of copatterns such that $\Sigma; \Delta \mid u : A \vdash [\bar{q}] : C$ where u and A are closed (i.e. do not depend on Δ). We call \bar{q} *respectful* in signature Σ if the following holds: for any signature extension $\Sigma \subseteq \Sigma'$ and any eliminations $\Sigma'; \Gamma \mid u : A \vdash \bar{e} : C$ such that $\Sigma' \vdash [\bar{e} / \bar{q}] \searrow \sigma$ and Σ' is respectful for any $\Sigma' \vdash s \searrow t$ used in the derivation of $\Sigma' \vdash [\bar{e} / \bar{q}] \searrow \sigma$, we have $\Sigma'; \Gamma \mid u : A \vdash \bar{q}\sigma = \bar{e} : C$.

Being respectful is stable under signature extension by definition: if \bar{q} is respectful in Σ and $\Sigma \subseteq \Sigma'$, then \bar{q} is also respectful in Σ' .

Lemma 20 (Signatures with respectful clauses are respectful). *If Σ is a well-formed signature such that all clauses in Σ have respectful copatterns in Σ , then Σ is respectful.*

PROOF. By induction on the derivation of $\Sigma \vdash u \searrow v$. Assume **clause** $\Delta \vdash f \bar{q} \hookrightarrow v : C \in \Sigma$ and $\Sigma \vdash [\bar{e} / \bar{q}] \searrow \sigma$ for well-typed eliminations $\Sigma; \Gamma \mid f : C \vdash \bar{e} : A$, then we have to prove that $\Sigma; \Gamma \vdash f \bar{e} = v\sigma : A$. By induction, Σ is respectful for any $\Sigma \vdash s \searrow t$ used in the derivation of $\Sigma \vdash [\bar{e} / \bar{q}] \searrow \sigma$. Since \bar{q} is respectful, this implies that $\Sigma; \Gamma \mid f : C \vdash \bar{q}\sigma = \bar{e} : A$. It follows that $\Sigma; \Gamma \vdash f \bar{q}\sigma = f \bar{e} : A$, hence also $\Sigma; \Gamma \vdash f \bar{e} = v\sigma : A$ by the β -rule for definitional equality. \square

Lemma 21 (Well-typed case trees have respectful clauses). *Consider a respectful signature Σ_0 and a case tree Q such that $\Sigma_0; \Gamma \vdash f \bar{q} := Q : C \rightsquigarrow \Sigma$ and \bar{q} is respectful in Σ_0 . Then all clauses in $\Sigma \setminus \Sigma_0$ have respectful patterns in Σ .*

PROOF. By induction on the derivation of $\Sigma_0; \Gamma \vdash f \bar{q} := Q : C \rightsquigarrow \Sigma$.

- In case $Q = v$, we have a single new clause **clause** $\Gamma \vdash f \bar{q} \hookrightarrow v : C$. Since \bar{q} is respectful in Σ_0 by assumption, it is also respectful in $\Sigma = \Sigma_0$, **clause** $\Gamma \vdash f \bar{q} \hookrightarrow v : C$.
- In case $Q = \lambda x. Q'$, we know from the typing rule of λx . that $\Sigma_0; \Gamma \vdash C = (x : A') \rightarrow B' : \mathbf{Set}_\ell$. and $\Sigma_0; \Gamma(x : A) \vdash f \bar{q} x := Q' : B \rightsquigarrow \Sigma$. Since \bar{q} is respectful, it follows that $\bar{q} x$ is also respectful, so the result follows from the induction hypothesis.
- In case $Q = \mathbf{case}_x\{c_1 \hat{\Delta}'_1 \mapsto Q_1; \dots; c_n \hat{\Delta}'_n \mapsto Q_n\}$, the typing rule for $\mathbf{case}_x\{\}$ tells us that $\Gamma = \Gamma_1(x : A)\Gamma_2$ and $\Sigma_0; \Gamma_1 \vdash A = D \bar{v} : \mathbf{Set}_\ell$. We also get that $\Sigma_{i-1}; \Gamma_1 \Delta'_i \Gamma_2 \rho_i \vdash f \bar{q} \rho_i := Q_i : C \rho_i \rightsquigarrow \Sigma_i$ where **constructor** $c_i \Delta_i : D \Delta \in \Sigma_0$ and $\Delta'_i = \Delta_i[\bar{v} / \Delta]$ and $\rho_i = [c_i \hat{\Delta}'_i / x]$. Since \bar{q} is respectful, so is $\bar{q} \rho_i$, so the result follows from the induction hypothesis.
- In case $Q = \mathbf{record}\{\pi_1 \mapsto Q_1; \dots; \pi_n \mapsto Q_n\}$, the typing rule for $\mathbf{record}\{\}$ tells us that $\Sigma_0; \Gamma \vdash C = R \bar{v} : \mathbf{Set}_\ell$. We also get that $\Sigma_{i-1}; \Gamma \vdash f \bar{q} .\pi_i := Q_i : A_i[\bar{v} / \Delta, f [\bar{q}] / x] \rightsquigarrow \Sigma_i$ where **projection** $x : R \Delta \vdash .\pi_i : A_i \in \Sigma_0$. Since \bar{q} is respectful, so is $\bar{q} .\pi_i$, so the result follows from the induction hypothesis.
- In case $Q = \mathbf{case}_x\{\mathbf{refl} \mapsto^\tau Q'\}$, the the typing rule tells us that $\Gamma = \Gamma_1(x : A)\Gamma_2$ and $\Sigma_0; \Gamma_1 \vdash A = s \equiv_E t : \mathbf{Set}_\ell$. We also have that $\Sigma_0; \Gamma_1 \vdash_x s =^\tau t : E \Rightarrow \mathbf{YES}(\Gamma'_1, \rho, \tau)$ and $\Sigma_0; \Gamma'_1 \Gamma_2 \rho \vdash f \bar{q} \rho' := Q' : C \rho' \rightsquigarrow \Sigma$ where $\rho' = \rho \uplus \mathbb{1}_{\Gamma_2}$. Since \bar{q} is respectful and ρ is a

strong unifier (Definition 13), $\bar{q}\rho'$ is also respectful, so the result follows from the induction hypothesis.

- The typing rule for $Q = \text{case}_e\{\}$ does not add any new clauses.

□

Theorem 22 (Type preservation). *If all functions in a signature Σ are given by well-typed case trees, then Σ is respectful.*

PROOF. This is a direct consequence of the previous two lemmata. □

Definition 23. A term u is *normalising* in a signature Σ if $\Sigma \vdash u \searrow w$, and additionally, if $w = \mathbf{c} \bar{v}$ then all \bar{v} are also normalising.

An elimination e is normalising if it is either a projection $\cdot\pi$ or a normalising term u . A substitution σ is normalising if $x\sigma$ is normalising for all variables x in $\text{dom}(\sigma)$.

The definition of a normalising term (and the proof of the following lemma) would be somewhat more complicated for a language with eta-equality for record types, such as the one used by Cockx et al. [2016]. In particular, all projections of a normalising expression of record type should also be normalising.

Lemma 24. *Suppose $\Sigma; \Gamma \vdash_x u =^? v : A \Rightarrow \text{YES}(\Gamma', \rho, \tau)$ and $\Sigma \vdash \sigma_0 : \Gamma$ such that $\Sigma \vdash u\sigma_0 = v\sigma_0 : A\sigma_0$. If σ_0 is normalising, then so is $\tau; \sigma_0$.*

PROOF. By construction, $y\tau = y$ for each variable y in Γ' (see Remark 15), so it follows trivially that $\tau; \sigma_0$ is normalising. □

Theorem 25 (Coverage). *Let Q be a case tree such that $\Sigma_0; \Gamma \vdash \mathbf{f} \bar{q} := Q : C \rightsquigarrow \Sigma$. Let further $\Sigma \vdash \sigma_0 : \Gamma$ be a (closed) substitution and $\Sigma \mid \mathbf{f} \bar{q}\sigma_0 : C\sigma_0 \vdash \bar{e} : B$ be (closed) eliminations such that σ_0 and \bar{e} are normalizing in Σ and B is not definitionally equal to a function type or a record type. Then $\Sigma \vdash Q\sigma_0 \bar{e} \longrightarrow v$ for some v .*

In particular, this theorem tells us that if $\Sigma_0 \vdash \mathbf{f} := Q : C \rightsquigarrow \Sigma$ and the eliminations $\Sigma \mid \mathbf{f} : C \vdash \bar{e} : A$ are normalising, then $\Sigma \vdash Q[\bar{e}] \longrightarrow v$. Thus evaluation of a function defined by a well-typed case tree applied to closed arguments can never get stuck.

PROOF. By induction of the case tree Q :

- If $Q = v$, we have $\Sigma \vdash Q\sigma_0 \bar{e} \longrightarrow v\sigma_0 \bar{e}$.
- If $Q = \lambda x. Q'$, we have $\Sigma_0; \Gamma \vdash C = (x : A') \rightarrow B' : \text{Set}_\ell$ and $\Sigma_0; \Gamma(x : A') \vdash \mathbf{f} \bar{q} x := Q' : B' \rightsquigarrow \Sigma$ from the typing rule of $\lambda x..$ Hence we have $\Sigma \vdash C\sigma_0 = (x : A'\sigma_0) \rightarrow B'\sigma_0 : \text{Set}_\ell$, so $\bar{e} = w \bar{e}'$ for some term $\Sigma \vdash w : A'\sigma_0$ and eliminations $\Sigma \mid \mathbf{f} \bar{q}\sigma_0 w : B'(\sigma_0 \uplus [w/x]) \vdash \bar{e}' : B$. By induction we now have that there exists some v such that $\Sigma \vdash Q'(\sigma_0 \uplus [w/x]) \bar{e}' \longrightarrow v$, hence also $\Sigma \vdash Q\sigma_0 \bar{e} \longrightarrow v$.
- If $Q = \text{case}_x\{\mathbf{c}_1 \hat{\Delta}'_1 \mapsto Q_1; \dots; \mathbf{c}_n \hat{\Delta}'_n \mapsto Q_n\}$, we have $x : \mathbf{D} \bar{v} \in \Gamma$, hence $\Sigma \vdash x\sigma_0 \searrow \mathbf{c}_i \bar{u}$ for some constructor \mathbf{c}_i of \mathbf{D} . By induction we have a v such that $\Sigma \vdash Q_i(\sigma_0 \uplus [\bar{u}/\Delta_i\sigma]) \bar{e} \longrightarrow v$, hence also $\Sigma \vdash Q\sigma_0 \bar{e} \longrightarrow v$.
- If $Q = \text{record}\{\pi_1 \mapsto Q_1; \dots; \pi_n \mapsto Q_n\}$, we have $\Sigma_0; \Gamma \vdash C = \mathbf{R} \bar{v} : \text{Set}_\ell$. Hence we have $\Sigma_0 \vdash C\sigma_0 = \mathbf{R} \bar{v}\sigma_0 : \text{Set}_\ell$, so $\bar{e} = \cdot\pi_i \bar{e}'$ for some field π_i of \mathbf{R} . By induction we get a v such that $\Sigma \vdash Q_i\sigma_0 \bar{e}' \longrightarrow v$, hence also $\Sigma \vdash Q\sigma_0 \bar{e} \longrightarrow v$.
- If $Q = \text{case}_x\{\text{refl} \mapsto^r Q'\}$, we have $x : u \equiv_E v \in \Gamma$, so $\Sigma \vdash x\sigma_0 \searrow \text{refl}$. Since σ_0 is normalising, $\tau; \sigma_0$ is also normalising (see Lemma 24). Now it follows from the inductive hypothesis that $\Sigma \vdash Q'(\tau; \sigma_0) \bar{e} \longrightarrow v$, hence also $\Sigma \vdash Q\sigma_0 \bar{e} \longrightarrow v$.

$$\begin{array}{c}
1128 \quad \boxed{\Sigma \vdash \text{decl} \rightsquigarrow \Sigma'} \quad \text{Presupposes: } \vdash \Sigma. \quad \text{Entails: } \vdash \Sigma'. \\
1129 \\
1130 \quad \frac{\Sigma \vdash \Delta \quad (\Sigma, \text{data } D \Delta : \text{Set}_\ell) \mid D \Delta : \text{Set}_\ell \vdash \overline{\text{con}} \rightsquigarrow \Sigma'}{\Sigma \vdash (\text{data } D \Delta : \text{Set}_\ell \text{ where } \overline{\text{con}}) \rightsquigarrow \Sigma'} \\
1131 \\
1132 \\
1133 \quad \frac{\Sigma \vdash \Delta \quad (\Sigma, \text{record } R \Delta : \text{Set}_\ell) \mid \text{self} : R \Delta : \text{Set}_\ell \vdash \overline{\text{field}} \rightsquigarrow \Sigma'}{\Sigma \vdash (\text{record } \text{self} : R \Delta : \text{Set}_\ell \text{ where } \overline{\text{field}}) \rightsquigarrow \Sigma'} \\
1134 \\
1135 \\
1136 \\
1137 \quad \frac{\Sigma \vdash A \quad (\Sigma, \text{definition } f : A) \vdash P \mid f := Q : A \rightsquigarrow \Sigma'}{\Sigma \vdash (\text{definition } f : A \text{ where } P) \rightsquigarrow \Sigma'} \\
1138 \\
1139 \\
1140 \quad \boxed{\Sigma \mid D \Delta : \text{Set}_\ell \vdash \overline{\text{con}} \rightsquigarrow \Sigma'} \quad \text{Presupposes: } \vdash \Sigma \text{ and } \Sigma \vdash D : \Delta \rightarrow \text{Set}_\ell. \quad \text{Entails: } \vdash \Sigma'. \\
1141 \\
1142 \quad \frac{\Sigma; \Delta \vdash_\ell \Delta_c \quad (\Sigma, \text{constructor } c \Delta_c : D \Delta) \mid D \Delta : \text{Set}_\ell \vdash \overline{\text{con}} \rightsquigarrow \Sigma'}{\Sigma \mid D \Delta : \text{Set}_\ell \vdash \epsilon \rightsquigarrow \Sigma} \quad \Sigma \mid D \Delta : \text{Set}_\ell \vdash c \Delta_c, \overline{\text{con}} \rightsquigarrow \Sigma' \\
1143 \\
1144 \\
1145 \quad \boxed{\Sigma \mid \text{self} : R \Delta : \text{Set}_\ell \vdash \overline{\text{field}} \rightsquigarrow \Sigma'} \quad \text{Presupposes: } \vdash \Sigma \text{ and } \Sigma \vdash R : \Delta \rightarrow \text{Set}_\ell. \quad \text{Entails: } \vdash \Sigma'. \\
1146 \\
1147 \quad \frac{\Sigma; \Delta(\text{self} : R \hat{\Delta}) \vdash A : \text{Set}_{\ell'} \quad \ell' \leq \ell \quad (\Sigma, \text{projection } \text{self} : R \Delta \vdash \cdot \pi : A) \mid \text{self} : R \Delta : \text{Set}_\ell \vdash \overline{\text{field}} \rightsquigarrow \Sigma'}{\Sigma \mid \text{self} : R \Delta : \text{Set}_\ell \vdash \epsilon \rightsquigarrow \Sigma} \quad \Sigma \mid \text{self} : R \Delta : \text{Set}_\ell \vdash \pi : A, \overline{\text{field}} \rightsquigarrow \Sigma' \\
1148 \\
1149 \\
1150 \\
1151 \\
1152 \\
1153 \\
1154 \\
1155 \\
1156 \\
1157 \\
1158 \\
1159 \\
1160 \\
1161 \\
1162 \\
1163 \\
1164 \\
1165 \\
1166 \\
1167 \\
1168 \\
1169 \\
1170 \\
1171 \\
1172 \\
1173 \\
1174 \\
1175 \\
1176
\end{array}$$

Fig. 9. Rules for checking declarations of data types, record types, and defined symbols.

- If $Q = \text{case}_x\{\}$, we have $x : u \equiv_E v \in \Gamma$, so $\Sigma \vdash x\sigma_0 \searrow \text{refl}$. But $u \equiv_E v$ is equivalent to the empty type by unification, so this case is impossible.

□

5 ELABORATION

In the previous two sections, we have described a core language with inductive datatypes, coinductive records, identity types, and functions defined by well-typed case trees. On the other hand, we also have a surface language consisting of declarations of datatypes, record types, and functions by dependent (co)pattern matching. In this section we show how to elaborate a program in this surface language to a well-formed signature in the core language.

The main goal of this section is to describe the elaboration of a definition given by a set of (unchecked) clauses to a well-typed case tree, and prove that this translation (if it succeeds) preserves the first-match semantics of the given clauses. Before we dive into this, we first describe the elaboration for data and record types.

5.1 Elaborating data and record types

Figure 9 gives the rules for checking declarations, constructors and projections. These rules are designed to correspond closely to those for signature extension in Fig. 4. Consequentially, if $\vdash \Sigma$ and $\Sigma \vdash \text{decl} \rightsquigarrow \Sigma'$, then also $\vdash \Sigma'$.

5.2 From clauses to a case tree

In Section 2 we showed how our elaboration algorithm works in a number of examples, here we describe it in general. The inputs to the algorithm are the following:

- A signature Σ containing previous declarations, as well as clauses for the branches of the case tree that have already been checked.
- A context Γ containing the types of the pattern variables: $\text{dom}(\Gamma) = \text{PV}(\bar{q})$.
- The function f currently being checked.
- The copatterns \bar{q} for the current branch of the case tree.
- The refined target type C of the current branch.
- The user input P , which is described below.

The outputs of the algorithm are a signature Σ' extending Σ with new clauses and a well-typed case tree Q such that $\Sigma; \Gamma \vdash f \bar{q} := Q : C \rightsquigarrow \Sigma'$.

We represent the user input P to the algorithm as an (ordered) list of partially decomposed clauses, called a left-hand side problem or *lhs problem* for short. Each partially decomposed clause is of the form $[E]\bar{q} \hookrightarrow rhs$ where E is an (unordered) set of constraints $\{w_k \int^? p_k : A_k \mid k = 1 \dots l\}$ between a pattern p_k and a term w_k , \bar{q} is a list of copatterns, and rhs is a right-hand side. In the special case E is empty, we have a complete clause written as $\bar{q} \hookrightarrow rhs$.

Elaboration of an lhs problem to a well-typed case tree $\Sigma; \Gamma \vdash P \mid f \bar{q} := Q : C \rightsquigarrow \Sigma'$ is defined in Fig. 10. This judgement is designed as an algorithmic version of the typing judgement for case trees $\Sigma; \Gamma \vdash f \bar{q} := Q : C \rightsquigarrow \Sigma'$, where the extra user input P guides the construction of the case tree. Each of the rules in Fig. 10 is a refined version of one of the rules in Fig. 7, so any case tree produced by this elaboration is well-typed by construction.

To check a definition of $f : A$ with clauses $\bar{q}_i \hookrightarrow rhs_i$ for $i = 1 \dots n$, the algorithm starts with $\Gamma = \epsilon$, $u = f$, and $P = \{\bar{q}_i \hookrightarrow rhs_i \mid i = 1 \dots n\}$. If we obtain $\Sigma; \Gamma \vdash P \mid f := Q : A \rightsquigarrow \Sigma'$, then the function f can be implemented using the case tree Q .

During elaboration, the algorithm maintains the invariants that $\vdash \Sigma$ is a well-formed signature, $\Sigma \vdash \Gamma$ is a well-formed context, and $\Sigma; \Gamma \vdash f [\bar{q}] : C$. It also maintains the invariant that for each constraint $w_k \int^? p_k : A_k$ in the lhs problem, we have $\Sigma; \Gamma \vdash w_k : A_k$.

The rules for $\Sigma; \Gamma \vdash P \mid f \bar{q} := Q : C \rightsquigarrow \Sigma'$ make use of some auxiliary operations for manipulating lhs problems:

- After each step, the algorithm uses $\Sigma; \Gamma \vdash E \Rightarrow \text{SOLVED}(\sigma)$ (Fig. 11) to check if the first user clause has no more (co)patterns, and all its constraints are solved. If this is the case, it returns a substitution σ assigning a well-typed value to each of the user-written pattern variables.
- After introducing a new variable, the algorithm uses $P(x : A)$ (Fig. 12) to remove the first application pattern from each of the user clauses and to introduce a new constraint between the variable and the pattern.
- After a copattern split on a record type, the algorithm uses $P.\pi$ (Fig. 13) to partition the clauses in the lhs problem according to the projection they belong to.
- After a case split on a datatype or an equality proof, the algorithm uses $\Sigma \vdash P\sigma \Rightarrow P'$ (Fig. 14) to refine the constraints in the lhs problem. It uses judgements $\Sigma \vdash v \int^? p : A \Rightarrow E_\perp$ and $\Sigma \vdash \bar{v} \int^? \bar{p} : \Delta \Rightarrow E_\perp$ (Fig. 15) to simplify the constraints if possible, and to filter out the clauses that definitely do not match the current branch.
- To check an absurd pattern \emptyset , the algorithm uses $\Sigma; \Gamma \vdash \emptyset : A$ (Fig. 16) to ensure that the type of the pattern is a *caseless type* [Goguen et al. 2006], i.e. a type that is empty and cannot

1226 $\boxed{\Sigma; \Gamma \vdash P \mid \mathbf{f} \bar{q} := Q : C \rightsquigarrow \Sigma'}$ In all rules $P = \{[E_i]\bar{q}_i \hookrightarrow rhs_i \mid i = 1 \dots m\}$.

1227 Presupposes: $\Sigma; \Gamma \vdash \mathbf{f} [\bar{q}] : C$ and $\text{dom}(\Gamma) = \text{PV}(\bar{q})$. Entails: $\Sigma; \Gamma \vdash \mathbf{f} \bar{q} := Q : C \rightsquigarrow \Sigma'$.

1228

1229

1230
$$\frac{\bar{q}_1 = \epsilon \quad \Sigma; \Gamma \vdash E_1 \Rightarrow \text{SOLVED}(\sigma) \quad rhs_1 = v \quad \Sigma; \Gamma \vdash v\sigma : C}{\Sigma; \Gamma \vdash P \mid \mathbf{f} \bar{q} := v\sigma : C \rightsquigarrow \Sigma, \text{clause } \Gamma \vdash \mathbf{f} \bar{q} \hookrightarrow v\sigma : C} \text{DONE}$$

1231

1232

1233
$$\frac{\bar{q}_1 = p \bar{q}'_1 \quad \Sigma \vdash C \searrow (x : A) \rightarrow B \quad \Sigma; \Gamma(x : A) \vdash P(x : A) \mid \mathbf{f} \bar{q} x := Q : B \rightsquigarrow \Sigma'}{\Sigma; \Gamma \vdash P \mid \mathbf{f} \bar{q} := \lambda x. Q : C \rightsquigarrow \Sigma'} \text{INTRO}$$

1234

1235

1236

1237
$$\frac{\bar{q}_1 = \pi_1 \bar{q}'_1 \quad \Sigma \vdash C \searrow R \bar{v} \quad \text{record self} : R \Delta : \text{Set}_\ell \text{ where } \overline{\pi_1 : A_i} \in \Sigma_0}{(\Sigma_{i=1} \vdash P \pi_1 \mid \mathbf{f} \bar{q} \pi_i := Q_i : A_i[\bar{v} / \Delta, \mathbf{f} [\bar{q}] / \text{self}] \rightsquigarrow \Sigma_i)_{i=1 \dots n}} \text{COSPLIT}$$

1238

1239
$$\Sigma; \Gamma \vdash P \mid \mathbf{f} \bar{q} := \text{record}\{\pi_1 \mapsto Q_1; \dots; \pi_n \mapsto Q_n\} : C \rightsquigarrow \Sigma_n$$

1240

1241
$$\frac{\bar{q}_1 = \emptyset \quad m = 1 \quad \Sigma \vdash C \searrow R \bar{v} \quad \text{record } _ : R \Delta : \text{Set}_\ell \text{ where } \epsilon \in \Sigma \quad rhs_1 = \text{impossible}}{\Sigma; \Gamma \vdash P \mid \mathbf{f} \bar{q} := \text{record}\{\} : C \rightsquigarrow \Sigma} \text{COSPLITEMPTY}$$

1242

1243

1244

1245
$$(x \ /^? \ c_j \bar{p} : A) \in E_1 \quad \Sigma \vdash A \searrow D \bar{v} \quad \Gamma = \Gamma_1(x : A)\Gamma_2$$

1246
$$\text{data } D \Delta : \text{Set}_\ell \text{ where } c_i \hat{\Delta}_i \in \Sigma_0$$

1247
$$\left(\frac{\Delta'_i = \Delta_i[\bar{v} / \Delta] \quad \rho_i = \mathbb{1}_{\Gamma_1} \uplus [c_i \hat{\Delta}'_i / x] \quad \rho'_i = \rho_i \uplus \mathbb{1}_{\Gamma_2}}{\Sigma_n \vdash P \rho'_i \Rightarrow P_i \quad (\Sigma_{i=1} \vdash \Gamma_1 \Delta'_i(\Gamma_2 \rho_i) \vdash P_i \mid \mathbf{f} \bar{q} \rho'_i := Q_i : C \rho'_i \rightsquigarrow \Sigma_i)_{i=1 \dots n}} \text{SPLITCON}$$

1248
$$\Sigma_0; \Gamma \vdash P \mid \mathbf{f} \bar{q} := \text{case}_x\{c_1 \hat{\Delta}'_1 \mapsto Q_1; \dots; c_n \hat{\Delta}'_n \mapsto Q_n\} : C \rightsquigarrow \Sigma_n$$

1249

1250

1251
$$(x \ /^? \ \text{refl} : A) \in E_1 \quad \Sigma \vdash A \searrow u \equiv_B v \quad \Gamma = \Gamma_1(x : A)\Gamma_2$$

1252
$$\Sigma; \Gamma_1 \vdash_x u =^? v : B \Rightarrow \text{YES}(\Gamma'_1, \rho, \tau) \quad \rho' = \rho \uplus \mathbb{1}_{\Gamma_2} \quad \tau' = \tau \uplus \mathbb{1}_{\Gamma_2}$$

1253
$$\Sigma \vdash P \rho' \Rightarrow P' \quad \Sigma; \Gamma'_1(\Gamma_2 \rho) \vdash P' \mid \mathbf{f} \bar{q} \rho' := Q : C \rho' \rightsquigarrow \Sigma'$$

1254
$$\Sigma; \Gamma \vdash P \mid \mathbf{f} \bar{q} := \text{case}_x\{\text{refl} \mapsto^{\tau'} Q\} : C \rightsquigarrow \Sigma'$$

1255

1256

1257
$$\frac{(x \ /^? \ \emptyset : A) \in E_1 \quad \Sigma; \Gamma \vdash \emptyset : A \quad rhs_1 = \text{impossible}}{\Sigma; \Gamma \vdash P \mid \mathbf{f} \bar{q} := \text{case}_x\{\} : C \rightsquigarrow \Sigma} \text{SPLITEMPTY}$$

1258

1259

Fig. 10. Rules for checking a list of clauses and elaborating them to a well-typed case tree.

1260

1261

1262 $\boxed{\Sigma; \Gamma \vdash E \Rightarrow \text{SOLVED}(\sigma)}$

1263

1264
$$\frac{(\Sigma \vdash [w_k / p_k] \searrow \sigma_k)_{k=1 \dots n} \quad \sigma = \uplus_k \sigma_k \quad (\Sigma; \Gamma \vdash [p_k] \sigma = w_k : A_k)_{k=1 \dots n}}{\Sigma; \Gamma \vdash \{w_k \ /^? \ p_k : A_k \mid k = 1 \dots n\} \Rightarrow \text{SOLVED}(\sigma)}$$

1265

1266

1267

Fig. 11. Rule for constructing the final substitution and checking all constraints when splitting is done.

1270 even contain constructor-headed terms in an *open* context. Our language has two kinds

1271 of case-style types: datatypes $D \bar{v}$ with no constructors, and identity types $u \equiv_A v$ where

1272 $\Sigma; \Gamma \vdash_x u =^? v : A \Rightarrow \text{NO}$.

1273

1274

1275 $\boxed{P(x : A)}$ Replace the first application pattern p in each clause by the constraint $x /? p : A$.

$$1276 \quad \epsilon(x : A) = \epsilon$$

$$1277 \quad ([E]p \bar{q} \hookrightarrow rhs, P)(x : A) = ([E \cup \{x /? p : A\}]\bar{q} \hookrightarrow rhs), P(x : A)$$

1280 Fig. 12. Partially decomposed clauses after introducing a new variable (partial function).

1282 $\boxed{P.\pi}$ Keep only clauses with copattern π , with this copattern removed.

$$1283 \quad \epsilon.\pi = \epsilon$$

$$1284 \quad ([E].\pi \bar{q} \hookrightarrow rhs, P).\pi = ([E]\bar{q} \hookrightarrow rhs), P.\pi$$

$$1285 \quad ([E].\pi' \bar{q} \hookrightarrow rhs, P).\pi = P.\pi \quad \text{if } \pi \neq \pi'$$

1288 Fig. 13. Partially decomposed clauses after a copattern split (partial function).

1290 $\boxed{\Sigma \vdash P\sigma \Rightarrow P'}$ (Σ fixed, dropped from rules.)

$$1291 \quad \frac{}{\epsilon\sigma \Rightarrow \epsilon} \quad \frac{(v /? p : A) \in E \quad v\sigma /? p : A\sigma \Rightarrow \perp \quad P\sigma \Rightarrow P'}{([E]\bar{q} \hookrightarrow rhs, P)\sigma \Rightarrow P'}$$

$$1292 \quad \frac{E = \{w_k /? p_k : A_k \mid k = 1 \dots n\} \quad (w_k\sigma /? p_k : A_k\sigma \Rightarrow E_i)_{k=1 \dots n} \quad P\sigma \Rightarrow P'}{([E]\bar{q} \hookrightarrow rhs, P)\sigma \Rightarrow ([\cup_i E_i]\bar{q} \hookrightarrow rhs), P'}$$

1298 Fig. 14. Rules for transforming partially decomposed clauses after refining the pattern with a case split.

1300 $\boxed{\Sigma \vdash v /? p : A \Rightarrow E_\perp}$ and $\boxed{\Sigma \vdash \bar{v} /? \bar{p} : \Delta \Rightarrow E_\perp}$ (Σ fixed, dropped from rules)

$$1301 \quad \frac{v \searrow c \bar{v} \quad A \searrow D \bar{u} \quad \text{constructor } c \Delta_c : D \Delta \in \Sigma \quad \bar{v} /? \bar{p} : \Delta_c[\bar{u} / \Delta] \Rightarrow E_\perp}{v /? c \bar{p} : A \Rightarrow E_\perp}$$

$$1302 \quad \frac{v \searrow \text{refl} \quad A \searrow u \equiv_B u'}{v /? \text{refl} : A \Rightarrow \{\}} \quad \frac{v \searrow c' \bar{v} \quad c \neq c'}{v /? c \bar{p} : A \Rightarrow \perp} \quad \frac{}{v /? p : A \Rightarrow \{v /? p : A\}}$$

$$1303 \quad \frac{}{\epsilon /? \epsilon : \epsilon \Rightarrow \{\}} \quad \frac{v /? p : A \Rightarrow E_\perp \quad \bar{v} /? \bar{p} : \Delta[v/x] \Rightarrow E'_\perp}{v \bar{v} /? p \bar{p} : (x : A)\Delta \Rightarrow E_\perp \cup E'_\perp}$$

1312 Fig. 15. Rules for simplifying the constraints of a partially decomposed clause.

1314 $\boxed{\Sigma; \Gamma \vdash \emptyset : A}$ (Σ fixed, dropped from rules)

$$1315 \quad \frac{A \searrow D \bar{v} \quad \text{data } D \Delta : \text{Set}_\ell \text{ where } \epsilon \in \Sigma}{\Gamma \vdash \emptyset : A} \quad \frac{A \searrow u \equiv_B v \quad \Gamma \vdash_x u =? v : B \Rightarrow \text{NO}}{\Gamma \vdash \emptyset : A}$$

1320 Fig. 16. Rules for caseless types.

The following rules constitute the elaboration algorithm $\Sigma; \Gamma \vdash P \mid f \bar{q} := Q : C \rightsquigarrow \Sigma'$:

DONE applies when the first user clause in P has no more copatterns and all its constraints are solved according to $\Sigma; \Gamma \vdash E \Rightarrow \text{SOLVED}(\sigma)$. If this is the case, then construction of the case tree is finished, adding the clause **clause** $\Gamma \vdash f \bar{q} \hookrightarrow v \sigma : C$ to the signature.

INTRO applies when C is a function type and all the user clauses have at least one application copattern. It constructs the case tree $\lambda x. Q$, using $P(x : A)$ to construct the subtree Q .

COSPLIT applies when C is a record type and all the user clauses have at least one projection copattern. It constructs the case tree **record** $\{\pi_1 \mapsto Q_1; \dots; \pi_n \mapsto Q_n\}$, using $P.\pi_i$ to construct the branch Q_i corresponding to projection π_i .

COSPLITEMPTY applies when C is a record type with no projections and the first clause starts with an absurd pattern. It then constructs the case tree **record** $\{\}$.

SPLITCON applies when the first clause has a constraint of the form $x \ /? c_i \bar{p}$ and the type of x in Γ is a datatype. For each constructor c_i of this datatype, it constructs a pattern substitution ρ_i replacing x by c_i applied to fresh variables. It then constructs the case tree **case** $_x\{c_1 \hat{\Delta}'_1 \mapsto Q_1; \dots; c_n \hat{\Delta}'_n \mapsto Q_n\}$, using $\Sigma \vdash P\rho_i \Rightarrow P_i$ to construct the branches Q_i .

SPLITEQ applies when the first clause has a constraint of the form $x \ /? \text{refl}$ and the type of x in Γ is an identity type $u \equiv_A v$. It tries to unify u with v , expecting a positive success. If unification succeeds with output (Γ'_1, ρ, τ) , it constructs the case tree **case** $_x\{\text{refl} \mapsto^{r'} Q\}$, using $\Sigma \vdash P\rho' \Rightarrow P'$ to construct the subtree Q . Here ρ' and τ' are lifted versions of ρ and τ over the part of the context that is untouched by unification.

SPLITEMPTY applies when the first clause has a constraint of the form $x \ /? \emptyset$, and the type of x is a caseless type according to $\Sigma; \Gamma \vdash \emptyset : A$. It then produces the case tree **case** $_x\{\}$.

Remark 26 (Limitations). The algorithm does not detect unreachable clauses, we left that aspect out of the formal description. Further, **SPLITEMPTY** may leave some user patterns uninspected, which may then be ill-typed. However, an easy check whether the whole lhs $f[\bar{q}]$ is well-typed as a term can rule out ill-typed patterns.

5.3 Preservation of first-match semantics

Now that we have described the elaboration algorithm from a list of clauses to a well-typed case tree, we can state and prove our main correctness theorem. We already know that elaboration always produces a well-typed case tree by construction (if it succeeds), and that well-typed case trees are type preserving (**Theorem 22**) and cover all cases (**Theorem 25**). Now we prove that the case tree we get is the right one, i.e. that it corresponds to the definition written by the user.

To prove this theorem, we assume that the clauses we get from the user have already been scope checked, i.e. each variable in the right-hand side of a clause is bound somewhere in the patterns on the left.

Definition 27. A partially decomposed clause $[E]\bar{q} \hookrightarrow v$ is *well-scoped* if every free variable in v occurs at least once as a pattern variable in either \bar{q} or in p for some constraint $(w \ /? p : A) \in E$.

Theorem 28. Let $P = \{\bar{q}_i \hookrightarrow \text{rhs}_i \mid i = 1 \dots n\}$ be a list of well-scoped clauses such that $\Sigma_0 \vdash P \mid f := Q : C \rightsquigarrow \Sigma$ and let $\Sigma; \Gamma \mid f : C \vdash \bar{e} : B$ be eliminations. Suppose there is an index i such that:

- $\Sigma \vdash [\bar{e} / \bar{q}_j] \searrow \perp$ for $j = 1 \dots i - 1$.
- $\Sigma \vdash [\bar{e} / \bar{q}_i] \searrow \sigma$.

Then $\text{rhs}_i = u_i$ is not **impossible** and $\Sigma \vdash Q \mid f \bar{e} \longrightarrow u_i \sigma$.

For the proof, we first need two basic properties of the auxiliary judgement $\Sigma \vdash v \ /? p : A \Rightarrow E$.

Lemma 29. If $\Sigma \vdash v \ /? p : A \Rightarrow E$ where $E = \{w_k \ /? p_k : B_k \mid k = 1 \dots l\}$, then for any substitution σ we also have $\Sigma \vdash v \sigma \ /? p : A\sigma \Rightarrow E'$ where $E' = \{w_k \sigma \ /? p_k : B_k \sigma \mid k = 1 \dots l\}$.

PROOF. This follows directly from the rules of matching (Fig. 6) and simplification (Fig. 15). \square

Lemma 30. *Let σ be a substitution and suppose $\Sigma \vdash v \int^? p : A \Rightarrow E$. Then the following hold:*

- $\Sigma \vdash [v\sigma / p] \searrow \sigma'$ if and only if for each $(w_k \int^? p_k : A_k) \in E$, we have $\Sigma \vdash [w_k\sigma / p_k] \searrow \sigma_k$, and $\sigma' = \biguplus_k \sigma_k$.
- $\Sigma \vdash [v\sigma / p] \searrow \perp$ if and only if for some $(w_k \int^? p_k : A_k) \in E$, we have $\Sigma \vdash [w_k\sigma / p_k] \searrow \perp$.

PROOF. This follows directly from the rule of matching (Fig. 6) and simplification (Fig. 15). \square

The following lemma is the main component of the proof. It generalizes the statement of **Theorem 28** to the case where the left-hand side has already been refined to $\mathbf{f} \bar{q}$ and the user clauses have been partially decomposed. From this lemma, the main theorem follows directly by taking $\bar{q} = \epsilon$ and $E_i = \{\}$ for $i = 1 \dots n$.

Lemma 31. *Let $P = \{[E_i]\bar{q}_i \hookrightarrow \text{rhs}_i \mid i = 1 \dots n\}$ be a list of well-scoped partially decomposed clauses such that $\Sigma_0; \Gamma_0 \vdash P \mid \mathbf{f} \bar{q} := Q : C \rightsquigarrow \Sigma$, and suppose $\Gamma \vdash \sigma_0 : \Gamma_0$ and $\Sigma; \Gamma \mid \mathbf{f} \bar{q}\sigma_0 : C\sigma_0 \vdash \bar{e} : B$. If there is an index i such that:*

- For each $j = 1 \dots i - 1$ and each constraint $(w_k \int^? p_k : A_k) \in E_j$, either $\Sigma \vdash [w_k\sigma_0 / p_k] \searrow \theta_{jk}$ or $\Sigma \vdash [w_k\sigma_0 / p_k] \searrow \perp$.
- For each $j = 1 \dots i - 1$, either $\Sigma \vdash [\bar{e} / \bar{q}_j] \searrow \theta_{j0}$ or $\Sigma \vdash [\bar{e} / \bar{q}_j] \searrow \perp$.
- For each $j = 1 \dots i - 1$, either $\Sigma \vdash [w_k\sigma_0 / p_k] \searrow \perp$ for some constraint $(w_k \int^? p_k : A_k) \in E_j$, or $\Sigma \vdash [\bar{e} / \bar{q}_j] \searrow \perp$.
- For each $(w_k \int^? p_k : A_k) \in E_i$, we have $\Sigma \vdash [w_k\sigma_0 / p_k] \searrow \theta_k$.
- $\Sigma \vdash [\bar{e} / \bar{q}_i] \searrow \theta_0$.

Then $\text{rhs}_i = v_i$ is not **impossible** and $\Sigma \vdash Q\sigma_0 \bar{e} \longrightarrow v_i\theta$ where $\theta = \theta_0 \uplus (\biguplus_k \theta_k)$.

PROOF. By induction on the derivation of $\Sigma_0; \Gamma_0 \vdash P \mid \mathbf{f} \bar{q} := Q : C \rightsquigarrow \Sigma$:

- For the **DONE** rule where $Q = v_1\sigma$ and $\Sigma_0 = \Sigma$ we have $\bar{q}_1 = \epsilon$ and $\text{rhs}_1 = v_1$ (i.e. rhs_1 is not **impossible**). We also get that $\sigma = \biguplus_k \sigma_k$ is a substitution such that $\Sigma; \Gamma_0 \vdash v_1\sigma : C$ and $\Sigma \vdash [w_k / p_k] \searrow \sigma_k$ and $\Sigma; \Gamma_0 \vdash [p_k]\sigma = w_k : A_k$ for each $(w_k \int^? p_k : A_k) \in E_1$. Because $Q = v_1\sigma$, we have $\Sigma \vdash Q\sigma_0 \epsilon \longrightarrow v_1\sigma\sigma_0$, so what's left to prove is that $v_1\sigma\sigma_0 = v_i\theta$ (syntactically). First we show that $i = 1$, i.e. the first clause matches. Since $\Sigma \vdash [w_k / p_k] \searrow \sigma_k$ we cannot have $\Sigma \vdash [w_k\sigma_0 / p_k] \searrow \perp$, and since $\bar{q}_1 = \epsilon$, we also cannot have $\Sigma \vdash [\bar{e} / \bar{q}_1] \searrow \perp$. The only remaining possibility is that i is 1. This means we have $\Sigma \vdash [w_k\sigma_0 / p_k] \searrow \theta_k$ for each $(w_k \int^? p_k : A_k) \in E_1$ and $\Sigma \vdash [\bar{e} / \bar{q}_1] \searrow \theta_0$. Since $\bar{q}_1 = \epsilon$ we also have $\bar{e} = \epsilon$ and $\theta_0 = []$. To finish this case, we show that $\sigma\sigma_0 = (\biguplus_k \sigma_k)\sigma_0$ and $\theta = \biguplus_k \theta_k$ coincide on all free variables in v . Since the clause $[E_1] \hookrightarrow v_1$ is well-scoped, for each free variable x in v_1 there is at least one constraint $(w_k \int^? p_k : A_k) \in E_1$ such that x is a pattern variable of p_k . Since we have both $\Sigma \vdash [w_k / p_k] \searrow \sigma_k$ and $\Sigma \vdash [w_k\sigma_0 / p_k] \searrow \theta_k$, we have $x\sigma_k\sigma_0 = x\theta_k$. This holds for any free variable x in v_1 , so we have $v_1\sigma\sigma_0 = v_1\theta$, finishing the proof for the base case.
- For the **INTRO** rule we have $Q = \lambda x. Q'$ where $\Sigma_0 \vdash C \searrow (x : A) \rightarrow B$ and $\bar{q}_i = p_i \bar{q}'_i$ for $i = 1 \dots n$. We also know that $\Sigma_0; \Gamma_0(x : A) \vdash P' \mid \mathbf{f} \bar{q} x := Q' : B \rightsquigarrow \Sigma$ where $P' = P(x : A)$. Since we have either $\Sigma \vdash [\bar{e} / p_1 \bar{q}'_1] \searrow \perp$ or $\Sigma \vdash [\bar{e} / p_1 \bar{q}'_1] \searrow \theta_0$, we have $\bar{e} = t \bar{e}'$ for some term t . Now we apply the induction hypothesis to show that $\text{rhs}_i = v_i$ is not **impossible** and $\Sigma \vdash Q'(\sigma_0 \uplus [t / x]) \bar{e}' \longrightarrow v_i\theta$, hence also $\Sigma \vdash Q\sigma_0 \bar{e} \longrightarrow v_i\theta$.
- For the **COSPLIT** rule where $Q = \text{record}\{\pi_1 \mapsto Q_1; \dots; \pi_n \mapsto Q_n\}$, we have $\Sigma_0 \vdash C \searrow \mathbf{R} \bar{v}$ and $\bar{q}_1 = .\pi_\alpha \bar{q}'_1$ where **projection** $x : \mathbf{R} \Delta \vdash .\pi_\alpha : A_\alpha \in \Sigma_0$. We have either $\Sigma \vdash [\bar{e} / .\pi_\alpha \bar{q}'_1] \searrow \theta_{10}$ or $\Sigma \vdash [\bar{e} / .\pi_\alpha \bar{q}'_1] \searrow \perp$, so $\bar{e} = .\text{pi}_\beta \bar{e}'$ for some **projection** $x : \mathbf{R} \Delta \vdash .\pi_\beta : A_\beta \in \Sigma_0$. We then

- 1422 have $\Sigma_{\beta-1}; \Gamma \vdash P \cdot \pi_\beta \mid f \bar{q} \cdot \text{pi}_\beta := Q_\beta : A_\beta[\bar{v} / \Delta; u / x] \rightsquigarrow \Sigma_\beta$. By induction we have that
 1423 rhs_i is not **impossible** and $\Sigma_\beta \vdash Q_\beta \sigma_0 \bar{e}' \longrightarrow v\theta$, hence also $\Sigma \vdash Q\sigma_0 \bar{e} \longrightarrow v\theta$.
- 1424 • For the **COSPLITEMPTY** rule, we have $\bar{q}_1 = \emptyset \bar{q}'_1$. Since there are no rules for $[\pi / \emptyset] \searrow \theta_\perp$,
 1425 this case is impossible.
- 1426 • For the **SPLITCON** rule we know that $Q = \text{case}_x\{c_1 \hat{\Delta}'_1 \mapsto Q_1; \dots; c_n \hat{\Delta}'_n \mapsto Q_n\}$ where $n \geq$
 1427 1 , $\Gamma = \Gamma_1(x : A)\Gamma_2$ and $\Sigma_0 \vdash A \searrow \mathbf{D} \bar{v}$. Since $(x \ /? c_\alpha \bar{p} : A) \in E_1$, we either have
 1428 $\Sigma \vdash [x\sigma_0 / c_\alpha \bar{p}] \searrow \theta_{1k}$ or $\Sigma \vdash [x\sigma_0 / c_\alpha \bar{p}] \searrow \perp$ (this is the case both if $i = 1$ and if $i > 1$).
 1429 In either case, we have $\Sigma \vdash x\sigma_0 \searrow c_\beta \bar{u}$ for some **constructor** $c_\beta \Delta_\beta : \mathbf{D} \Delta \in \Sigma_0$. Let $\Delta'_\beta =$
 1430 $\Delta_\beta[\bar{v} / \Delta]$ and $\rho_\beta = [c_\beta \hat{\Delta}'_\beta / x]$, then we have $\Sigma \vdash P\rho_\beta \Rightarrow P_\beta$ and $\Sigma_{\beta-1}; \Gamma_1 \Delta'_\beta \Gamma_2 \rho_\beta \vdash P_\beta \mid f \bar{q}\rho_\beta :=$
 1431 $Q_\beta : C\rho_\beta \rightsquigarrow \Sigma_\beta$. We now apply the induction hypothesis to get that $rhs_i = v_i$ is not
 1432 **impossible** and $\Sigma_\beta \vdash Q_\beta(\sigma_0 \uplus [\bar{u} / \Delta'_\beta \sigma_0]) \bar{e} \longrightarrow v_i\theta$, hence also $\Sigma \vdash Q\sigma_0 \bar{e} \longrightarrow v_i\theta$.
- 1433 • For the **SPLITEQ** rule where $Q = \text{case}_x\{\text{refl} \mapsto^r Q'\}$, we know that $\Gamma = \Gamma_1(x : A)\Gamma_2$ and
 1434 $\Sigma_0 \vdash A \searrow u \equiv_A v$. Since $(x \ /? \text{refl} : A) \in E_1$, we either have $\Sigma \vdash [x\sigma_0 / \text{refl}] \searrow \theta_{1k}$ or
 1435 $\Sigma \vdash [x\sigma_0 / \text{refl}] \searrow \perp$. However, the latter case is impossible since **refl** is the only constructor
 1436 of the identity type, so we have $\Sigma \vdash x\sigma_0 \searrow \text{refl}$ and $\theta_{1k} = []$. We moreover have $\Sigma_0; \Gamma_1 \vdash_x$
 1437 $u =^? v : B \Rightarrow \text{YES}(\Gamma'_1, \rho, \tau)$ and $\Sigma_0; \Gamma'_1 \Gamma_2 \rho \vdash P' \mid f \bar{q}\rho' := Q' : C\rho' \rightsquigarrow \Sigma$ where $\rho' = \rho \uplus \mathbb{1}_{\Gamma_2}$
 1438 and $\Sigma_0 \vdash P\rho \Rightarrow P'$. By induction (and using **Definition 13** to show that $\rho'; \tau; \sigma_0 = \sigma_0$), we get
 1439 that $rhs_i = v_i$ is not **impossible** and $\Sigma \vdash Q'(\tau; \sigma_0) \bar{e} \longrightarrow \theta$, hence also $\Sigma \vdash Q\sigma_0 \bar{e} \longrightarrow \theta$.
- 1440 • For the **SPLITEMPTY** rule we know that $Q = \text{case}_x\{\}$ and $(x \ /? \emptyset : A) \in E_1$ where $\Sigma_0; \Gamma \vdash \emptyset : A$.
 1441 We either have $\Sigma \vdash [x\sigma_0 / \emptyset] \searrow \theta_x$ or $\Sigma \vdash [x\sigma_0 / \emptyset] \searrow \perp$. However, there are no rules for
 1442 $\Sigma \vdash [v / \emptyset] \searrow \theta_\perp$ so this case is impossible.

□

6 RELATED WORK

1447 Dependent pattern matching was introduced in the seminal work by [Coquand \[1992\]](#). It is used
 1448 in the implementation of various dependently typed languages such as Agda [\[Norell 2007\]](#), Idris
 1449 [\[Brady 2013\]](#), the Equations package for Coq [\[Sozeau 2010\]](#), and Lean [\[de Moura et al. 2015\]](#).

1450 Previous work by [Norell \[2007\]](#), [Sozeau \[2010\]](#), and [Cockx \[2017\]](#) also describe elaborations
 1451 from clauses to a case tree, but in much less detail than presented here, and they do not support
 1452 copatterns or provide a correctness proof. In cases where both our current algorithm and these
 1453 previous algorithms succeed, we expect there is no difference between the resulting case trees.
 1454 However, our current algorithm is much more flexible in the placement of dot patterns, so it accepts
 1455 more definitions than was possible before (see [Example 4](#)).

1456 The translation from a case tree to primitive datatype eliminators was pioneered by [McBride](#)
 1457 [\[2000\]](#) and further detailed by [Goguen et al. \[2006\]](#) for type theory with uniqueness of identity
 1458 proofs and [Cockx \[2017\]](#) in a theory without.

1459 Forced patterns, as well as forced constructors, were introduced by [Brady et al. \[2003\]](#). [Brady](#)
 1460 [et al.](#) focus mostly on the compilation process and the possibility to erase arguments and construc-
 1461 tor tags, while we focus more on the process of typechecking a definition by pattern matching and
 1462 the construction of a case tree.

1463 Copatterns were introduced in the simply-typed setting by [Abel et al. \[2013\]](#) and subsequently
 1464 used for unifying corecursion and recursion in System F^ω [\[Abel and Pientka 2013\]](#). In the context of
 1465 Isabelle/HOL, [Blanchette et al. \[2017\]](#) use copatterns as syntax for mixed recursive-corecursive defi-
 1466 nitions. [Setzer et al. \[2014\]](#) give an algorithm for elaborating a definition by mixed pattern/copattern
 1467 matching to a nested case expression, yet only for a simply typed language. [Thibodeau et al. \[2016\]](#)
 1468 present a language with deep (co)pattern matching and a restricted form of dependent types. In
 1469 their language, types can only depend on a user-defined domain with decidable equality and the

1470

1471 types of record fields cannot depend on each other, thus, a *self* value is not needed for checking pro-
 1472 jections. They feature indexed data and record types in the surface language which are elaborated
 1473 into non-indexed types via equality types, just as in our core language.

1474 The connection between focusing [Andreoli 1992] and pattern matching has been systematically
 1475 explored by Zeilberger [2009]. In Licata et al. [2008] copatterns (“destructor patterns”) also appear
 1476 in the context of simple typing with connectives from linear logic. Krishnaswami [2009] boils the
 1477 connection to focusing down to usual non-linear types; however, he has no copatterns as he only
 1478 considers the product type as multiplicative (tensor), not additive. Thibodeau et al. [2016] extend
 1479 the connection to copatterns for indexed record types.

1480 Elaborating a definition by pattern matching to a case tree [Augustsson 1985] simultaneously
 1481 typechecks the clauses and checks their coverage, so our algorithm has a lot in common with
 1482 coverage checking algorithms. For example, Norell [2007] views the construction of a case tree as
 1483 a part of coverage checking. Oury [2007] presents a similar algorithm for coverage checking and
 1484 detecting useless cases in definitions by dependent pattern matching.

1485

1486 7 FUTURE WORK AND CONCLUSION

1487 In this paper, we give a description of an elaboration algorithm for definitions by dependent co-
 1488 pattern matching that is at the same time elegant enough to be intuitively understandable, simple
 1489 enough to study formally, and detailed enough to serve as the basis for a practical implementation.

1490 The implementation of our algorithm as part of the Agda typechecker is at the moment of
 1491 writing still work in progress. In fact, the main reason to write this paper was to get a clear idea
 1492 of what exactly should be implemented. For instance, while working on the proof of Theorem 28,
 1493 we were quite surprised to discover that it did not hold at first: matching was performed lazily
 1494 from left to right, but the case tree produced by elaboration may not agree on this order! This
 1495 problem was not just theoretical, but also manifested itself in the implementation of Agda as a
 1496 violation of subject reduction [Agda issue 2018a]. Removing the shortcut rule from the definition
 1497 of matching removed this behavioral divergence mismatch. The complete formalization of the
 1498 elaboration algorithm in this paper lets us continue the implementation with confidence.

1499 Agda also has a number of features that are not described in this paper, such as nonrecursive
 1500 record types with η equality and general indexed datatypes (not just the identity type). The imple-
 1501 mentation also has to deal with the insertion of implicit arguments, the presence of metavariables
 1502 in the syntax, and reporting understandable errors when the algorithm fails. Based on our practi-
 1503 cal experience, we are confident that the algorithm presented here can be extended to deal with
 1504 all of these features.

1505

1506 REFERENCES

- 1507 Andreas Abel and Brigitte Pientka. 2013. Wellfounded Recursion with Copatterns: A Unified Approach to Termination
 1508 and Productivity. In *Proceedings of the Eighteenth ACM SIGPLAN International Conference on Functional Programming,*
 1509 *ICFP’13, Boston, MA, USA, September 25–27, 2013*, Greg Morrisett and Tarmo Uustalu (Eds.). ACM Press, 185–196. <https://doi.org/10.1145/2500365.2500591>
- 1510 Andreas Abel, Brigitte Pientka, David Thibodeau, and Anton Setzer. 2013. Copatterns: Programming Infinite Structures by
 1511 Observations. In *The 40th Annual ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL’13,*
 1512 *Rome, Italy, January 23 - 25, 2013*, Roberto Giacobazzi and Radhia Cousot (Eds.). ACM Press, 27–38. <http://dl.acm.org/citation.cfm?id=2429069>
- 1513 Agda development team. 2017. *Agda 2.5.3 documentation*. <http://agda.readthedocs.io/en/v2.5.3/>
- 1514 Agda issue. 2017a. Disambiguation of type based on pattern leads to non-unique meta solution. (2017). <https://github.com/agda/agda/issues/2834> (on the Agda bug tracker).
- 1515 Agda issue. 2017b. Internal error in src/full/Agda/TypeChecking/Coverage/Match.hs:312. (2017). <https://github.com/agda/agda/issues/2874> (on the Agda bug tracker).
- 1516 Agda issue. 2017c. Panic: unbound variable. (2017). <https://github.com/agda/agda/issues/2856> (on the Agda bug tracker).

1519

- 1520 Agda issue. 2017d. Record constructor is accepted, record pattern is not. (2017). <https://github.com/agda/agda/issues/2850>
 1521 (on the Agda bug tracker).
- 1522 Agda issue. 2018a. Mismatch between order of matching in clauses and case tree; subject reduction broken. (2018). <https://github.com/agda/agda/issues/2964> (on the Agda bug tracker).
- 1523 Agda issue. 2018b. Unifier throws away pattern. (2018). <https://github.com/agda/agda/issues/2896> (on the Agda bug
 1524 tracker).
- 1525 Jean-Marc Andreoli. 1992. Logic Programming with Focusing Proofs in Linear Logic. *Journal of Logic and Computation* 2,
 1526 3 (1992), 297–347. <https://doi.org/10.1093/logcom/2.3.297>
- 1527 Lennart Augustsson. 1985. Compiling Pattern Matching. In *Functional Programming Languages and Computer Architecture, FPCA 1985, Nancy, France, September 16-19, 1985, Proceedings (Lecture Notes in Computer Science)*, Jean-Pierre Jouannaud
 1528 (Ed.), Vol. 201. Springer, 368–381. https://doi.org/10.1007/3-540-15975-4_48
- 1529 Jasmin Christian Blanchette, Aymeric Bouzy, Andreas Lochbihler, Andrei Popescu, and Dmitriy Traytel. 2017. Friends with
 1530 Benefits - Implementing Corecursion in Foundational Proof Assistants. In *Programming Languages and Systems - 26th
 1531 European Symposium on Programming, ESOP 2017, Held as Part of the European Joint Conferences on Theory and Practice
 1532 of Software, ETAPS 2017, Uppsala, Sweden, April 22-29, 2017, Proceedings (Lecture Notes in Computer Science)*, Hongseok
 1533 Yang (Ed.), Vol. 10201. Springer, 111–140. https://doi.org/10.1007/978-3-662-54434-1_5
- 1534 Edwin Brady. 2013. Idris, a general-purpose dependently typed programming language: Design and implementation. *Journal
 1535 of Functional Programming* 23, 5 (2013), 552–593. <https://doi.org/10.1017/S095679681300018X>
- 1536 Edwin Brady, Conor McBride, and James McKinna. 2003. Inductive Families Need Not Store Their Indices. In *Types for
 1537 Proofs and Programs, International Workshop, TYPES 2003, Torino, Italy, April 30 - May 4, 2003, Revised Selected Papers
 1538 (Lecture Notes in Computer Science)*, Stefano Berardi, Mario Coppo, and Ferruccio Damiani (Eds.), Vol. 3085. Springer,
 1539 115–129. https://doi.org/10.1007/978-3-540-24849-1_8
- 1540 Luca Cardelli. 1984. Compiling a Functional Language. In *Proceedings of the 1984 ACM Conference on LISP and Functional
 1541 Programming, August 5-8, 1984, Austin, Texas, USA*. ACM Press, 208–217. [http://lucacardelli.name/Papers/CompilingML.
 1542 A4.pdf](http://lucacardelli.name/Papers/CompilingML.A4.pdf)
- 1543 Jesper Cockx. 2017. *Dependent pattern matching and proof-relevant unification*. Ph.D. Dissertation. KU Leuven.
- 1544 Jesper Cockx, Dominique Devriese, and Frank Piessens. 2016. Unifiers as equivalences: proof-relevant unification of de-
 1545 pendently typed data, See [Garrigue et al. 2016], 270–283. <https://doi.org/10.1145/2951913.2951917>
- 1546 Thierry Coquand. 1992. Pattern Matching with Dependent Types. In *Proceedings of the 1992 Workshop on Types for Proofs
 1547 and Programs, Båstad, Sweden, June 1992*, Bengt Nordström, Kent Pettersson, and Gordon Plotkin (Eds.), 71–83. <http://www.cse.chalmers.se/~coquand/pattern.ps>
- 1548 Leonardo Mendonça de Moura, Soonho Kong, Jeremy Avigad, Floris van Doorn, and Jakob von Raumer. 2015. The Lean
 1549 Theorem Prover (System Description). In *Automated Deduction - CADE-25 - 25th International Conference on Automated
 1550 Deduction, Berlin, Germany, August 1-7, 2015, Proceedings (Lecture Notes in Computer Science)*, Amy P. Felty and Aart
 1551 Middeldorp (Eds.), Vol. 9195. Springer, 378–388. https://doi.org/10.1007/978-3-319-21401-6_26
- 1552 Jacques Garrigue, Gabriele Keller, and Eijiro Sumii (Eds.). 2016. *Proceedings of the 21st ACM SIGPLAN International Con-
 1553 ference on Functional Programming, ICFP 2016, Nara, Japan, September 18-22, 2016*. ACM Press. [https://doi.org/10.1145/
 1554 2951913](https://doi.org/10.1145/2951913)
- 1555 Healfdene Goguen, Conor McBride, and James McKinna. 2006. Eliminating Dependent Pattern Matching. In *Algebra,
 1556 Meaning, and Computation, Essays Dedicated to Joseph A. Goguen on the Occasion of His 65th Birthday (Lecture Notes
 1557 in Computer Science)*, Kokichi Futatsugi, Jean-Pierre Jouannaud, and José Meseguer (Eds.), Vol. 4060. Springer, 521–540.
 1558 https://doi.org/10.1007/11780274_27
- 1559 INRIA. 2017. *The Coq Proof Assistant Reference Manual* (version 8.7 ed.). INRIA. <http://coq.inria.fr/>
- 1560 Neelakantan R. Krishnaswami. 2009. Focusing on pattern matching. In *Proceedings of the 36th ACM SIGPLAN-SIGACT
 1561 Symposium on Principles of Programming Languages, POPL 2009, Savannah, GA, USA, January 21-23, 2009*, Zhong Shao
 1562 and Benjamin C. Pierce (Eds.). ACM Press, 366–378.
- 1563 Daniel R. Licata, Noam Zeilberger, and Robert Harper. 2008. Focusing on Binding and Computation. In *Proceedings of the
 1564 Twenty-Third Annual IEEE Symposium on Logic in Computer Science, LICS 2008, 24-27 June 2008, Pittsburgh, PA, USA*, Frank
 1565 Pfenning (Ed.). IEEE Computer Society Press, 241–252. [http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=
 1566 4557886](http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=4557886) Long version available as Technical Report CMU-CS-08-101.
- 1567 Luc Maranget. 1992. Compiling Lazy Pattern Matching. In *LISP and Functional Programming*. 21–31. [https://doi.org/10.
 1568 1145/141471.141499](https://doi.org/10.1145/141471.141499)
- 1569 Conor McBride. 2000. *Dependently typed functional programs and their proofs*. Ph.D. Dissertation. University of Edinburgh.
- 1570 Ulf Norell. 2007. *Towards a practical programming language based on dependent type theory*. Ph.D. Dissertation. Chalmers
 1571 University of Technology.
- 1572 Nicolas Oury. 2007. Pattern matching coverage checking with dependent types using set approximations. In *Proceedings of
 1573 the ACM Workshop Programming Languages meets Program Verification, PLPV 2007, Freiburg, Germany, October 5, 2007*,

- 1569 Aaron Stump and Hongwei Xi (Eds.). ACM Press, 47–56. <https://doi.org/10.1145/1292597.1292606>
- 1570 Robert Pollack. 1998. How to Believe a Machine-Checked Proof. In *Twenty Five Years of Constructive Type Theory*, Giovanni Sambin and Jan Smith (Eds.). Oxford University Press. <http://www.brics.dk/RS/97/18/BRICS-RS-97-18.pdf>
- 1571 Anton Setzer, Andreas Abel, Brigitte Pientka, and David Thibodeau. 2014. Unnesting of Copatterns. In *Rewriting and Typed Lambda Calculi - Joint International Conference, RTA-TLCA 2014, Held as Part of the Vienna Summer of Logic, VSL 2014, Vienna, Austria, July 14-17, 2014. Proceedings (Lecture Notes in Computer Science)*, Gilles Dowek (Ed.), Vol. 8560. Springer, 31–45. https://doi.org/10.1007/978-3-319-08918-8_3
- 1572 Matthieu Sozeau. 2010. Equations: A Dependent Pattern-Matching Compiler. In *Interactive Theorem Proving, First International Conference, ITP 2010, Edinburgh, UK, July 11-14, 2010. Proceedings (Lecture Notes in Computer Science)*, Matt Kaufmann and Lawrence C. Paulson (Eds.), Vol. 6172. Springer, 419–434. https://doi.org/10.1007/978-3-642-14052-5_29
- 1573 David Thibodeau, Andrew Cave, and Brigitte Pientka. 2016. Indexed codata types, See [Garrigue et al. 2016], 351–363. <https://doi.org/10.1145/2951913.2951929>
- 1574 Noam Zeilberger. 2008. Focusing and higher-order abstract syntax. In *Proceedings of the 35th ACM SIGPLAN-SIGACT Symposium on Principles of Programming Languages, POPL 2008, San Francisco, California, USA, January 7-12, 2008*, George C. Necula and Philip Wadler (Eds.). ACM Press, 359–369.
- 1575 Noam Zeilberger. 2009. *The Logical Basis of Evaluation Order and Pattern-Matching*. Ph.D. Dissertation. Carnegie Mellon University. <http://software.imdea.org/~noam.zeilberger/thesis.pdf>
- 1576
- 1577
- 1578
- 1579
- 1580
- 1581
- 1582
- 1583
- 1584
- 1585
- 1586
- 1587
- 1588
- 1589
- 1590
- 1591
- 1592
- 1593
- 1594
- 1595
- 1596
- 1597
- 1598
- 1599
- 1600
- 1601
- 1602
- 1603
- 1604
- 1605
- 1606
- 1607
- 1608
- 1609
- 1610
- 1611
- 1612
- 1613
- 1614
- 1615
- 1616
- 1617

A INFERENCE RULES

$\vdash \Gamma$

$$\frac{}{\vdash \epsilon} \quad \frac{\Gamma \vdash A \quad x \notin \text{dom}(\Gamma)}{\vdash \Gamma(x : A)}$$

Fig. 17. The typing rules for valid contexts.

$\Gamma \vdash u = v : A$

$$\frac{\Gamma \vdash u : A}{\Gamma \vdash u = u : A} \quad \frac{\Gamma \vdash u_1 = u_2 : A}{\Gamma \vdash u_2 = u_1 : A} \quad \frac{\Gamma \vdash u_1 = u_2 : A \quad \Gamma \vdash u_2 = u_3 : A}{\Gamma \vdash u_1 = u_3 : A}$$

$$\frac{\Gamma \vdash u_1 = u_2 : A_1 \quad \Gamma \vdash A_1 = A_2}{\Gamma \vdash u_1 = u_2 : A_2} \quad \frac{\Gamma \vdash A_1 = A_2 : \text{Set}_\ell \quad \Gamma(x : A_1) \vdash B_1 = B_2 : \text{Set}_{\ell'}}{\Gamma \vdash (x : A_1) \rightarrow B_1 = (x : A_2) \rightarrow B_2 : \text{Set}_{\max(\ell, \ell')}}}$$

$$\frac{\vdash \Gamma \quad x : A \in \Gamma \quad \Gamma \mid x : A \vdash \bar{e}_1 = \bar{e}_2 : B}{\Gamma \vdash x \bar{e}_1 = x \bar{e}_2 : B}$$

$$\frac{\Gamma \vdash A_1 = A_2 : \text{Set}_\ell \quad \Gamma \vdash u_1 = u_2 : A_1 \quad \Gamma \vdash v_1 = v_2 : A_1}{\Gamma \vdash (u_1 \equiv_{A_1} v_1) = (u_2 \equiv_{A_2} v_2) : \text{Set}_\ell}$$

$$\frac{\text{data } D \Delta : \text{Set}_\ell \in \Sigma \quad \Gamma \vdash \bar{u}_1 = \bar{u}_2 : \Delta}{\Gamma \vdash D \bar{u}_1 = D \bar{u}_2 : \text{Set}_\ell} \quad \frac{\text{record } R \Delta : \text{Set}_\ell \in \Sigma \quad \Gamma \vdash \bar{u}_1 = \bar{u}_2 : \Delta}{\Gamma \vdash R \bar{u}_1 = R \bar{u}_2 : \text{Set}_\ell}$$

$$\frac{\text{constructor } c \Delta_c : D \Delta \in \Sigma \quad \Gamma \vdash \bar{u} : \Delta \quad \Gamma \vdash \bar{v}_1 = \bar{v}_2 : \Delta_c[\bar{u} / \Delta]}{\Gamma \vdash c \bar{v}_1 = c \bar{v}_2 : D \bar{u}}$$

$$\frac{\text{definition } f : A \in \Sigma \quad \Gamma \mid f : A \vdash \bar{e}_1 = \bar{e}_2 : B}{\Gamma \vdash f \bar{e}_1 = f \bar{e}_2 : B} \quad \frac{\text{clause } \Delta \vdash f \bar{q} \leftrightarrow v : B \in \Sigma \quad \Gamma \vdash \sigma : \Delta}{\Gamma \vdash f [\bar{q}] \sigma = v \sigma : B \sigma}$$

Fig. 18. The conversion rules for terms.

$$\begin{array}{c}
1667 \\
1668 \\
1669 \\
1670 \\
1671 \\
1672 \\
1673 \\
1674 \\
1675 \\
1676 \\
1677 \\
1678
\end{array}
\quad
\boxed{\Gamma \mid u : A \vdash \bar{e}_1 = \bar{e}_2 : B}
\quad
\frac{\Gamma \vdash v_1 = v_2 : A \quad \Gamma \mid u \ v_1 : B[v_1/x] \vdash \bar{e}_1 = \bar{e}_2 : C}{\Gamma \mid u : (x : A) \rightarrow B \vdash v_1 \bar{e}_1 = v_2 \bar{e}_2 : C}$$

$$\frac{\text{projection } x : \mathbf{R} \ \Delta \vdash .\pi : A \in \Sigma \quad \Gamma \mid u \ .\pi : A[\bar{v}/\Delta, u/x] \vdash \bar{e}_1 = \bar{e}_2 : C}{\Gamma \mid u : \mathbf{R} \ \bar{v} \vdash .\pi \ \bar{e}_1 = .\pi \ \bar{e}_2 : C}$$

$$\frac{\Gamma \mid u : A \vdash \bar{e}_1 = \bar{e}_2 : B \quad \Gamma \vdash A = A' \quad \Gamma \vdash B = B'}{\Gamma \mid u : A' \vdash \bar{e}_1 = \bar{e}_2 : B'}$$

Fig. 19. The conversion rules for eliminations.

$$\begin{array}{c}
1681 \\
1682 \\
1683 \\
1684 \\
1685 \\
1686 \\
1687 \\
1688 \\
1689 \\
1690 \\
1691 \\
1692 \\
1693 \\
1694 \\
1695 \\
1696 \\
1697 \\
1698 \\
1699 \\
1700 \\
1701 \\
1702 \\
1703 \\
1704 \\
1705 \\
1706 \\
1707 \\
1708 \\
1709 \\
1710 \\
1711 \\
1712 \\
1713 \\
1714 \\
1715
\end{array}
\quad
\boxed{\Gamma \vdash \bar{u} = \bar{v} : \Delta}
\quad
\frac{\vdash \Gamma \quad \Gamma \vdash u_1 = u_2 : A \quad \Gamma \vdash \bar{u}_1 = \bar{u}_2 : \Delta[u_1/x]}{\Gamma \vdash \epsilon = \epsilon : \epsilon \quad \Gamma \vdash u_1 \ \bar{u}_1 = u_2 \ \bar{u}_2 : (x : A)\Delta}$$

Fig. 20. The conversion rules for lists of terms.