

Chick: A Dependently-Typed List Language

Abd-El-Aziz Zayed Louis Hildebrand

McGill University

April 10, 2025

Motivation

Many list-related functions have constraints related to the list length.¹ Can we check them statically?

Module List

```
module List: sig .. end
```

List operations.

```
val tl : 'a list -> 'a list
```

Return the given list without its first element.

Raises `Failure` if the list is empty.

```
val nth : 'a list -> int -> 'a
```

Return the `n`-th element of the given list. The first element (head of the list) is at position 0.

Raises

- `Failure` if the list is too short.

```
val map2 : ('a -> 'b -> 'c) -> 'a list -> 'b list -> 'c list
```

`map2 f [a1; ...; an] [b1; ...; bn] is [f a1 b1; ...; f an bn].`

Raises `Invalid_argument` if the two lists are determined to have different lengths.

¹<https://ocaml.org/manual/5.1/api/List.html>

Outline

Basic Language

Better Pattern Matching

Extending the Language

Conclusion

Syntax of Types

Type	A	$::=$	Nat	Natural numbers
			$ \text{Vec } \ell$	Fixed-size list of nats
			$ \Pi(x: A_1).A_2$	Dependent functions
Length	ℓ	$::=$	$n \mid x \mid \ell_1 + \dots + \ell_k$	
N numeral	n	\in	\mathbb{N}	

Note

$$A_1 \rightarrow A_2 \triangleq \Pi(-: A_1).A_2$$

Syntax of Terms

Syn $s ::= x \mid s \ t$
| $\text{head } \ell \ s \mid \text{tail } \ell \ s$

Chk $t ::= s \mid \lambda x. t \mid \text{fix } x. t$
| $n \mid t_1 + \dots + t_k$
| $(\text{match } s \text{ with } \mid 0 \rightarrow t_1 \mid x + 1 \rightarrow t_2)$
| $\text{nil} \mid \text{cons } \ell \ t_1 \ t_2$
| $(\text{match } s \text{ with } \mid \text{nil} \rightarrow t_1 \mid \text{cons } x_1 \ x_2 \ x_3 \rightarrow t_2)$

Example

`cons` : $\prod (n : \text{Nat}) \text{Nat} \rightarrow \text{Vec } n \rightarrow \text{Vec } (n + 1)$

i.e., `cons` : $\prod (n : \text{Nat}). \prod (_ : \text{Nat}). \prod (_ : \text{Vec } n). \text{Vec } (n + 1)$

e.g., `[10]` \triangleq `cons` `(n = 0)` 10 `nil`

Example

$\text{cons} : \Pi(n : \text{Nat}). \text{Nat} \rightarrow \text{Vec } n \rightarrow \text{Vec } (n + 1)$

i.e., $\text{cons} : \Pi(n : \text{Nat}). \Pi(_ : \text{Nat}). \Pi(_ : \text{Vec } n). \text{Vec } (n + 1)$

e.g., $[10] \triangleq \text{cons } (n = 0) \ 10 \ \text{nil}$

Example

$\text{cons} : \Pi(n : \text{Nat}). \text{Nat} \rightarrow \text{Vec } n \rightarrow \text{Vec } (n + 1)$

i.e., $\text{cons} : \Pi(n : \text{Nat}). \Pi(_ : \text{Nat}). \Pi(_ : \text{Vec } n). \text{Vec } (n + 1)$

e.g., $[10] \triangleq \text{cons } (n = 0) \ 10 \ \text{nil}$

Example

$\text{cons} : \Pi(n : \text{Nat}). \text{Nat} \rightarrow \text{Vec } n \rightarrow \text{Vec } (n + 1)$

i.e., $\text{cons} : \Pi(n : \text{Nat}). \Pi(_ : \text{Nat}). \Pi(_ : \text{Vec } n). \text{Vec } (n + 1)$

e.g., $[10] \triangleq \text{cons } (n = 0) \ 10 \ \text{nil}$

One More Example

e.g., $\text{drop } (k = 2) (n = 3) [9, 8, 7, 6, 5] = [7, 6, 5]$

// Discard first k elements

```
drop :  $\Pi(k : \text{Nat}). \Pi(n : \text{Nat}). \text{Vec } (k + n) \rightarrow \text{Vec } n =$   
  fix drop.  $\lambda k. \lambda n. \lambda v.$   
    match  $k$  with  
    | 0  $\rightarrow$  nil  
    |  $k' + 1 \rightarrow \text{drop } k \ n \ (\text{tail } (k' + n) \ v)$ 
```

One More Example

e.g., $\text{drop } (k = 2) (n = 3) [9, 8, 7, 6, 5] = [7, 6, 5]$

Can you spot the mistakes?

// Discard first k elements

```
drop :  $\Pi(k : \text{Nat}). \Pi(n : \text{Nat}). \text{Vec } (k + n) \rightarrow \text{Vec } n =$   
  fix drop.  $\lambda k. \lambda n. \lambda v.$   
    match  $k$  with  
    | 0  $\rightarrow$  nil  
    |  $k' + 1 \rightarrow \text{drop } k \ n \ (\text{tail } (k' + n) \ v)$ 
```

One More Example

e.g., $\text{drop } (k = 2) (n = 3) [9, 8, 7, 6, 5] = [7, 6, 5]$

Can you spot the mistakes?

// Discard first k elements

```
drop :  $\Pi(k : \text{Nat}). \Pi(n : \text{Nat}). \text{Vec } (k + n) \rightarrow \text{Vec } n =$   
  fix drop.  $\lambda k. \lambda n. \lambda v.$   
    match  $k$  with  
    | 0  $\rightarrow$  nil      nil does not have type  $\text{Vec } n$   
    |  $k' + 1 \rightarrow$  drop  $k$   $n$  (tail  $(k' + n)$   $v$ )
```

One More Example

e.g., $\text{drop } (k = 2) (n = 3) [9, 8, 7, 6, 5] = [7, 6, 5]$

Can you spot the mistakes?

// Discard first k elements

```
drop :  $\Pi(k : \text{Nat}). \Pi(n : \text{Nat}). \text{Vec } (k + n) \rightarrow \text{Vec } n =$   
  fix drop.  $\lambda k. \lambda n. \lambda v.$   
    match  $k$  with  
    |  $0 \rightarrow v$        $v : \text{Vec } (0 + n)$   
    |  $k' + 1 \rightarrow \text{drop } k \ n \ (\text{tail } (k' + n) \ v)$ 
```

One More Example

e.g., $\text{drop } (k = 2) (n = 3) [9, 8, 7, 6, 5] = [7, 6, 5]$

Can you spot the mistakes?

// Discard first k elements

$\text{drop} : \Pi(k : \text{Nat}). \Pi(n : \text{Nat}). \text{Vec } (k + n) \rightarrow \text{Vec } n =$

$\text{fix drop. } \lambda k. \lambda n. \lambda v.$

$\text{match } k \text{ with}$

$| 0 \rightarrow v \quad v : \text{Vec } (0 + n)$

$| k' + 1 \rightarrow \text{drop } k' n (\text{tail } (k' + n) v)$

$\text{tail } (k' + n) v$ does not have type $\text{Vec } (k + n)$

One More Example

e.g., $\text{drop } (k = 2) (n = 3) [9, 8, 7, 6, 5] = [7, 6, 5]$

All good now :)

// Discard first k elements

$\text{drop} : \Pi(k : \text{Nat}). \Pi(n : \text{Nat}). \text{Vec } (k + n) \rightarrow \text{Vec } n =$

$\text{fix drop. } \lambda k. \lambda n. \lambda v.$

$\text{match } k \text{ with}$

$| 0 \rightarrow v \quad v : \text{Vec } (0 + n)$

$| k' + 1 \rightarrow \text{drop } k' n (\text{tail } (k' + n) v)$

$v : \text{Vec } (k' + 1 + n), \text{tail } \dots : \text{Vec } (k' + n), \text{drop } \dots : \text{Vec } n$

Judgements

$$\Gamma \vdash s \Rightarrow A$$

(In context Γ , s synthesizes type A)

$$\Gamma \vdash t \Leftarrow A$$

(In context Γ , check that t has type A)

$$A_1 \equiv A_2$$

(Types are equivalent)

$$\ell_1 \equiv \ell_2$$

(Lengths are equivalent)

Rules: Checking

$$\boxed{\Gamma \vdash t \Leftarrow A}$$
$$\overline{\Gamma \vdash \text{nil} \Leftarrow \text{Vec } 0}$$

Rules: Checking

$$\Gamma \vdash t \Leftarrow A$$
$$\overline{\Gamma \vdash \text{nil} \Leftarrow \text{Vec } 0}$$

What about $\Gamma \vdash \text{nil} \Leftarrow \text{Vec } (0 + 0)$?

Rules: Checking

$$\boxed{\Gamma \vdash t \Leftarrow A}$$

$$\frac{\ell \equiv 0}{\Gamma \vdash \text{nil} \Leftarrow \text{Vec } \ell}$$

Rules: Checking

$$\Gamma \vdash t \Leftarrow A$$

$$\frac{\ell \equiv 0}{\Gamma \vdash \text{nil} \Leftarrow \text{Vec } \ell}$$

$$\frac{\begin{array}{c} \Gamma \vdash x \Longrightarrow \text{Nat} \quad \Gamma \vdash t_0 \Leftarrow A \\ \Gamma' = (\Gamma, y : \text{Nat}) \quad \Gamma' \vdash t_1 \Leftarrow A \end{array}}{\Gamma \vdash (\text{match } x \text{ with } | 0 \rightarrow t_0 | y + 1 \rightarrow t_1) \Leftarrow A}$$

Rules: Checking

$$\boxed{\Gamma \vdash t \Leftarrow A}$$

$$\frac{\ell \equiv 0}{\Gamma \vdash \text{nil} \Leftarrow \text{Vec } \ell}$$

$$\frac{\begin{array}{c} \Gamma \vdash x \Longrightarrow \text{Nat} \quad \Gamma \vdash t_0 \Leftarrow A \\ \Gamma' = (\Gamma, y : \text{Nat}) \quad \Gamma' \vdash t_1 \Leftarrow A \end{array}}{\Gamma \vdash (\text{match } x \text{ with } | 0 \rightarrow t_0 | y + 1 \rightarrow t_1) \Leftarrow A}$$

Rules: Checking

$$\boxed{\Gamma \vdash t \Leftarrow A}$$

$$\frac{\ell \equiv 0}{\Gamma \vdash \text{nil} \Leftarrow \text{Vec } \ell}$$

$$\frac{\begin{array}{c} \Gamma \vdash x \Longrightarrow \text{Nat} \quad \Gamma \vdash t_0 \Leftarrow A \\ \Gamma' = (\Gamma, y : \text{Nat}) \quad \Gamma' \vdash t_1 \Leftarrow A \end{array}}{\Gamma \vdash (\text{match } x \text{ with } | 0 \rightarrow t_0 | y + 1 \rightarrow t_1) \Leftarrow A}$$

Rules: Checking

$$\boxed{\Gamma \vdash t \Leftarrow A}$$

$$\frac{\ell \equiv 0}{\Gamma \vdash \text{nil} \Leftarrow \text{Vec } \ell}$$

$$\frac{\begin{array}{c} \Gamma \vdash x \Longrightarrow \text{Nat} \quad \Gamma \vdash t_0 \Leftarrow A \\ \Gamma' = (\Gamma, y : \text{Nat}) \quad \Gamma' \vdash t_1 \Leftarrow A \end{array}}{\Gamma \vdash (\text{match } x \text{ with } | 0 \rightarrow t_0 | y + 1 \rightarrow t_1) \Leftarrow A}$$

Rules: Checking

$$\boxed{\Gamma \vdash t \Leftarrow A}$$

$$\frac{\ell \equiv 0}{\Gamma \vdash \text{nil} \Leftarrow \text{Vec } \ell}$$

$$\frac{\begin{array}{c} \Gamma \vdash x \Rightarrow \text{Nat} \quad [0/x]\Gamma \vdash [0/x]t_0 \Leftarrow [0/x]A \\ \Gamma' = (\Gamma, y : \text{Nat}) \quad \Gamma' \vdash t_1 \Leftarrow A \end{array}}{\Gamma \vdash (\text{match } x \text{ with } | 0 \rightarrow t_0 | y + 1 \rightarrow t_1) \Leftarrow A}$$

Rules: Checking

$$\boxed{\Gamma \vdash t \Leftarrow A}$$

$$\frac{\ell \equiv 0}{\Gamma \vdash \text{nil} \Leftarrow \text{Vec } \ell}$$

$$\frac{\begin{array}{l} \Gamma \vdash x \Longrightarrow \text{Nat} \quad [0/x]\Gamma \vdash [0/x]t_0 \Leftarrow [0/x]A \\ \Gamma' = (\Gamma, y : \text{Nat}) \quad [y + 1/x]\Gamma' \vdash [y + 1/x]t_1 \Leftarrow [y + 1/x]A \end{array}}{\Gamma \vdash (\text{match } x \text{ with } | 0 \rightarrow t_0 \mid y + 1 \rightarrow t_1) \Leftarrow A}$$

Rules: Checking

$$\frac{\Gamma \vdash x \implies \text{Nat} \quad [0/x]\Gamma \vdash [0/x]t_0 \Leftarrow [0/x]A \quad \Gamma' = (\Gamma, y : \text{Nat}) \quad [y+1/x]\Gamma' \vdash [y+1/x]t_1 \Leftarrow [y+1/x]A}{\Gamma \vdash (\text{match } x \text{ with } | 0 \rightarrow t_0 \mid y+1 \rightarrow t_1) \Leftarrow A}$$

```
drop : Π(k : Nat).Π(n : Nat).Vec (k + n) → Vec n =
  fix drop.λk.λn.λv.
    match k with
    | 0 → v      v : Vec (0 + n)
    | k' + 1 → drop k' n (tail (k' + n) v)
v : Vec (k' + 1 + n), tail ... : Vec (k' + n)
```

Rules: Equality

$$A_1 \equiv A_2$$

$$\begin{cases} \text{Nat} \equiv \text{Nat} \\ \Pi(x : A_1). B_1 \equiv \Pi(x : A_2). B_2 & \text{if } A_1 \equiv A_2 \text{ and } B_1 \equiv B_2 \\ \text{Vec } \ell_1 \equiv \text{Vec } \ell_2 & \text{if } \ell_1 \equiv \ell_2 \end{cases}$$

$$\ell_1 \equiv \ell_2$$

- Normalize and compare syntactically.

$$\text{Vec } (\lambda x. \text{Vec } (\lambda y. x)) \rightarrow \text{Vec } (\lambda x. x)$$

$$\text{Vec } (\lambda x. \text{Vec } (\lambda y. x)) \rightarrow \text{Vec } (\lambda x. x)$$

Rules: Equality

$$A_1 \equiv A_2$$

$$\begin{cases} \text{Nat} \equiv \text{Nat} \\ \Pi(x : A_1). B_1 \equiv \Pi(x : A_2). B_2 & \text{if } A_1 \equiv A_2 \text{ and } B_1 \equiv B_2 \\ \text{Vec } \ell_1 \equiv \text{Vec } \ell_2 & \text{if } \ell_1 \equiv \ell_2 \end{cases}$$

$$\ell_1 \equiv \ell_2$$

- Normalize and compare syntactically.
 - $1 + n + (2 + m) \longrightarrow 3 + n + m$
 - $(n + 0) + 3 + m \longrightarrow 3 + n + m$

Outline

Basic Language

Better Pattern Matching

Extending the Language

Conclusion

An Unfortunate Restriction

- Subject of `match` must be a variable `x` or synthesize `Vec x`
 - Needed for substitution
- This is why `head` and `tail` must be built-in!

`tail` : $\Pi(n : \text{Nat}). \text{Vec } (n + 1) \rightarrow \text{Vec } n =$
 $\lambda n. \lambda v.$

`match` v with $v : \text{Vec } (n + 1)$
| `cons` $n' \ x \ v' \rightarrow v'$

A Better Approach

```
tail :  $\Pi(n: \text{Nat}). \text{Vec } (n + 1) \rightarrow \text{Vec } n =$   
       $\lambda n. \lambda v.$   
        match v with  
        | cons  $n' \_ v' \rightarrow v'$ 
```

A Better Approach

`tail : $\Pi(n: \text{Nat}). \text{Vec } (n + 1) \rightarrow \text{Vec } n =$
 $\lambda n. \lambda v.$`

`match v with`

`| cons n' _ v' $\rightarrow v'$`

Is the cons branch correct? (i.e., $\text{Vec } n \equiv \text{Vec } n'$?)

$\forall n, n' \in \mathbb{N}. (n' + 1 = n + 1) \implies n = n' \quad \checkmark$

A Better Approach

`tail` : $\Pi(n : \text{Nat}). \text{Vec } (n + 1) \rightarrow \text{Vec } n =$
 $\lambda n. \lambda v.$

`match v with`

`| cons n' _ v' → v'`

Is the `cons` branch correct? (i.e., $\text{Vec } n \equiv \text{Vec } n'$?)

$$\forall n, n' \in \mathbb{N}. (n' + 1 = n + 1) \implies n = n' \quad \checkmark$$

Is the `nil` branch unreachable? (i.e., $\text{Vec } (n + 1) \not\equiv \text{Vec } 0$?)

$$\neg (\exists n \in \mathbb{N}. n + 1 = 0) \quad \checkmark$$

In General

- Add a new context $\Delta ::= \cdot \mid \Delta, \ell_1 = \ell_2$
 - When entering a `match`, extend Δ
 - $\ell \equiv \ell'$: make a \forall proposition
 - Reachability : make a \exists proposition

$$\Gamma \mid \Delta \vdash s \Longrightarrow A$$

$$\Gamma \mid \Delta \vdash t \Longleftarrow A$$

$$\Delta \vdash A_1 \equiv A_2$$

$$\Delta \vdash \ell_1 \equiv \ell_2$$

- To decide whether propositions are true, use Z3 solver ²
- Propositions in Presburger arithmetic are always decidable! ³

²<https://github.com/Z3Prover/z3>

³Presburger 1929

In General

- Add a new context $\Delta ::= \cdot \mid \Delta, \ell_1 = \ell_2$
 - When entering a match, extend Δ
 - $\ell \equiv \ell'$: make a \forall proposition
 - Reachability : make a \exists proposition

$$\Gamma \mid \Delta \vdash s \Longrightarrow A$$

$$\Gamma \mid \Delta \vdash t \Longleftarrow A$$

$$\Delta \vdash A_1 \equiv A_2$$

$$\Delta \vdash \ell_1 \equiv \ell_2$$

- To decide whether propositions are true, use Z3 solver ²
- Propositions in Presburger arithmetic are always decidable! ³

²<https://github.com/Z3Prover/z3>

³Presburger 1929

Outline

Basic Language

Better Pattern Matching

Extending the Language

Conclusion

Booleans

- Updated syntax:

Type $A ::= \dots \mid \text{Bool}$

Chk $t ::= \dots$
 $\mid \text{true} \mid \text{false}$
 $\mid (\text{match } s \text{ with } \mid \text{true} \rightarrow t_1 \mid \text{false} \rightarrow t_2)$

- New typing rules:

$$\frac{}{\Gamma \mid \Delta \vdash \text{true} \Leftarrow \text{Bool}} \quad \frac{}{\Gamma \mid \Delta \vdash \text{false} \Leftarrow \text{Bool}}$$

$$\frac{\Gamma \mid \Delta \vdash s \Longrightarrow \text{Bool} \quad \Gamma \mid \Delta \vdash t_1 \Leftarrow A \quad \Gamma \mid \Delta \vdash t_2 \Leftarrow A}{\Gamma \mid \Delta \vdash (\text{match } s \text{ with } \mid \text{true} \rightarrow t_1 \mid \text{false} \rightarrow t_2) \Leftarrow A}$$

- New case for type equality : $\Delta \vdash \text{Bool} \equiv \text{Bool}$
- All pretty straightforward!

Booleans

- Updated syntax:

Type $A ::= \dots \mid \text{Bool}$

Chk $t ::= \dots$
 $\mid \text{true} \mid \text{false}$
 $\mid (\text{match } s \text{ with } \mid \text{true} \rightarrow t_1 \mid \text{false} \rightarrow t_2)$

- New typing rules:

$$\frac{}{\Gamma \mid \Delta \vdash \text{true} \Leftarrow \text{Bool}} \quad \frac{}{\Gamma \mid \Delta \vdash \text{false} \Leftarrow \text{Bool}}$$

$$\frac{\Gamma \mid \Delta \vdash s \Rightarrow \text{Bool} \quad \Gamma \mid \Delta \vdash t_1 \Leftarrow A \quad \Gamma \mid \Delta \vdash t_2 \Leftarrow A}{\Gamma \mid \Delta \vdash (\text{match } s \text{ with } \mid \text{true} \rightarrow t_1 \mid \text{false} \rightarrow t_2) \Leftarrow A}$$

- New case for type equality : $\Delta \vdash \text{Bool} \equiv \text{Bool}$
- All pretty straightforward!

Outline

Basic Language

Better Pattern Matching

Extending the Language

Conclusion

Limitations

- Vectors can only contain natural numbers
 - Can add polymorphism (using Hindley-Milner)
- Impossible to implement `filter`—what should the output length be?
 - Can do this using *dependent pairs*
- Explicit parameters are not ergonomic
 - Could add implicit parameters (as in Agda)
- Only addition allowed in lengths
 - Cannot implement `split`, `join`, etc.

Conclusion

- You can check simple constraints on list sizes using dependent types!
 - Enforce constraints at compile time
 - Catch some (but obviously not all) common typos
- How to check equivalence of lengths? How to prevent infinite loops in type checker?
 - Presburger arithmetic + Z3

Polymorphism

We implement the Hindley-Milner algorithm.

Type scheme	\hat{A}	$::=$	\dots	$ \forall \alpha_1, \alpha_2, \dots, \alpha_N. A$
Type	A	$::=$	\dots	$ \alpha$
Context	Γ	$::=$	\cdot	$ \Gamma, x : \hat{A}$
Type subst.	σ	$::=$	\cdot	$ \sigma, A/\alpha$

$$\frac{\Gamma(x) = \hat{A} \quad \Gamma \vdash \hat{A} \sqsubseteq A}{\Gamma \vdash x \Rightarrow A}$$

$$\frac{\Gamma, x : \forall _ . A \vdash t \Leftarrow B}{\Gamma \vdash \lambda x. t \Leftarrow \Pi(x : A). B}$$

- Generalization: $\alpha \rightarrow \alpha$ generalizes to $\forall \alpha. \alpha \rightarrow \alpha$.
- Instantiation: $\forall \alpha. \alpha \rightarrow \alpha$ can be instantiated to $\text{Nat} \rightarrow \text{Nat}$.
- Unification: $\text{unify}(\alpha \rightarrow \alpha, \text{Nat} \rightarrow \text{Nat}) = \{\text{Nat}/\alpha\}$

Polymorphism

We implement the Hindley-Milner algorithm.

Type scheme	\hat{A}	$::=$	\dots	$ \forall \alpha_1, \alpha_2, \dots, \alpha_N. A$
Type	A	$::=$	\dots	$ \alpha$
Context	Γ	$::=$	\cdot	$ \Gamma, x : \hat{A}$
Type subst.	σ	$::=$	\cdot	$ \sigma, A/\alpha$

$$\frac{\Gamma(x) = \hat{A} \quad \Gamma \vdash \hat{A} \sqsubseteq A}{\Gamma \vdash x \Longrightarrow A} \qquad \frac{\Gamma, x : \forall _ . A \vdash t \Leftarrow B}{\Gamma \vdash \lambda x. t \Leftarrow \Pi(x : A). B}$$

- Generalization: $\alpha \rightarrow \alpha$ generalizes to $\forall \alpha. \alpha \rightarrow \alpha$.
- Instantiation: $\forall \alpha. \alpha \rightarrow \alpha$ can be instantiated to $\text{Nat} \rightarrow \text{Nat}$.
- Unification: $\text{unify}(\alpha \rightarrow \alpha, \text{Nat} \rightarrow \text{Nat}) = \{\text{Nat}/\alpha\}$