Hazelnut: A Bidirectionally Typed Structure Editor Calculus

Cyrus Omar

Ian Voysey Michael Hilton Jonathan Aldrich Matthew Hammer Carnegie Mellon University Carnegie Mellon University Oregon State University Carnegie Mellon University University of Colorado Boulder

POPL 2017

What is it that program editors reason about?

```
fun summary_stats(m : matrix) =
  { mean = stats.mean(m, ColumnWise),
    std = stats.std(m,
    median =
```

syntactically malformed program text

```
fun summary_stats(m : matrix) =
  { mean = stats.mean(m, ColumnWise),
    std = stats.std(m, □),
    median = □ }
```

syntactically malformed program text \rightarrow term with holes

[Kats et al., OOPSLA 2009; Amorim et al., SLE 2016]

```
fun summary_stats(m : matrix) =
  { mean = stats.mean(m, ColumnWise),
    std = stats.std(m, □),
    median = □ }
```

syntactically malformed program text → term with holes

[Teitelbaum and Reps, Comm. ACM 1981; many others since]

```
fun summary_stats(m : matrix) =
  { mean = stats.mean(m, ColumnWise),
    std = stats.std(m, □),
    median = □ }
```

Below to **reason statically** about terms with holes?

What type is synthesized for the function as a whole?

```
fun summary_stats(m : matrix) =
  { mean = stats.mean(m, ColumnWise),
    std = stats.std(m, □),
    median = □ }
```

How to **reason statically** about terms with holes?

```
matrix \rightarrow
                                                       { mean : vec,
What type is synthesized for the function as a whole?
    fun summary_stats(m : matrix) =
       { mean = stats.mean(m, ColumnWise),
         std = stats.std(m, \Box),
         median = \Box }
```

std : vec, median : \Box }

How to **reason statically** about terms with type errors?

What type is synthesized for the function as a whole?

```
fun summary_stats(m : matrix) =
  { mean = stats.mean(m, ColumnWise),
    std = stats.std(m, "oops"),
    median = □ }
```

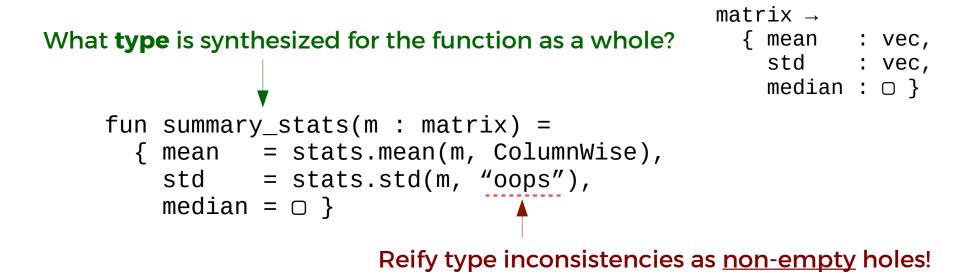
How to **reason statically** about terms with type errors?

What type is synthesized for the function as a whole?

```
fun summary_stats(m : matrix) =
  { mean = stats.mean(m, ColumnWise),
   std = stats.std(m, "oops"),
   median = □ }
```

Reify type inconsistencies as <u>non-empty</u> holes!

How to reason statically about terms with type errors?



 $\dot{\Gamma} \vdash \dot{e} \Rightarrow \dot{\tau} \mid \dot{e} \text{ synthesizes } \dot{\tau}$

 $\dot{\Gamma} \vdash \dot{e} \Leftarrow \dot{\tau}$ \dot{e} analyzes against $\dot{\tau}$

 $\dot{\Gamma} \vdash \dot{e} \Rightarrow \dot{\tau} \quad \dot{e} \text{ synthesizes } \dot{\tau} \qquad \qquad \dot{\Gamma} \vdash \dot{e} \Leftarrow \dot{\tau} \quad \dot{e} \text{ analyzes against } \dot{\tau}$

$$\begin{split} \overline{\dot{\Gamma} \vdash (|\rangle} \Rightarrow (|\rangle) \\ \\ \frac{\dot{\Gamma} \vdash \dot{e} \Rightarrow \dot{\tau}}{\dot{\Gamma} \vdash (|\dot{e}|\rangle \Rightarrow (|\rangle)} \end{split}$$

...

 $\begin{array}{cccc} \dot{\Gamma} \vdash \dot{e} \Rightarrow \dot{\tau} & \dot{e} \text{ synthesizes } \dot{\tau} & & \dot{\Gamma} \vdash \dot{e} \Leftrightarrow \dot{\tau} & \dot{e} \text{ analyzes against } \dot{\tau} \\ & & & & & & & \\ \hline \dot{\Gamma} \vdash \langle \! \! & \rangle \rangle \Rightarrow \langle \! \! & \rangle \\ \hline \dot{\Gamma} \vdash \dot{e} \Rightarrow \dot{\tau} & & & \\ \hline \dot{\Gamma} \vdash \langle \! & \rangle \Rightarrow \langle \! \! & \rangle \end{array} & \begin{array}{cccc} \dot{\Gamma} \vdash \dot{e} \Leftrightarrow \dot{\tau} & & & \\ \hline \dot{\Gamma} \vdash \dot{e} \Leftrightarrow \dot{\tau} & & & \\ \hline \dot{\Gamma} \vdash \langle \! & \rangle \Rightarrow \langle \! & \rangle \end{array}$

$$\begin{split} & \overset{\cdots}{\bar{\Gamma} \vdash \langle \! | \! \rangle } \Rightarrow \langle \! | \! \rangle \\ & \frac{\dot{\Gamma} \vdash \dot{e} \Rightarrow \dot{\tau}}{\dot{\Gamma} \vdash \langle \! | \dot{e} \! \rangle \Rightarrow \langle \! | \! \rangle } \end{split}$$

 $\dot{\Gamma} \vdash \dot{e} \Rightarrow \dot{\tau} \dot{e}$ synthesizes $\dot{\tau}$

$$\begin{split} \vec{\Gamma} \vdash \vec{e} &\Leftarrow \vec{\tau} \quad \vec{t} \text{ analyzes against } \vec{\tau} \\ & \cdots \\ \vec{\Gamma} \vdash \vec{e} \Rightarrow \vec{\tau}' \quad \vec{\tau} \sim \vec{\tau}' \\ \hline \vec{\Gamma} \vdash \vec{e} &\Leftarrow \vec{\tau} \\ \end{split}$$
$$\begin{aligned} \vec{\tau} \sim \vec{\tau}' \quad \vec{\tau} \text{ and } \vec{\tau}' \text{ are consistent} \\ \hline \vec{\psi} \sim \vec{\tau} \quad \vec{\tau} \sim \vec{\psi} \quad \vec{\tau} \sim \vec{\tau} \quad \frac{\dot{\tau}_1 \sim \dot{\tau}_1' \quad \dot{\tau}_2 \sim \dot{\tau}_2'}{(\dot{\tau}_1 \rightarrow \dot{\tau}_2) \sim (\dot{\tau}_1' \rightarrow \dot{\tau}_2')} \end{split}$$

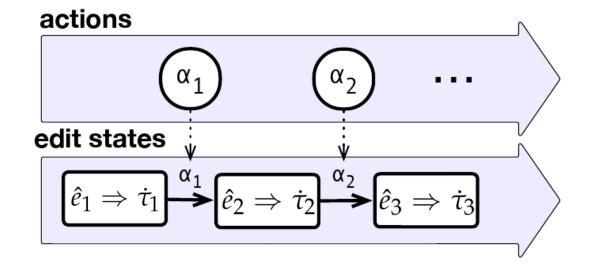
$$\begin{split} \overline{\dot{\Gamma} \vdash (\!\!\!|) \Rightarrow (\!\!\!|)} \\ \frac{\dot{\Gamma} \vdash \dot{e} \Rightarrow \dot{\tau}}{\dot{\Gamma} \vdash (\!\!\!|\dot{e}|\!\!\!) \Rightarrow (\!\!\!|)} \end{split}$$

 $\dot{\Gamma} \vdash \dot{e} \Rightarrow \dot{\tau} \mid \dot{e} \text{ synthesizes } \dot{\tau}$

...

$$\begin{split} \hline \dot{\Gamma} \vdash \dot{e} &\Leftarrow \dot{\tau} \\ \hline \dot{\Gamma} \vdash \dot{e} &\Leftarrow \dot{\tau} \\ & \vdots \\ \hline \vdots \\ \hline \dot{\Gamma} \vdash \dot{e} &\Rightarrow \dot{\tau}' & \dot{\tau} \sim \dot{\tau}' \\ \hline \dot{\Gamma} \vdash \dot{e} &\Leftarrow \dot{\tau} \\ \hline \dot{\tau} \vdash \dot{e} &\Leftarrow \dot{\tau} \\ \hline \hline \dot{\tau} \vdash \dot{e} &\Leftarrow \dot{\tau} \\ \hline \hline \dot{\tau} \sim \dot{\tau}' & \dot{\tau} \text{ and } \dot{\tau}' \text{ are consistent} \\ \hline \hline \hline \hline \psi \sim \dot{\tau} & \overline{\dot{\tau} \sim \psi} & \overline{\dot{\tau} \sim \dot{\tau}} & \frac{\dot{\tau}_1 \sim \dot{\tau}_1' & \dot{\tau}_2 \sim \dot{\tau}_2'}{(\dot{\tau}_1 \rightarrow \dot{\tau}_2) \sim (\dot{\tau}_1' \rightarrow \dot{\tau}_2')} \end{split}$$

coincides with **gradual typing** [Siek and Taha, 2006]



See http://hazelgrove.org/

 $\begin{array}{l} \mathsf{ZTyp} \ \hat{\tau} ::= \ \triangleright \dot{\tau} \triangleleft \mid (\hat{\tau} \to \dot{\tau}) \mid (\dot{\tau} \to \hat{\tau}) \\ \mathsf{ZExp} \ \hat{e} ::= \ \triangleright \dot{e} \triangleleft \mid (\lambda x. \hat{e}) \mid \hat{e}(\dot{e}) \mid \dot{e}(\hat{e}) \mid (\hat{e} + \dot{e}) \mid (\dot{e} + \hat{e}) \\ \mid \ \hat{e} : \dot{\tau} \mid \dot{e} : \hat{\tau} \mid \langle \hat{e} \rangle \end{array}$

$$\dot{\Gamma} \vdash \hat{e} \stackrel{\alpha}{\longrightarrow} \hat{e}' \Leftarrow \dot{\tau}$$

$$\dot{\Gamma} \vdash \hat{e} \Rightarrow \dot{\tau} \stackrel{\alpha}{\longrightarrow} \hat{e}' \Rightarrow \dot{\tau}'$$

$$\dot{\Gamma} \vdash \hat{e} \xrightarrow{\alpha} \hat{e}' \Leftarrow \dot{\tau}$$

$$\dot{\Gamma} \vdash \hat{e} \Rightarrow \dot{\tau} \stackrel{\alpha}{\longrightarrow} \hat{e}' \Rightarrow \dot{\tau}'$$

$$\dot{\Gamma} \vdash {\rhd} ({|\hspace{-.06cm}|\hspace{-.06cm}|} {\triangleleft} \Rightarrow ({|\hspace{-.06cm}|\hspace{-.06cm}|} \xrightarrow{\texttt{construct lit } n} {\succ} \underline{n} {\triangleleft} \Rightarrow \texttt{num}$$

$$\dot{\Gamma} \vdash \hat{e} \stackrel{\alpha}{\longrightarrow} \hat{e}' \Leftarrow \dot{\tau}$$

 $\dot{\Gamma} \vdash \hat{e} \Rightarrow \dot{\tau} \stackrel{\alpha}{\longrightarrow} \hat{e}' \Rightarrow \dot{\tau}'$

$$\begin{split} \dot{\Gamma} \vdash \triangleright () \triangleleft \Rightarrow () \xrightarrow{\text{construct lit } n} \triangleright \underline{n} \triangleleft \Rightarrow \texttt{num} \\ \\ \dot{\Gamma} \vdash \hat{e} \xrightarrow{\alpha} \hat{e}' \Leftarrow \texttt{num} \\ \\ \hline{\dot{\Gamma} \vdash (\hat{e} + \dot{e}) \Rightarrow \texttt{num}} \xrightarrow{\alpha} (\hat{e}' + \dot{e}) \Rightarrow \texttt{num}} \end{split}$$

$$\dot{\Gamma} \vdash \hat{e} \Rightarrow \dot{\tau} \stackrel{\alpha}{\longrightarrow} \hat{e}' \Rightarrow \dot{\tau}'$$

$$\begin{split} \underline{\dot{\Gamma} \vdash \hat{e} \xrightarrow{\alpha} \hat{e}' \Leftarrow \dot{\tau}} \\ \underline{\dot{\Gamma} \vdash \hat{e}^{\diamond} \Rightarrow \dot{\tau}' \qquad \dot{\Gamma} \vdash \hat{e} \Rightarrow \dot{\tau}' \xrightarrow{\alpha} \hat{e}' \Rightarrow \dot{\tau}'' \qquad \dot{\tau} \sim \dot{\tau}''} \\ \underline{\dot{\Gamma} \vdash \hat{e} \xrightarrow{\alpha} \hat{e}' \Leftarrow \dot{\tau}} \end{split}$$

$$\begin{split} \dot{\Gamma} \vdash \hat{e} \xrightarrow{\alpha} \hat{e}' \Leftarrow \texttt{num} \\ \dot{\Gamma} \vdash (\hat{e} + \dot{e}) \Rightarrow \texttt{num} \xrightarrow{\alpha} (\hat{e}' + \dot{e}) \Rightarrow \texttt{num} \end{split}$$

 $\dot{\Gamma} \vdash \triangleright () \triangleleft \Rightarrow ()) \xrightarrow{\text{construct lit } n} \triangleright n \triangleleft \Rightarrow \text{num}$

$$\begin{aligned} \mathsf{ZTyp} \ \hat{\tau} &::= | \flat \dot{\tau} \triangleleft | (\hat{\tau} \to \dot{\tau}) | (\dot{\tau} \to \hat{\tau}) \\ \mathsf{ZExp} \ \hat{e} &::= | \flat \dot{e} \triangleleft | (\lambda x. \hat{e}) | \hat{e}(\dot{e}) | \dot{e}(\hat{e}) | (\hat{e} + \dot{e}) | (\dot{e} + \hat{e}) \\ & | \hat{e} : \dot{\tau} | \dot{e} : \hat{\tau} | (\hat{e}) \end{aligned}$$

Action $\alpha ::= \text{move } \delta \mid \text{construct } \psi \mid \text{del} \mid \text{finish}$ Dir δ ::= child $n \mid$ parent Shape ψ ::= arrow | num $\operatorname{asc} |\operatorname{var} x| \operatorname{lam} x| \operatorname{ap} |\operatorname{lit} n| \operatorname{plus}$

$$\dot{\Gamma} \vdash \hat{e} \Rightarrow \dot{\tau} \xrightarrow{\alpha} \hat{e}' \Rightarrow \dot{\tau}'$$

 $\dot{\Gamma} \vdash (\hat{e} + \dot{e}) \Rightarrow \texttt{num} \xrightarrow{\alpha} (\hat{e}' + \dot{e}) \Rightarrow \texttt{num}$

 $\dot{\Gamma} \vdash \triangleright () \triangleleft \xrightarrow{\text{construct lit } n} (\triangleright n \triangleleft) \Leftarrow \dot{\tau}$

Every edit action leaves the edit state well-typed.

Theorem 1 (Action Sensibility). 1. If $\dot{\Gamma} \vdash \hat{e}^{\diamond} \Rightarrow \dot{\tau}$ and $\dot{\Gamma} \vdash \hat{e} \Rightarrow \dot{\tau} \xrightarrow{\alpha} \hat{e}' \Rightarrow \dot{\tau}'$ then $\dot{\Gamma} \vdash \hat{e}'^{\diamond} \Rightarrow \dot{\tau}'.$ 2. If $\dot{\Gamma} \vdash \hat{e}^{\diamond} \Leftarrow \dot{\tau}$ and $\dot{\Gamma} \vdash \hat{e} \xrightarrow{\alpha} \hat{e}' \Leftarrow \dot{\tau}$ then $\dot{\Gamma} \vdash \hat{e}'^{\diamond} \Leftarrow \dot{\tau}.$

The cursor can reach any position in the program.

Theorem 3 (Reachability).
1. If τ̂[◊] = τ̂'[◊] then there exists some ā such that ā movements and τ̂ → * τ̂'.
2. If Γ΄ ⊢ ê[◊] ⇒ τ and ê[◊] = ê'[◊] then there exists some ā such that ā movements and Γ΄ ⊢ ê ⇒ τ̂ → * ê' ⇒ τ̂.
3. If Γ΄ ⊢ ê[◊] ⇐ τ̂ and ê[◊] = ê'[◊] then there exists some ā such that ā movements and Γ΄ ⊢ ê → * ê' ⇐ τ̂.

Any well-typed expression can be constructed using edit actions.

Theorem 6 (Constructability).

1. For every $\dot{\tau}$ there exists $\bar{\alpha}$ such that $\triangleright (|) \triangleleft \xrightarrow{\bar{\alpha}} * \triangleright \dot{\tau} \triangleleft$. 2. If $\dot{\Gamma} \vdash \dot{e} \Rightarrow \dot{\tau}$ then there exists $\bar{\alpha}$ such that:

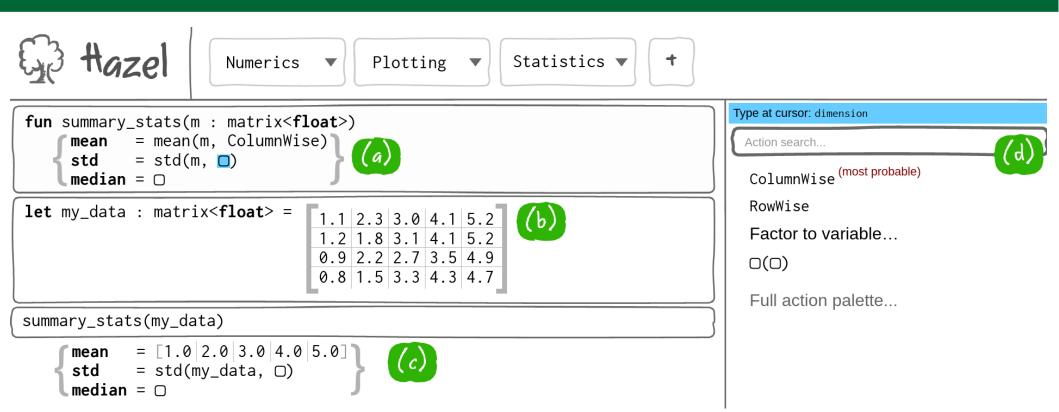
$$\dot{\Gamma} \vdash \triangleright () \triangleleft \Rightarrow () \xrightarrow{\bar{\alpha}} {}^* \triangleright \dot{e} \triangleleft \Rightarrow \dot{\tau}$$

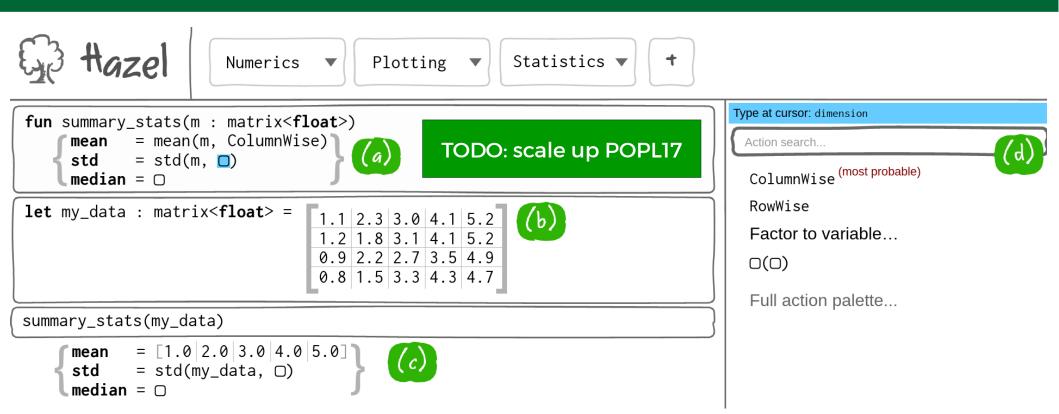
3. If $\dot{\Gamma} \vdash \dot{e} \Leftarrow \dot{\tau}$ *then there exists* $\bar{\alpha}$ *such that:*

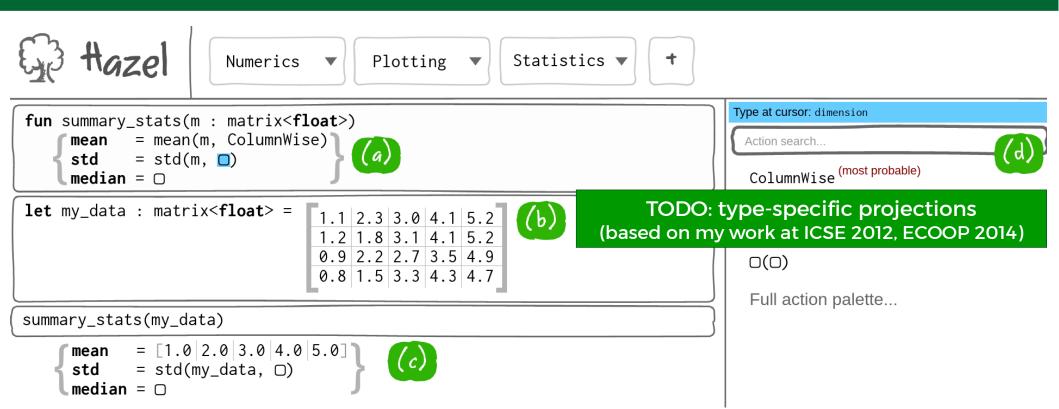
$$\dot{\Gamma} \vdash \operatorname{dim} \bar{\alpha} \xrightarrow{\bar{\alpha}} {}^* \operatorname{dim} \bar{\tau}$$

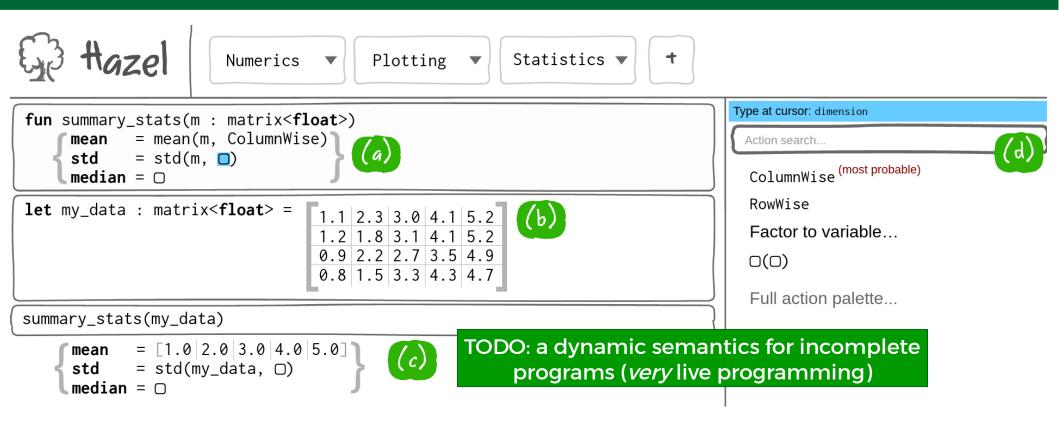
Summary: Hazelnut

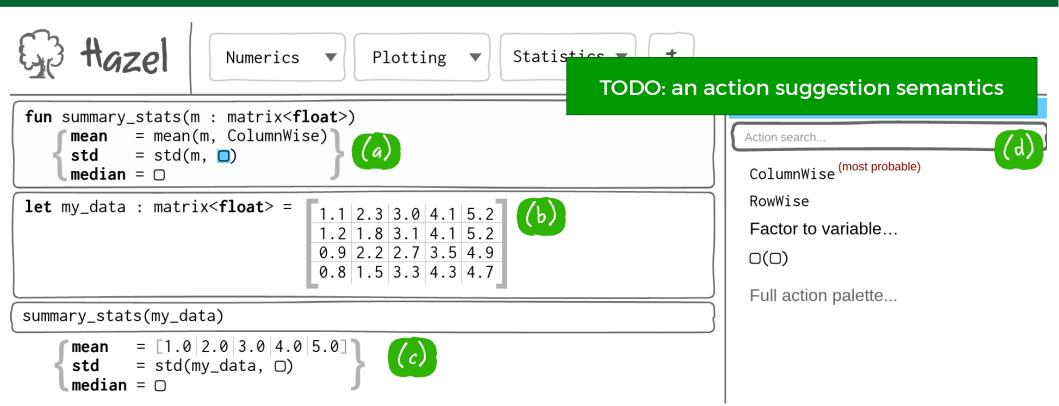
- A static semantics for terms with holes and type inconsistencies.
- An typed action semantics that maintains sensibility invariant.
 - **HZ**: A reference implementation written in OCaml React + js_of_ocaml.
- A rich metatheory that establishes the correctness of Hazelnut.
 - Mechanized using the **Agda** proof assistant.
 - Guides the definition of an extension (sum types see paper!)

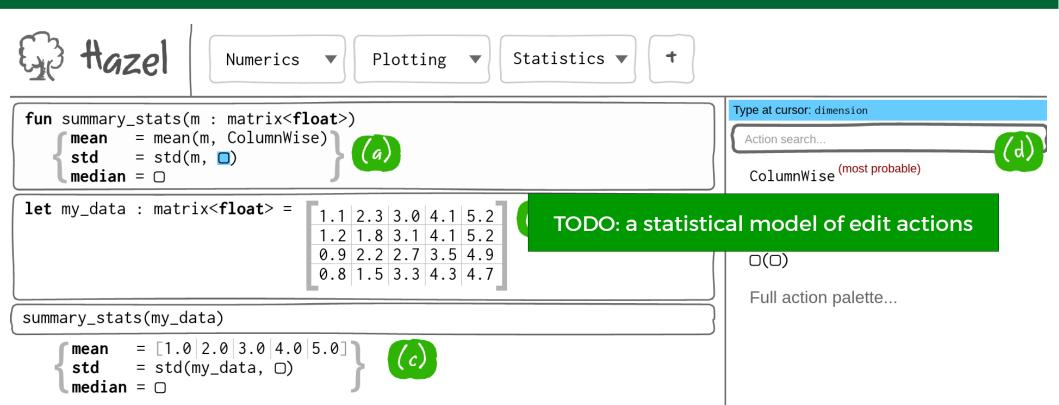


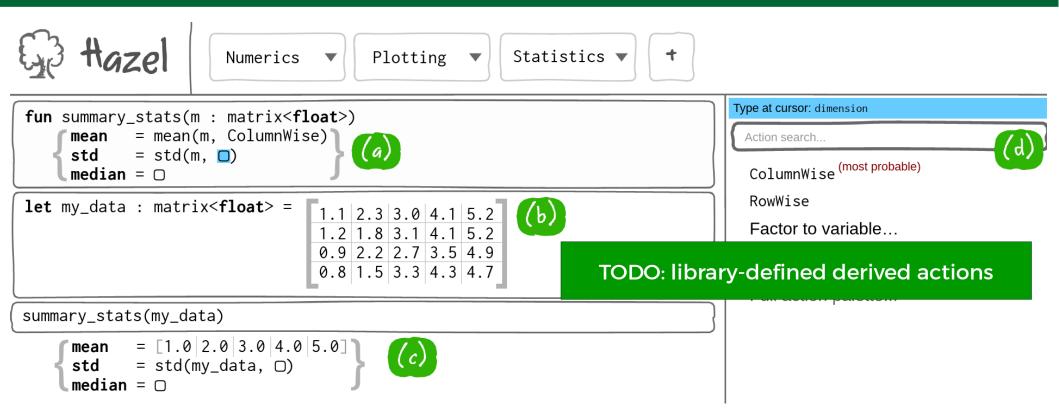


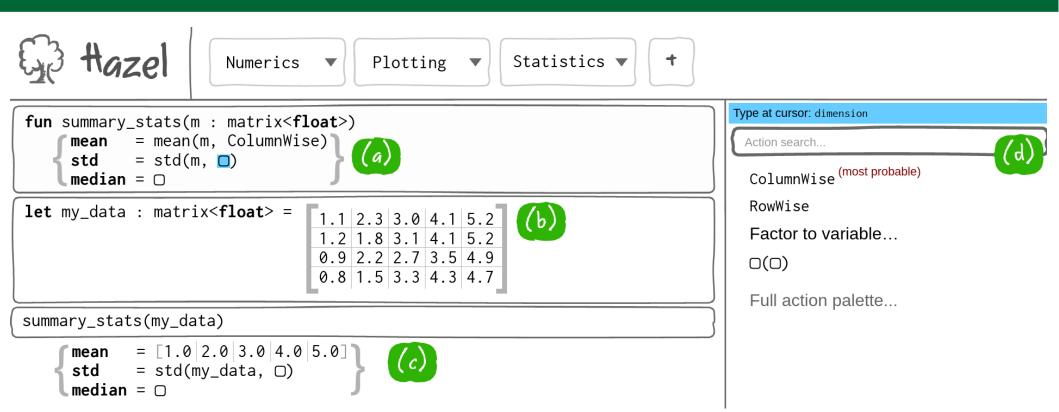


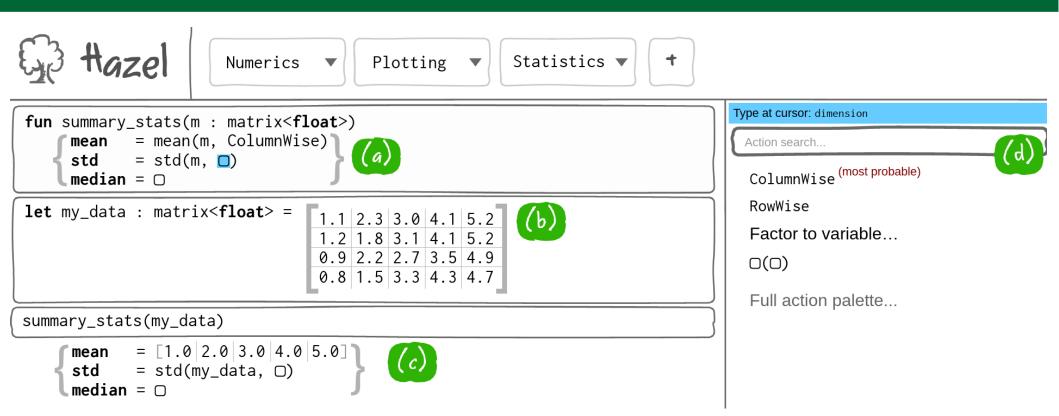












See http://www.hazelgrove.org/.

Thanks!