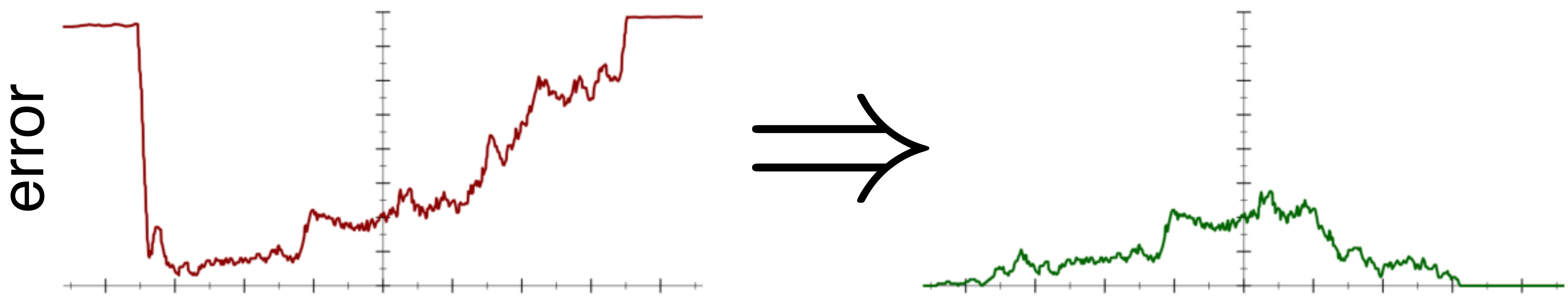
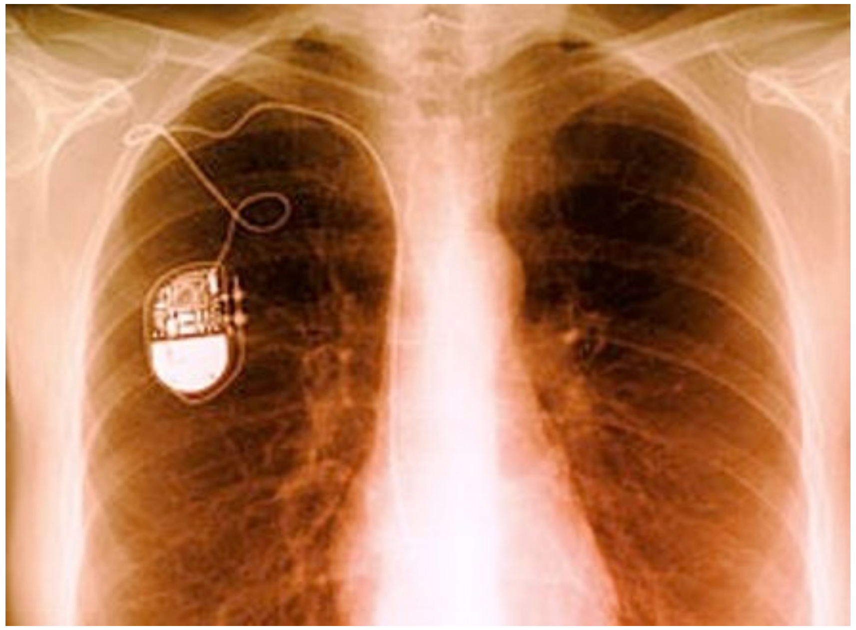
Automatically Improving Accuracy for Floating Point Expressions



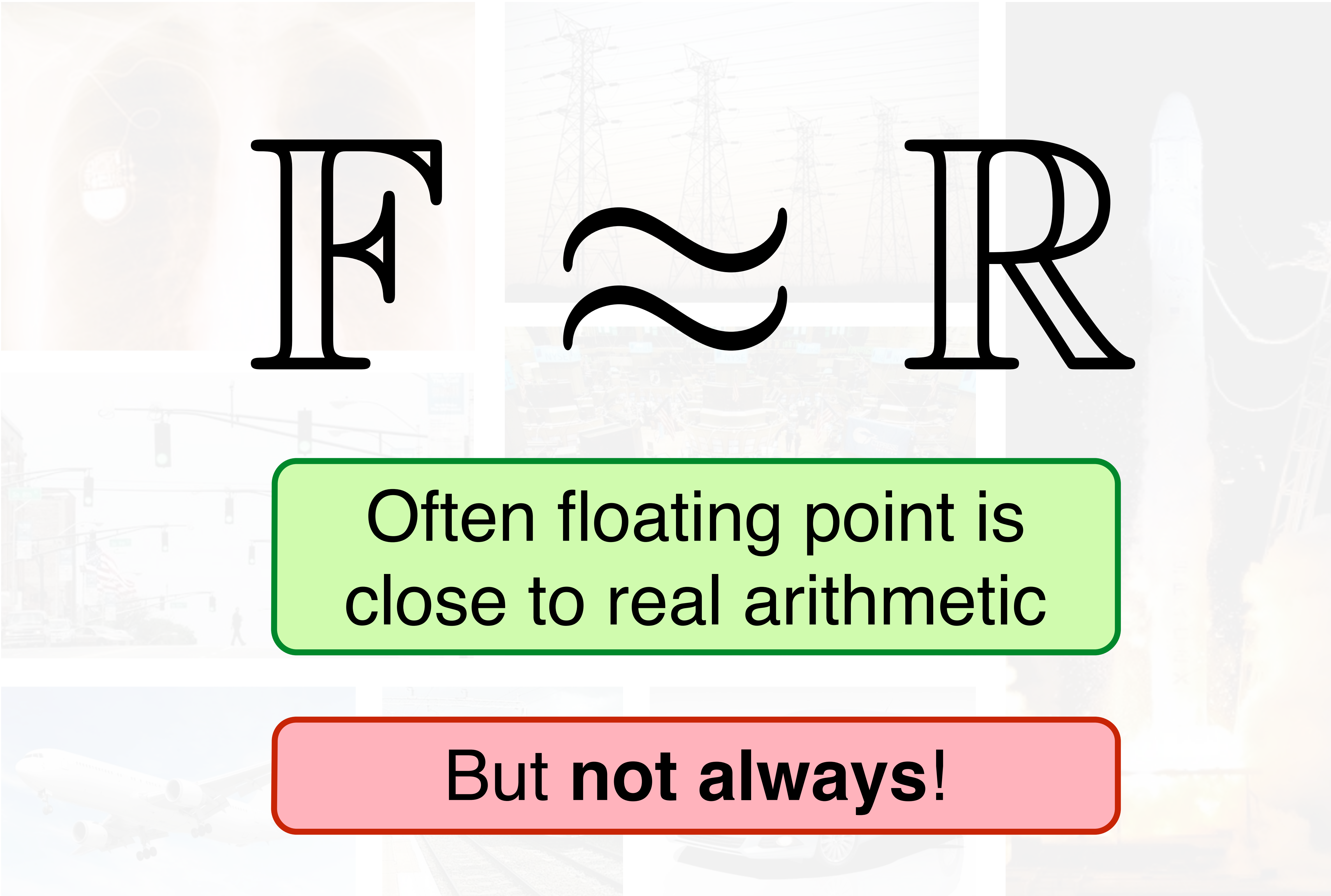



Pavel Panchekha
Alex Sanchez-Stern
James Wilcox
Zach Tatlock

Floating Point's Wild Success



Floating Point's Wild Success



$$\mathbb{F} \approx \mathbb{R}$$

Often floating point is close to real arithmetic

But **not always**!

Floating Point's Wild Success

But not always!

Numerous articles retracted *[Altman '99, '03]*

Financial regulations *[Euro '98]*

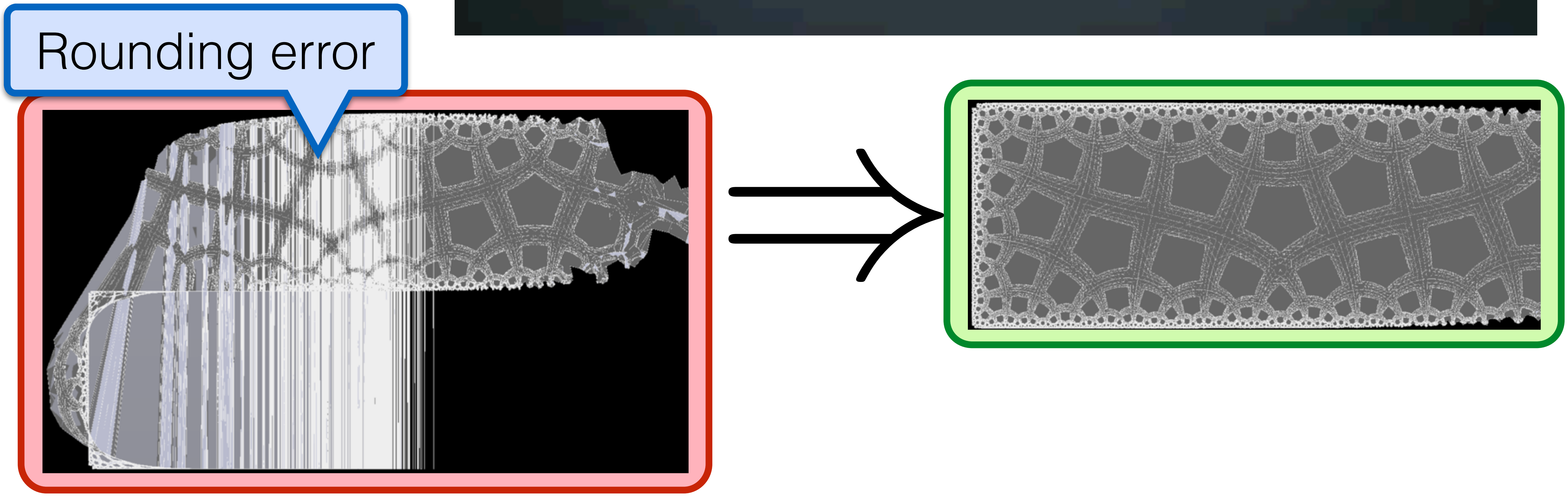
Market distortions *[McCullough '99, Quinn '83]*

Rounding Error in Sculpture

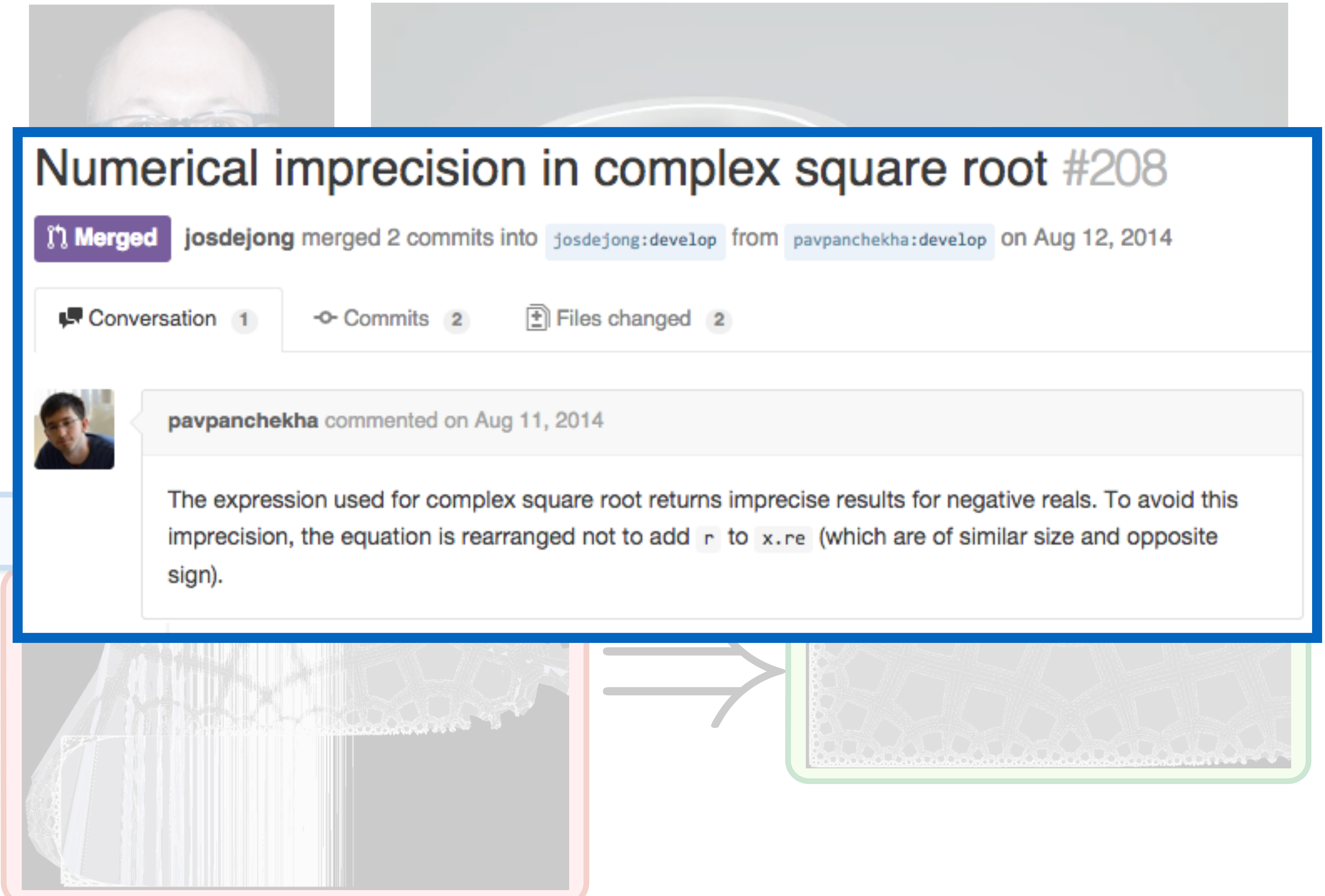
Blake Courter
@bcourter

Rounding Error in Sculpture

Blake Courter
@bcourter



Rounding Error in Sculpture

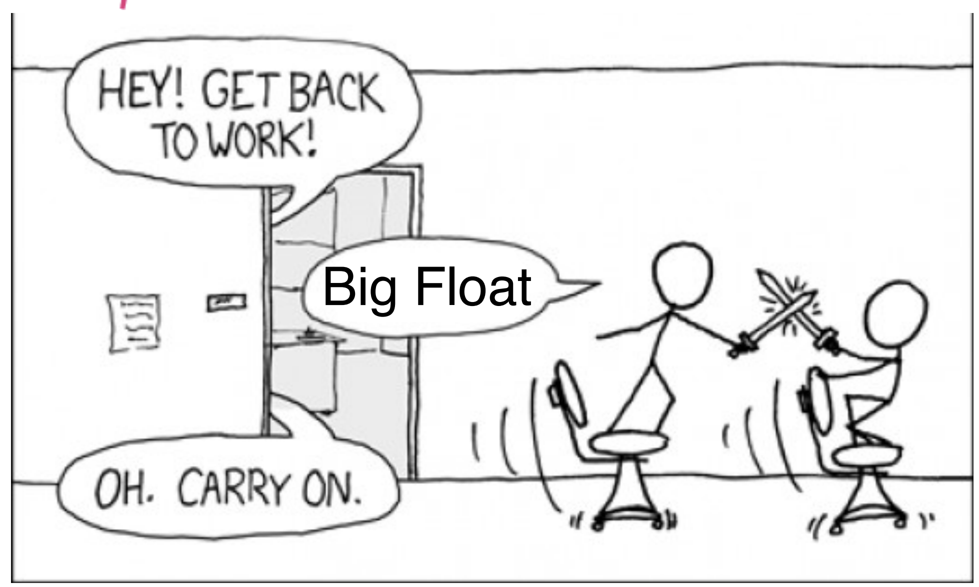


Existing options

- Unreliable

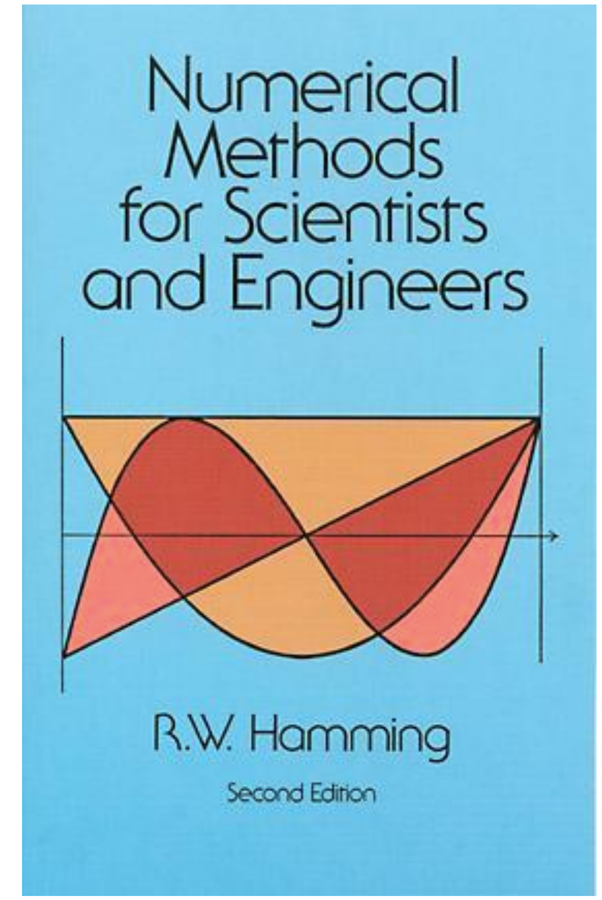
+ Fast Code

MPFR



+ More Reliable

- Slow Code



+ Reliable

+ Fast Code

- Expert Task

Heuristic search to find expert transformations

Heuristic search to find expert transformations

Worked Example

How Herbie Works

Evaluation

Heuristic search to find expert transformations

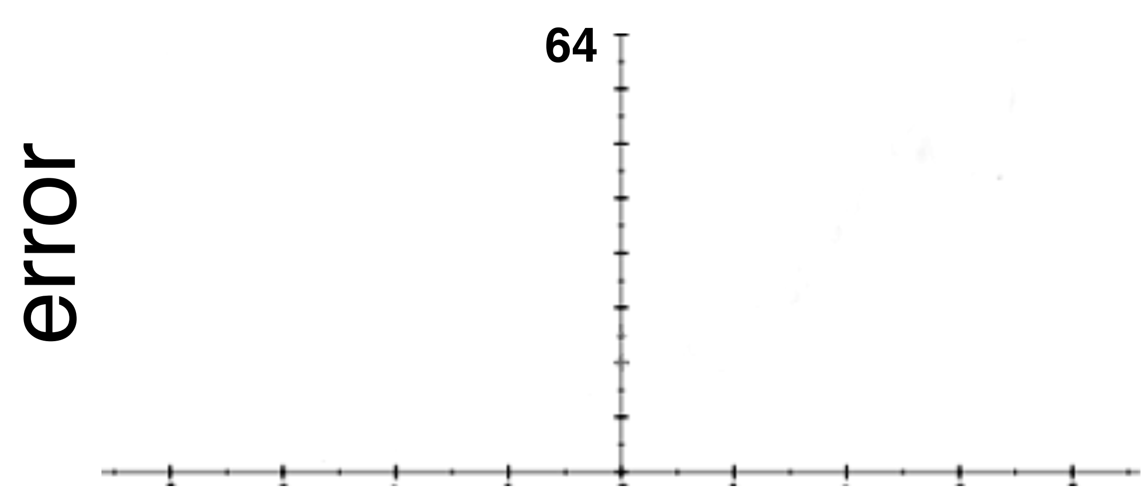
Worked Example

How Herbie Works

Evaluation

Rounding Error in Quadratic

$$\frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

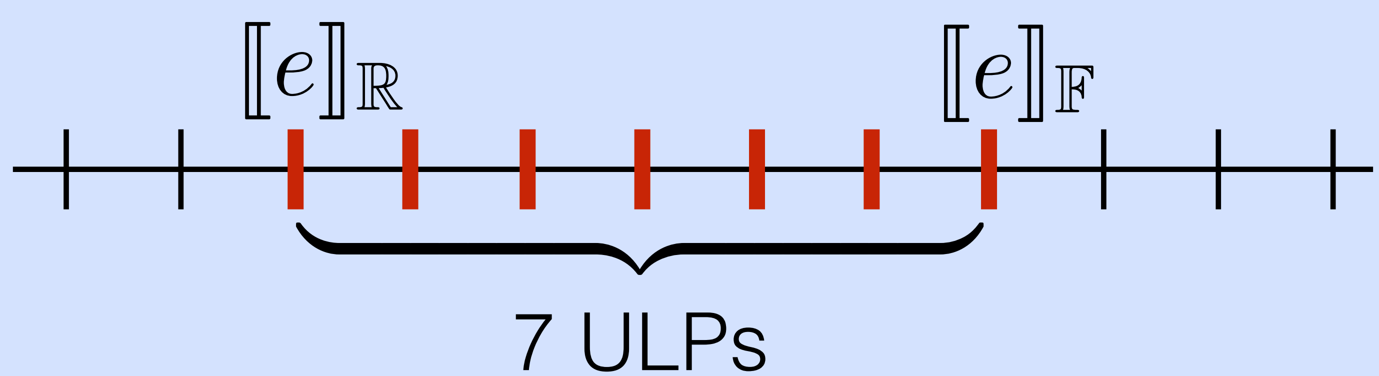


Rounding Error in Quadratic

$$\frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

What is rounding error?

exact computed



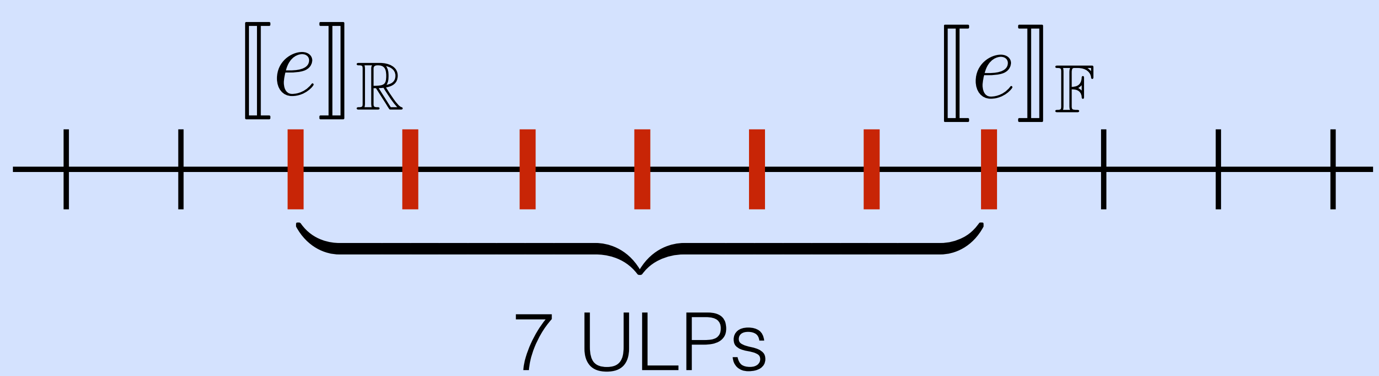
error

Rounding Error in Quadratic

$$\frac{-b+\sqrt{b^2-4ac}}{2a}$$

What is rounding error?

exact computed

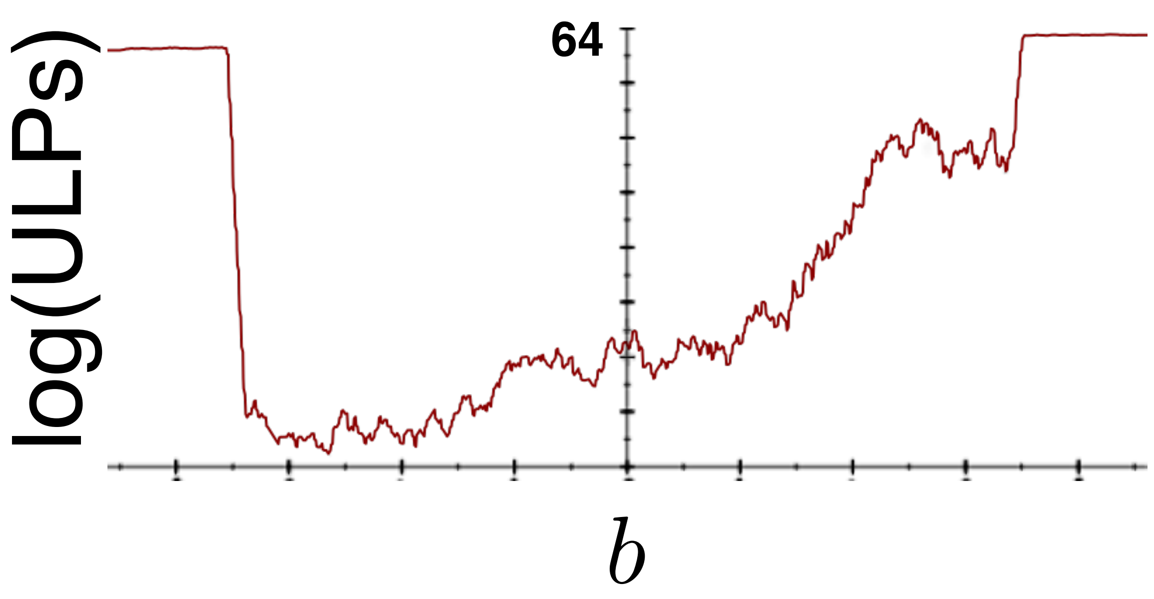


log(ULPs) estimates # of incorrect bits

log(ULPs)

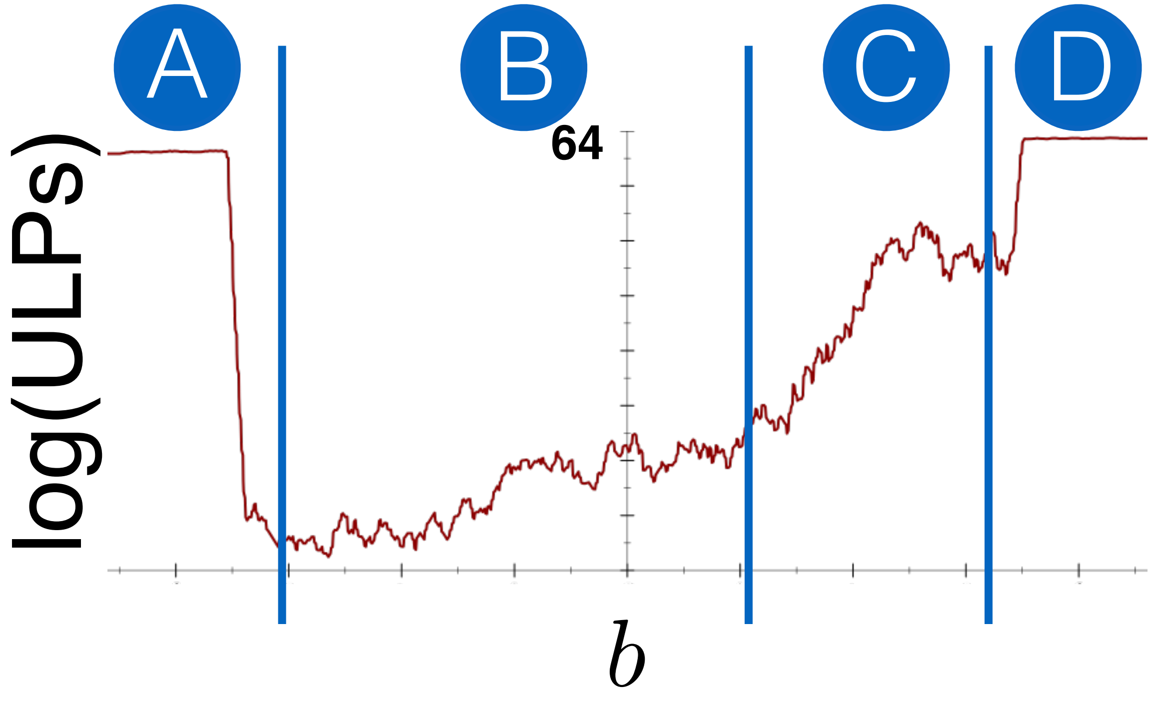
Rounding Error in Quadratic

$$\frac{-b + \sqrt{b^2 - 4ac}}{2a}$$



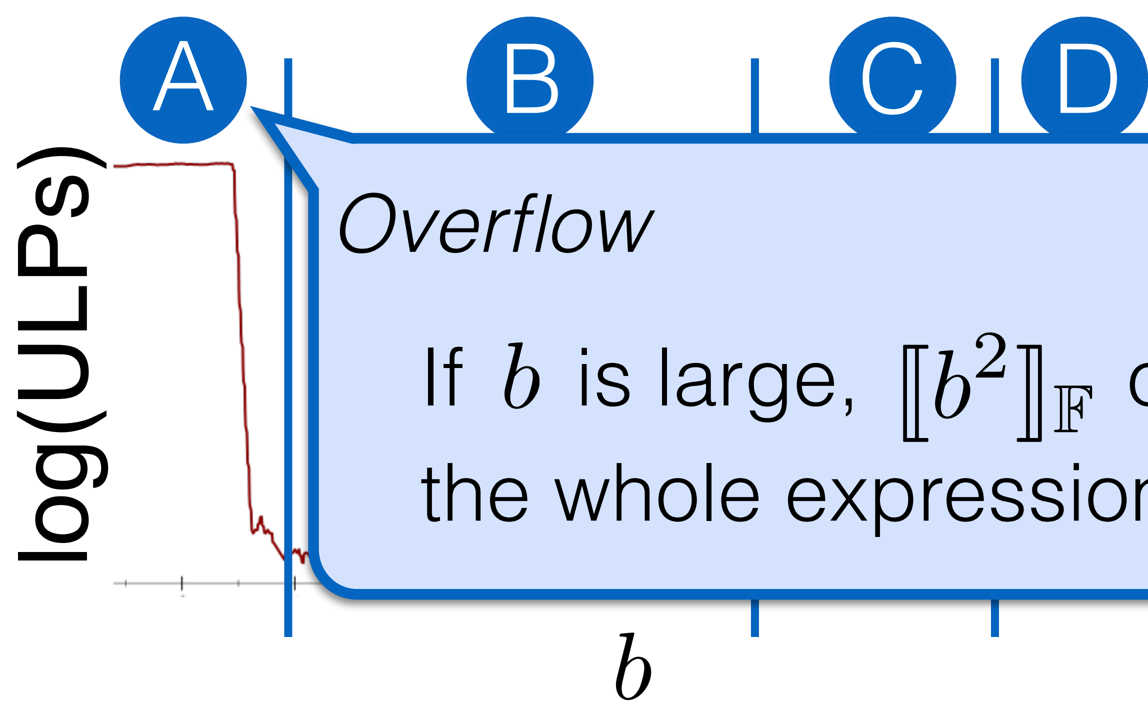
Rounding Error in Quadratic

$$\frac{-b + \sqrt{b^2 - 4ac}}{2a} \Rightarrow \left\{ \right.$$



Rounding Error in Quadratic

$$\frac{-b+\sqrt{b^2-4ac}}{2a} \Rightarrow \left\{ \right.$$

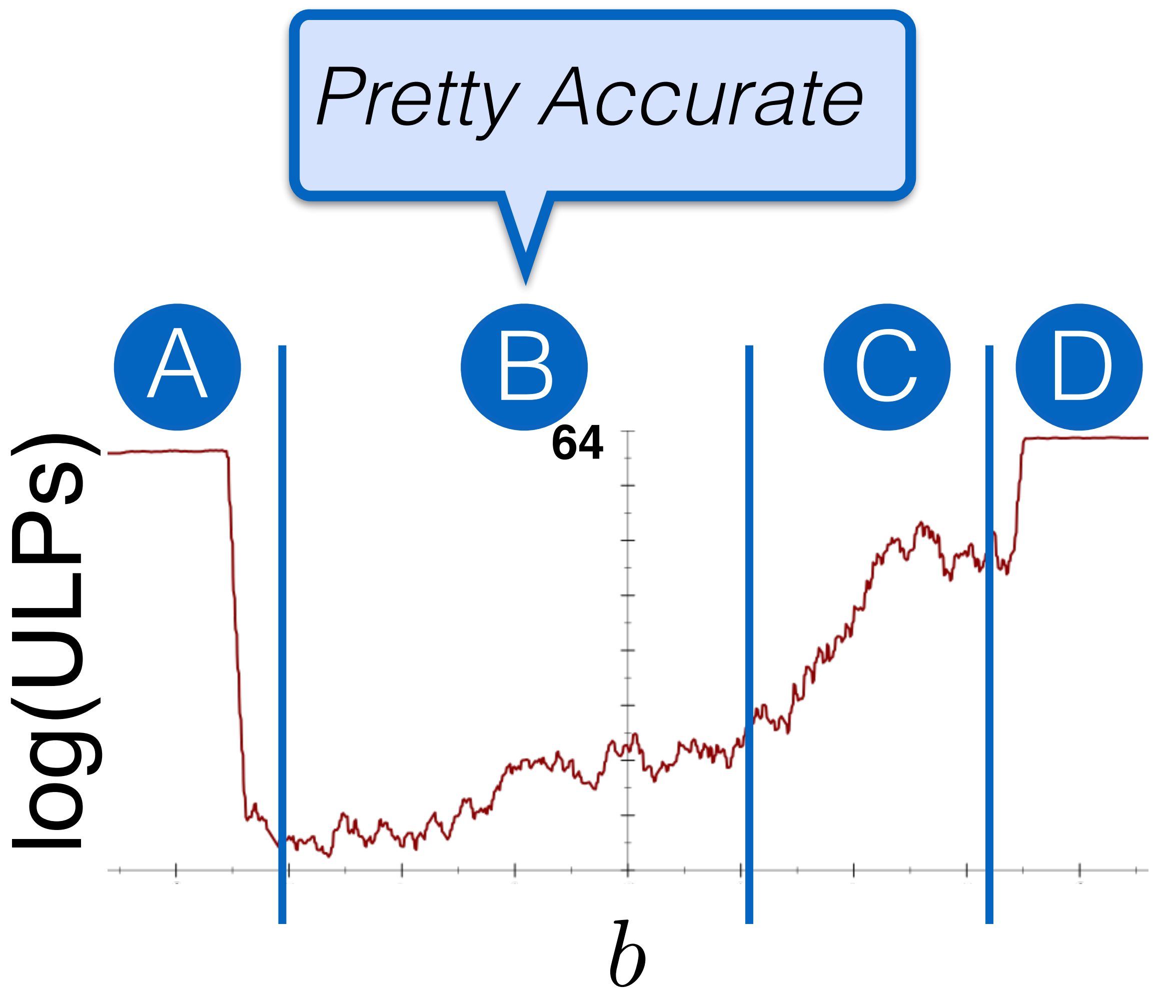


Overflow

If b is large, $[\![b^2]\!]_{\mathbb{F}}$ overflows and the the whole expression returns ∞.

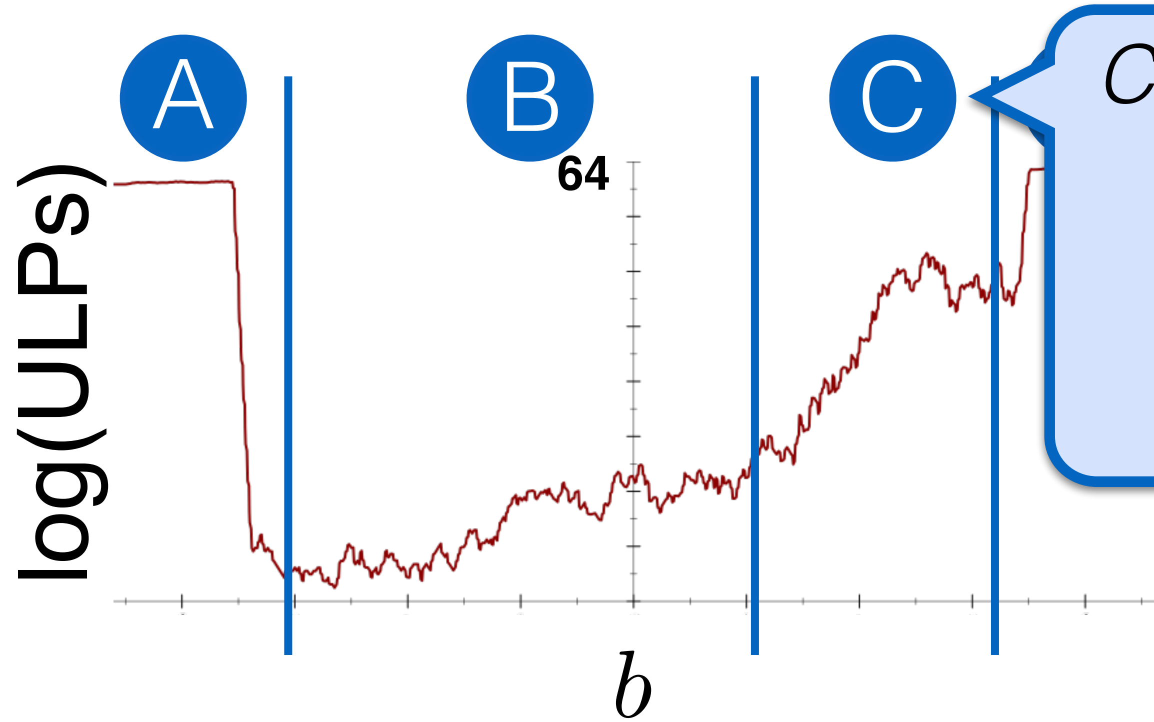
Rounding Error in Quadratic

$$\frac{-b + \sqrt{b^2 - 4ac}}{2a} \Rightarrow \begin{cases} \frac{c}{b} - \frac{b}{a} & \text{if } b \in \text{A} \\ & \\ & \end{cases}$$



Rounding Error in Quadratic

$$\frac{-b + \sqrt{b^2 - 4ac}}{2a} \Rightarrow \begin{cases} \frac{c}{b} - \frac{b}{a} & \text{if } b \in \text{A} \\ \frac{-b+\sqrt{b^2-4ac}}{2a} & \text{if } b \in \text{B} \end{cases}$$

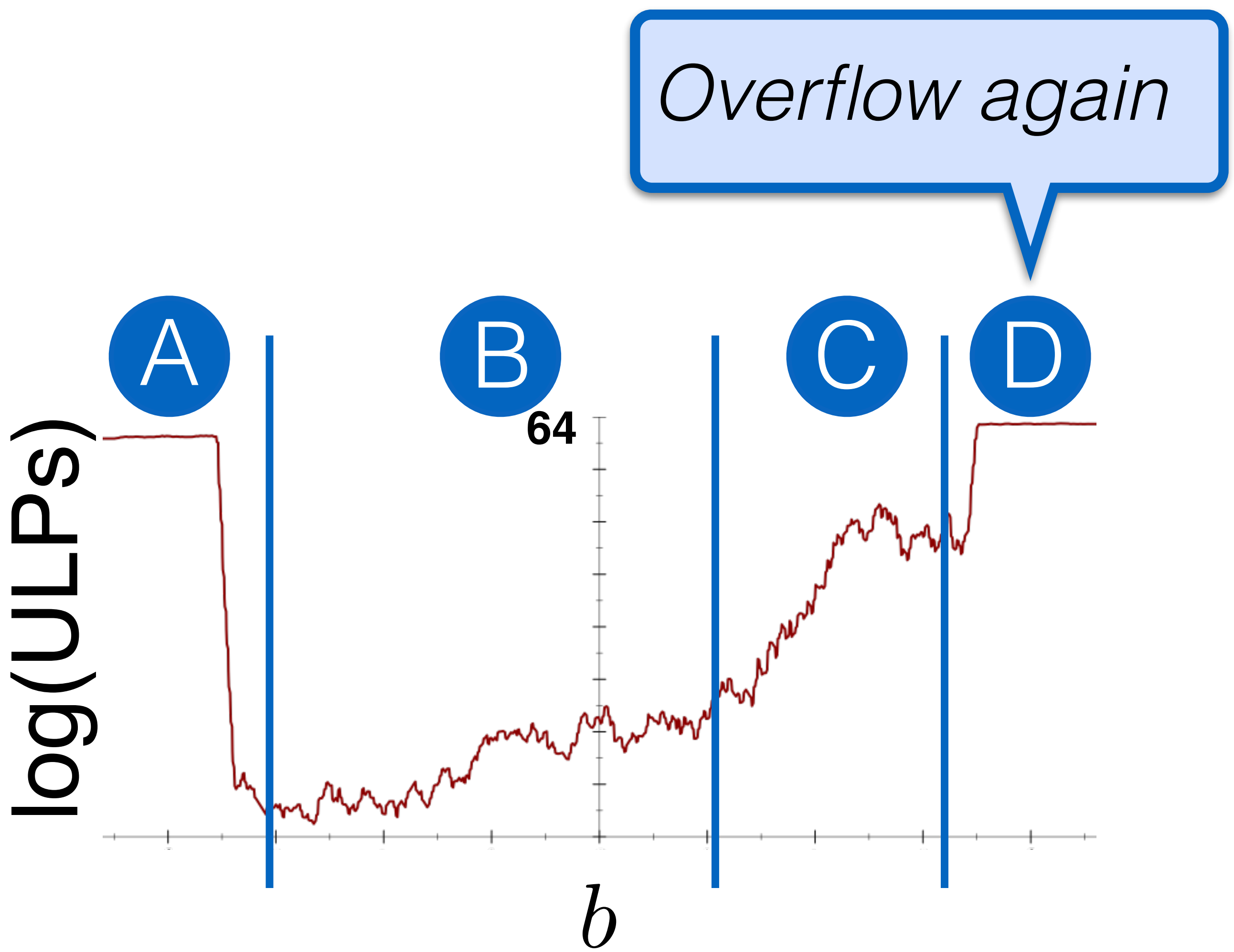


Catastrophic Cancellation

If b is large, but a and c are small, $b \approx \sqrt{b^2 - 4ac}$ and the difference is rounded off.

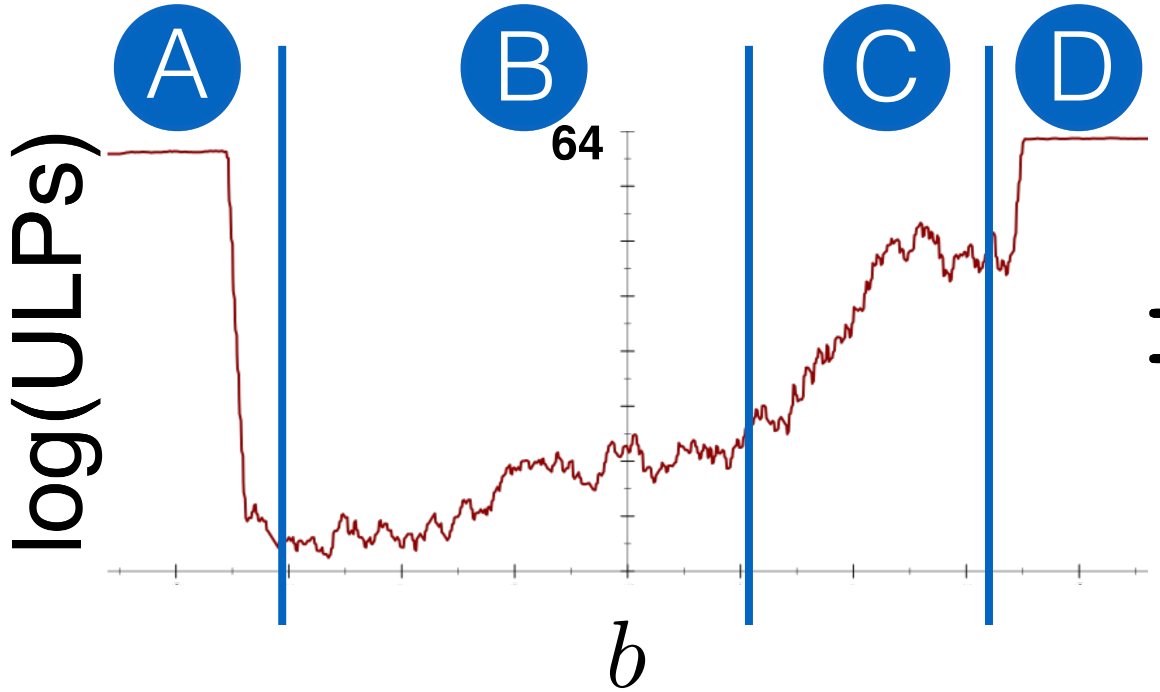
Rounding Error in Quadratic

$$\frac{-b + \sqrt{b^2 - 4ac}}{2a} \Longrightarrow \begin{cases} \frac{c}{b} - \frac{b}{a} & \text{if } b \in \text{A} \\ \frac{-b+\sqrt{b^2-4ac}}{2a} & \text{if } b \in \text{B} \\ \frac{2c}{-b-\sqrt{b^2-4ac}} & \text{if } b \in \text{C} \end{cases}$$

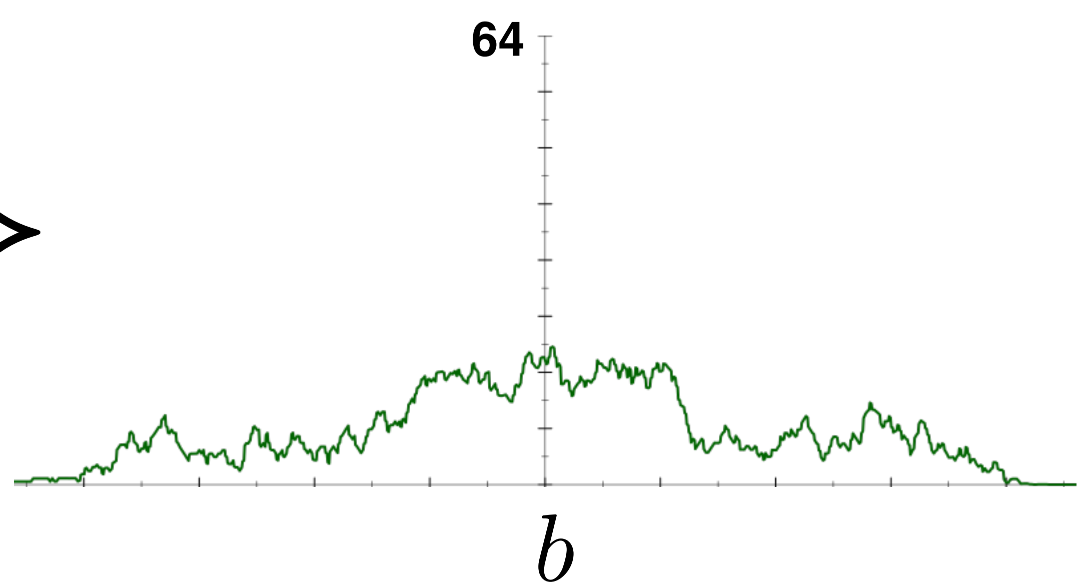


Rounding Error in Quadratic

$$\frac{-b+\sqrt{b^2-4ac}}{2a} \implies \begin{cases} \frac{c}{b} - \frac{b}{a} & \text{if } b \in \text{A} \\ \frac{-b+\sqrt{b^2-4ac}}{2a} & \text{if } b \in \text{B} \\ \frac{2c}{-b-\sqrt{b^2-4ac}} & \text{if } b \in \text{C} \\ -\frac{c}{b} & \text{if } b \in \text{D} \end{cases}$$



$\implies$



Heuristic search to find expert transformations

Worked Example

How Herbie Works

Evaluation

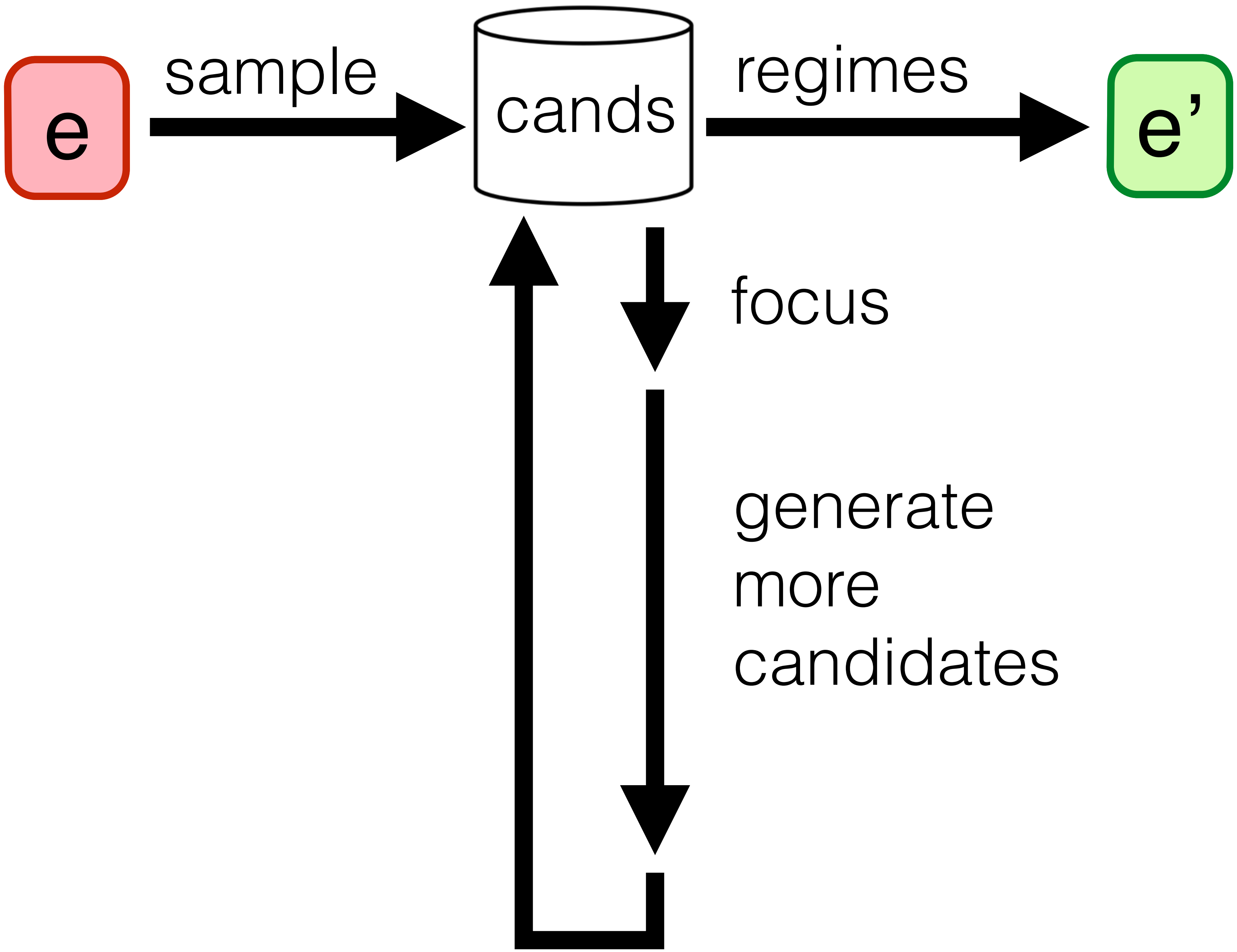
Heuristic search to find expert transformations

Worked Example

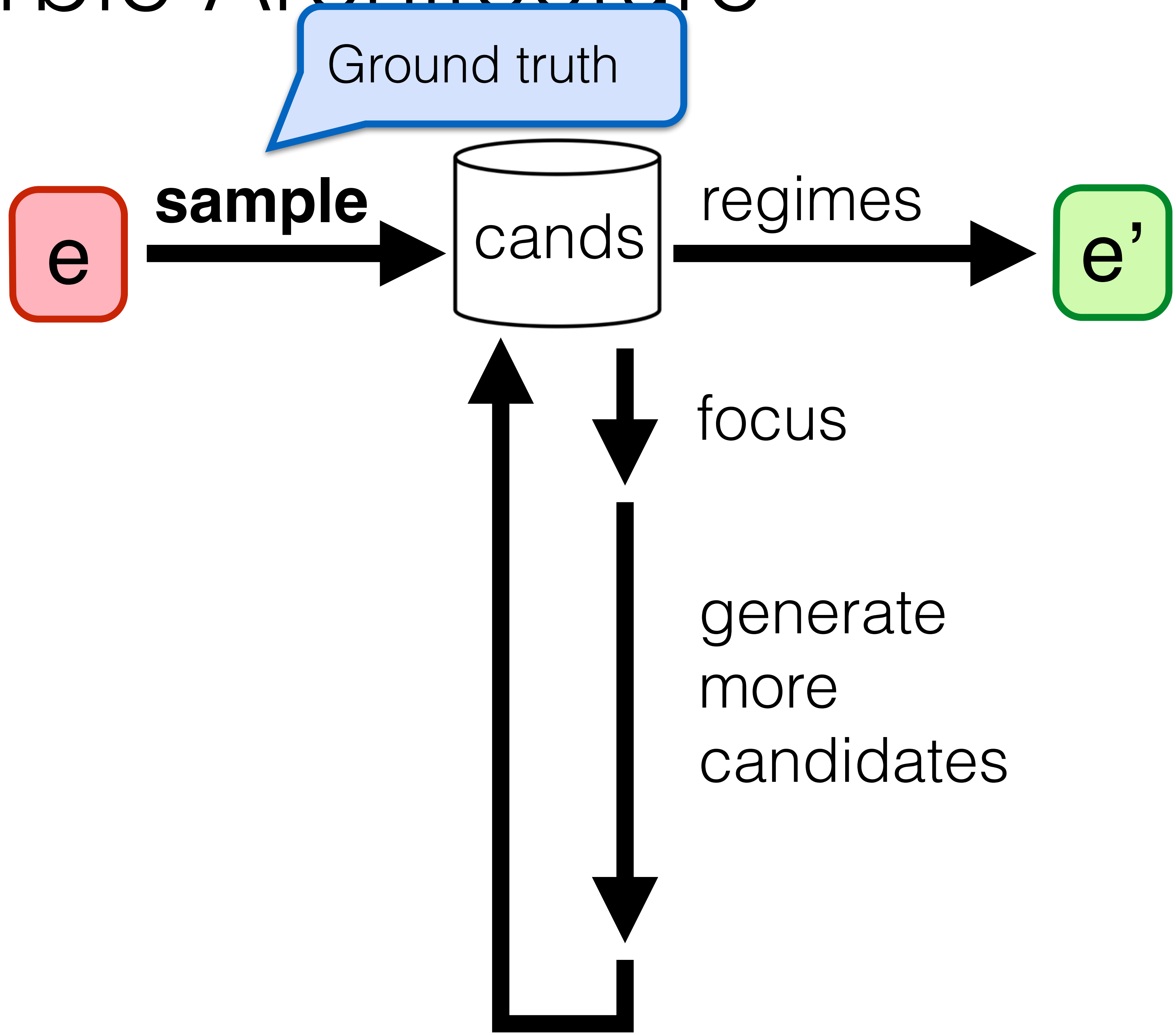
How Herbie Works

Evaluation

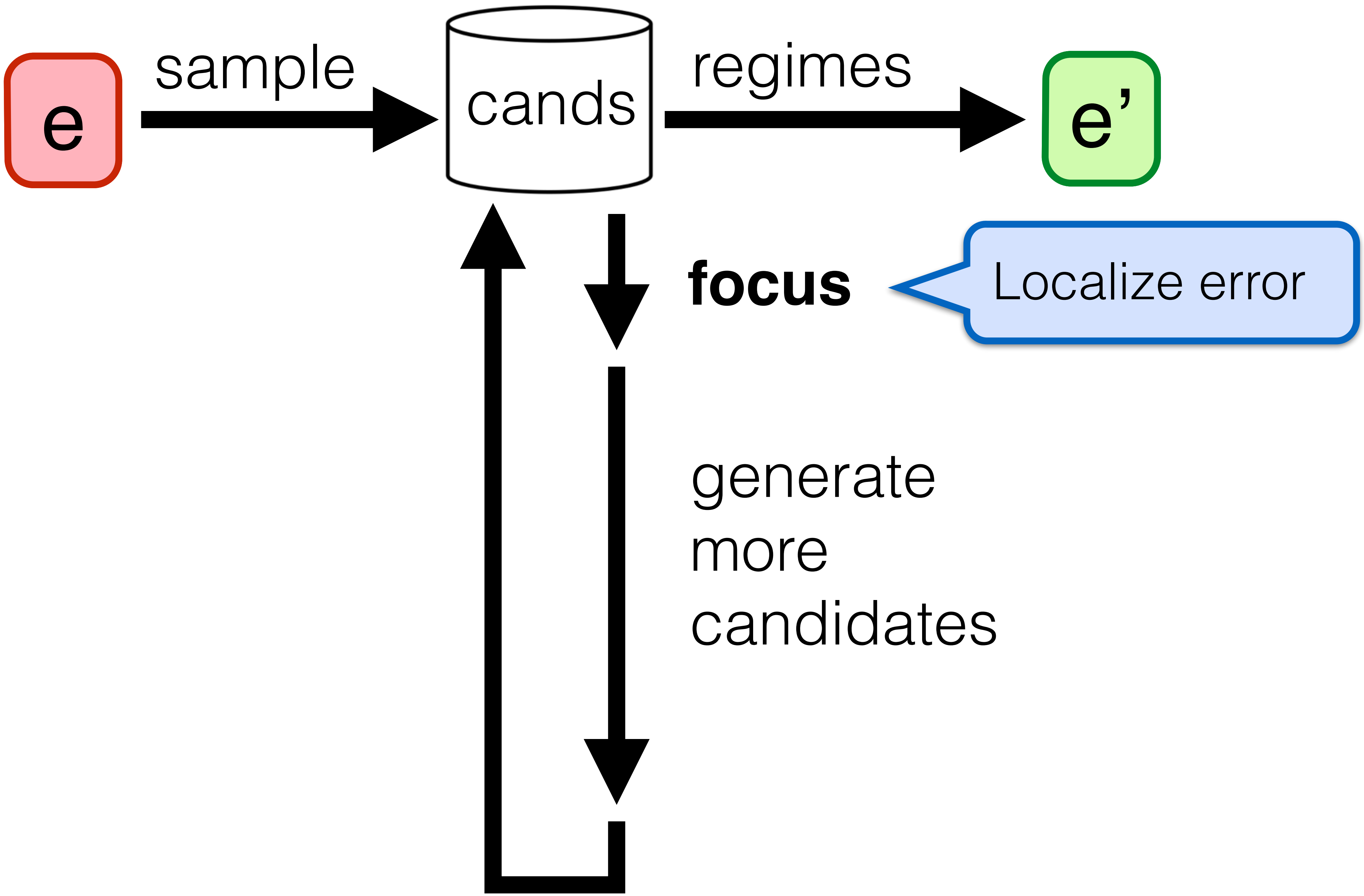
Herbie Architecture



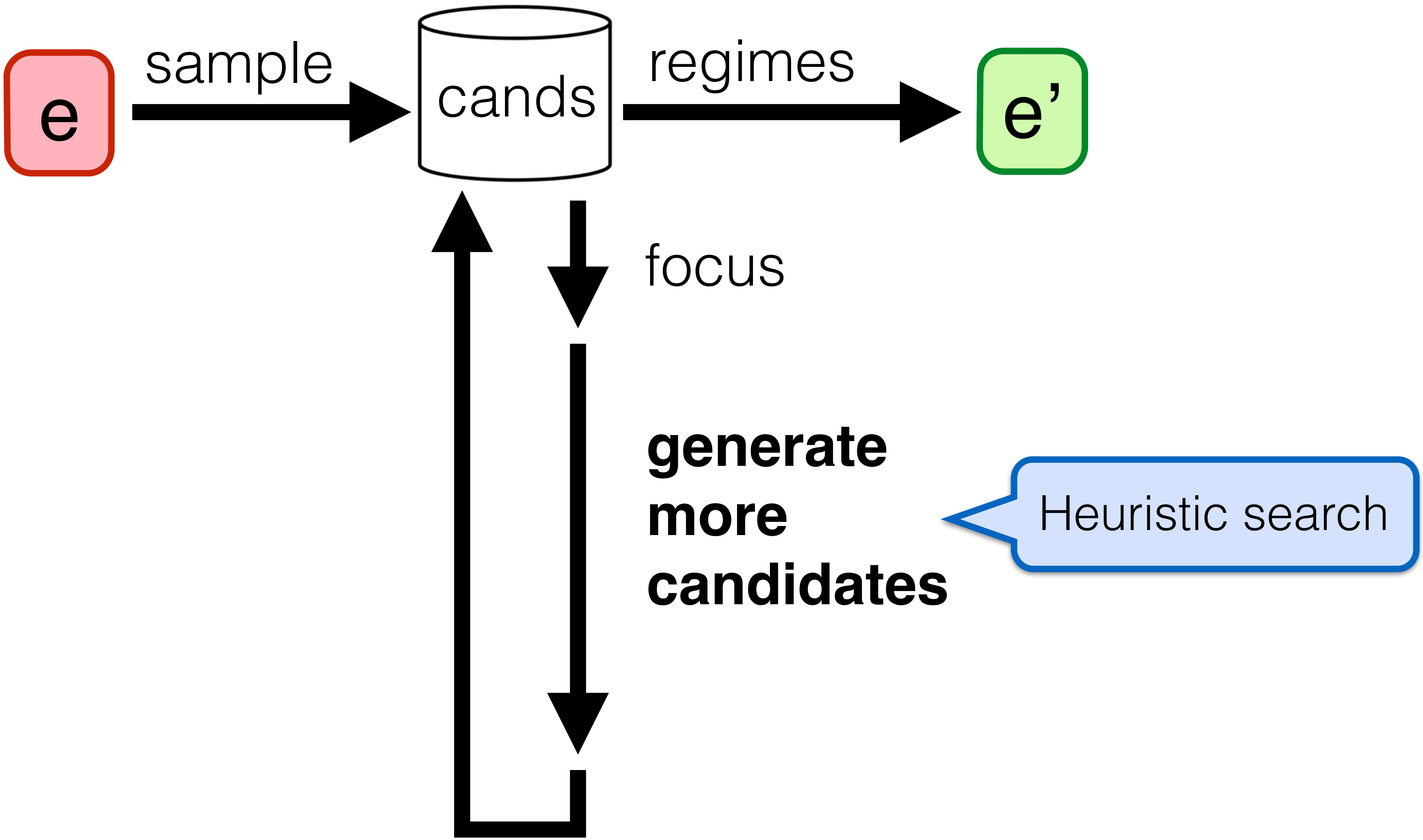
Herbie Architecture



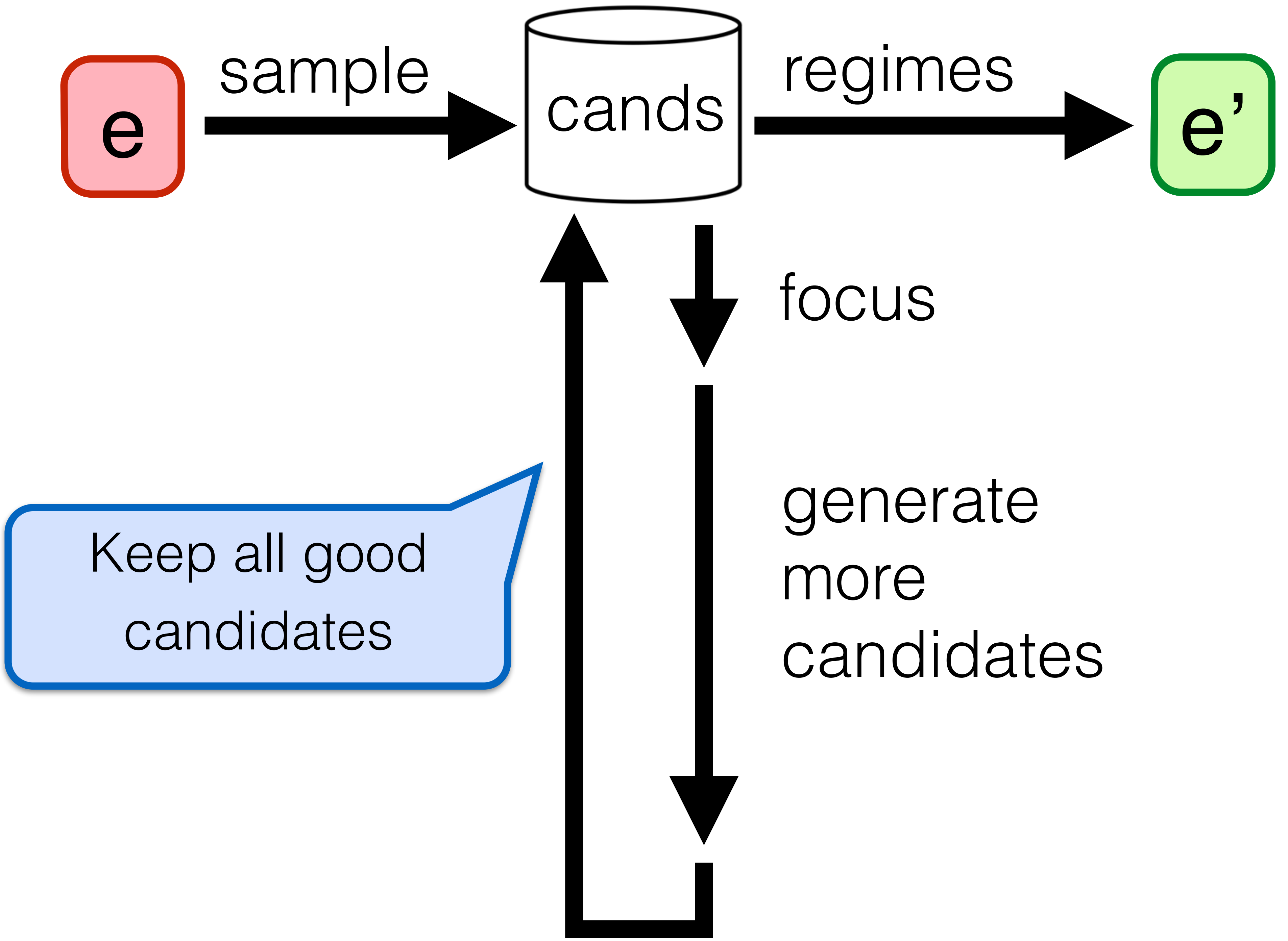
Herbie Architecture



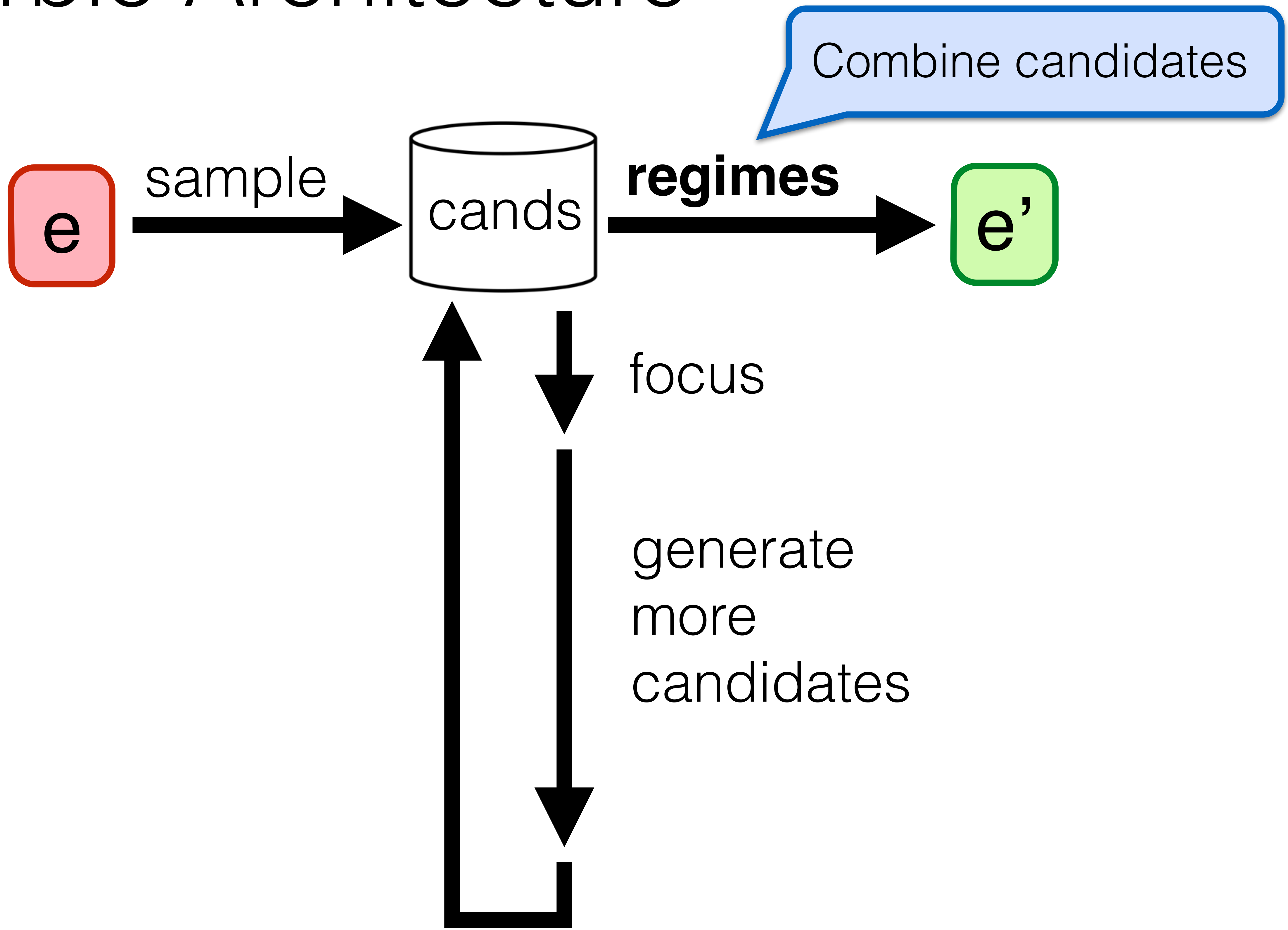
Herbie Architecture



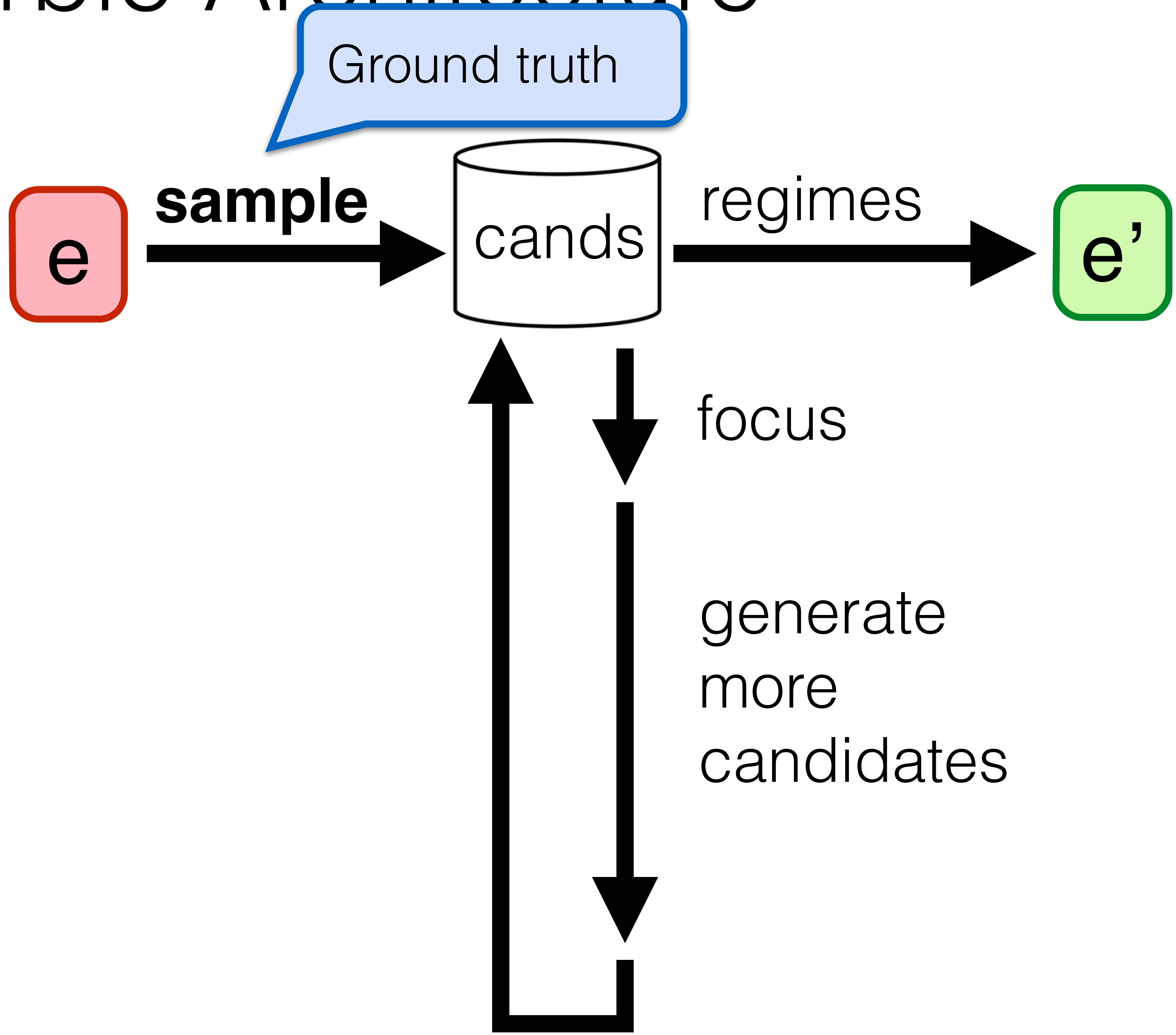
Herbie Architecture



Herbie Architecture



Herbie Architecture



Determine ground truth

$X = \textsf{sample}(\textsf{domain}(e))$

e.g. $X = \{1.2 \cdot 10^{-17}, -3.8 \cdot 10^{204}, 173.5, \ldots\}$

64 random bits

$\textsf{Round}(\llbracket e \rrbracket_{\mathbb{R}}(X))$

Get 64-bit prefix with MPFR.
Subtle! See paper.

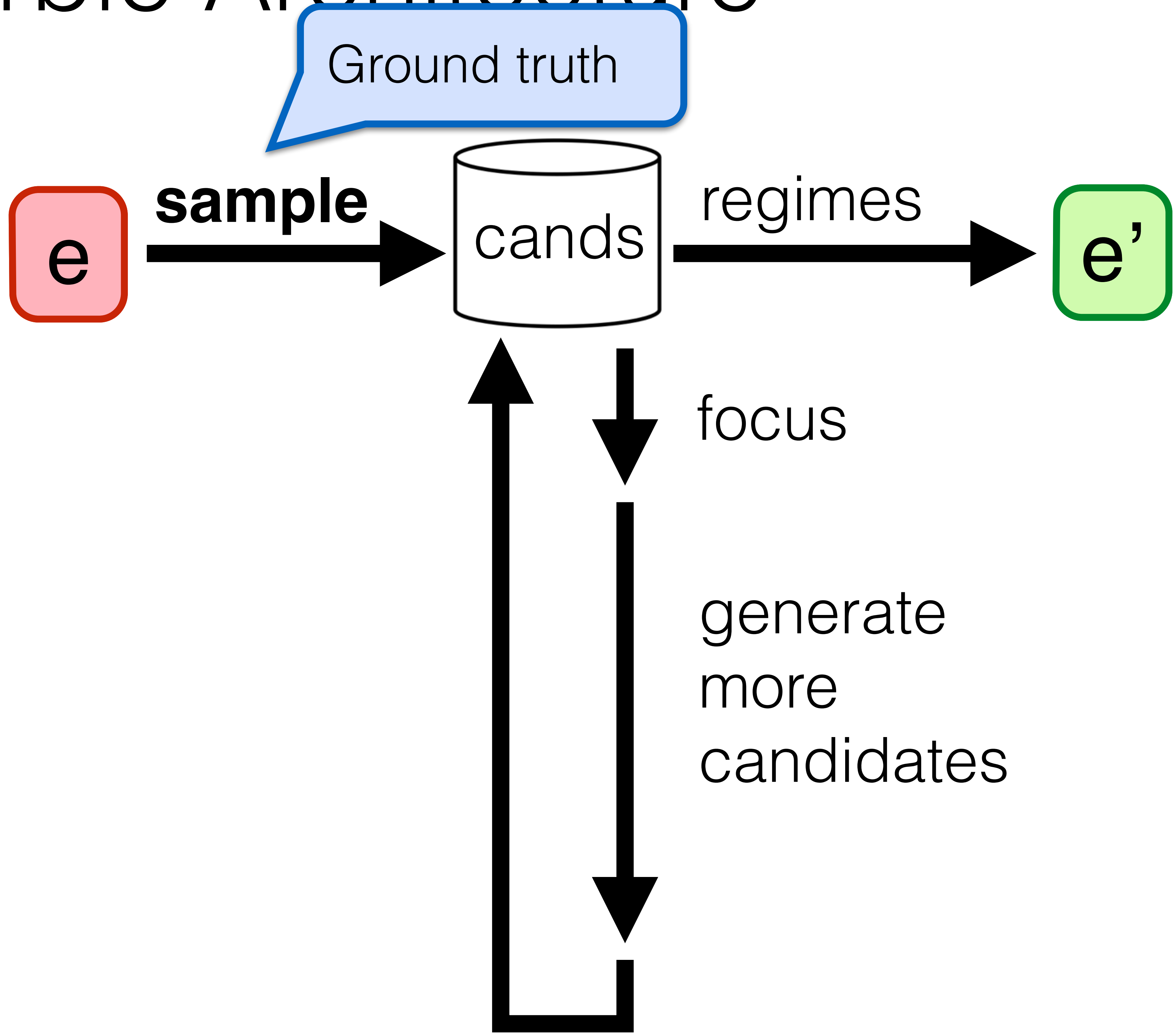
$\llbracket e \rrbracket_{\mathbb{F}}(X)$

Compute in $\mathbb{F}$

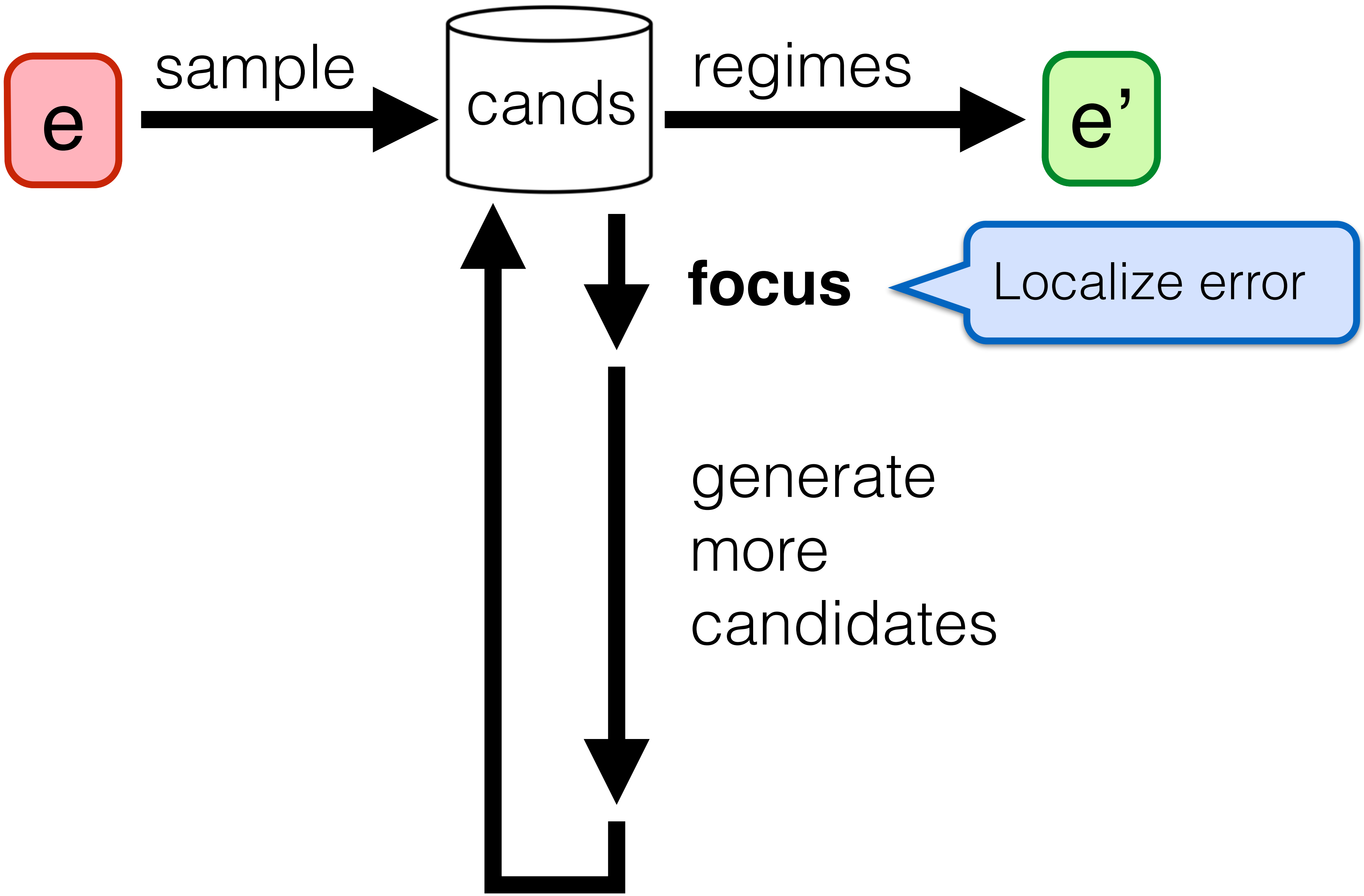
error

e.g. $\{13.2\text{b}, 51.7\text{b}, 1\text{b}, \ldots\}$

Herbie Architecture

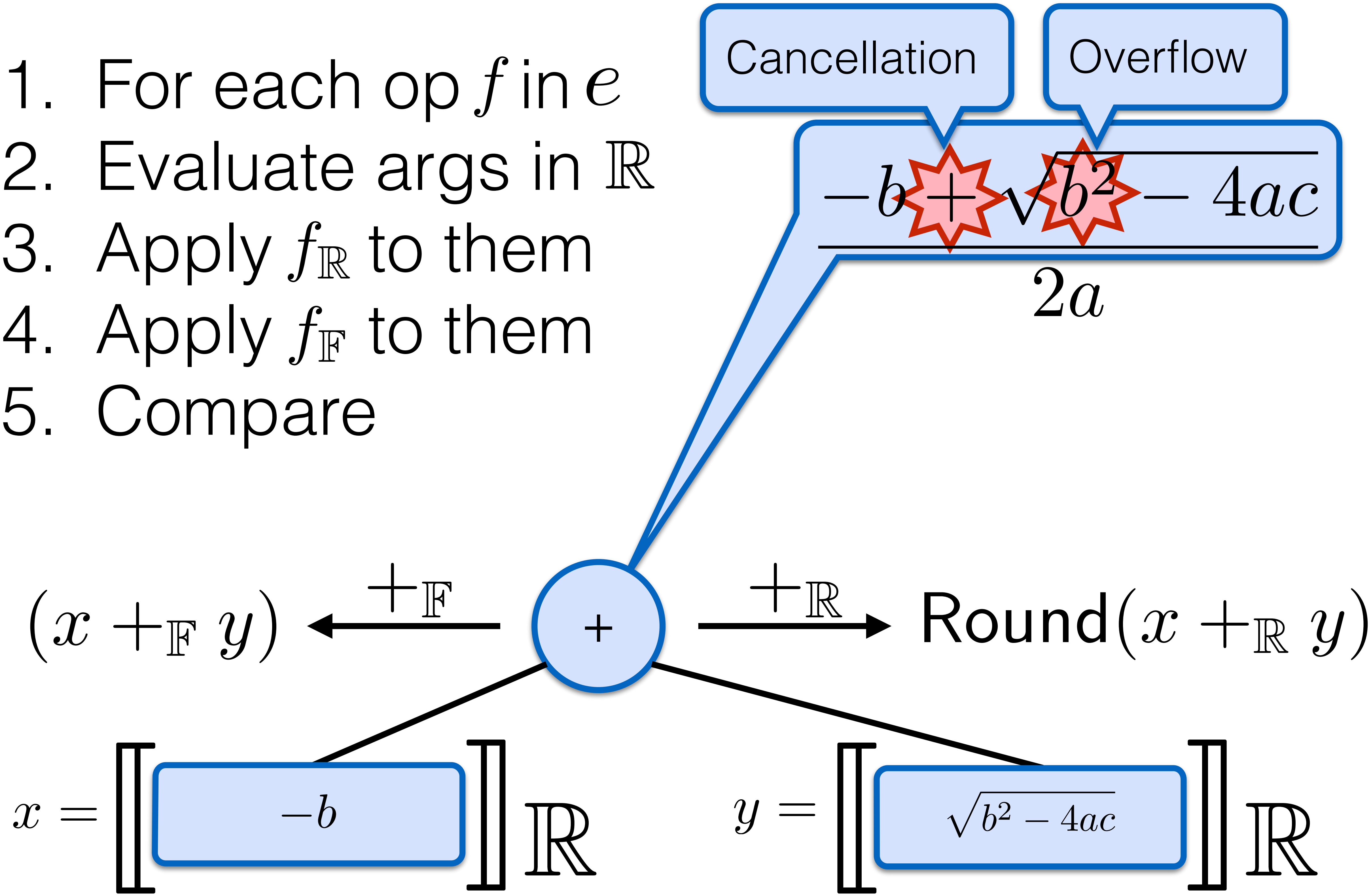


Herbie Architecture

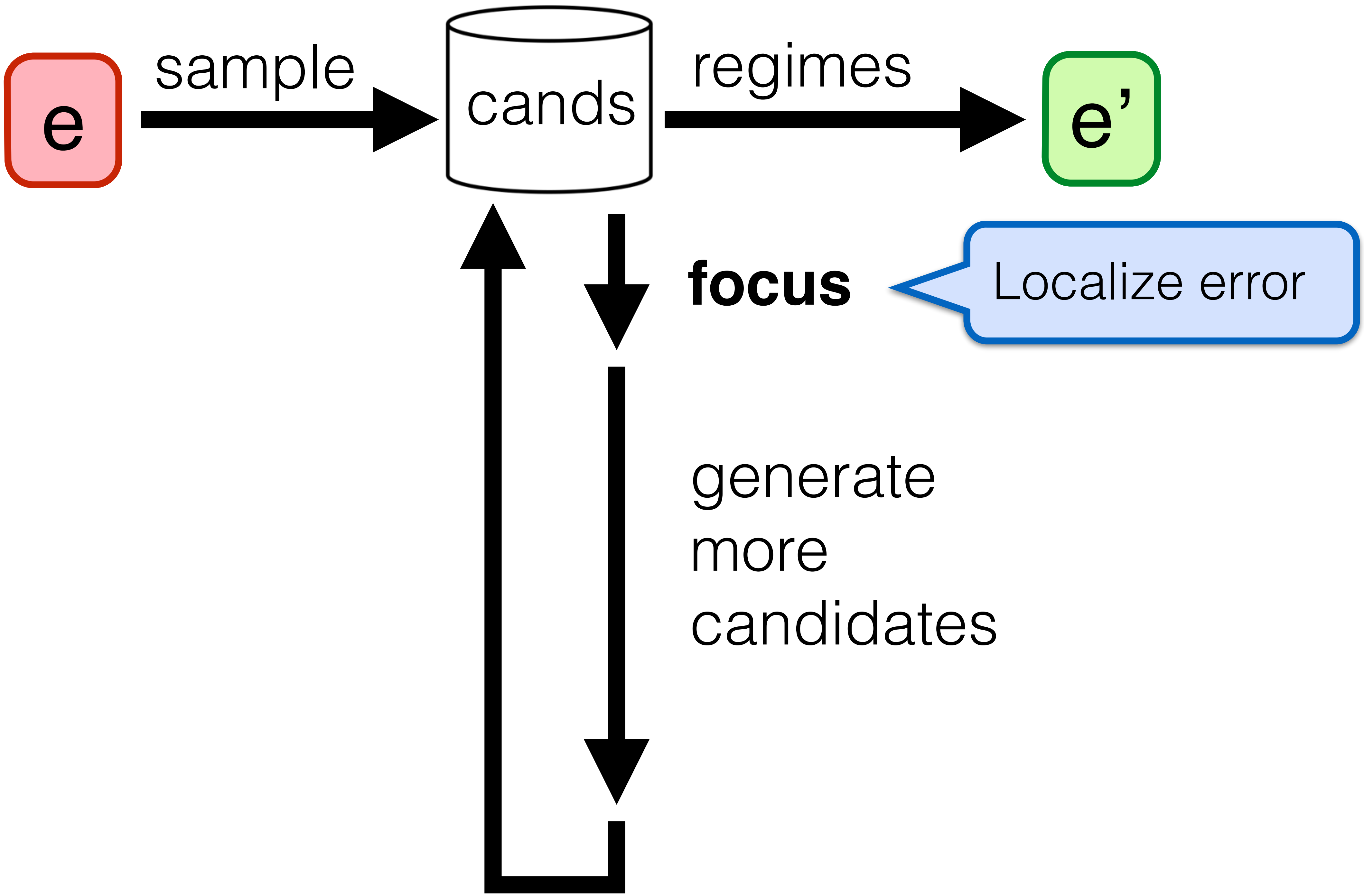


Focus: Estimate Error Source

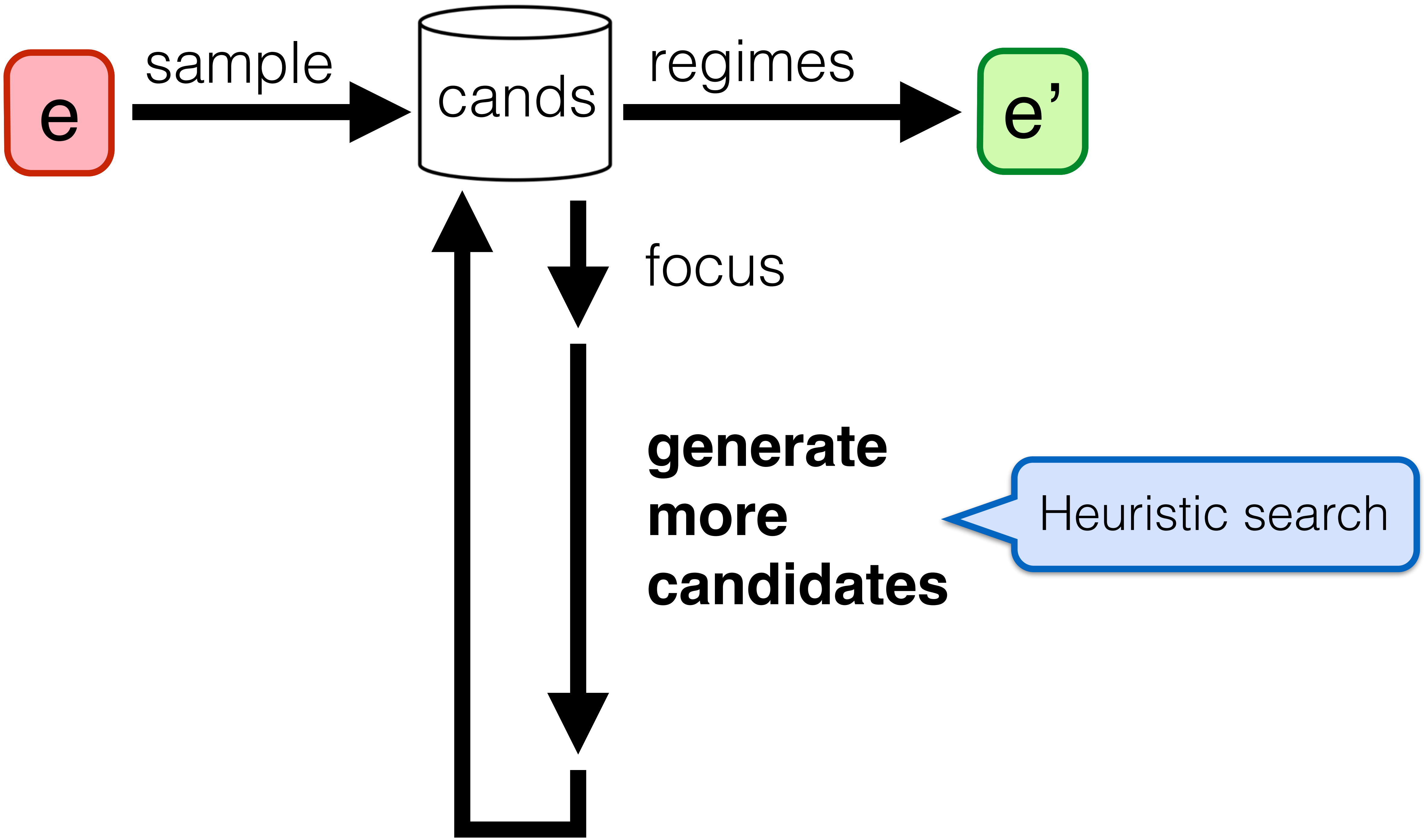
1. For each op f in e
2. Evaluate args in $\mathbb{R}$
3. Apply $f_{\mathbb{R}}$ to them
4. Apply $f_{\mathbb{F}}$ to them
5. Compare



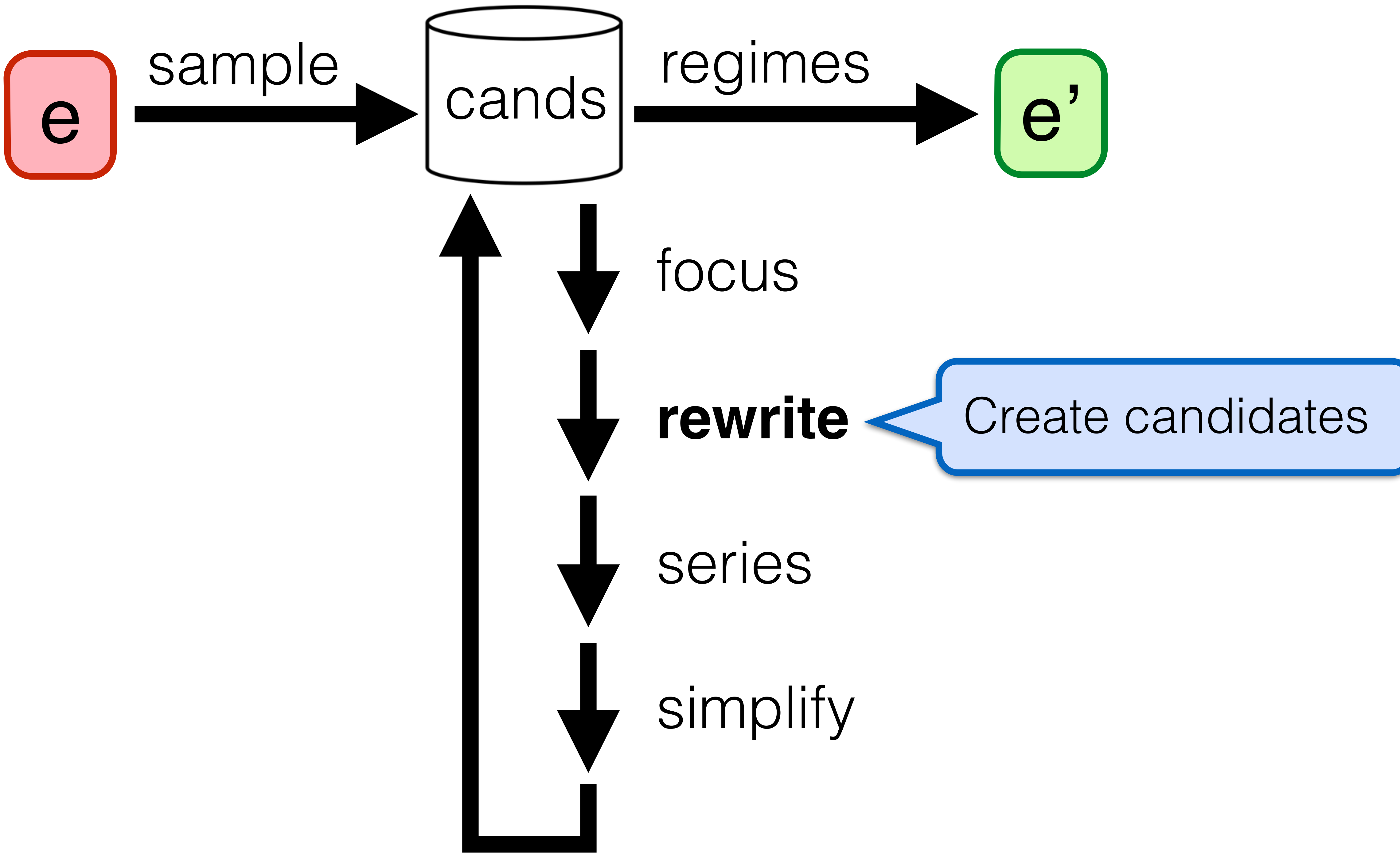
Herbie Architecture



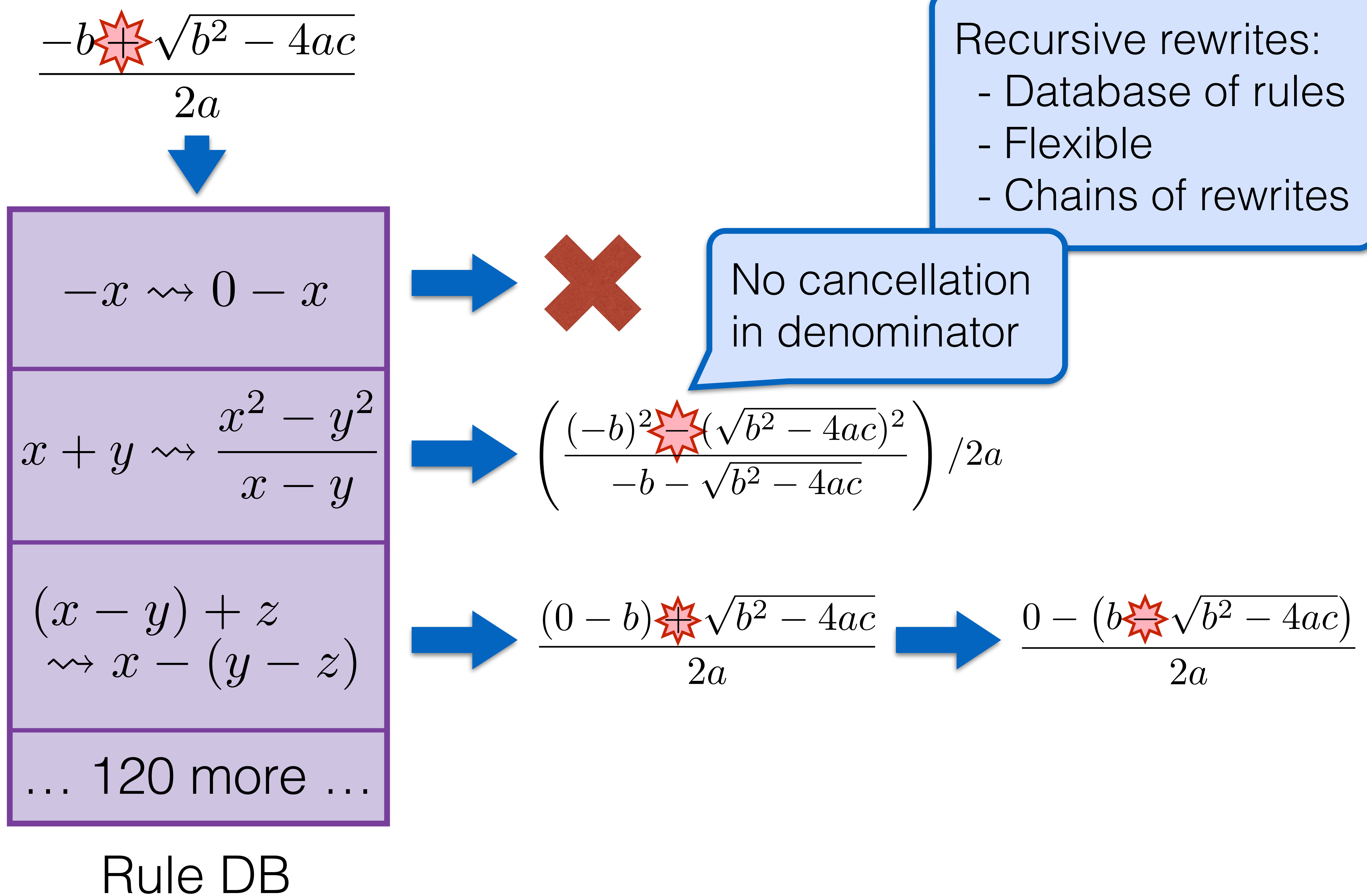
Herbie Architecture



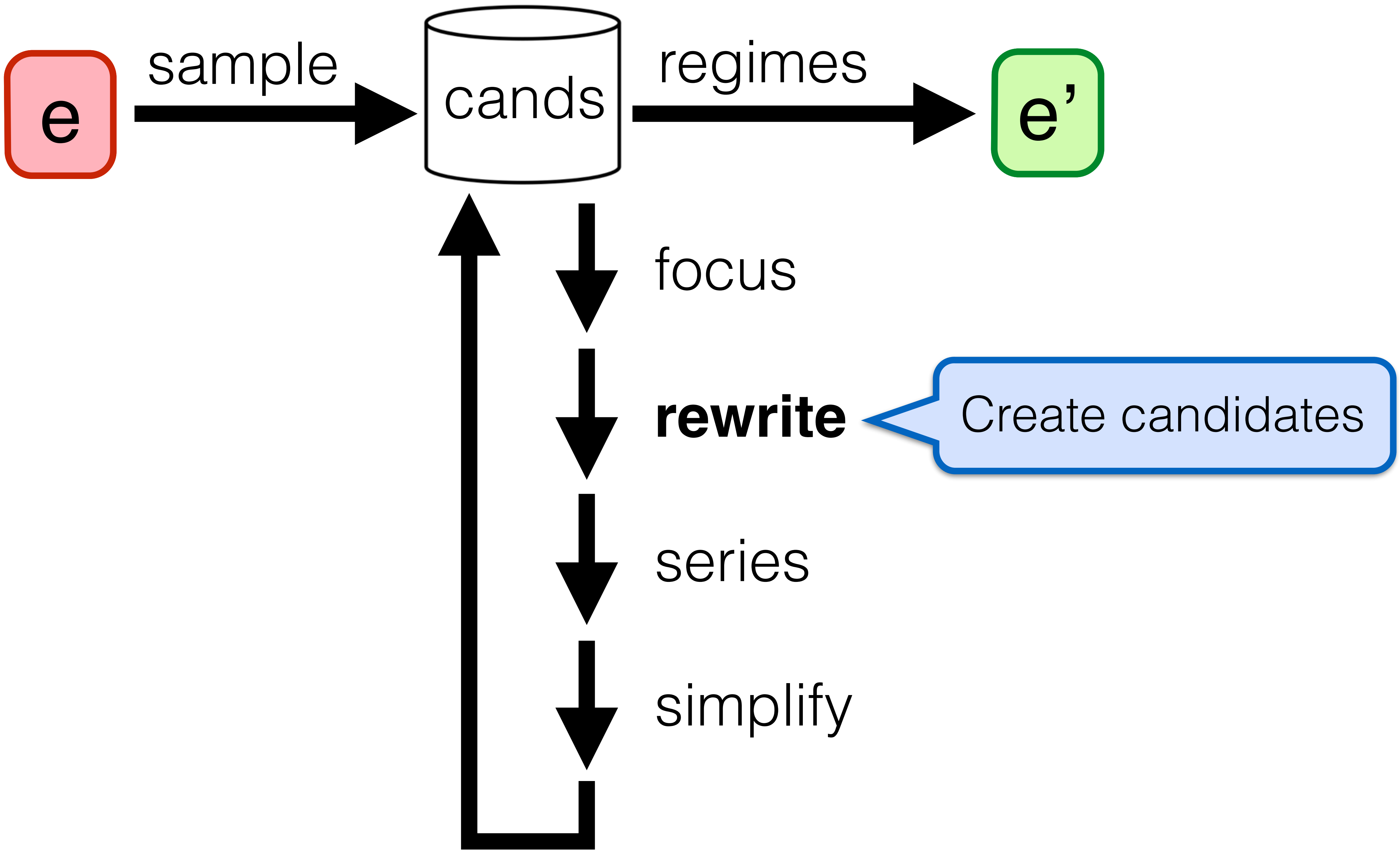
Herbie Architecture



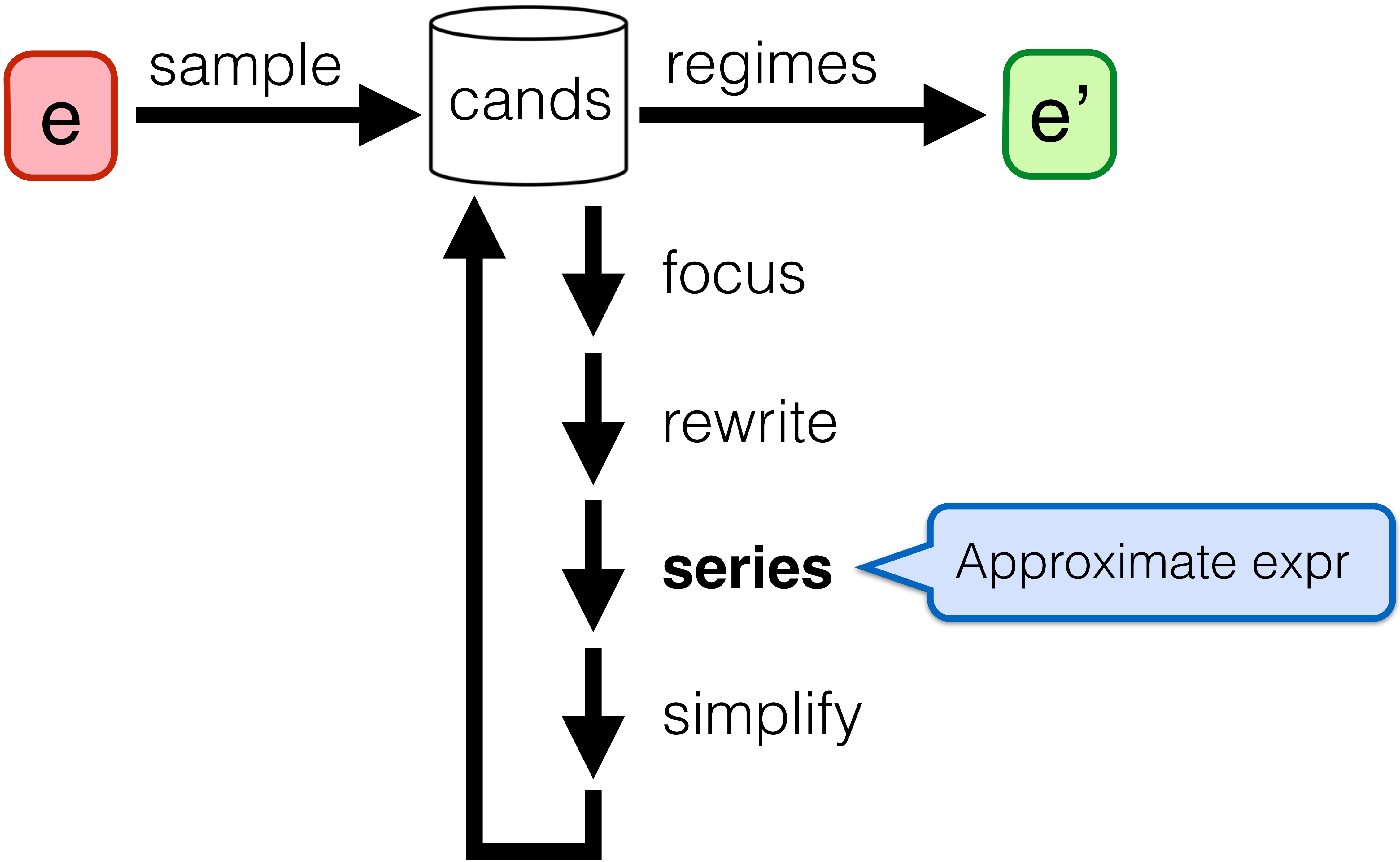
Apply rewrites to



Herbie Architecture



Herbie Architecture



Series Expansions

Idea: *near*-identities

$$\sqrt{1-x} \approx 1 - x/2$$

(for $x \approx 0$)

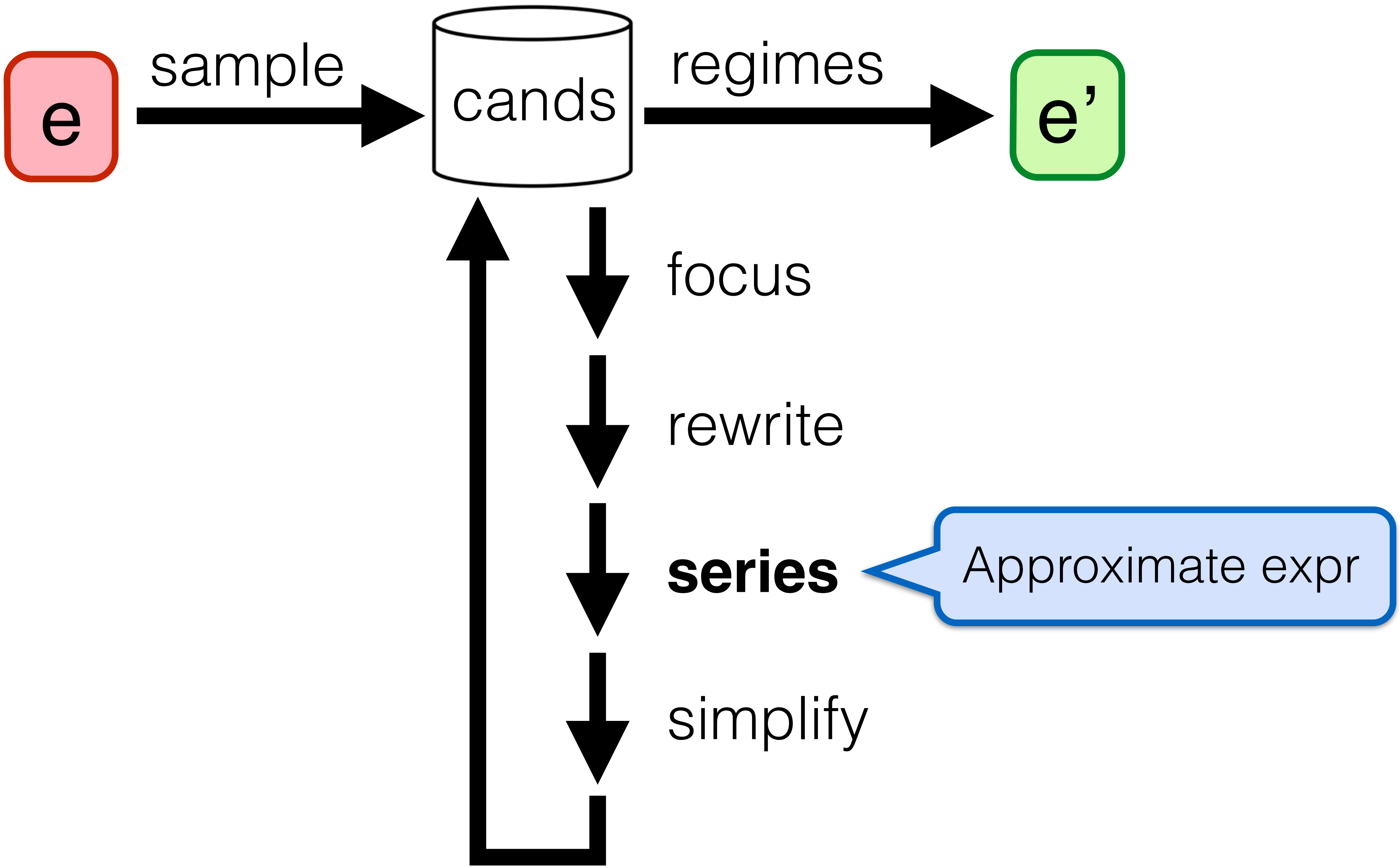
$$\frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

$$\frac{-b + b(1 - 4ac/2b^2)}{2a}$$

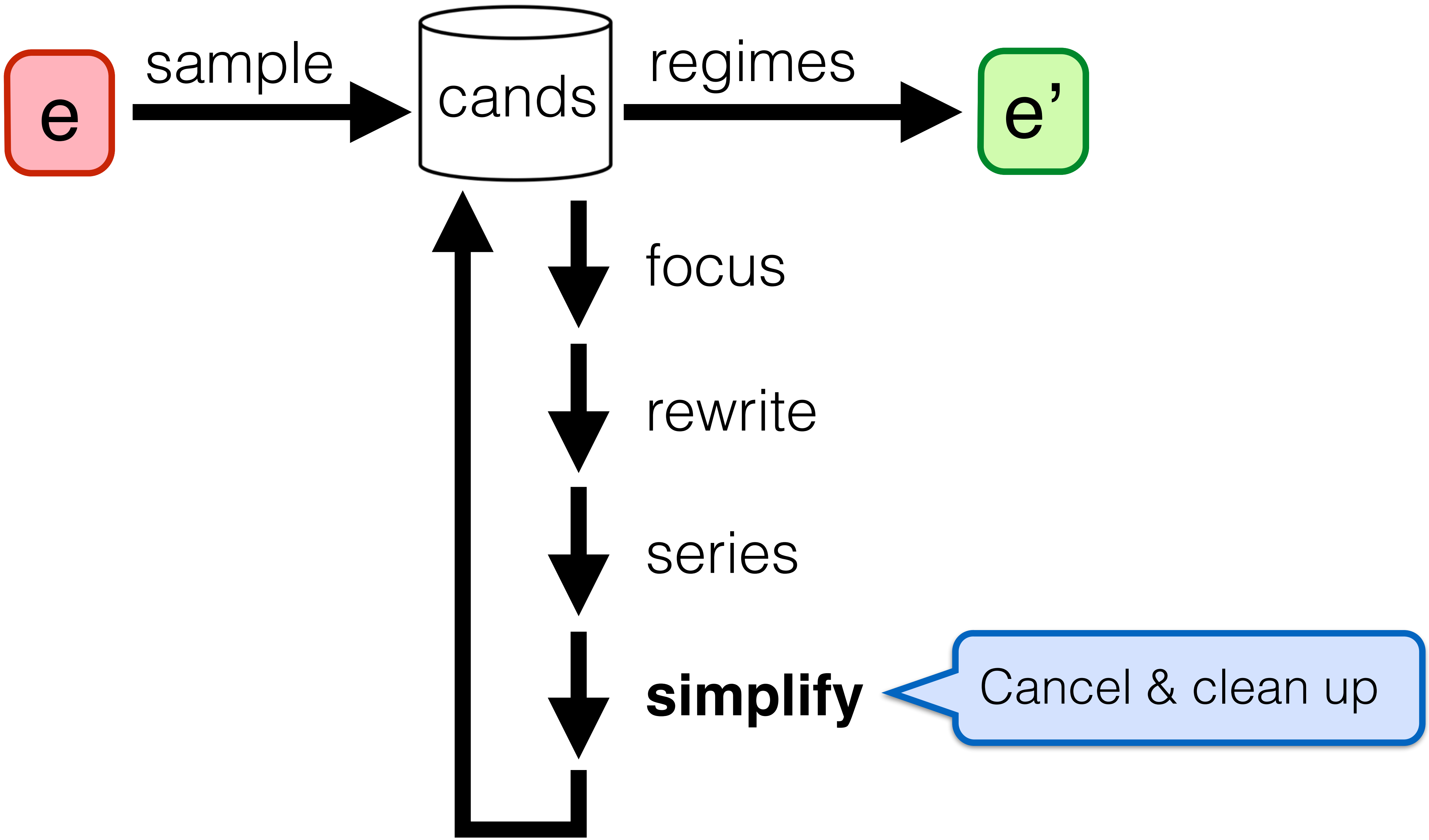
Bounded Laurent series:

- Transcendental functions
- Singularities
- Number of terms to take

Herbie Architecture



Herbie Architecture



Simplify Expressions

$$\left(\frac{(-b)^2 - (\sqrt{b^2 - 4ac})^2}{-b - \sqrt{b^2 - 4ac}}\right) / 2a$$

$$= \left(\frac{b^2 - (\sqrt{b^2 - 4ac})^2}{-b - \sqrt{b^2 - 4ac}}\right) / 2a$$

$$= \left(\frac{b^2 - (b^2 - 4ac)}{-b - \sqrt{b^2 - 4ac}}\right) / 2a$$

$$= \left(\frac{4ac}{-b - \sqrt{b^2 - 4ac}}\right) / 2a$$

$$= \frac{2c}{-b - \sqrt{b^2 - 4ac}}$$

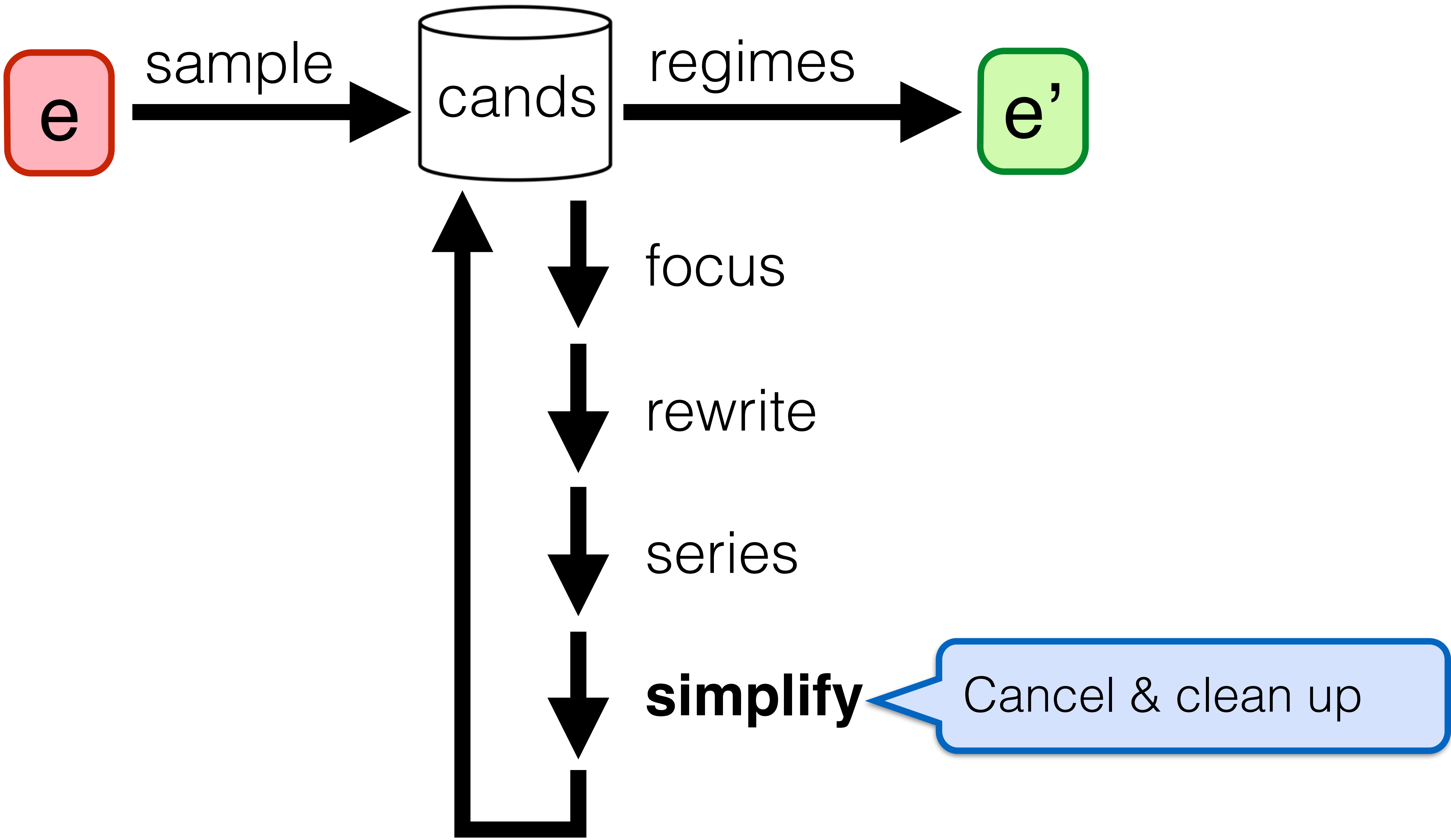
Difficult! *[Caviness '70]*

- many possible rewrites
- huge search space
- avoid undoing progress!

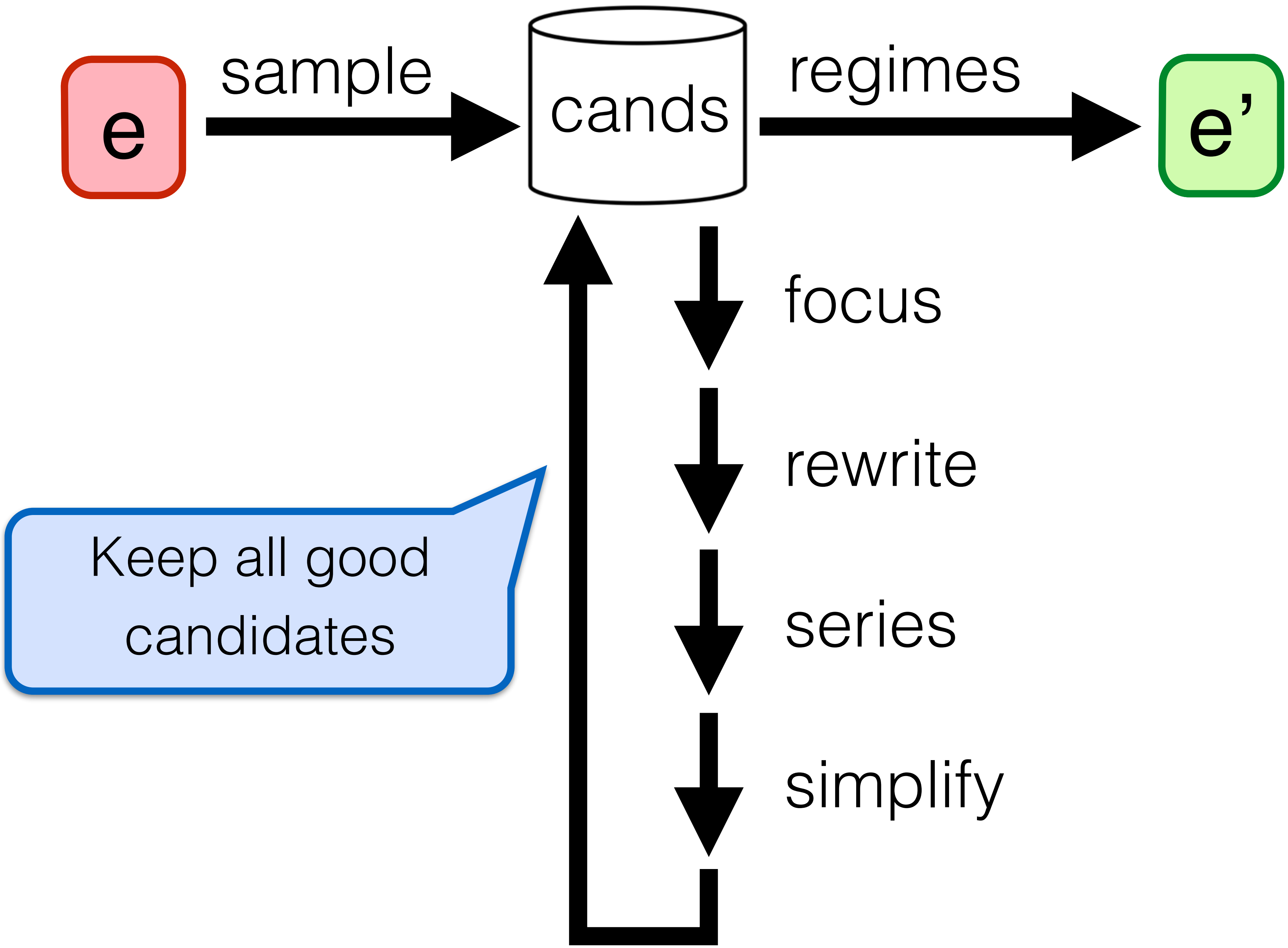
E-graphs *[Nelson '79]*

- Terminate early
- Prune useless nodes
- Restrict rewrites

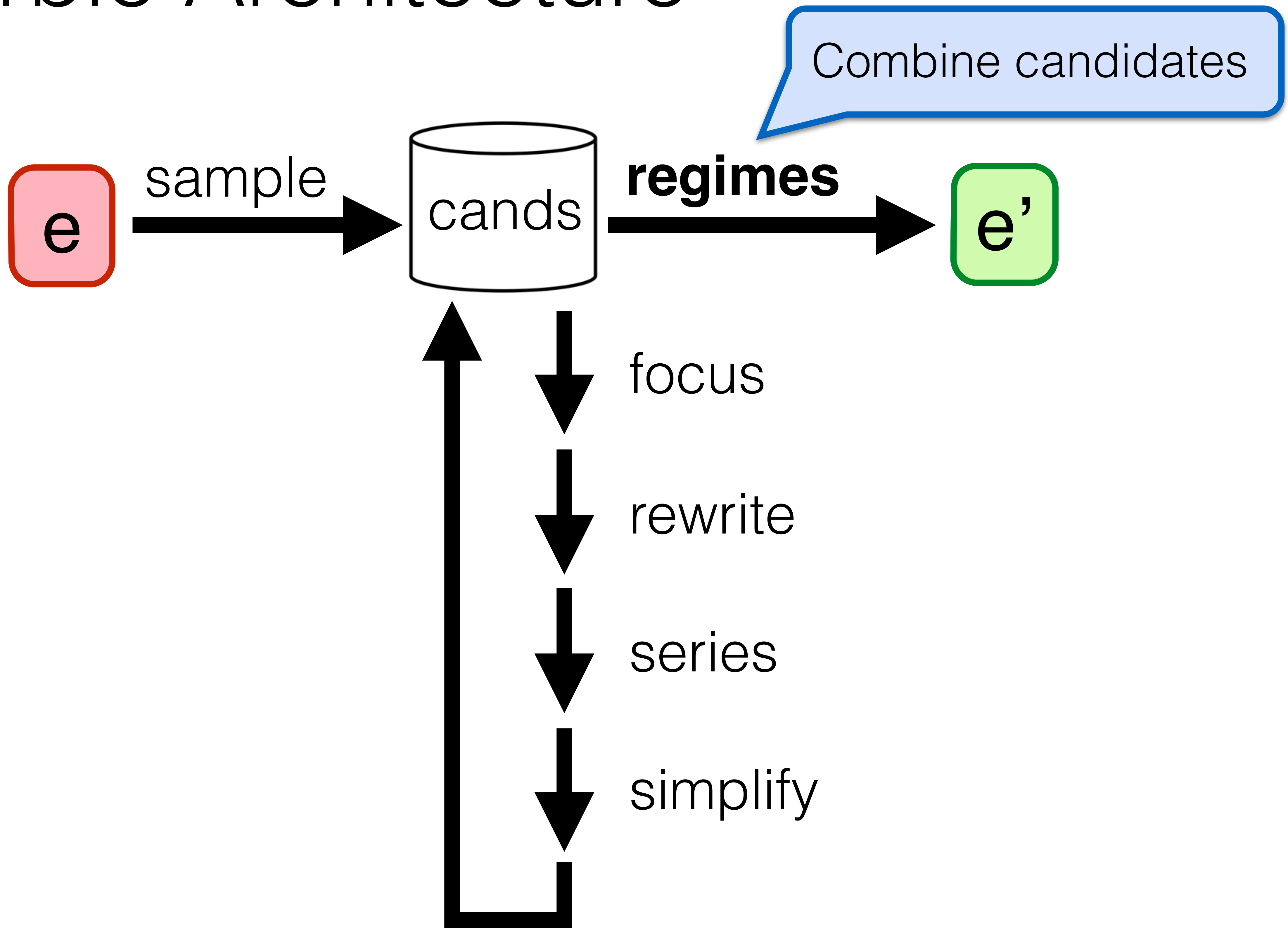
Herbie Architecture



Herbie Architecture



Herbie Architecture



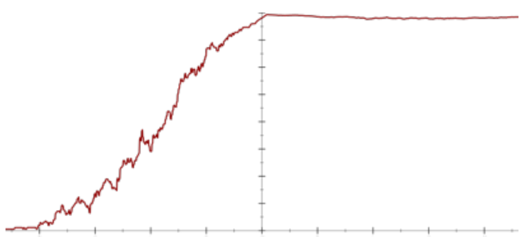
Regime Inference

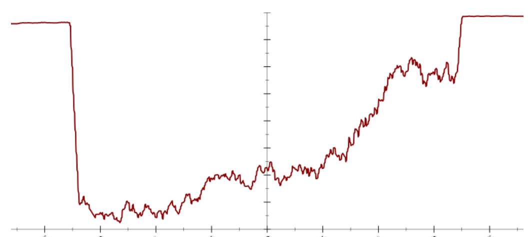
$$\frac{c}{b} - \frac{b}{a}$$

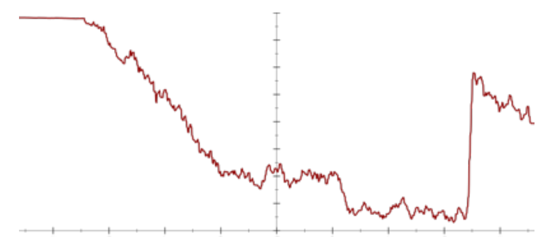
$$\frac{-b + \sqrt{b^2 - 4ac}}{2a}$$

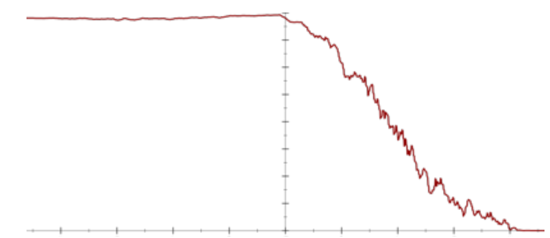
$$\frac{2c}{-b - \sqrt{b^2 - 4ac}}$$

$$-\frac{c}{b}$$





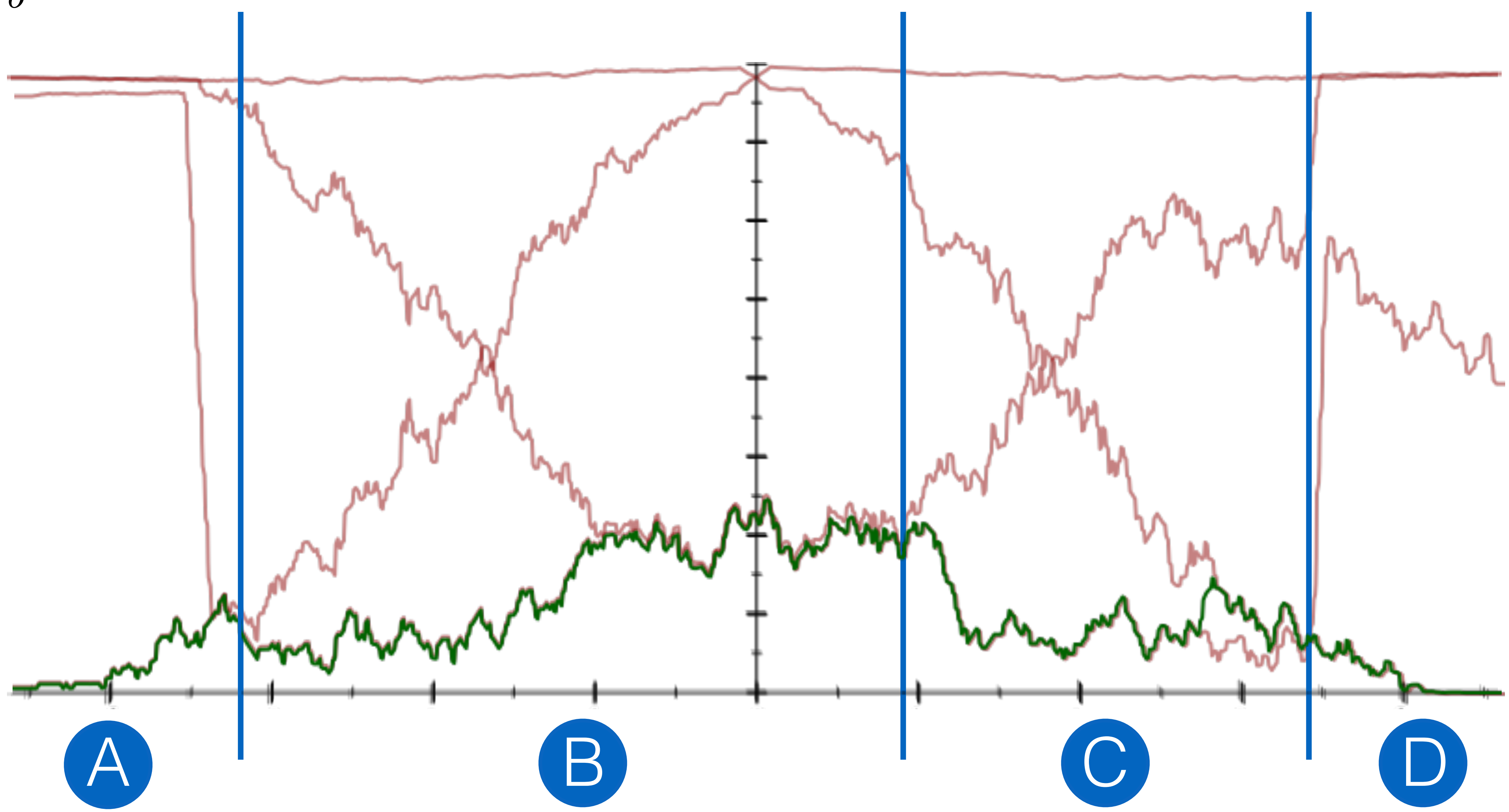




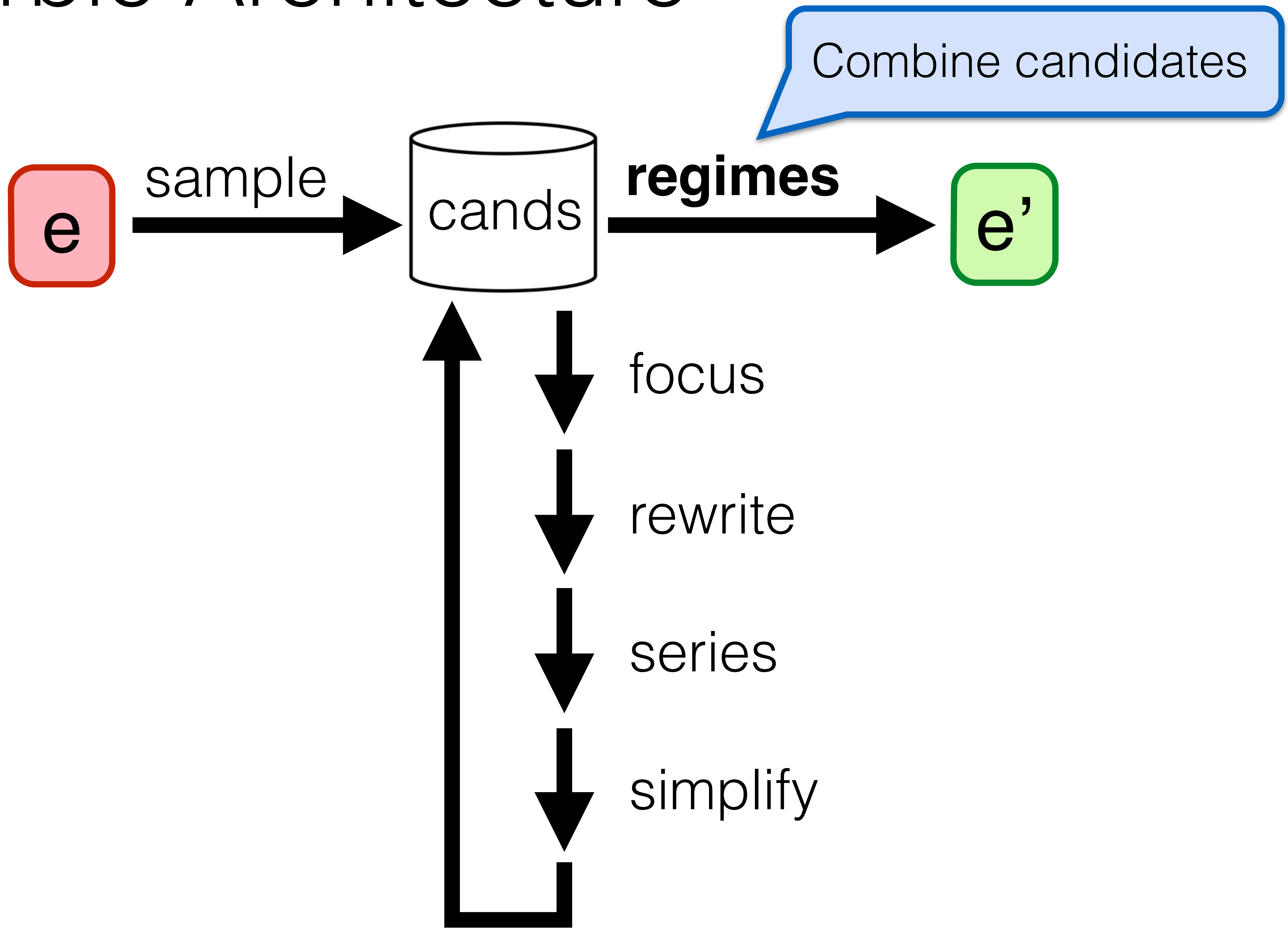
Regime Inference

$$\begin{cases} \frac{c}{b} - \frac{b}{a} \\ \frac{-b+\sqrt{b^2-4ac}}{2a} \\ \frac{2c}{-b-\sqrt{b^2-4ac}} \\ -\frac{c}{b} \end{cases}$$

Dynamic programming:
- Bounds quickly
- Tune: binary search
- Pick best variable



Herbie Architecture



Heuristic search to find expert transformations

Worked Example

How Herbie Works

Evaluation

Heuristic search to find expert transformations

Worked Example

How Herbie Works

Evaluation

Evaluating Herbie

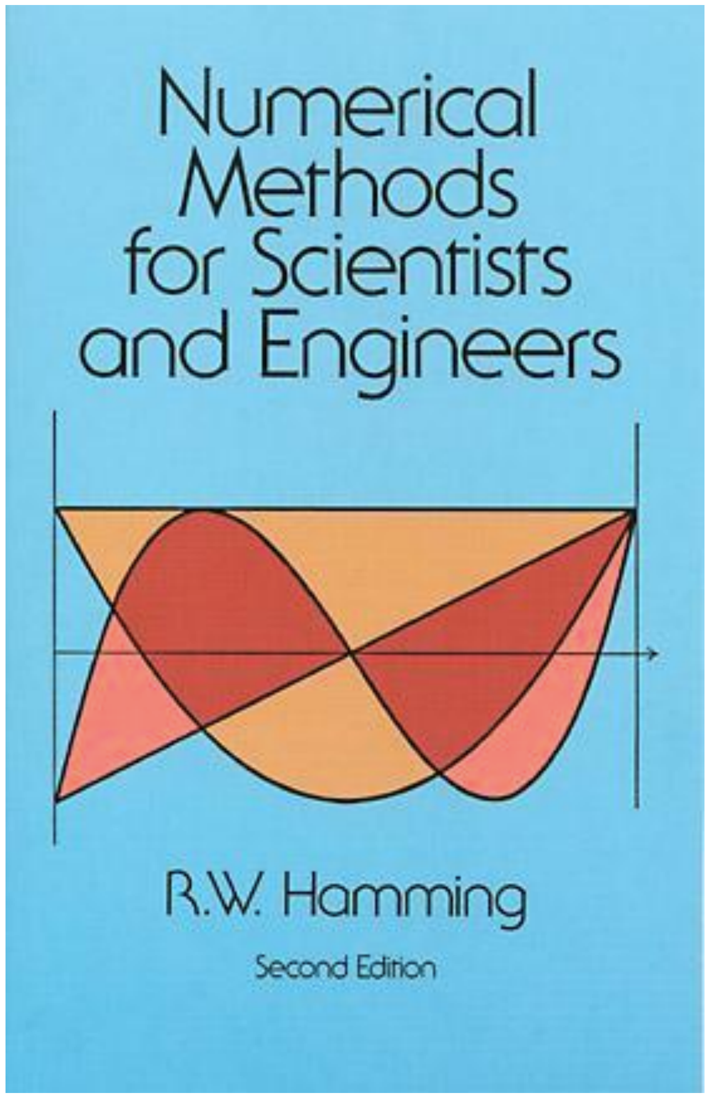
A. Does accuracy improve?

B. Does it reproduce expert transformations?

C. Is the output code fast?

D. Does it work in the real world?

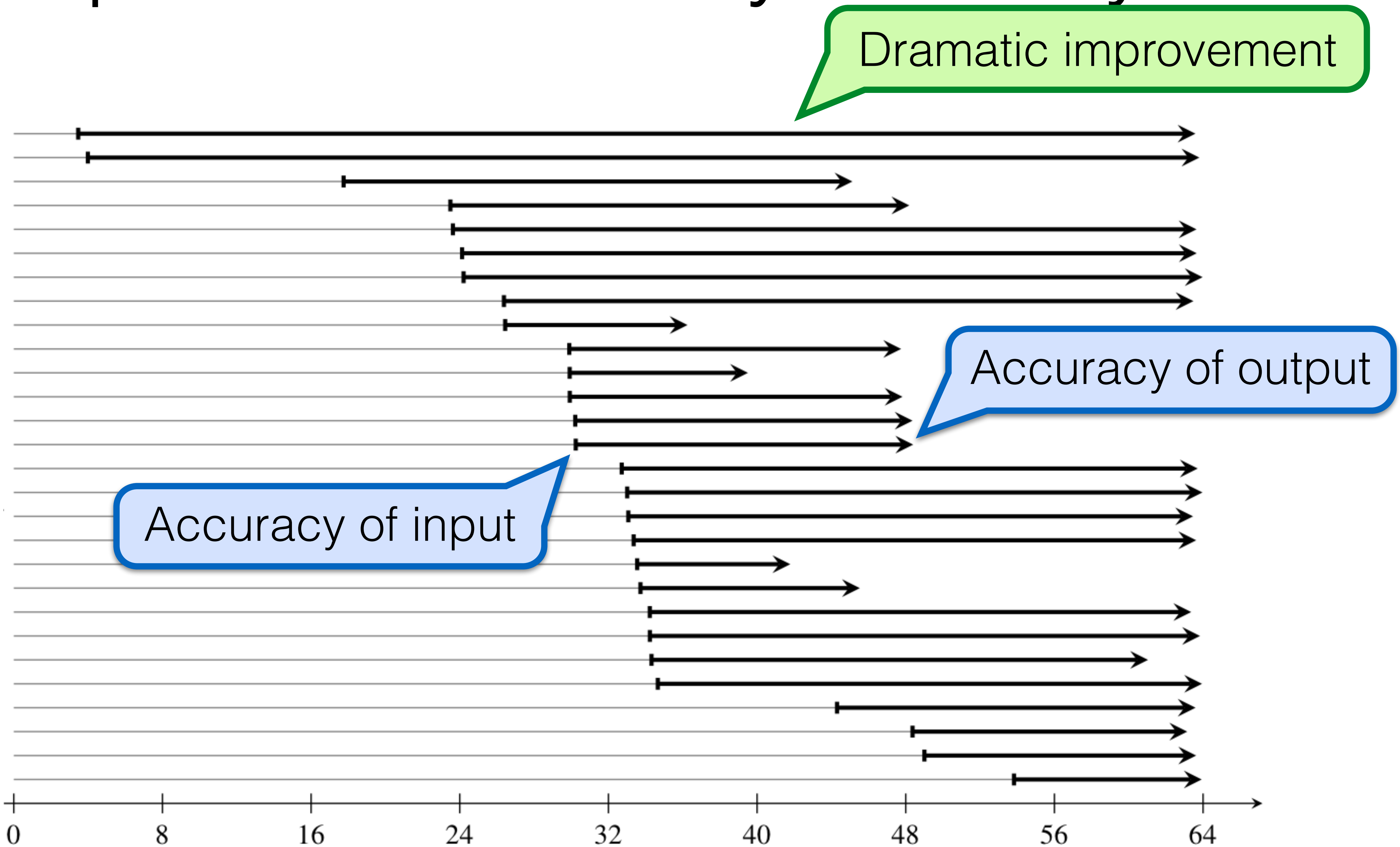
Examples from Hamming's *NMSE*



Chapter 3: Function evaluation
28 worked examples & problems

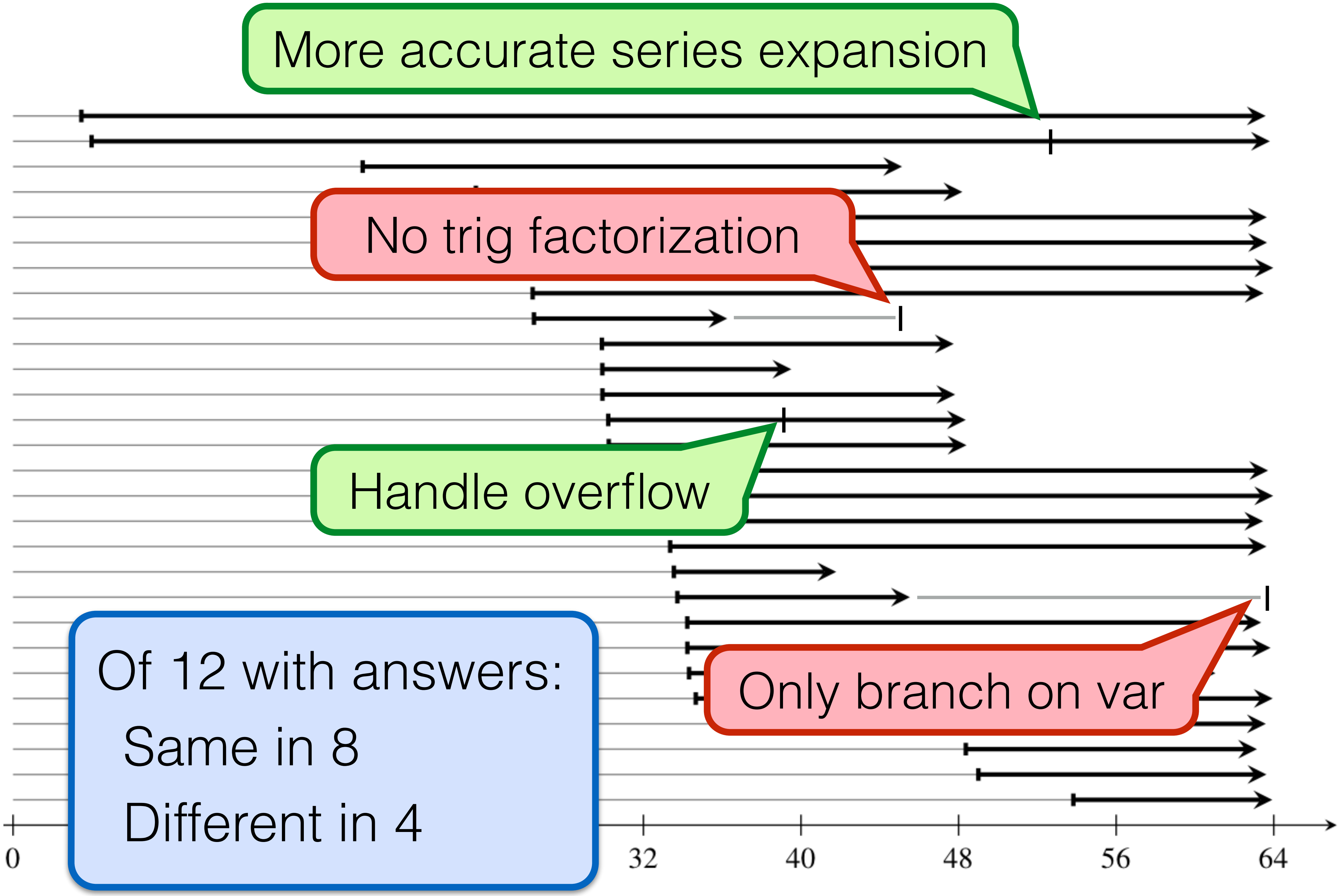
Quadratic formula (4)
Algebraic rearrangement (12)
Series expansion (12)
Branches and regimes (2)

A. Improves accuracy in every test



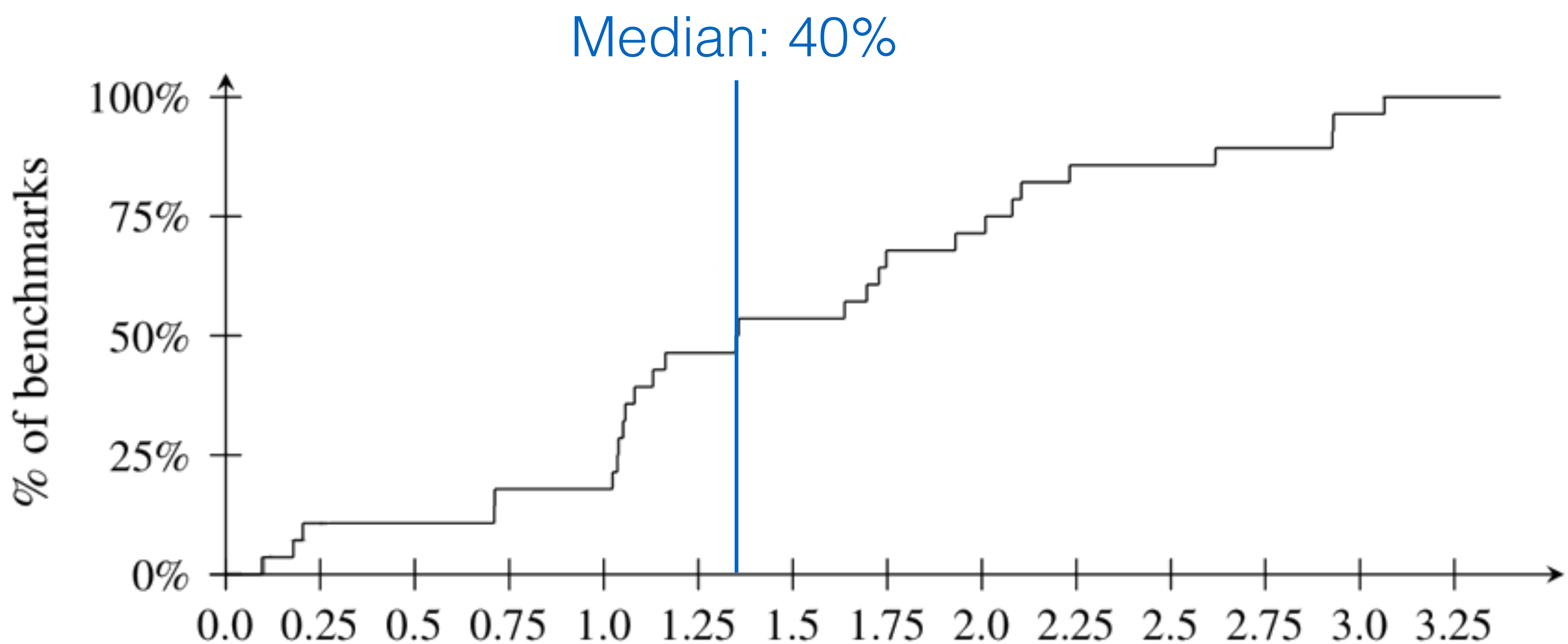
Average bits correct (longer is better)

B. Reproduces expert changes



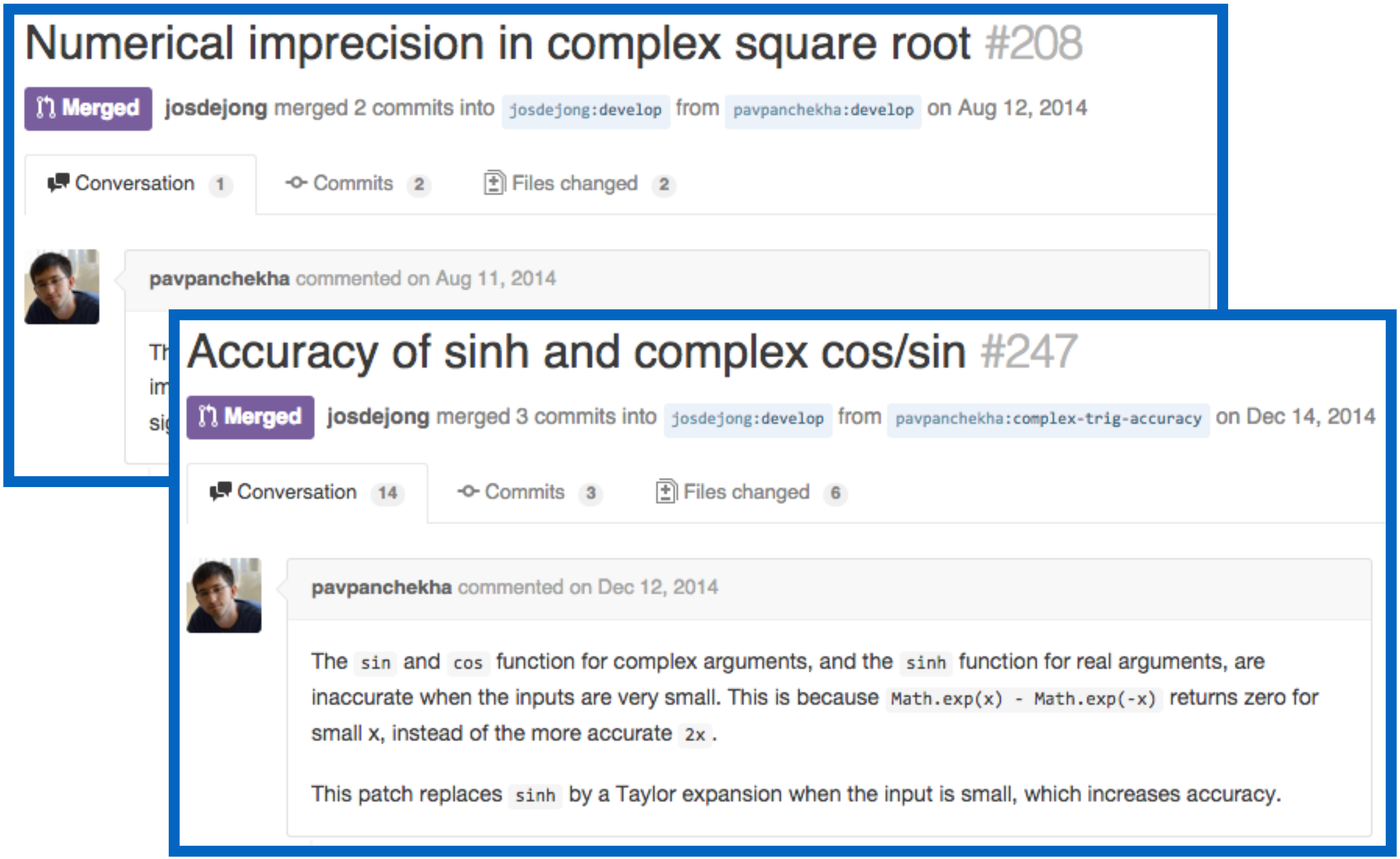
Average bits correct (longer is better)

C. Output code is fast



Overhead CDF
(left is better)

D. Two MathJS Patches Accepted



Numerical imprecision in complex square root #208

Merged josdejong merged 2 commits into `josdejong:develop` from `pavpanchekha:develop` on Aug 12, 2014

Conversation 1 | Commits 2 | Files changed 2

pavpanchekha commented on Aug 11, 2014

Th
im
si

Accuracy of sinh and complex cos/sin #247

Merged josdejong merged 3 commits into `josdejong:develop` from `pavpanchekha:complex-trig-accuracy` on Dec 14, 2014

Conversation 14 | Commits 3 | Files changed 6

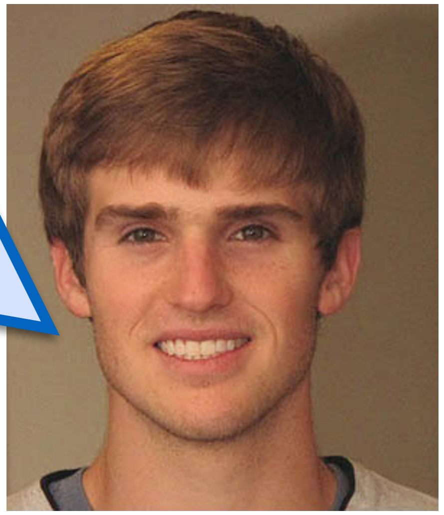
pavpanchekha commented on Dec 12, 2014

The `sin` and `cos` function for complex arguments, and the `sinh` function for real arguments, are inaccurate when the inputs are very small. This is because `Math.exp(x) - Math.exp(-x)` returns zero for small x, instead of the more accurate `2x`.

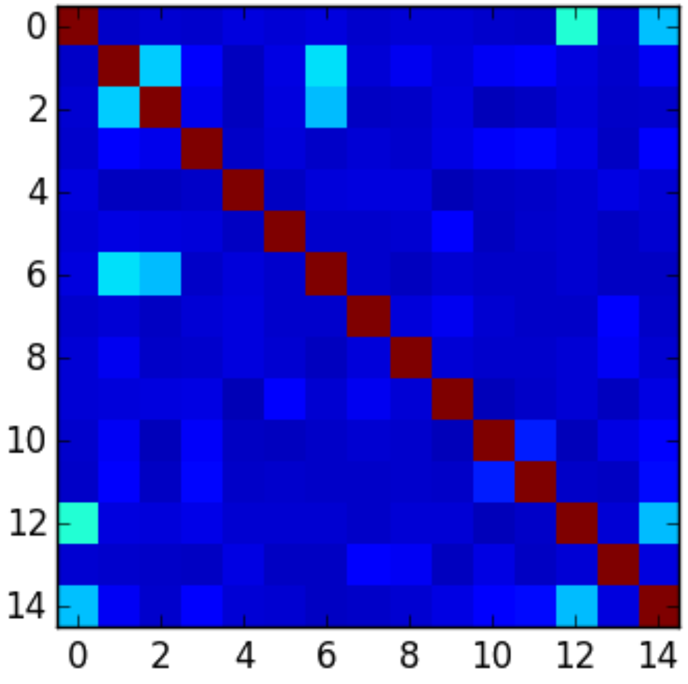
This patch replaces `sinh` by a Taylor expansion when the input is small, which increases accuracy.

D. Machine Learning Anecdote

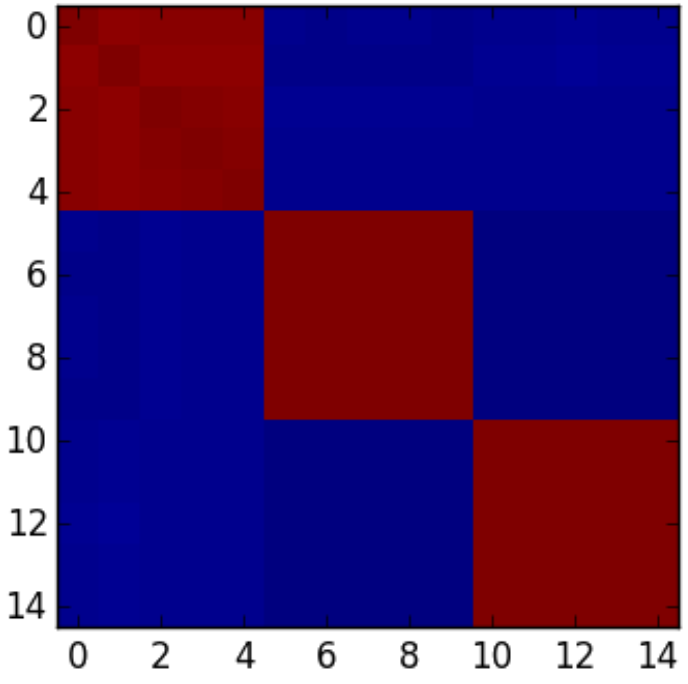
I wasn't sure how to best rewrite [my] equations. **Herbie found numerically stable versions of the formulas**, and fixed all the divide-by-zero errors.



Harley Montgomery



⇒



Clustering (bigger, darker blocks better)

Heuristic search to find expert transformations

Worked Example

How Herbie Works

Evaluation

Improve accuracy of floating point programs

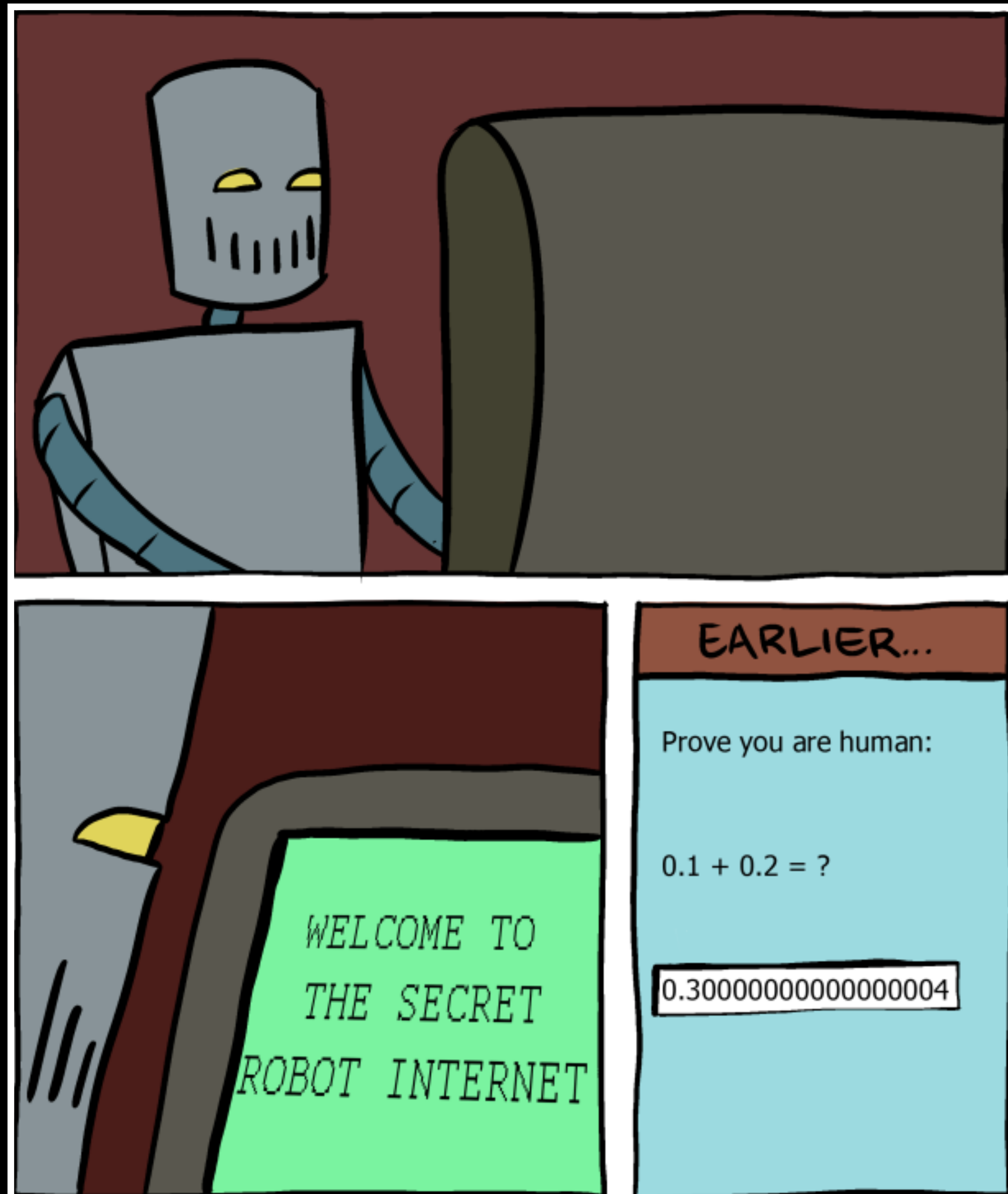
Sampling to estimate error

Reduce global error to per-operation error

Iterative rewriting highest-error operations

Different expressions for different inputs

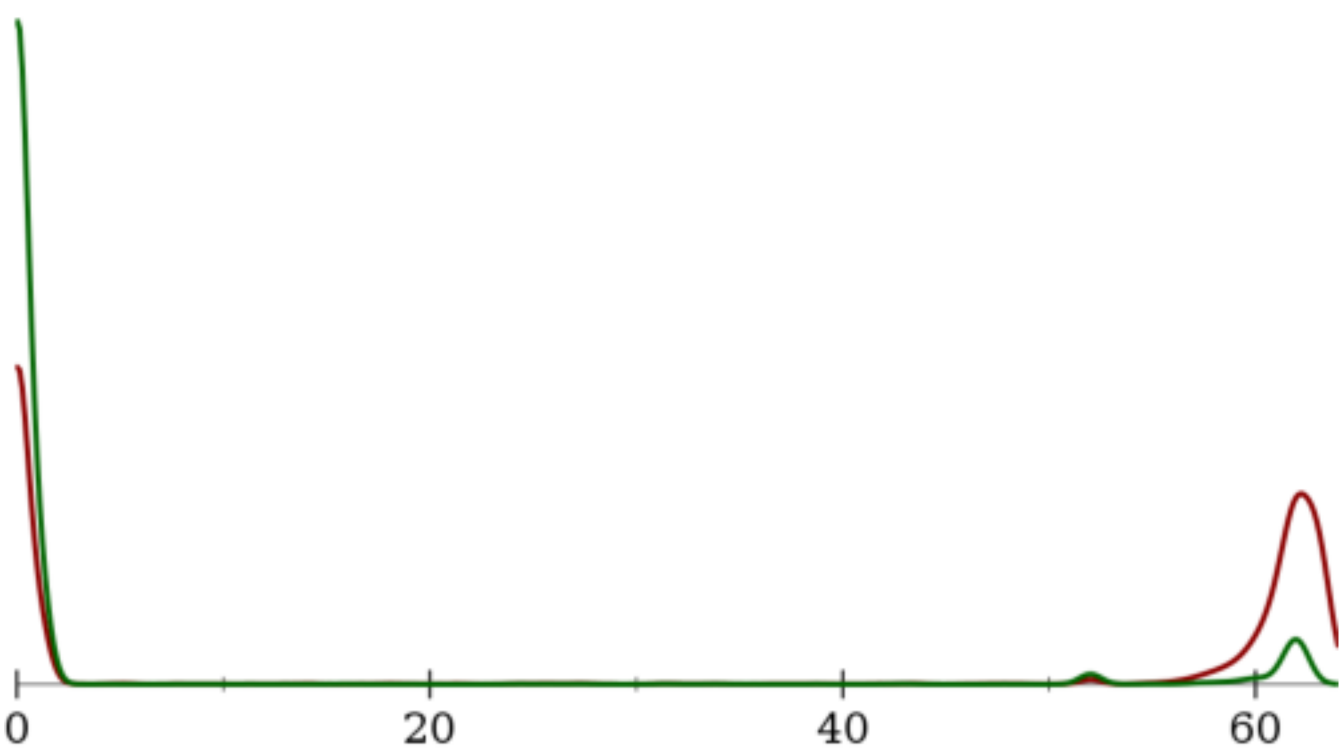
http://herbie.uwplse.org/


WELCOME TO THE SECRET ROBOT INTERNET
EARLIER...
Prove you are human:
0.1 + 0.2 = ?
0.30000000000000004

Herbie and Maximum Error

Often improved by Herbie:

- Improvements large (28b) and small (.5b)
- 1+b improvement for 10/28 programs



Bits error (histogram)

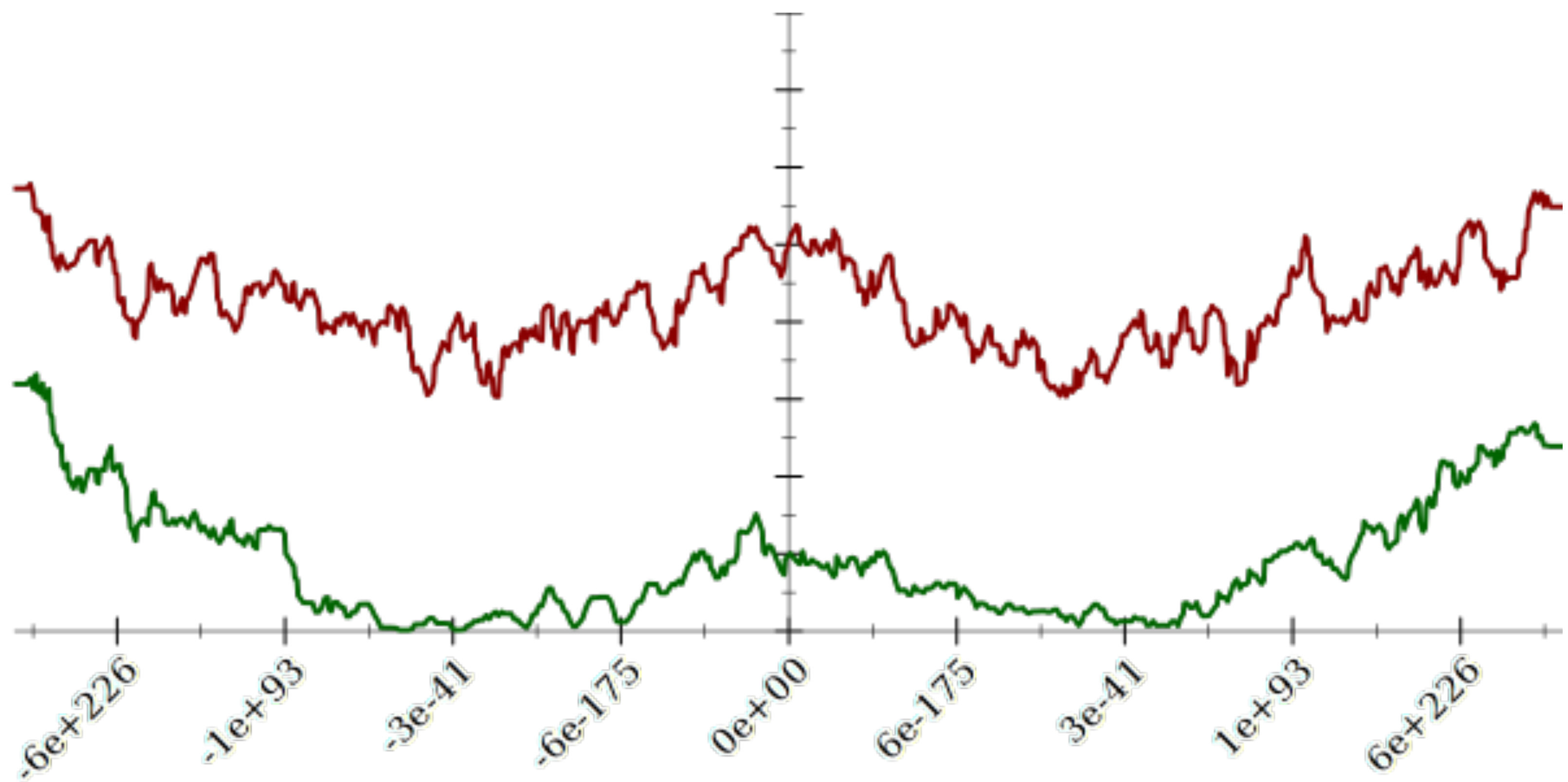
Fewer high-error pts, same max error.

Herbie as Part of a Pipeline

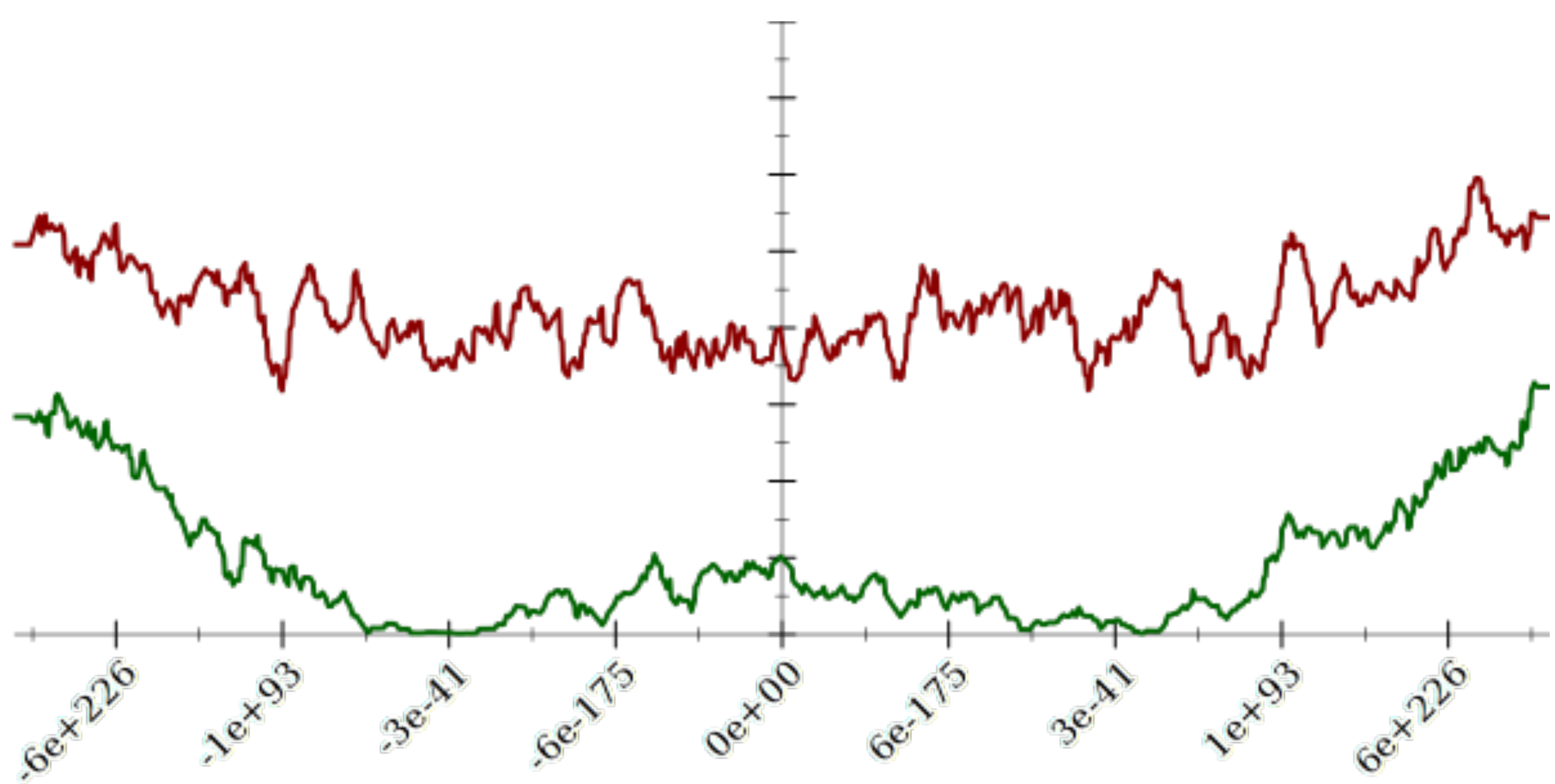
FPDebug	Find inaccurate expressions
Herbie	Improve accuracy
Rosa FPTaylor	Prove accuracy satisfactory
STOKE-FP	Optimize code

Error graphs along a and c

a



c



Finding the rewrite rules

Standard mathematical identities:

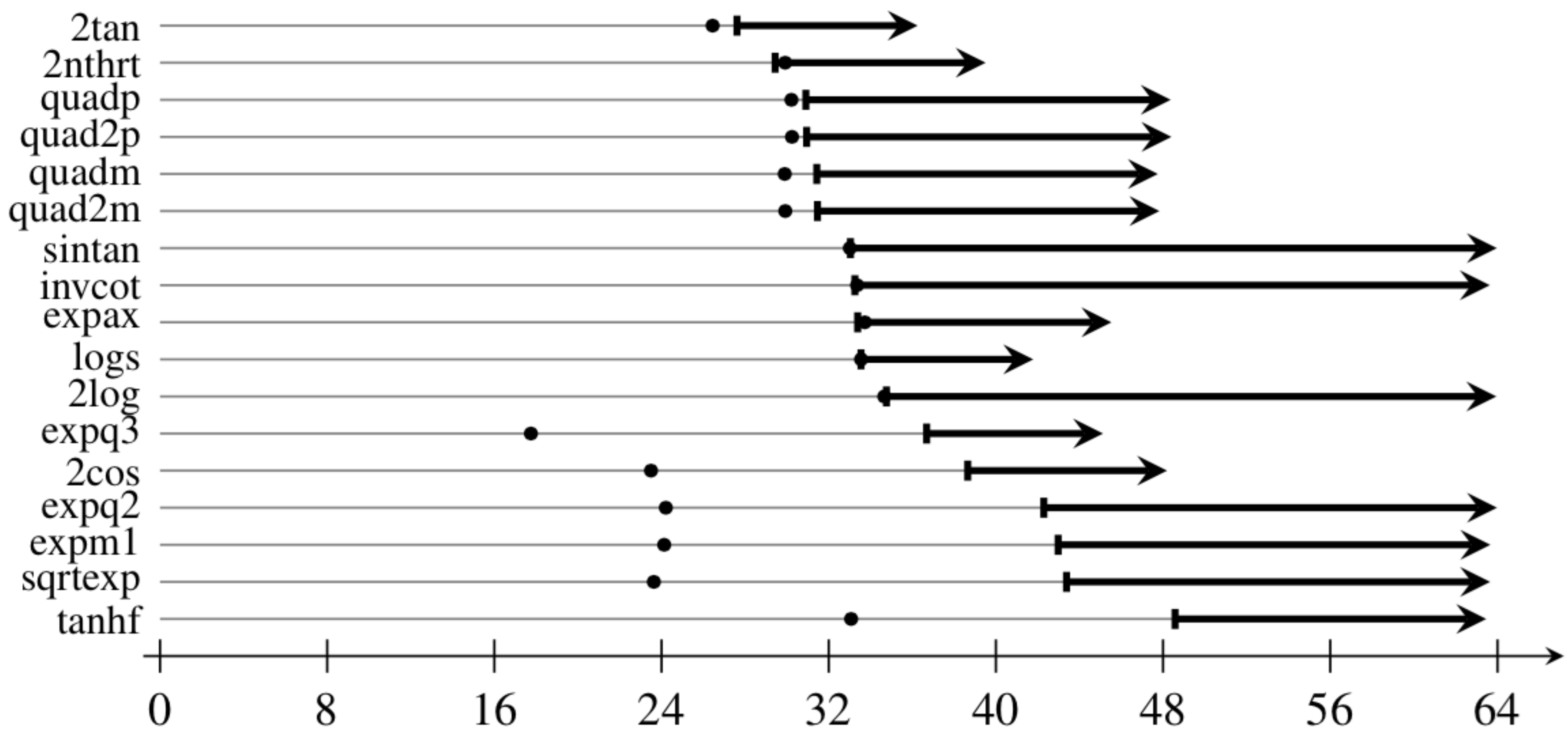
- Commutativity, inverses, fractions, trig identities

No numerical methods knowledge

Don’t need to be true identities

- False rules do not improve accuracy
- Herbie will ignore them

Regimes often gains ~15 bits



Improvement from regimes (longer is better)

Dot : input program average accurage

Bar : Herbie result w/out regimes