# Optimizing Synthesis with Metasketches



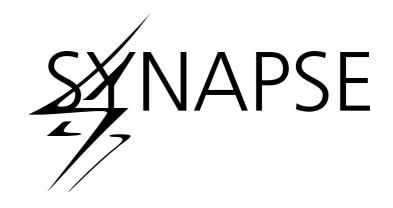
James Bornholt

Emina Torlak

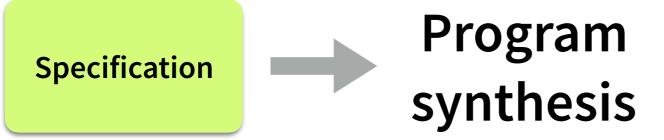
Dan Grossman

Luis Ceze

University of Washington



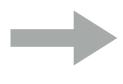
Program synthesis



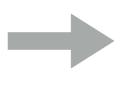




f(x) = 4x



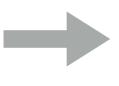
# Program synthesis



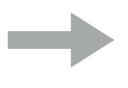
Program



f(x) = 4x



# Program synthesis



Program

 $\chi + \chi + \chi + \chi$ 

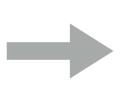
Compilation
[PLDI'14]

Data Structures [PLDI'15] End-user Programming
[POPL'11]



$$f(x) = 4x$$

# Program synthesis



**Program** 

X+X+X+X

Executable Biology [POPL'13]

Browser Layout
[PPOPP'13]

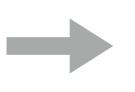
Cache Protocols [PLDI'13] Compilation
[PLDI'14]

Data Structures
[PLDI'15]

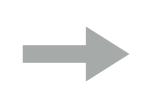
End-user Programming
[POPL'11]

Often looking for an optimal solution, not just any correct program

**Specification** 



Program synthesis



**Program** 

Executable Biology [POPL'13]

Browser Layout
[PPOPP'13]

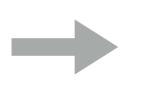
Cache Protocols [PLDI'13] Compilation
[PLDI'14]

Data Structures
[PLDI'15]

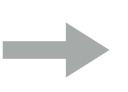
End-user Programming
[POPL'11]

Often looking for an optimal solution, not just any correct program

**Specification** 



Program synthesis



**Program** 

There are *many* programs, so tools must control search strategy

Executable Biology [POPL'13]

Browser Layout
[PPOPP'13]

Cache Protocols [PLDI'13]

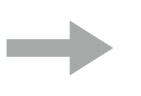




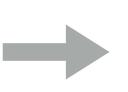


Often looking for an optimal solution, not just any correct program

**Specification** 



Program synthesis



**Program** 

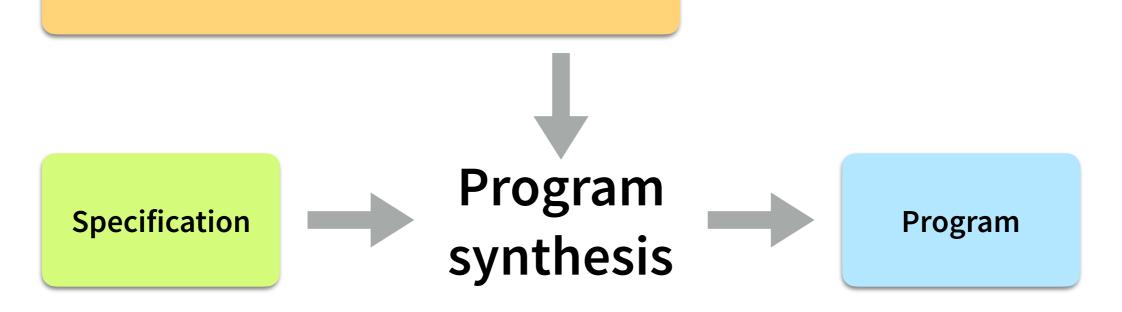
There are *many* programs, so tools must control search strategy



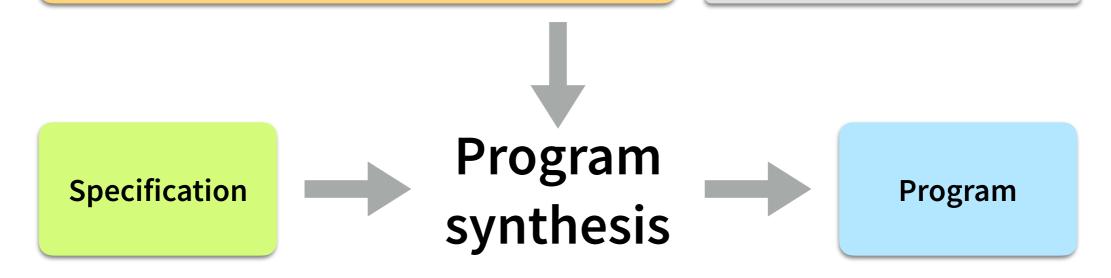








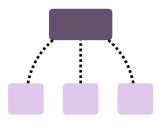
A framework that makes search strategy and optimality part of the problem definition





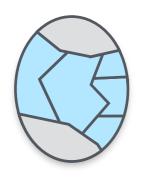


Design and structure

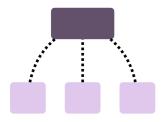


# **Synapse**

A metasketch solver



Design and structure



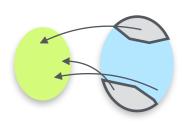
#### **Synapse**

A metasketch solver



Results

Better solutions, faster



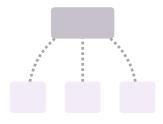
#### Background

Syntax-guided synthesis



#### Metasketches

Design and structure



#### Synapse

A metasketch solver



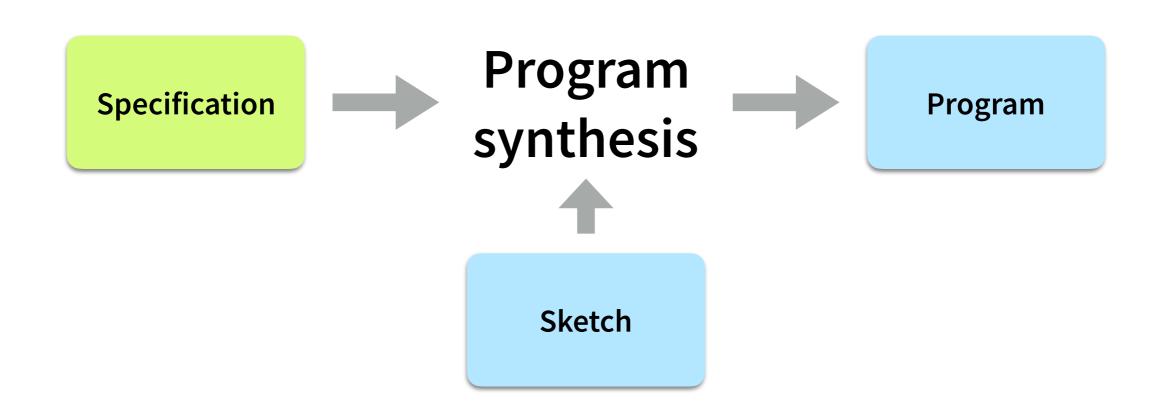
#### Results

Better solutions, faster

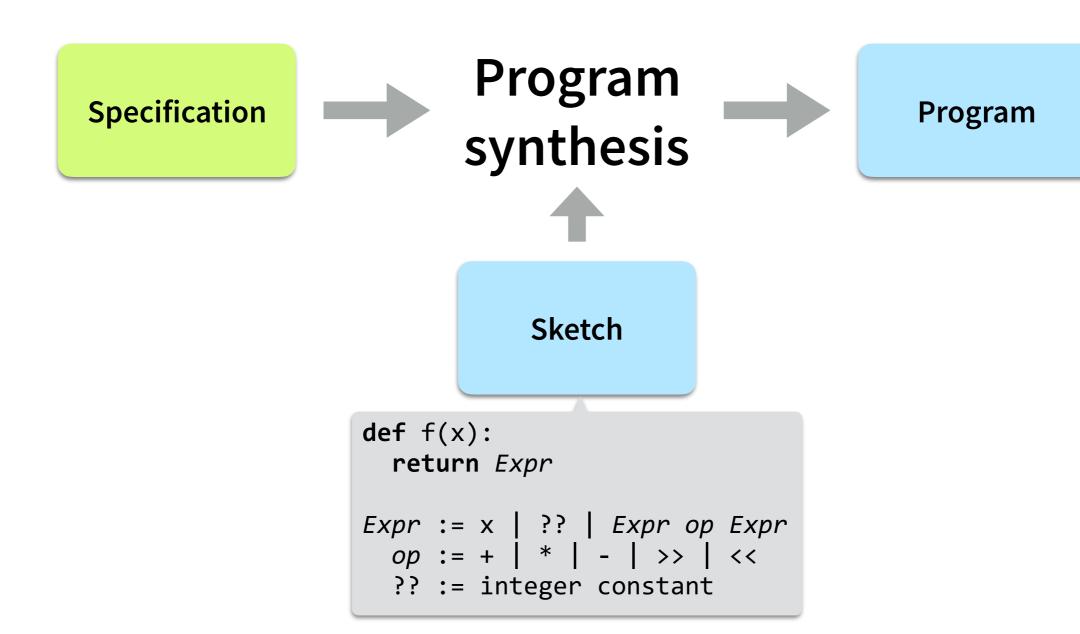
# Syntax-guided synthesis



# Syntax-guided synthesis



# Syntax-guided synthesis



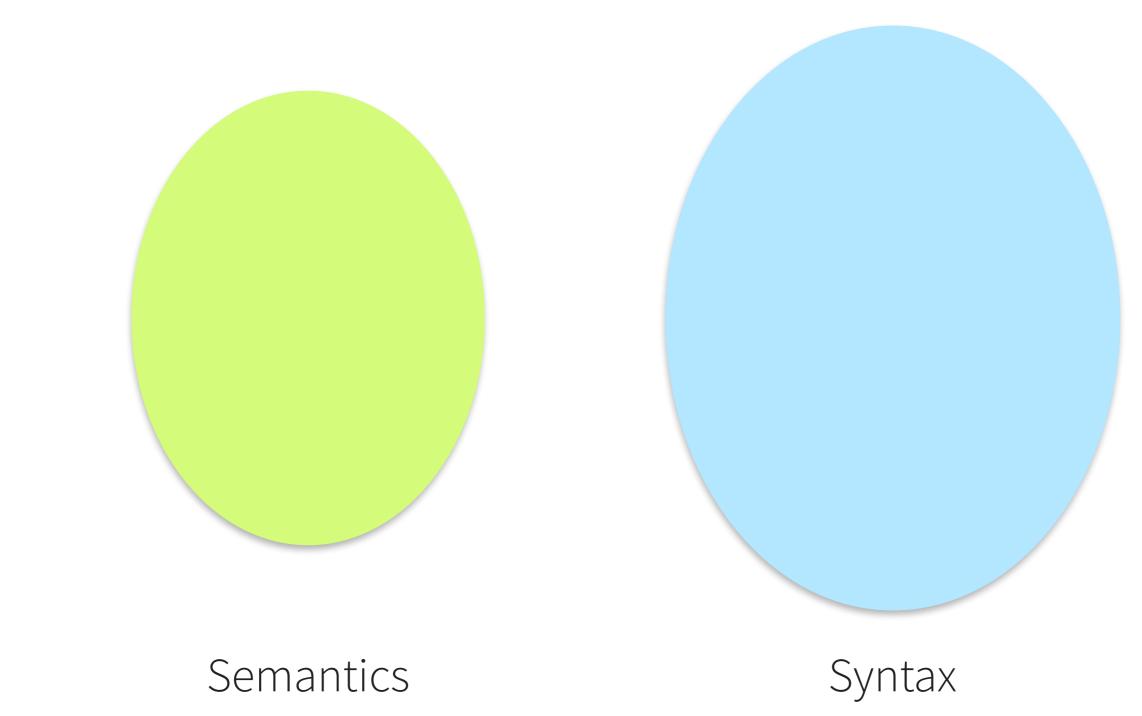
```
def f(x):
    return Expr

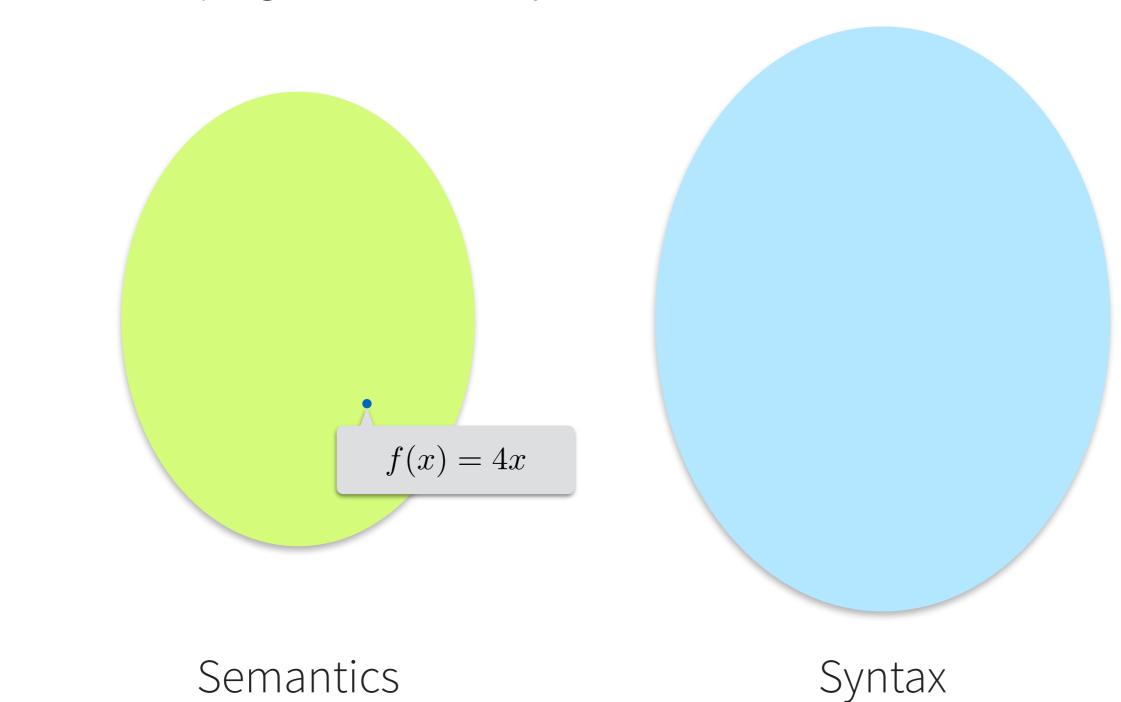
Expr := x | ?? | Expr op Expr
    op := + | * | - | >> | <<
        ?? := integer constant</pre>
```

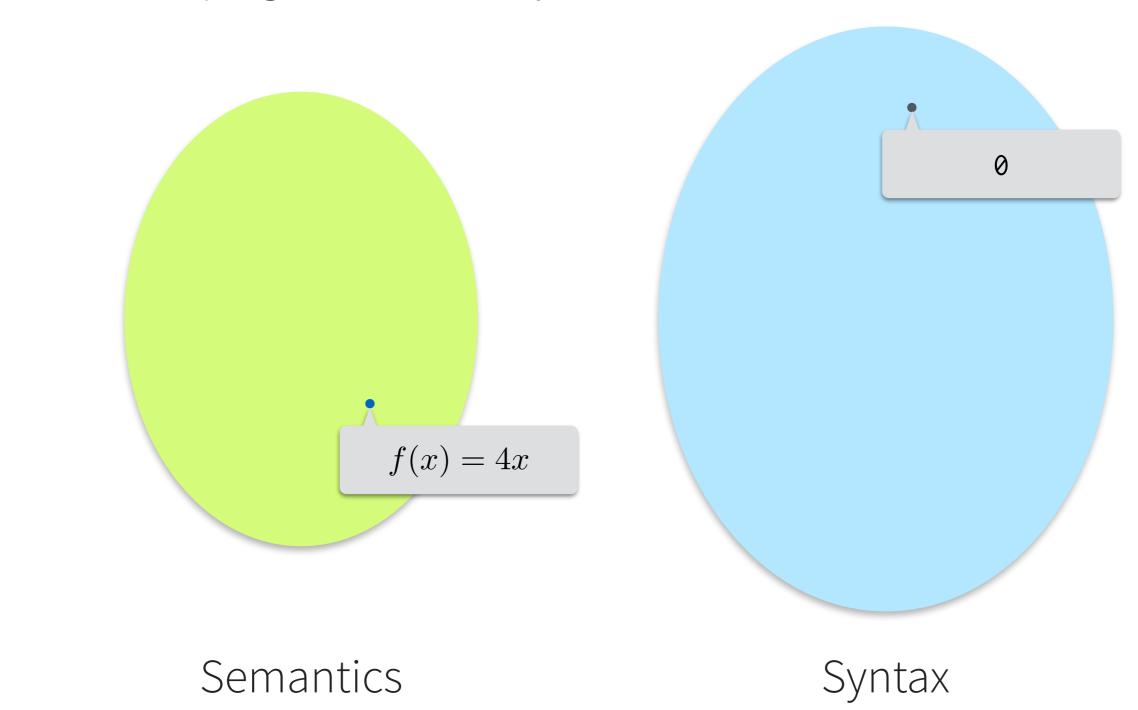
```
def f(x):
    return Expr

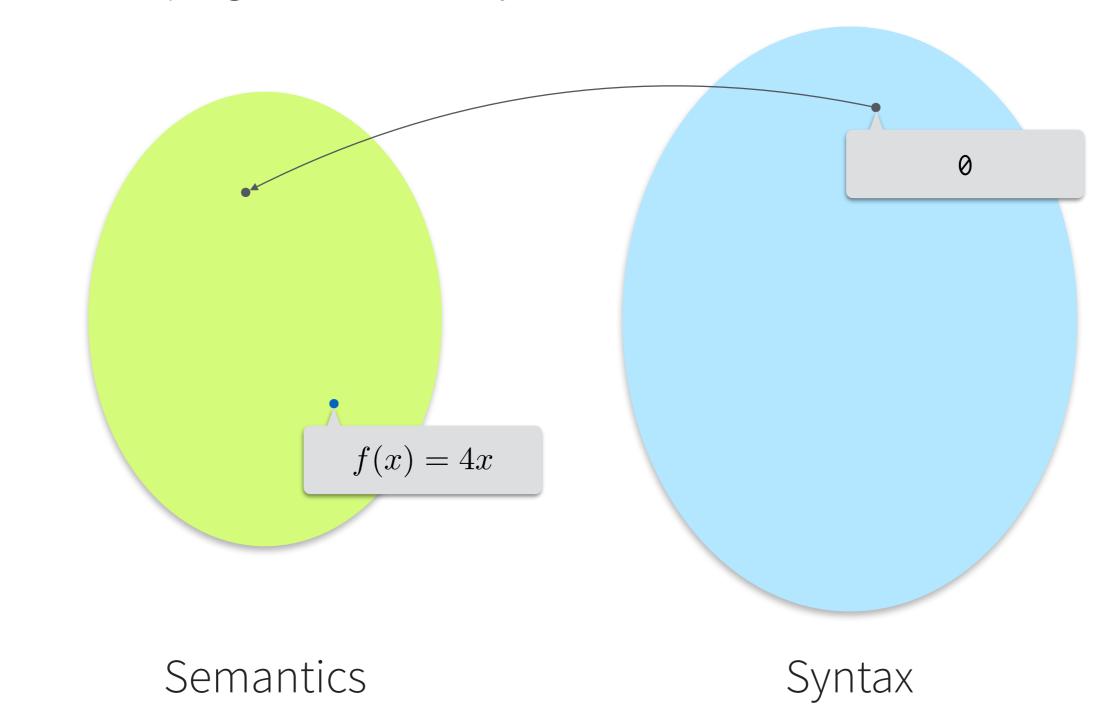
Expr := x | ?? | Expr op Expr
    op := + | * | - | >> | <<
        ?? := integer constant</pre>
```

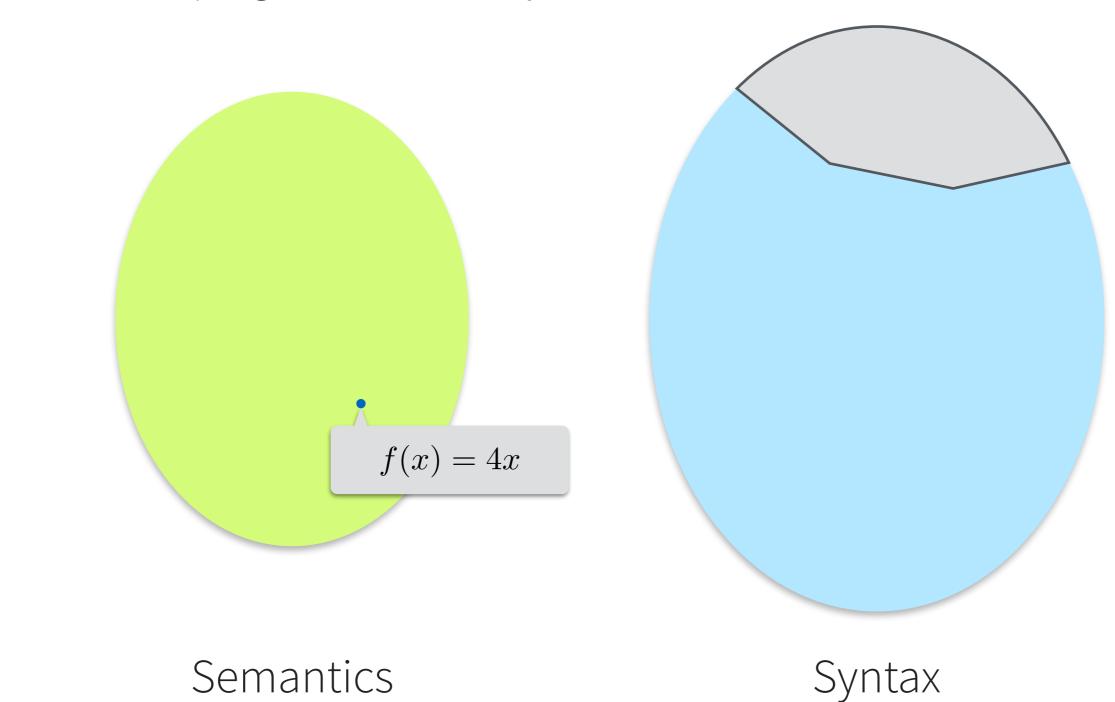


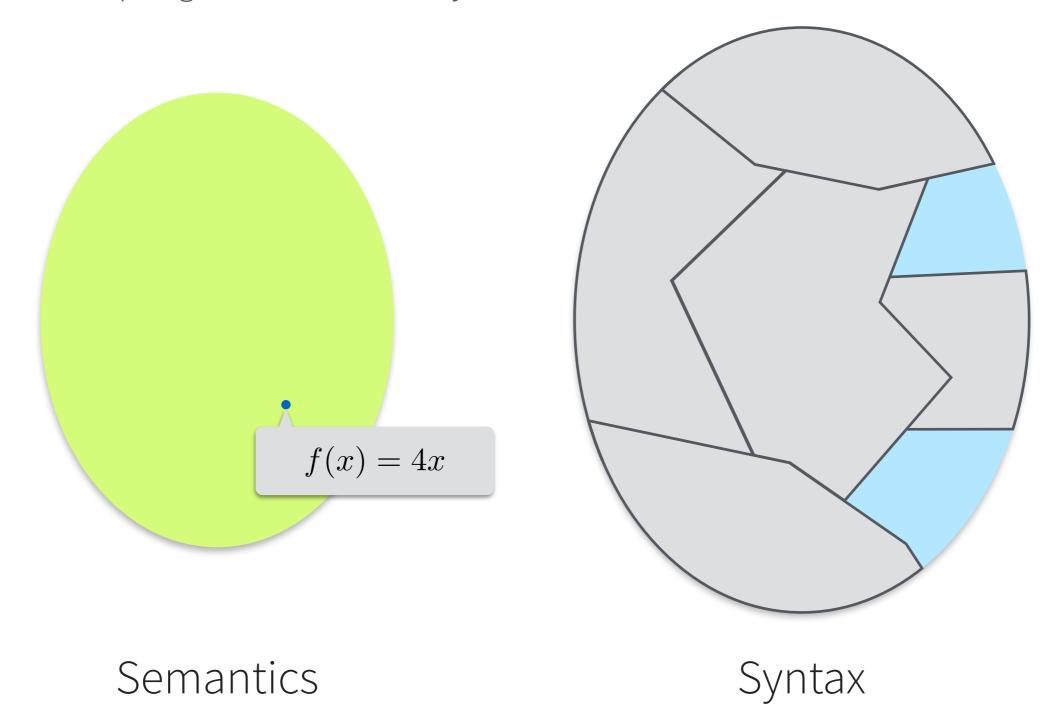


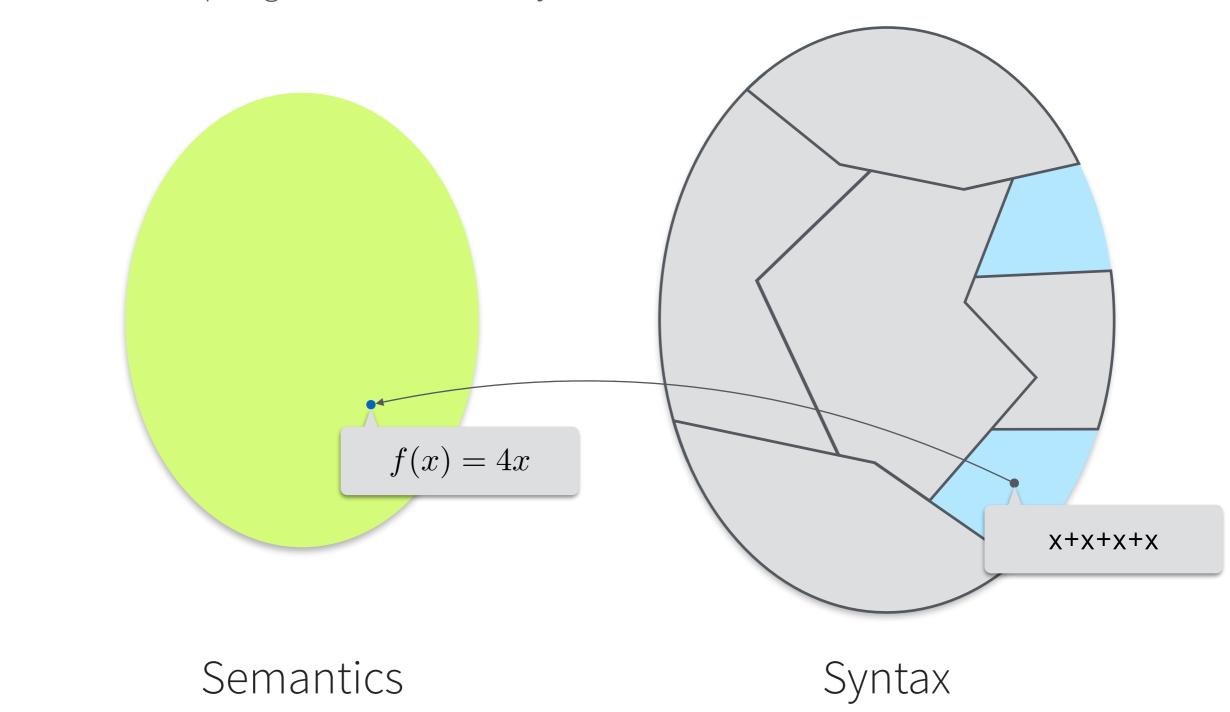






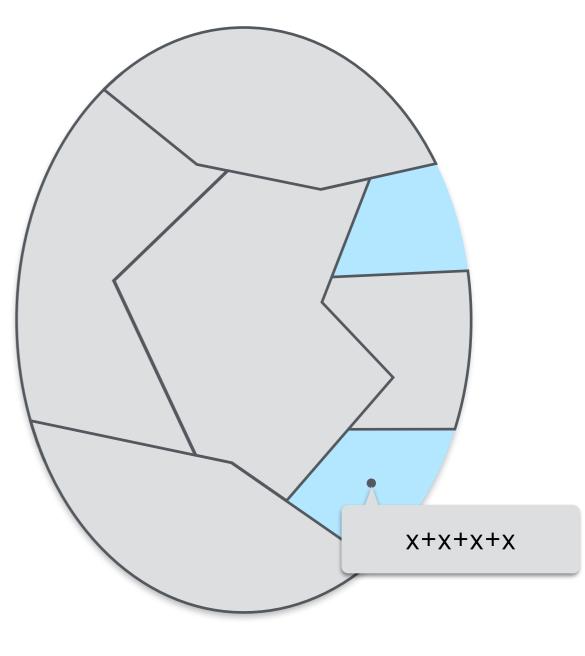






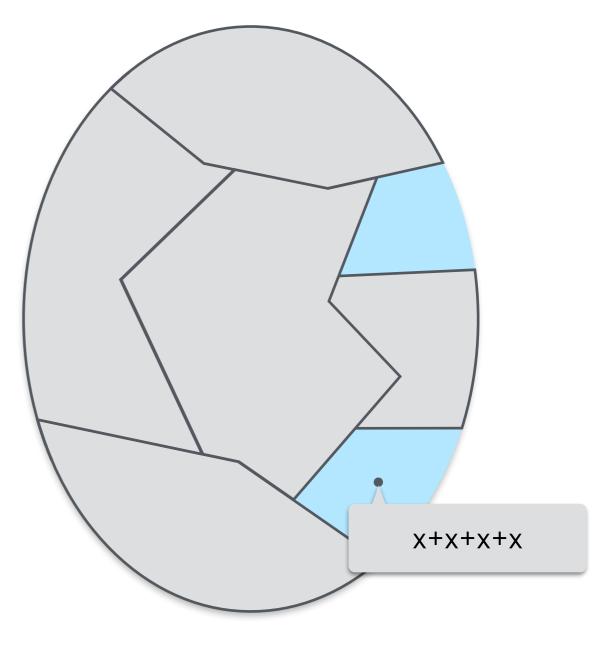
Counterexample-guided inductive synthesis [Solar-Lezama et al, 2006]

1. **Search order** is critical



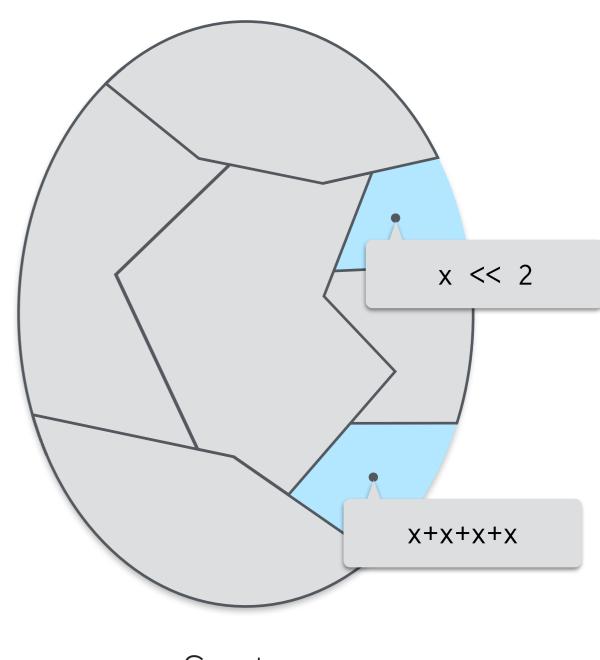
Syntax

- 1. **Search order** is critical
- 2. Desire **optimal** solutions

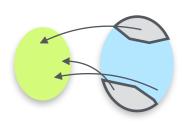


Syntax

- 1. **Search order** is critical
- 2. Desire **optimal** solutions



Syntax



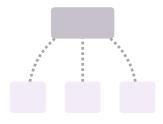
#### Background

Syntax-guided synthesis



#### Metasketches

Design and structure



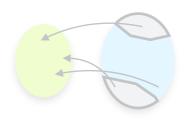
#### Synapse

A metasketch solver



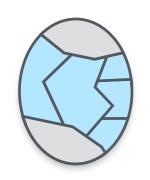
#### Results

Better solutions, faster



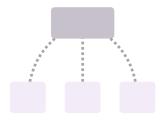
# Background

Syntax-guided synthesis



#### Metasketches

Design and structure



Synapse

A metasketch solver

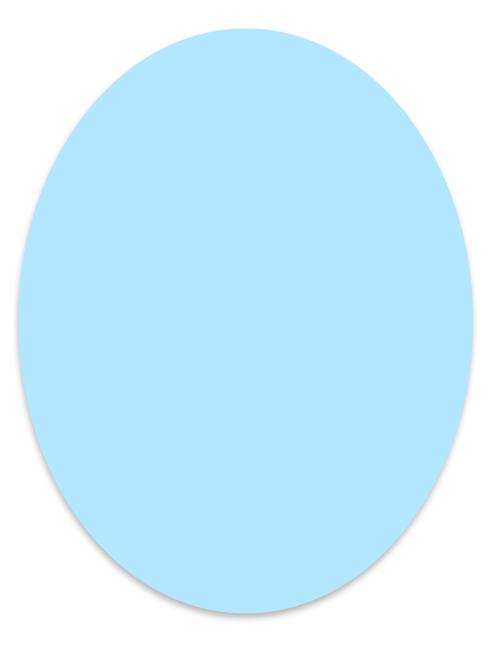


Results

Better solutions, faster

#### Metasketches express structure and strategy

- 1. **Search order** is critical
- 2. Desire **optimal** solutions



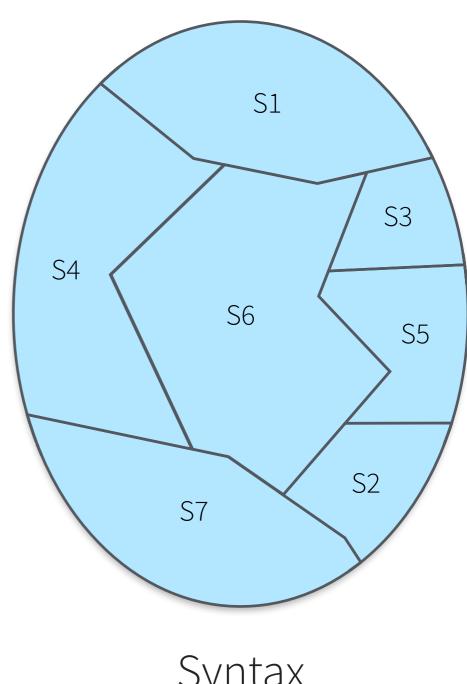
Syntax

# Metasketches express structure and strategy

- 1. **Search order** is critical
- 2. Desire **optimal** solutions

#### A metasketch contains:

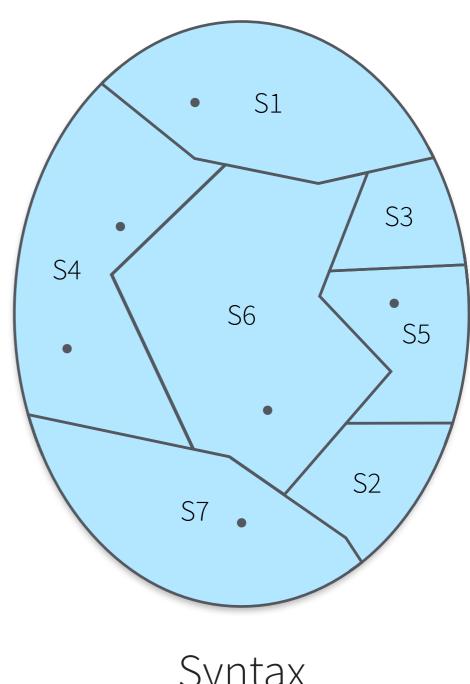
1. structured candidate space  $(S, \leq)$ 



Syntax

- 1. **Search order** is critical
- 2. Desire **optimal** solutions

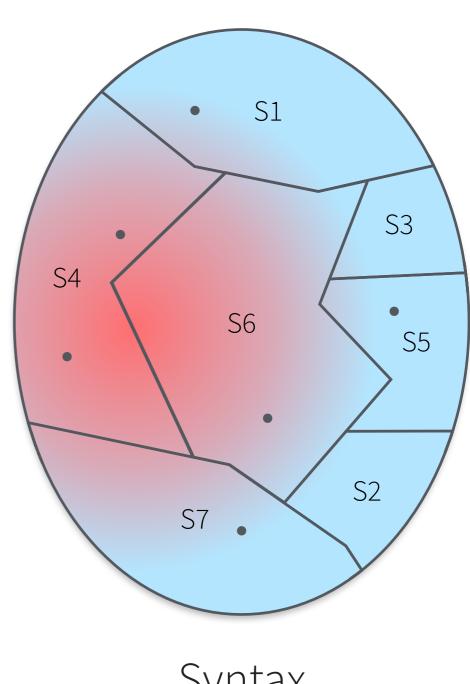
- 1. structured candidate space  $(S, \leq)$
- 2. cost function (κ)



Syntax

- 1. **Search order** is critical
- Desire **optimal** solutions

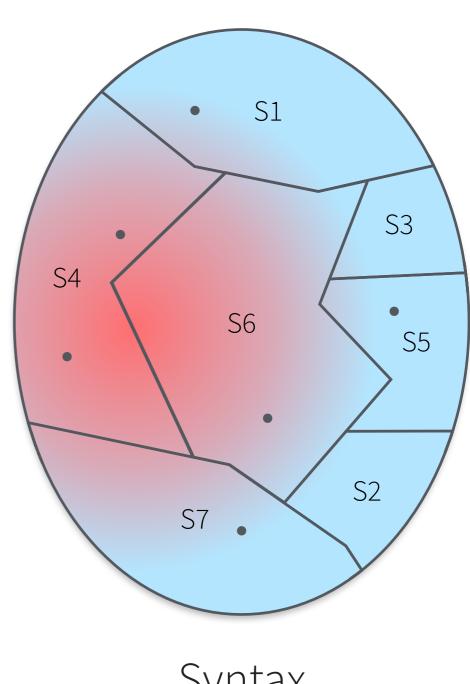
- 1. structured candidate space  $(S, \leq)$
- 2. cost function (k)
- 3. gradient function (g)



Syntax

- 1. **Search order** is critical
- Desire **optimal** solutions

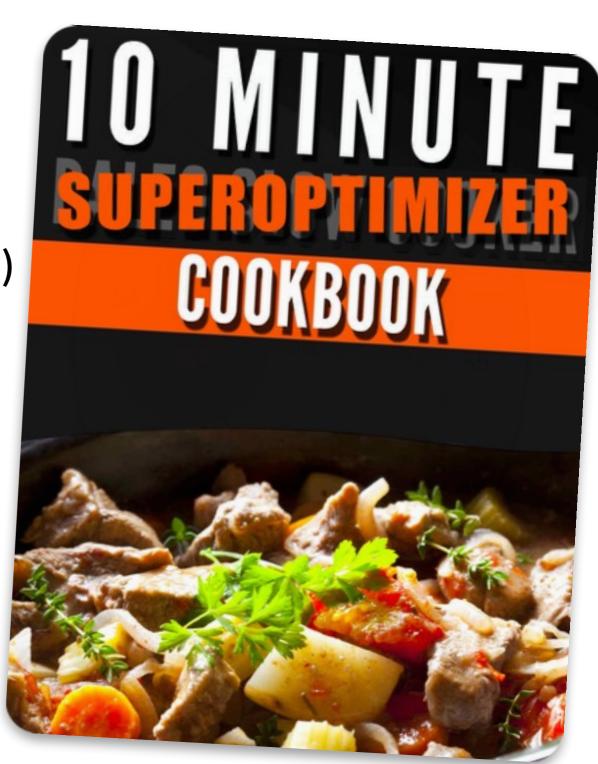
- 1. structured candidate space  $(S, \leq)$
- 2. cost function (k)
- 3. gradient function (g)



Syntax

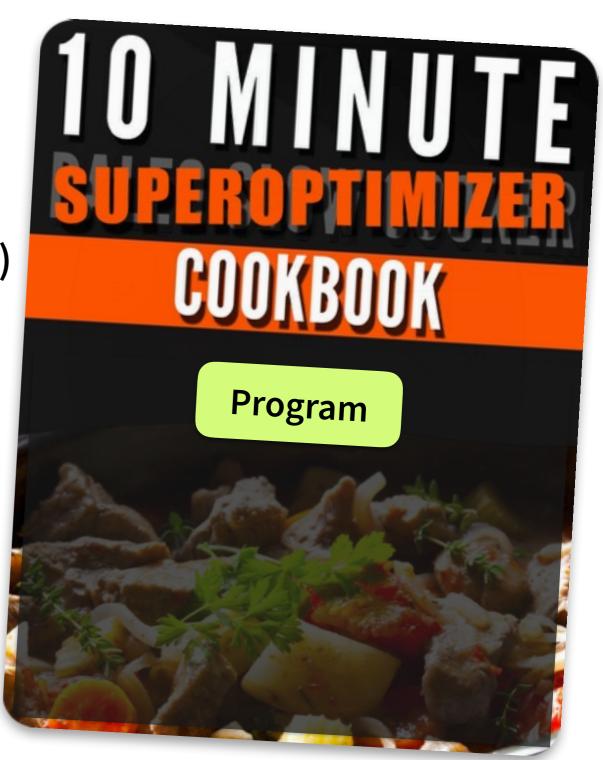
- 1. **Search order** is critical
- 2. Desire **optimal** solutions

- 1. structured candidate space  $(S, \leq)$
- 2. cost function (k)
- 3. gradient function (g)



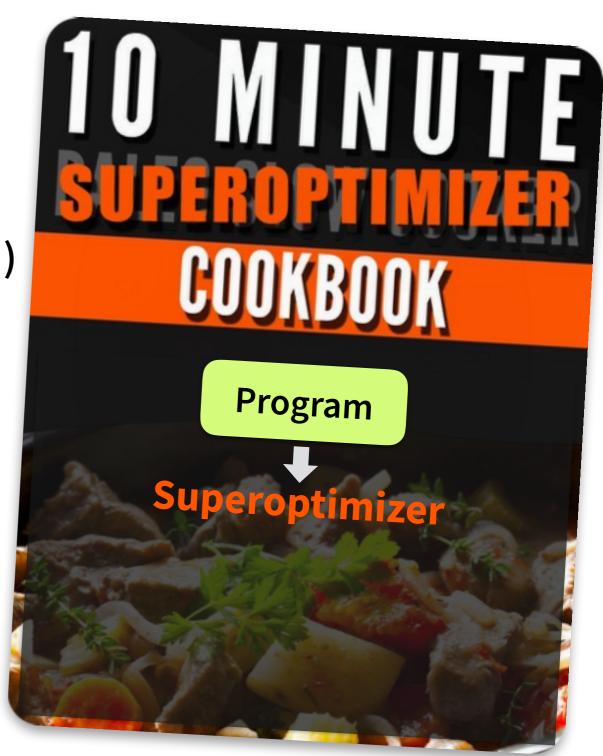
- 1. **Search order** is critical
- 2. Desire **optimal** solutions

- 1. structured candidate space  $(S, \leq)$
- 2. cost function (k)
- 3. gradient function (g)



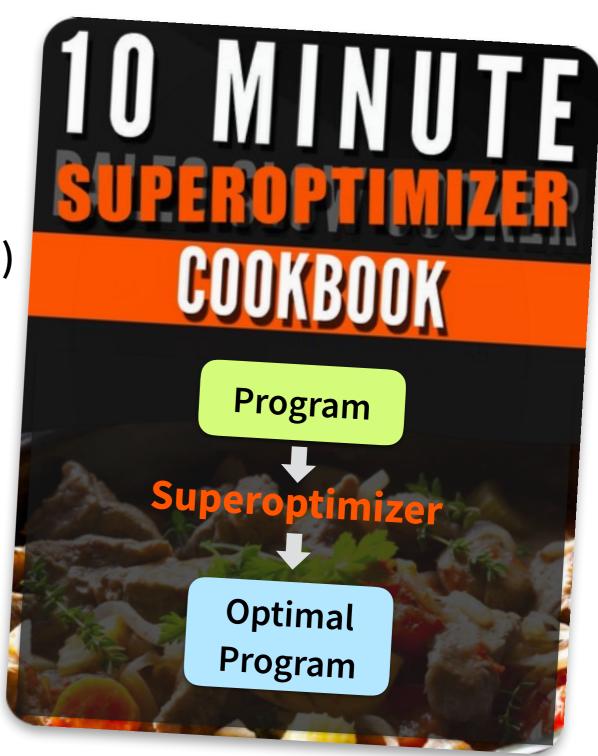
- 1. **Search order** is critical
- 2. Desire **optimal** solutions

- 1. structured candidate space  $(S, \leq)$
- 2. cost function (k)
- 3. gradient function (g)



- 1. **Search order** is critical
- 2. Desire **optimal** solutions

- 1. structured candidate space  $(S, \leq)$
- 2. cost function (k)
- 3. gradient function (g)



1. structured candidate space  $(S, \leq)$ 

- 2. cost function (K)
- 3. gradient function (g)

#### 1. structured candidate space $(S, \leq)$

A fragmentation of the candidate space, and an ordering on those fragments.

- 2. cost function (κ)
- 3. gradient function (g)

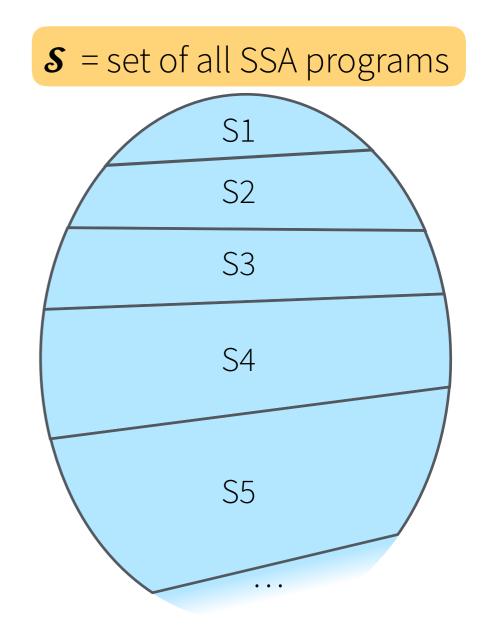
#### 1. structured candidate space $(S, \leq)$

A fragmentation of the candidate space, and an ordering on those fragments.

- 2. cost function (к)
- 3. gradient function (g)

#### 1. structured candidate space $(S, \leq)$

A fragmentation of the candidate space, and an ordering on those fragments.

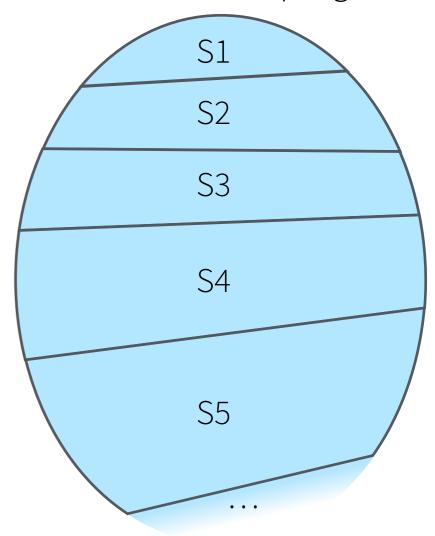


- 2. cost function (K)
- 3. gradient function (g)

#### 1. structured candidate space $(S, \leq)$

- a countable set  $\boldsymbol{\mathcal{S}}$  of sketches
- a total order ≤ on S

A fragmentation of the candidate space, and an ordering on those fragments.



- 2. cost function (к)
- 3. gradient function (g)

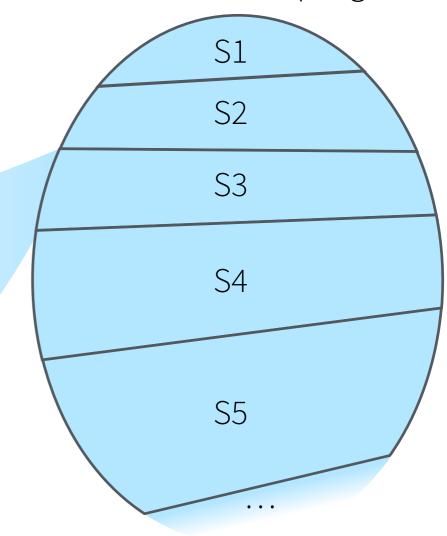
#### 1. structured candidate space $(S, \leq)$

- a countable set  $\boldsymbol{\mathcal{S}}$  of sketches
- a total order ≤ on S

A fragmentation of the candidate space, and an ordering on those fragments.

- 2. cost function (κ)
- 3. gradient function (g)





#### 1. structured candidate space $(S, \leq)$

- a countable set  $\boldsymbol{\mathcal{S}}$  of sketches
- a total order ≤ on S

A fragmentation of the candidate space, and an ordering on those fragments.

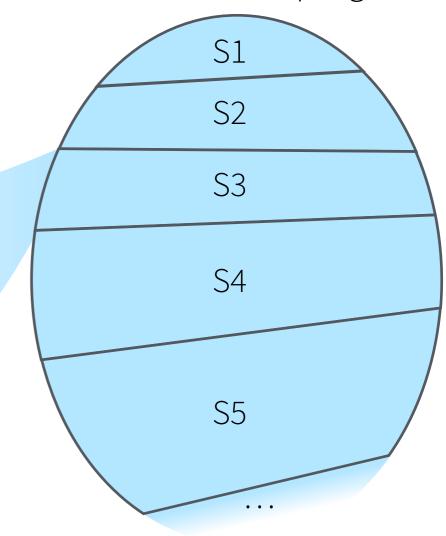
**S**<sub>3</sub> (SSA programs of length 3)   
**def** 
$$f(x)$$
:
$$r_1 = ??_{op}(??_{\{x\}})$$

$$r_2 = ??_{op}(??_{\{x,r_1\}})$$

$$r_3 = ??_{op}(??_{\{x,r_1,r_2\}})$$
**return**  $r_3$ 

#### 2. cost function (к)





#### 1. structured candidate space $(S, \leq)$

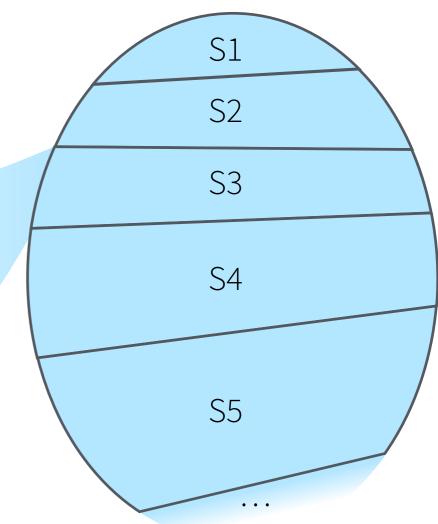
- a countable set **S** of sketches
- a total order ≤ on **S**

A fragmentation of the candidate space, and an ordering on those fragments.

**S**<sub>3</sub> (SSA programs of length 3)   
**def** 
$$f(x)$$
:
$$r_1 = ??_{op}(??_{\{x\}}) \quad \text{Vars \& constants}$$

$$r_2 = ??_{op}(??_{\{x,r_1\}})$$

$$r_3 = ??_{op}(??_{\{x,r_1,r_2\}})$$
**return**  $r_3$ 



- 2. cost function (K)
- 3. gradient function (g)

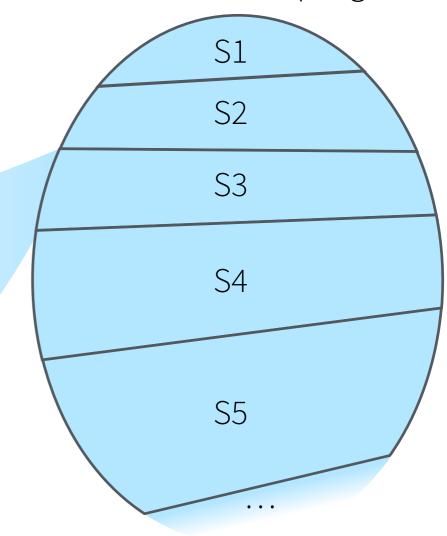
#### 1. structured candidate space $(S, \leq)$

- a countable set  $\boldsymbol{\mathcal{S}}$  of sketches
- a total order ≤ on S

A fragmentation of the candidate space, and an ordering on those fragments.

- 2. cost function (κ)
- 3. gradient function (g)





#### 1. structured candidate space $(S, \leq)$

- a countable set  $\boldsymbol{\mathcal{S}}$  of sketches
- a total order ≤ on S

A fragmentation of the candidate space, and an ordering on those fragments.

**S**<sub>3</sub> (SSA programs of length 3) **def** f(x):  $r_1 = ??_{op}(??_{\{x\}})$   $r_2 = ??_{op}(??_{\{x,r_1\}})$   $r_3 = ??_{op}(??_{\{x,r_1,r_2\}})$ **return**  $r_3$  Ordering expresses high-level search strategy.

**s** = set of all SSA programs S1 S2 **S**3 **S4 S5** 

- 2. cost function (к)
- 3. gradient function (g)

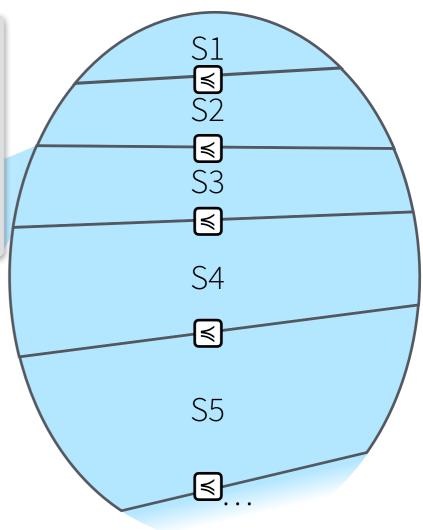
#### 1. structured candidate space $(S, \leq)$

- a countable set S of sketches
- a total order ≤ on S

A fragmentation of the candidate space, and an ordering on those fragments.

**S**<sub>3</sub> (SSA programs of length 3) **def** f(x):  $r_1 = ??_{op}(??_{\{x\}})$   $r_2 = ??_{op}(??_{\{x,r_1\}})$   $r_3 = ??_{op}(??_{\{x,r_1,r_2\}})$ **return**  $r_3$  Ordering expresses high-level search strategy.

Here, ≤ expresses iterative deepening.



- 2. cost function (к)
- 3. gradient function (g)

#### 1. structured candidate space $(S, \leq)$

- a countable set S of sketches
- a total order ≤ on S

A fragmentation of the candidate space, and an ordering on those fragments.

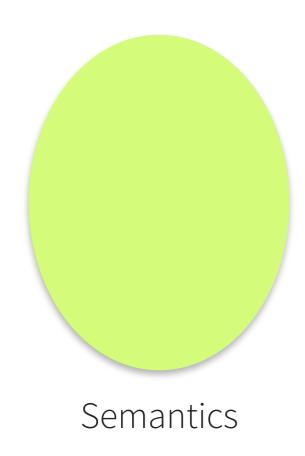
Implemented as a generator that returns the next sketch in the space

```
def f(x):
                                        S_1
   r1 = ??_{op}(??_{\{x\}})
   return r1
                                        S<sub>2</sub>
def f(x):
   r1 = ??_{op}(??_{\{x\}})
   r2 = ??_{op}(??_{\{x,r1\}})
   return r2
def f(x):
                                        S<sub>3</sub>
   r1 = ??_{op}(??_{\{x\}})
   r2 = ??_{op}(??_{\{x,r1\}})
   r3 = ??_{op}(??_{\{x,r1,r2\}})
   return r3
```

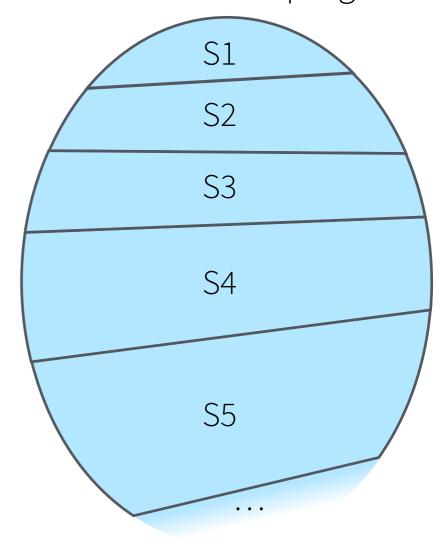
- 2. cost function (к)
- 3. gradient function (g)

#### 1. structured candidate space $(S, \leq)$

- a countable set  $\boldsymbol{\mathcal{S}}$  of sketches
- a total order ≤ on S

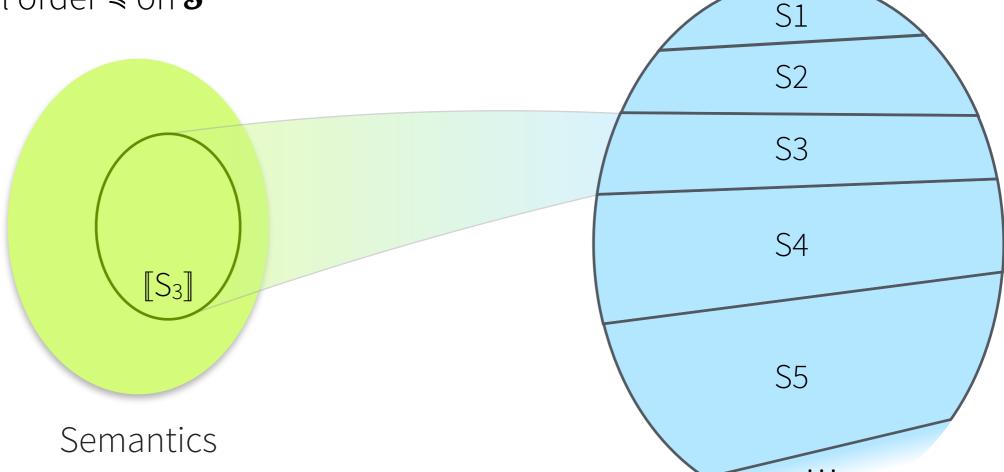


- 2. cost function (K)
- 3. gradient function (g)



#### 1. structured candidate space $(S, \leq)$

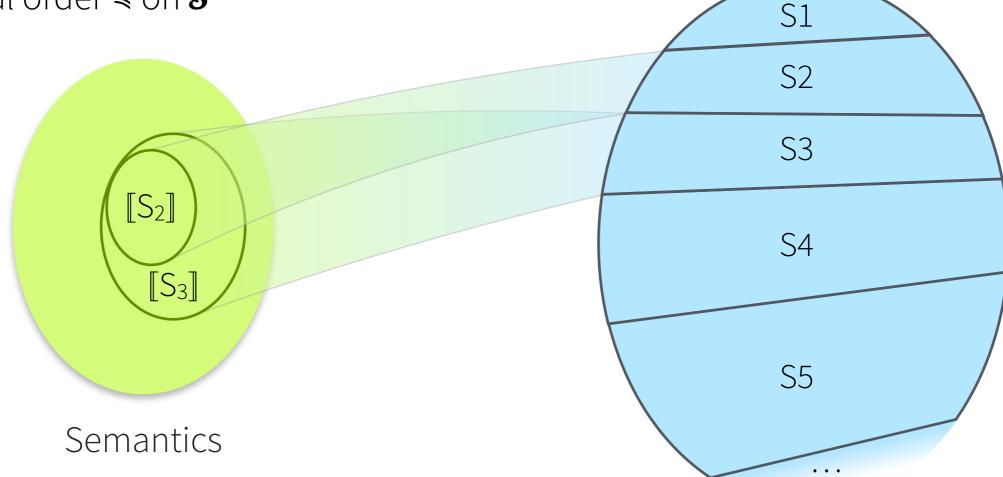
- a countable set S of sketches
- a total order ≤ on S



- 2. cost function (K)
- 3. gradient function (g)

#### 1. structured candidate space $(S, \leq)$

- a countable set S of sketches
- a total order ≤ on S



- 2. cost function (K)
- 3. gradient function (g)

#### 1. structured candidate space $(S, \leq)$

a countable set S of sketches

a total order ≤ on **S**Semantic redundancy in the

search space.

\$1 \$2 \$3 \$4 \$5

**S** = set of all SSA programs

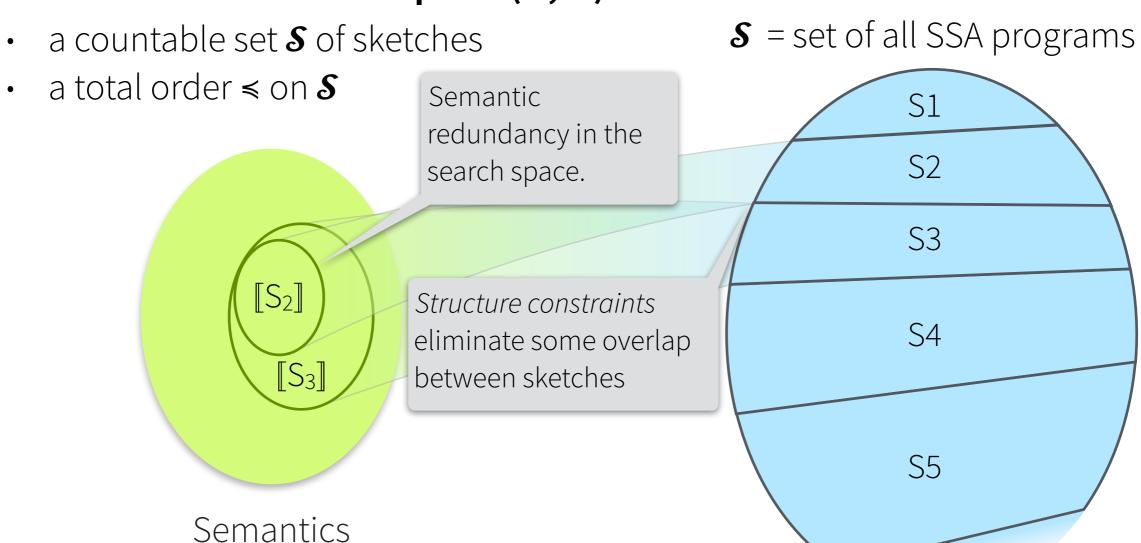
Semantics

 $\llbracket S_2 \rrbracket$ 

 $[S_3]$ 

- 2. cost function (к)
- 3. gradient function (g)

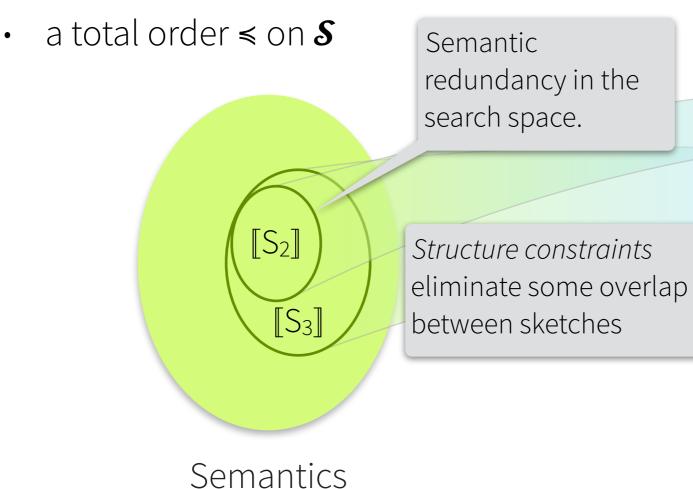
#### 1. structured candidate space $(S, \leq)$



- 2. cost function (κ)
- 3. gradient function (g)

#### 1. structured candidate space $(S, \leq)$

a countable set **S** of sketches



**S** = set of all SSA programs

S1 S2 **S**3 (SSA programs of length 3) def f(x):  $r_1 = ??_{op}(??_{\{x\}})$  $r_2 = ??_{op}(??_{\{x,r_1\}})$  $r_3 = ??_{op}(??_{\{x,r_1,r_2\}})$ return r<sub>3</sub>

- 2. cost function (κ)
- 3. gradient function (g)

#### 1. structured candidate space $(S, \leq)$

a countable set S of sketches

• a total order ≤ on **S** 

Semantic redundancy in the search space.

 $[S_2]$   $[S_3]$ 

Semantics

Structure constraints eliminate some overlap between sketches

Eliminate dead-code redundancy: assert that each  $\mathbf{r_i}$  is read

**S** = set of all SSA programs

\$1 \$2 \$3

**S**<sub>3</sub> (SSA programs of length 3)

def f(x):
 r<sub>1</sub> = ??<sub>op</sub>(??<sub>{x}</sub>)
 r<sub>2</sub> = ??<sub>op</sub>(??<sub>{x,r<sub>1</sub>}</sub>)
 r<sub>3</sub> = ??<sub>op</sub>(??<sub>{x,r<sub>1</sub>,r<sub>2</sub>}</sub>)
 return r<sub>3</sub>

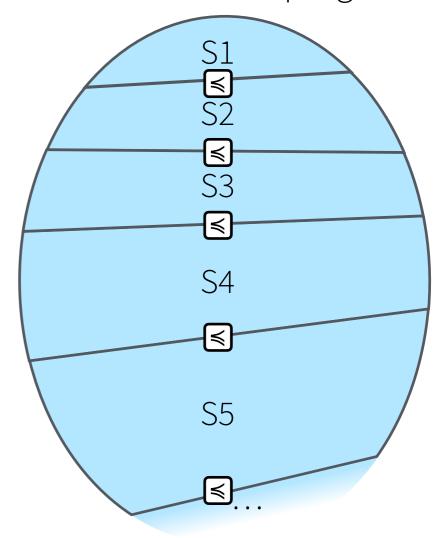
- 2. cost function (к)
- 3. gradient function (g)

## Cost functions rank candidate programs

- 1. structured candidate space  $(S, \leq)$
- 2. cost function (k)

 $\kappa: \mathcal{L} \to \mathbb{R}$  assigns a numeric cost to each program in the language  $\mathcal{L}$ 

**s** = set of all SSA programs



### Cost functions rank candidate programs

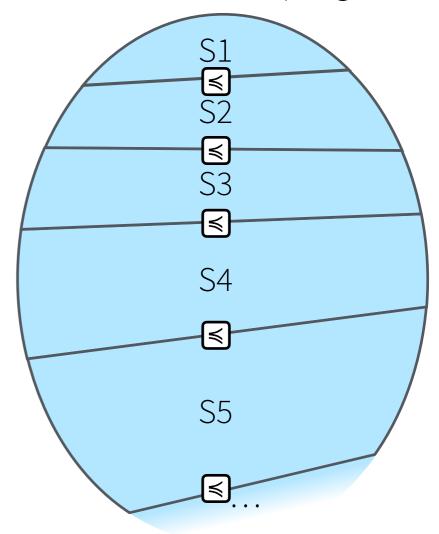
#### 1. structured candidate space $(S, \leq)$

#### 2. cost function (k)

 $\kappa: \mathcal{L} \to \mathbb{R}$  assigns a numeric cost to each program in the language  $\mathcal{L}$ 

Cost functions can be based on both syntax and semantics (dynamic behavior)

**s** = set of all SSA programs



## Cost functions rank candidate programs

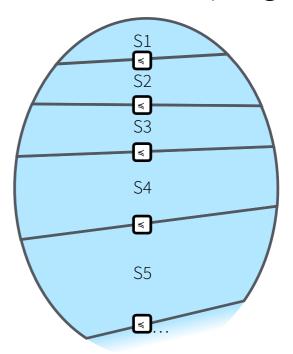
#### 1. structured candidate space $(S, \leq)$

#### 2. cost function (k)

 $\kappa: \mathcal{L} \to \mathbb{R}$  assigns a numeric cost to each program in the language  $\mathcal{L}$ 

Cost functions can be based on both syntax and semantics (dynamic behavior)

**s** = set of all SSA programs

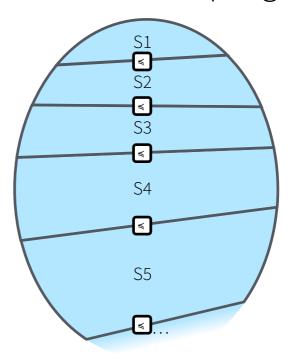


$$K(P) = i$$
 for  $P \in S_i \in \mathbf{S}$ 

The number of variables defined in P

- 1. structured candidate space  $(S, \leq)$
- 2. cost function (K)
- 3. gradient function (g)

 $g: \mathbb{R} \to 2^{\mathbf{S}}$ g(c) is the set of sketches in  $\mathbf{S}$  that may contain a solution P with  $\kappa(P) < c$ 

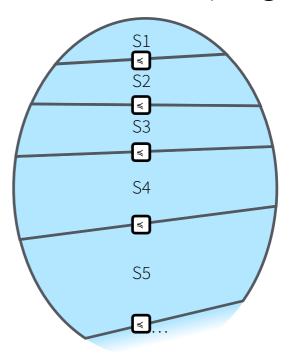


$$\kappa(P) = i \quad \text{for } P \in S_i \in \mathbf{S}$$

- 1. structured candidate space  $(S, \leq)$
- 2. cost function (к)
- 3. gradient function (g)

$$g: \mathbb{R} \to 2^{\mathbf{S}}$$
  
  $g(c)$  is the set of sketches in  $\mathbf{S}$  that  $may$  contain a solution  $P$  with  $\kappa(P) < c$ 

The gradient function overapproximates the behavior of  $\kappa$  on  $\boldsymbol{\mathcal{S}}$ 

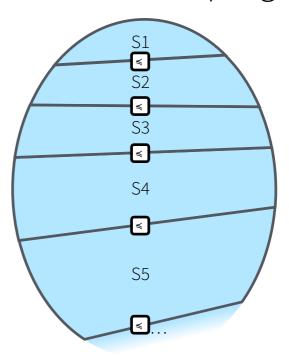


$$\kappa(P) = i \quad \text{for } P \in S_i \in \mathbf{S}$$

- 1. structured candidate space  $(S, \leq)$
- 2. cost function (K)
- 3. gradient function (g)

$$g: \mathbb{R} \to 2^{\mathcal{S}}$$
  
  $g(c)$  is the set of sketches in  $\mathcal{S}$  that  $may$  contain a solution  $P$  with  $\kappa(P) < c$ 

The gradient function overapproximates the behavior of  $\kappa$  on  $\boldsymbol{\mathcal{S}}$ 



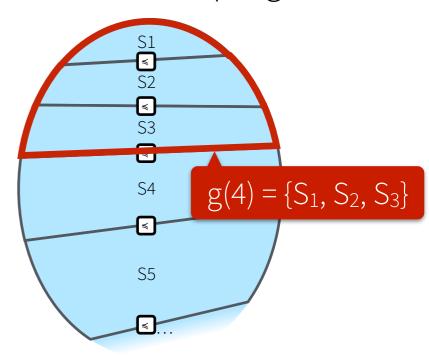
$$\kappa(P) = i$$
 for  $P \in S_i \in \mathbf{S}$ 

$$g(c) = \{ S_i \in S \mid i < c \}$$

- 1. structured candidate space  $(S, \leq)$
- 2. cost function (K)
- 3. gradient function (g)

 $g: \mathbb{R} \to 2^{\mathbf{S}}$ g(c) is the set of sketches in  $\mathbf{S}$  that may contain a solution P with  $\kappa(P) < c$ 

The gradient function overapproximates the behavior of  $\kappa$  on  $\boldsymbol{\mathcal{S}}$ 



$$K(P) = i$$
 for  $P \in S_i \in \mathbf{S}$ 

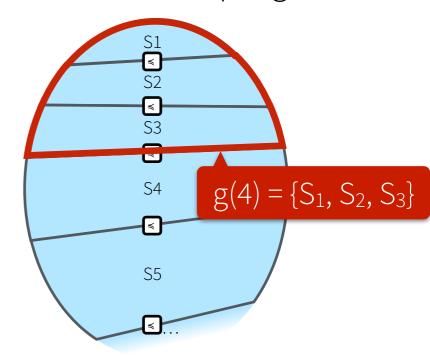
$$g(c) = \{ S_i \in S \mid i < c \}$$

- 1. structured candidate space  $(S, \leq)$
- 2. cost function (K)
- 3. gradient function (g)

$$g: \mathbb{R} \to 2^{\mathbf{S}}$$
  
  $g(c)$  is the set of sketches in  $\mathbf{S}$  that  $may$  contain a solution  $P$  with  $\kappa(P) < c$ 

The gradient function overapproximates the behavior of  $\kappa$  on  $\boldsymbol{\mathcal{S}}$ 

Always sound for g to return all of **S** if a tighter bound is unavailable.



$$\kappa(P) = i$$
 for  $P \in S_i \in \mathbf{S}$   
 $g(c) = \{ S_i \in \mathbf{S} \mid i < c \}$ 

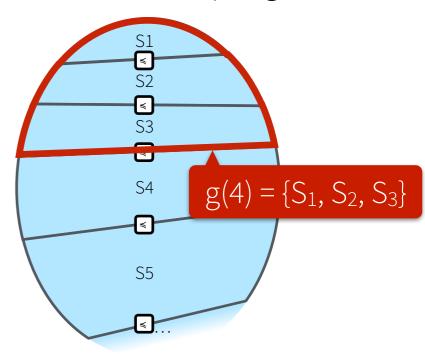
- 1. structured candidate space  $(S, \leq)$
- 2. cost function (K)
- 3. gradient function (g)

$$g: \mathbb{R} \to 2^{\mathbf{S}}$$
  
  $g(c)$  is the set of sketches in  $\mathbf{S}$  that  $may$  contain a solution  $P$  with  $\kappa(P) < c$ 

The gradient function overapproximates the behavior of  $\kappa$  on  $\boldsymbol{\mathcal{S}}$ 

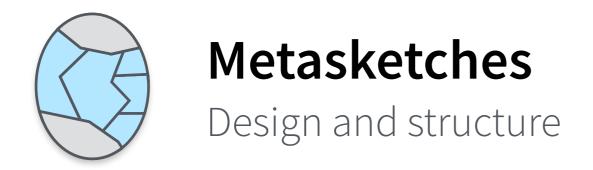
Always sound for g to return all of **S** if a tighter bound is unavailable.

g(c) always being finite is sufficient (not necessary) to guarantee termination.

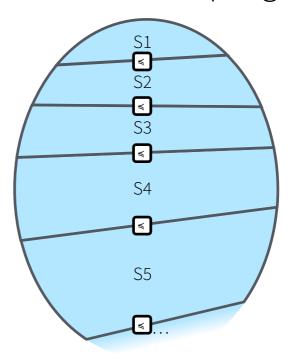


$$K(P) = i$$
 for  $P \in S_i \in \mathbf{S}$ 

$$g(c) = \{ S_i \in S \mid i < c \}$$

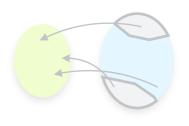


- 1. structured candidate space  $(S, \leq)$
- 2. cost function (K)
- 3. gradient function (g)



$$\kappa(P) = i$$
 for  $P \in S_i \in \mathbf{S}$ 

$$g(c) = \{ S_i \in S \mid i < c \}$$



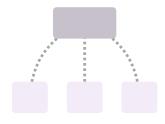
### Background

Syntax-guided synthesis



#### Metasketches

Design and structure



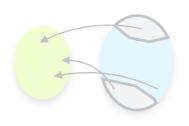
Synapse

A metasketch solver



Results

Better solutions, faster



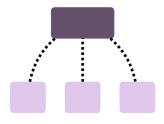
## Background

Syntax-guided synthesis



#### Metasketches

Design and structure



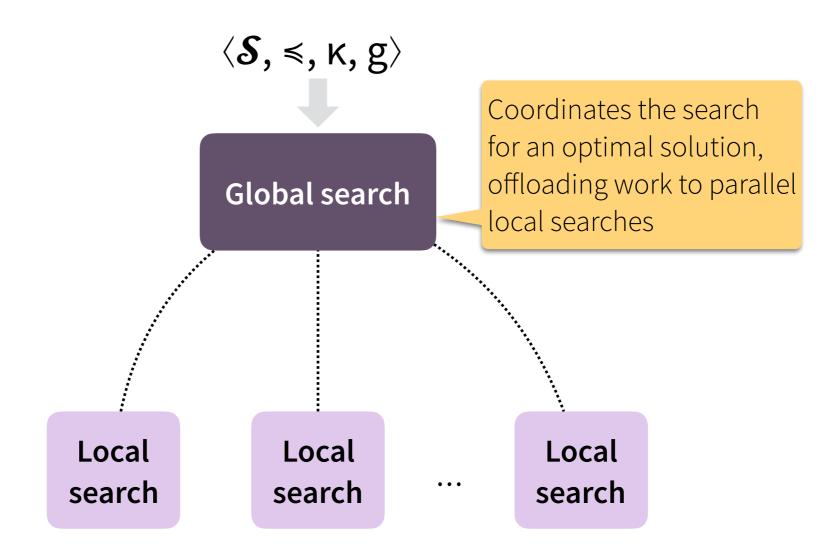
### **Synapse**

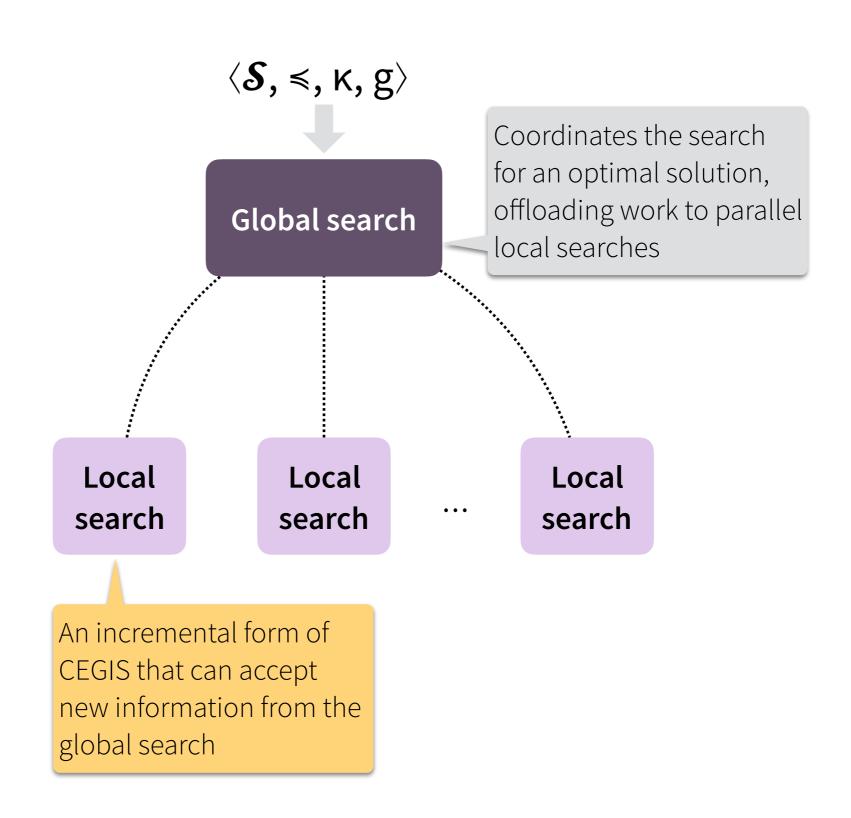
A metasketch solver

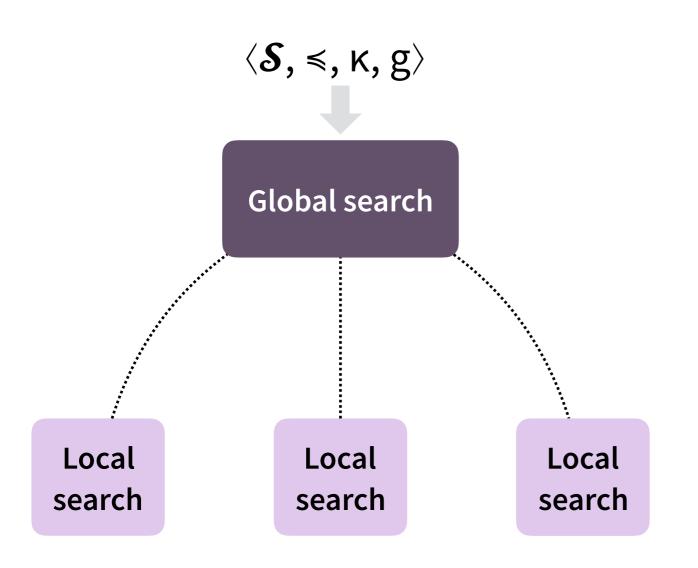


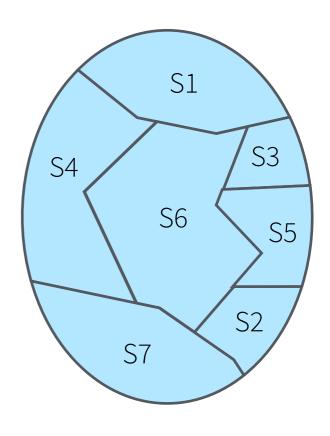
#### Results

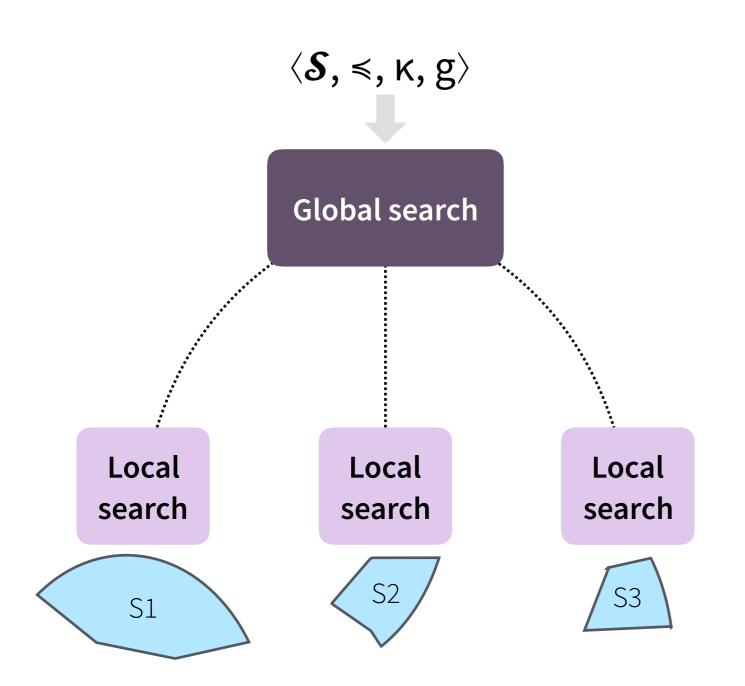
Better solutions, faster

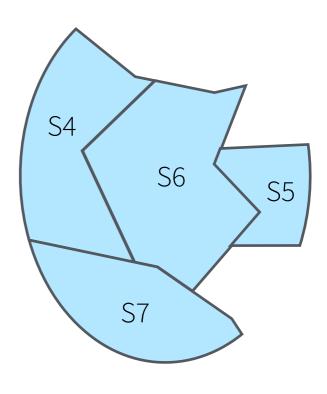


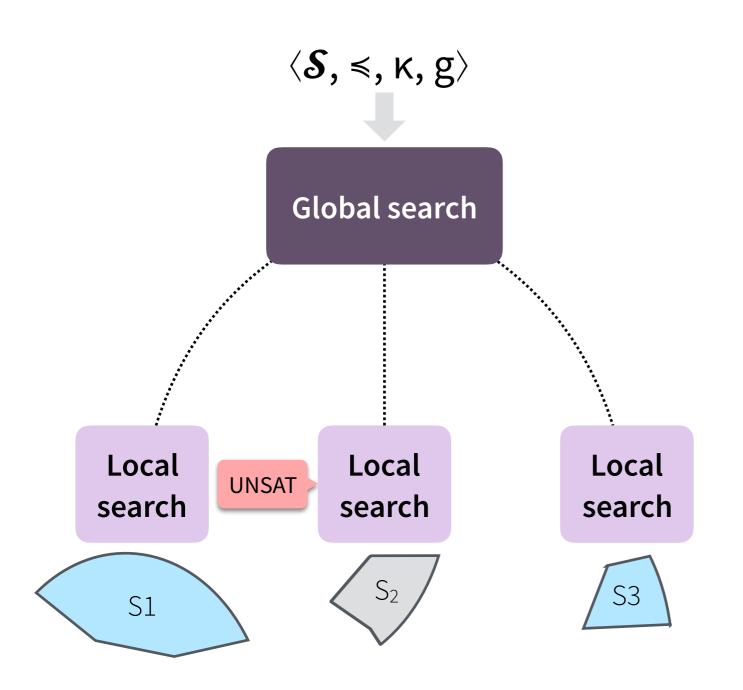


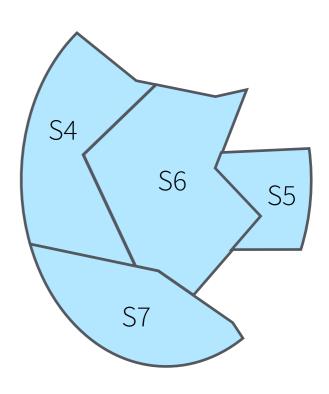


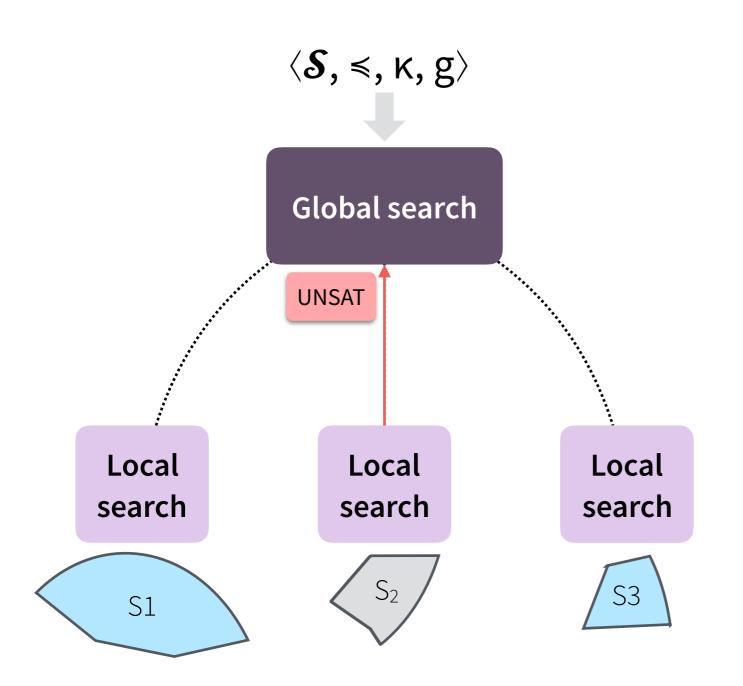


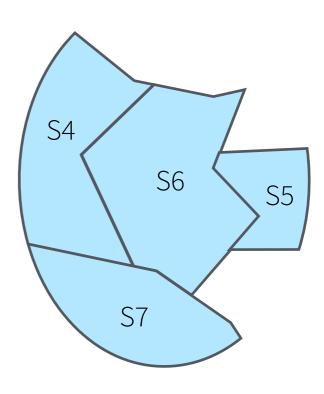


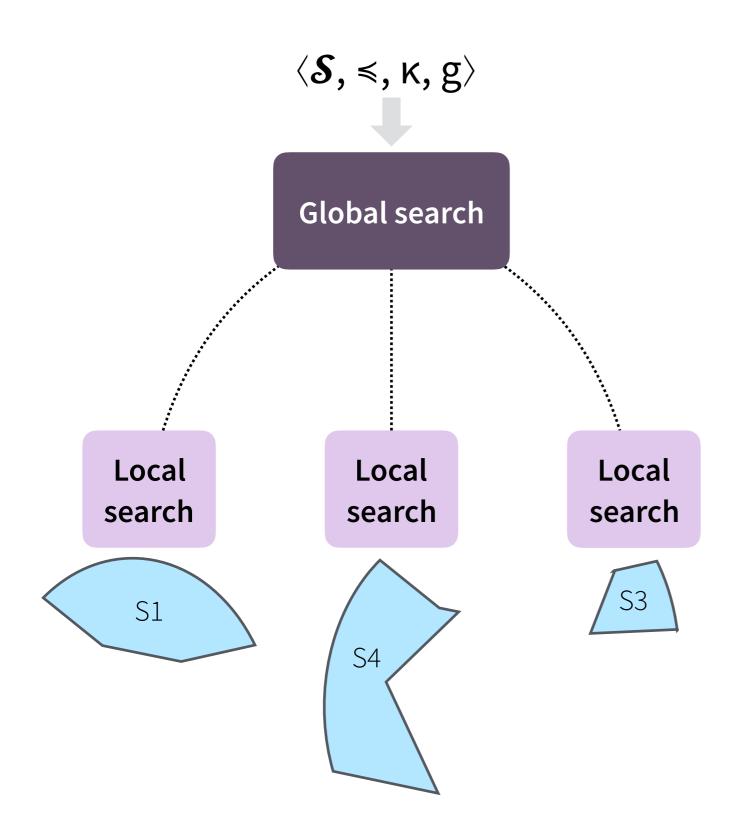


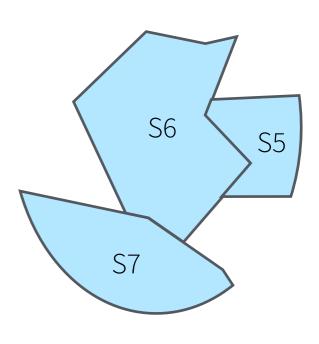


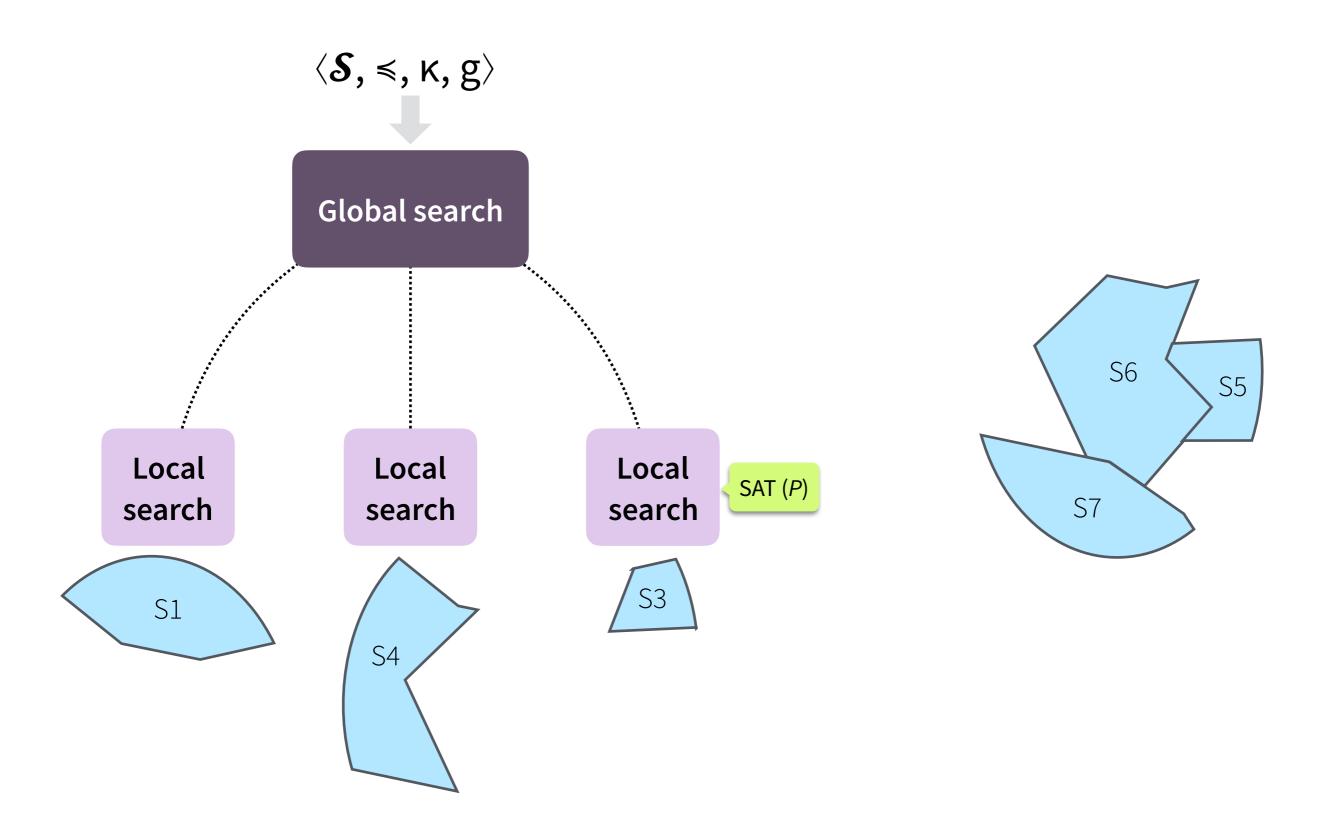


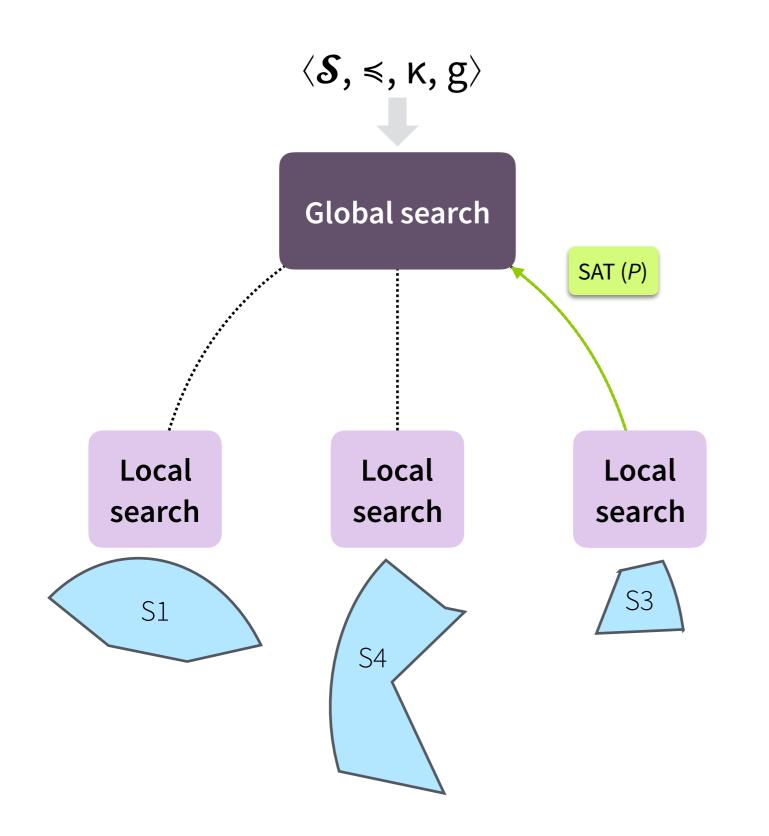


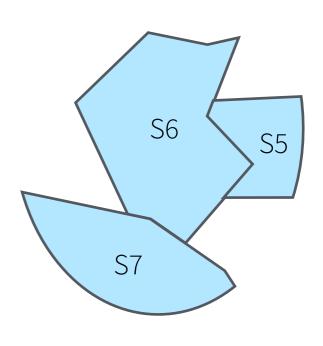


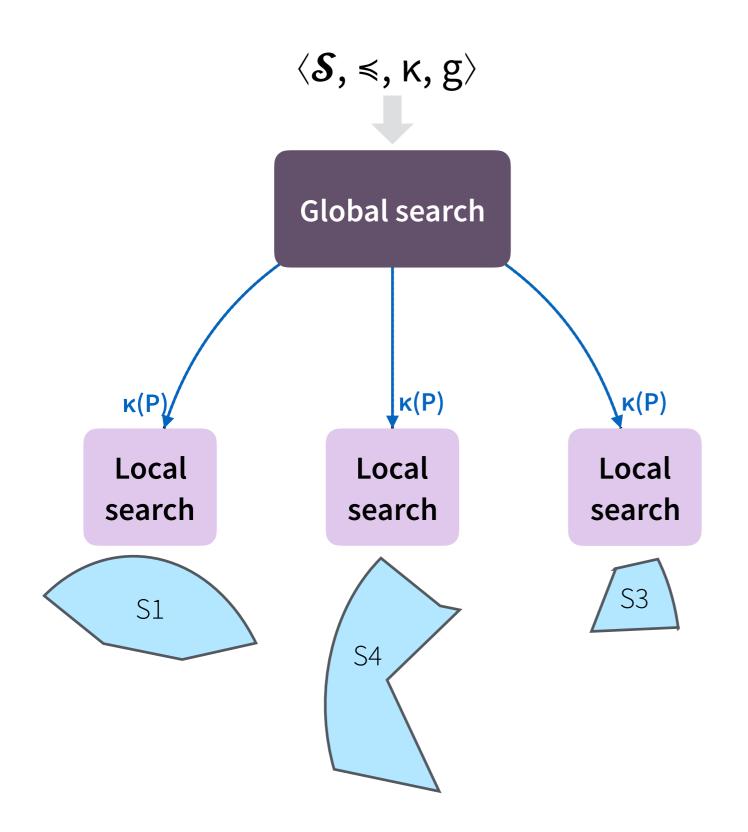


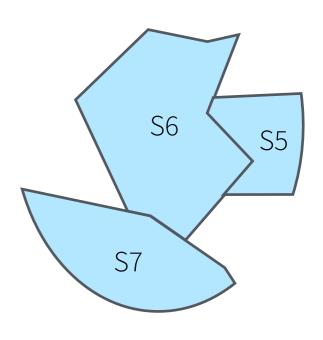


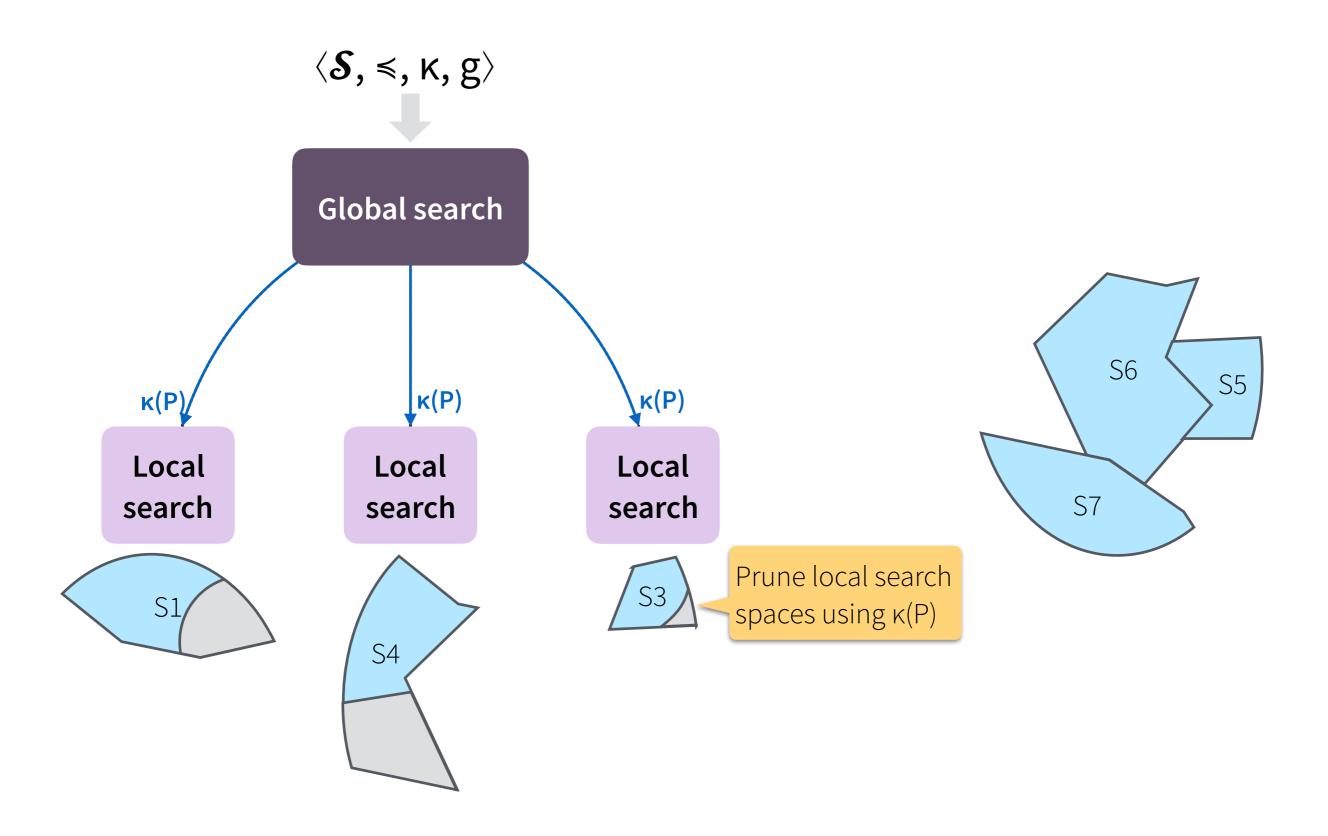


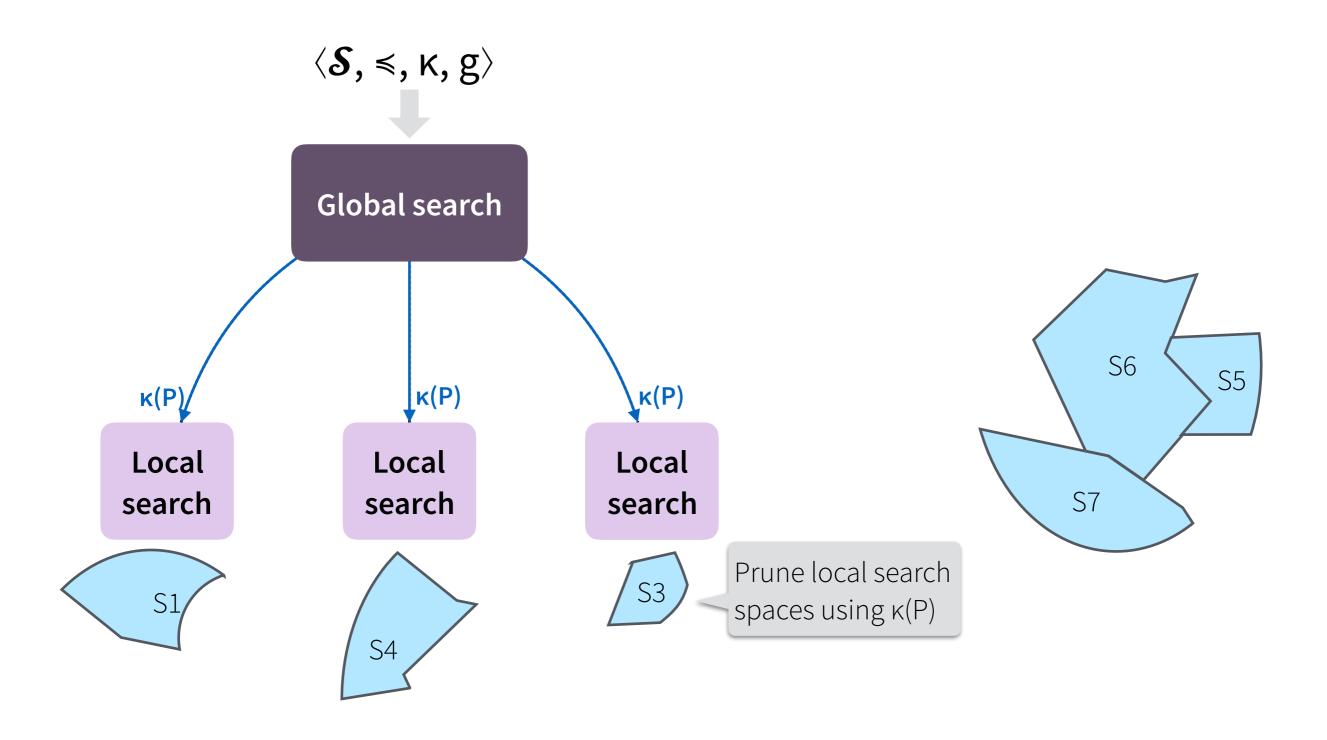


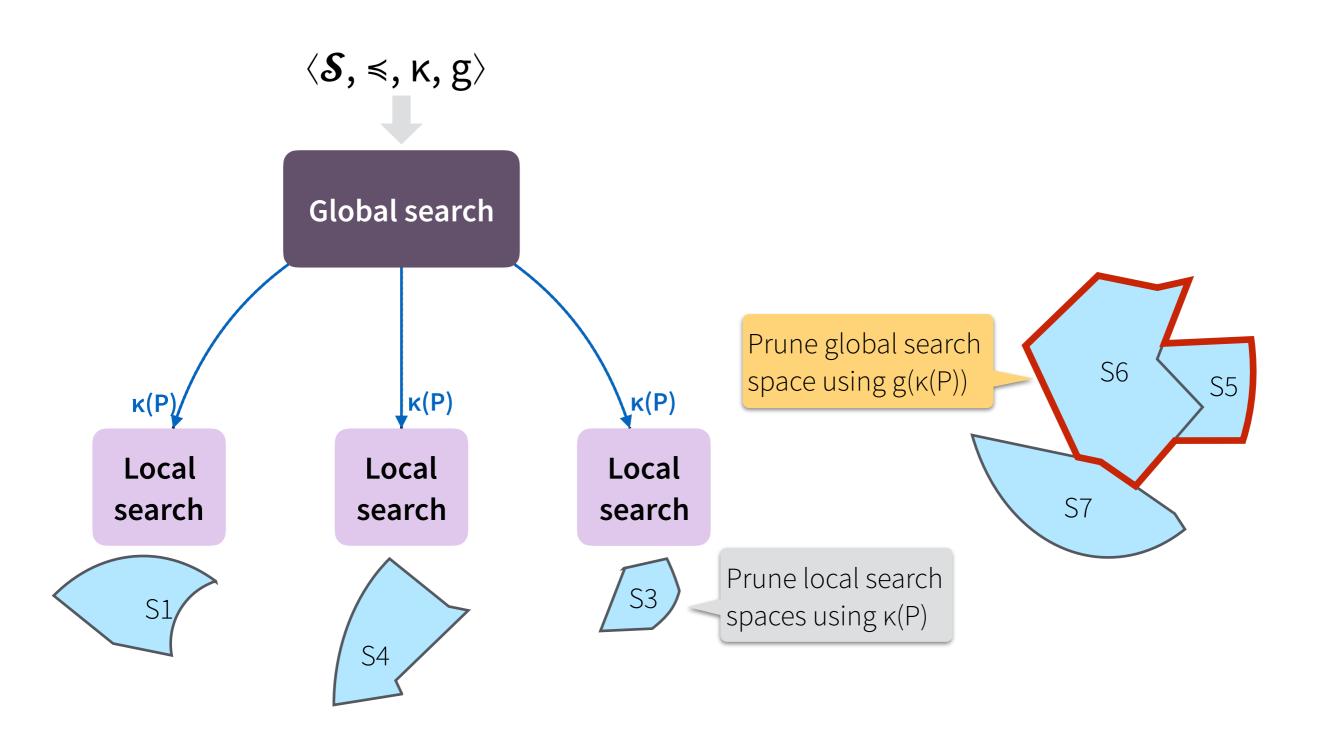


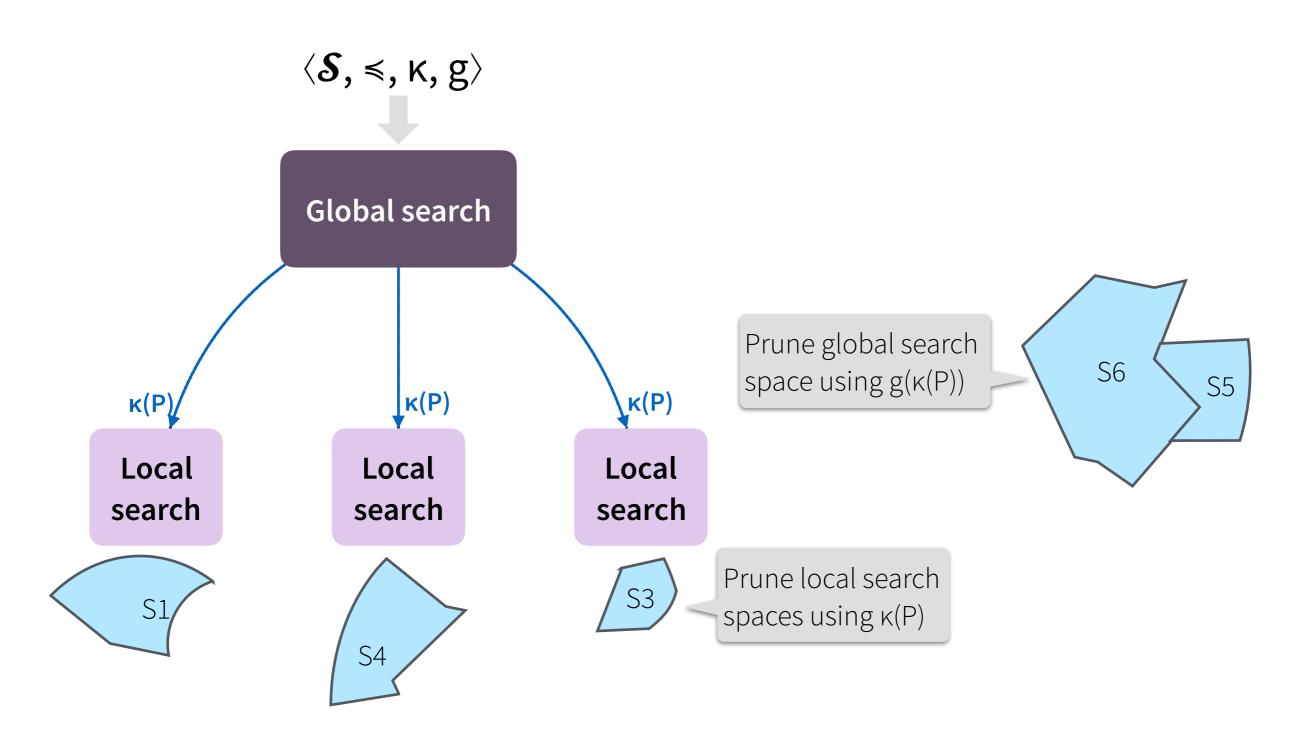


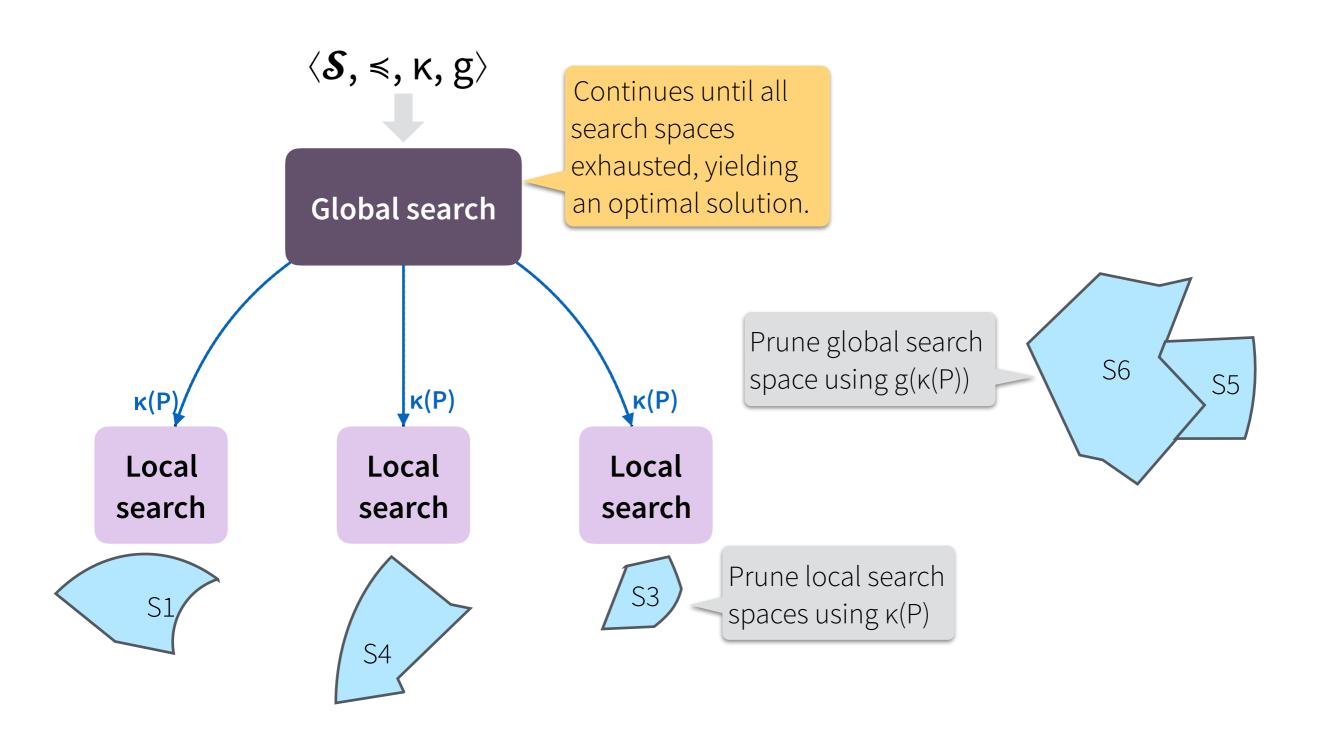




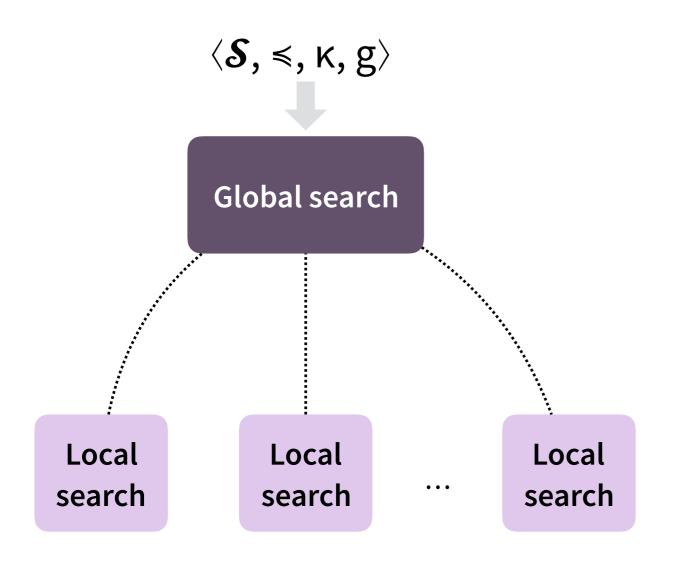






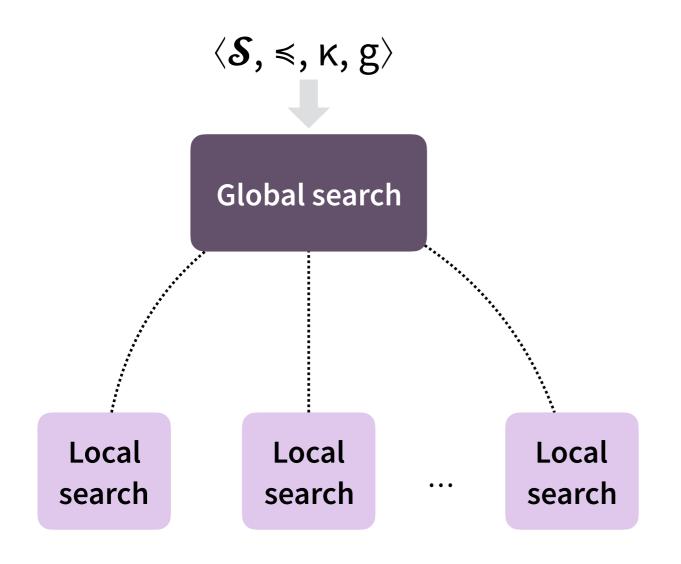


## Synapse implementation



Implemented in Rosette, a solver-aided extension of Racket

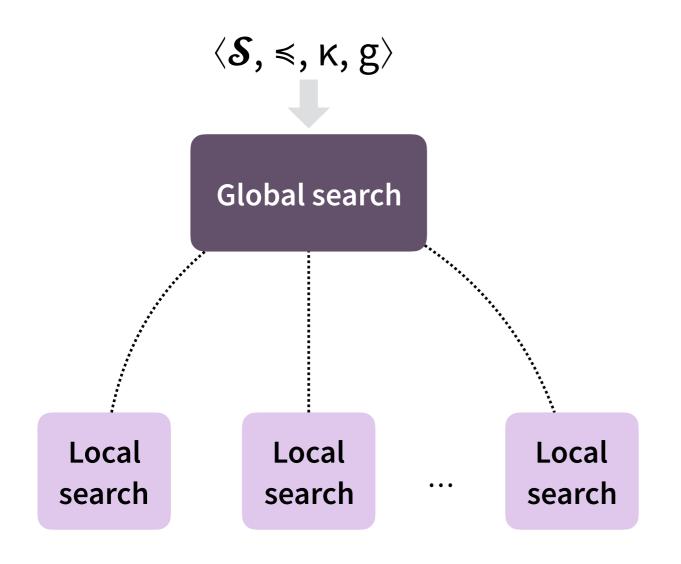
### Synapse implementation



Implemented in Rosette, a solver-aided extension of Racket

Local CEGIS searches can share counterexamples

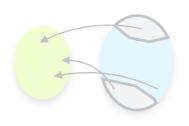
### Synapse implementation



Implemented in Rosette, a solver-aided extension of Racket

Local CEGIS searches can share counterexamples

Local searches can time out, which weakens optimality



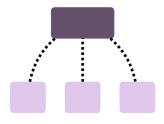
## Background

Syntax-guided synthesis



#### Metasketches

Design and structure



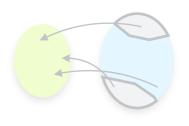
### **Synapse**

A metasketch solver



#### Results

Better solutions, faster



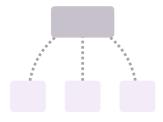
## Background

Syntax-guided synthesis



#### Metasketches

Design and structure



Synapse

A metasketch solver



Results

Better solutions, faster

### **Evaluation questions**

Is Synapse a practical approach to solving different kinds of synthesis problems?

Approximate computing, array programs

### **Evaluation questions**

Is Synapse a practical approach to solving different kinds of synthesis problems?

Approximate computing, array programs

Can Synapse reason about complex cost functions?

### **Evaluation questions**

Is Synapse a practical approach to solving different kinds of synthesis problems?

Approximate computing, array programs

Can Synapse reason about complex cost functions?

#### In the paper:

- Parallel speedup
- Optimizations (structure constraints, sharing)
- More kinds of problems
- More complex cost functions

### Synapse solves previously-intractable problems

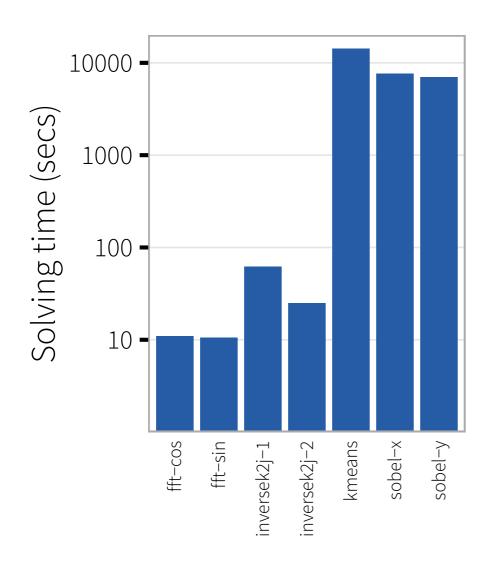
Parrot benchmarks from approximate computing [Esmaelizadeh et al., 2012]

Find the most efficient approximate program within an error bound

### Synapse solves previously-intractable problems

Parrot benchmarks from approximate computing [Esmaelizadeh et al., 2012]

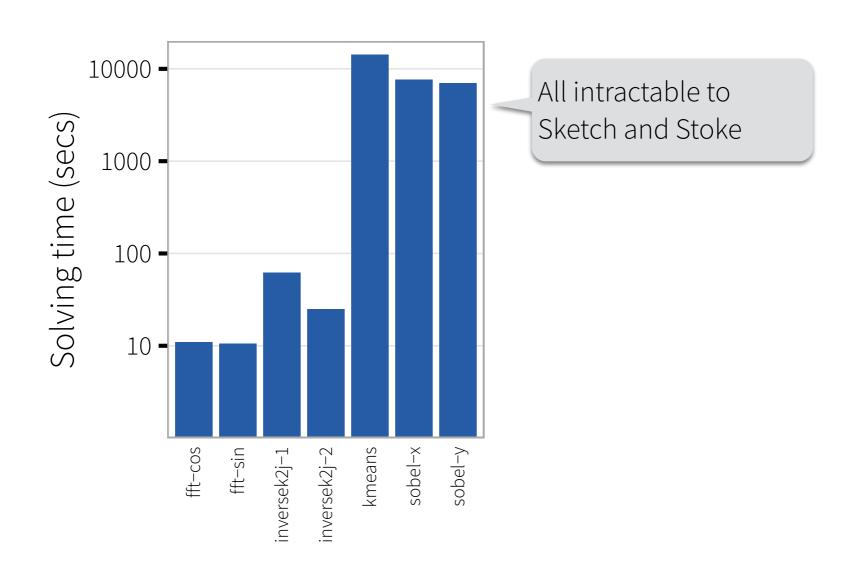
Find the most efficient approximate program within an error bound



## Synapse solves previously-intractable problems

Parrot benchmarks from approximate computing [Esmaelizadeh et al., 2012]

Find the most efficient approximate program within an error bound



### Synapse solves standard benchmarks optimally

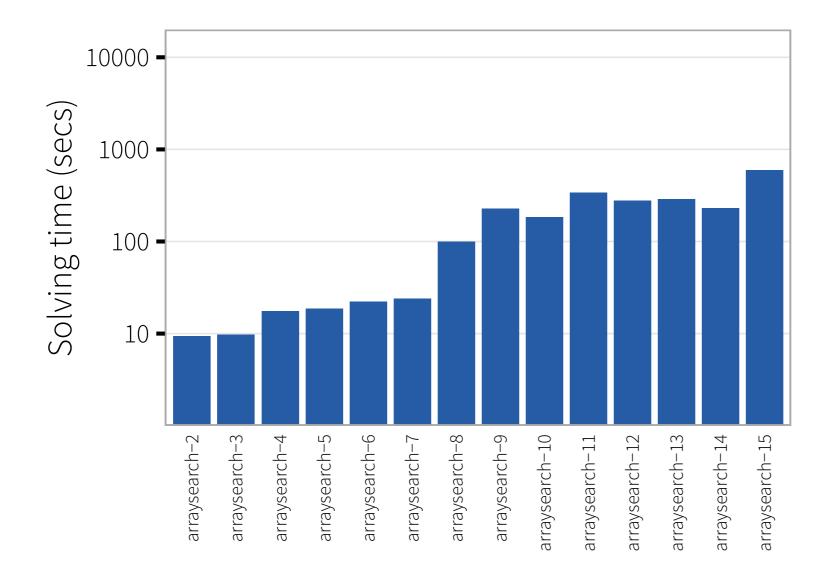
**Array Search** benchmarks from the syntax-guided synthesis (SyGuS) competition [Alur et al., 2015]

arraysearch-n: find program that searches lists of length n

### Synapse solves standard benchmarks optimally

**Array Search** benchmarks from the syntax-guided synthesis (SyGuS) competition [Alur et al., 2015]

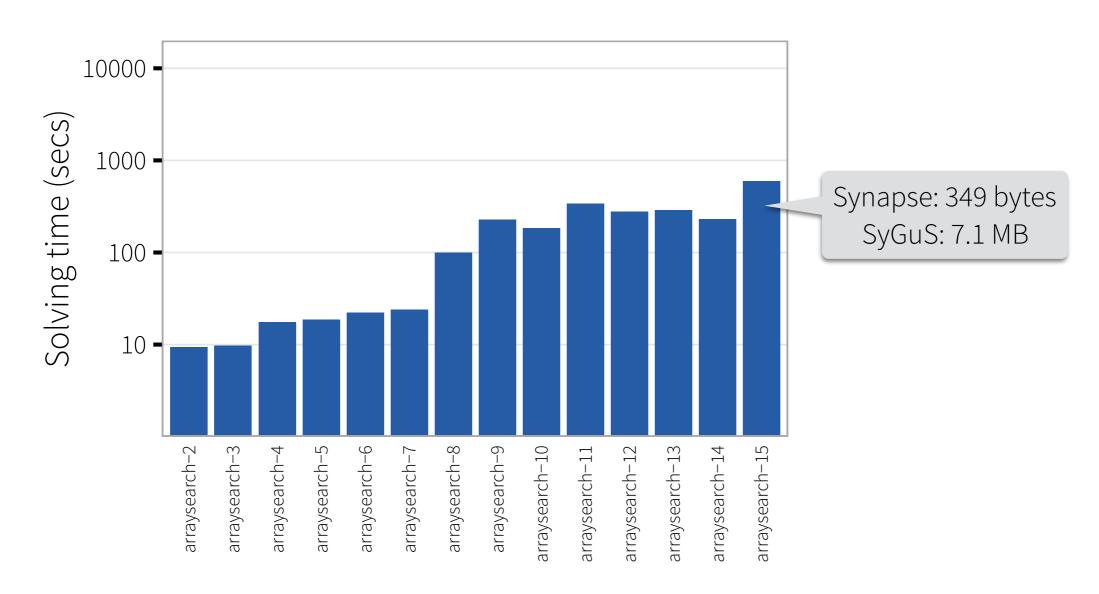
arraysearch-n: find program that searches lists of length n



### Synapse solves standard benchmarks optimally

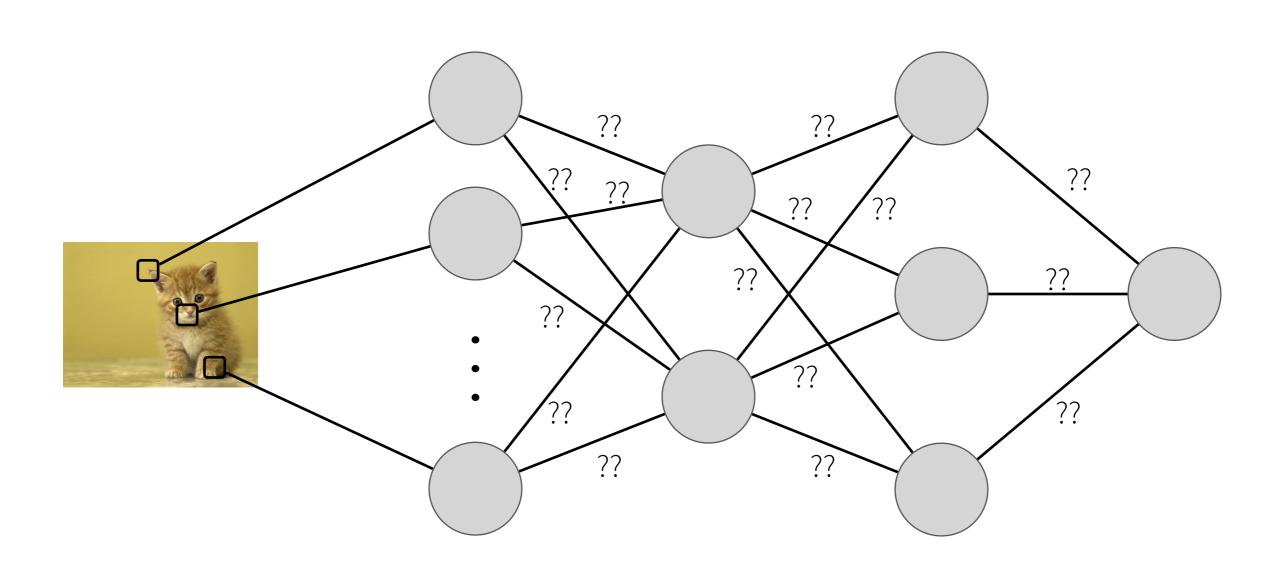
**Array Search** benchmarks from the syntax-guided synthesis (SyGuS) competition [Alur et al., 2015]

arraysearch-n: find program that searches lists of length n

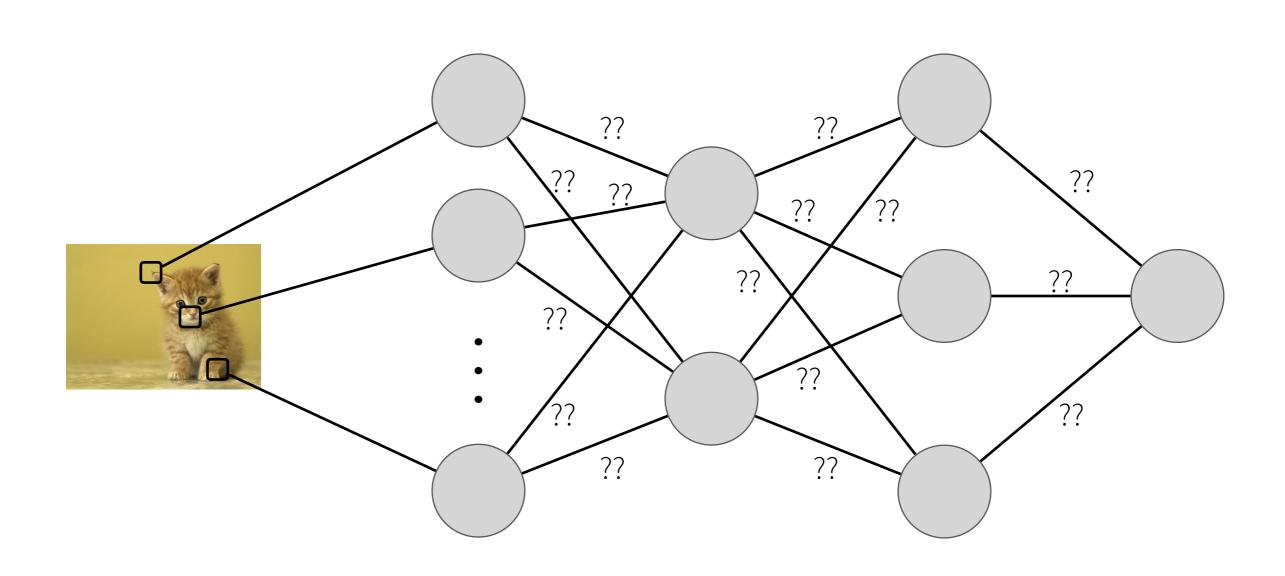




# Synapse reasons about complex costs

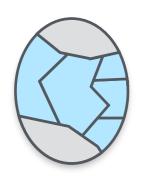


### Synapse reasons about complex costs



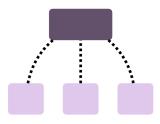
$$\kappa(P) = \sum_{i} |P(x_i) - y_i|$$

Classification error executes  $\kappa(P) = \sum_{i} |P(x_i) - y_i|$  the program for each point in the training set



#### Metasketches

Design and structure



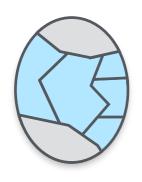
### **Synapse**

A metasketch solver



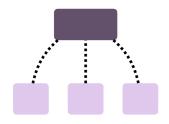
Results

Better solutions, faster



#### Metasketches

Design and structure



### **Synapse**

A metasketch solver

synapse.uwplse.org



Results

Better solutions, faster