

CS5733 Program Synthesis

#1. Introduction and Overview

Ashish Mishra

Instructor



Ashish Mishra

- Asst. Professor at CSE
- Before: Postdoc, Purdue CS, Programming Languages Group.
- Even before: IISc, CSA, PhD
- Research Goal: Help programmers write correct and trustworthy software.
- Areas: Programming Languages, Program Verification, Synthesis.

Logistics

- Lecture:
 - When: Tuesdays 2:30 - 3:55 pm, Fridays 4:00 - 5:25 pm
 - Where: C-LH5
- Course Website: Please register on Google Classroom link.
 - <https://aegis-iisc.github.io/cs5733/>
- Office Hours:
 - After the class OR with appointment a few hours before.

Goals and Activities

-
- The diagram illustrates the relationship between three goals and two activities. On the left, three goals are listed in boxes: 1. Understand what program synthesis can do and how, 2. Use existing synthesis tools, and 3. Contribute to synthesis techniques and tools towards a publication in an academic conference. To the right, two activities are listed: 'lectures' and 'read and discuss research papers', which are associated with the first two goals. Below them, a bracket groups the third goal under the activity 'project'.
1. Understand what program synthesis can do and how
 2. Use existing synthesis tools
 3. Contribute to synthesis techniques and tools towards a publication in an academic conference
- lectures
read and discuss research papers
- project

Evaluation

- Class Participation : 5%
 - Ask/answer questions in class
 - Participate in discussions on Classroom
- Paper Reviews : 25 %
 - 10 papers
- Midterm : 20 %
- Final Course Project : 50 %
 - Team formed by deadline: 5%
 - 1-page project proposal: 15%
 - Project presentation: 15%
 - Final report: 15%

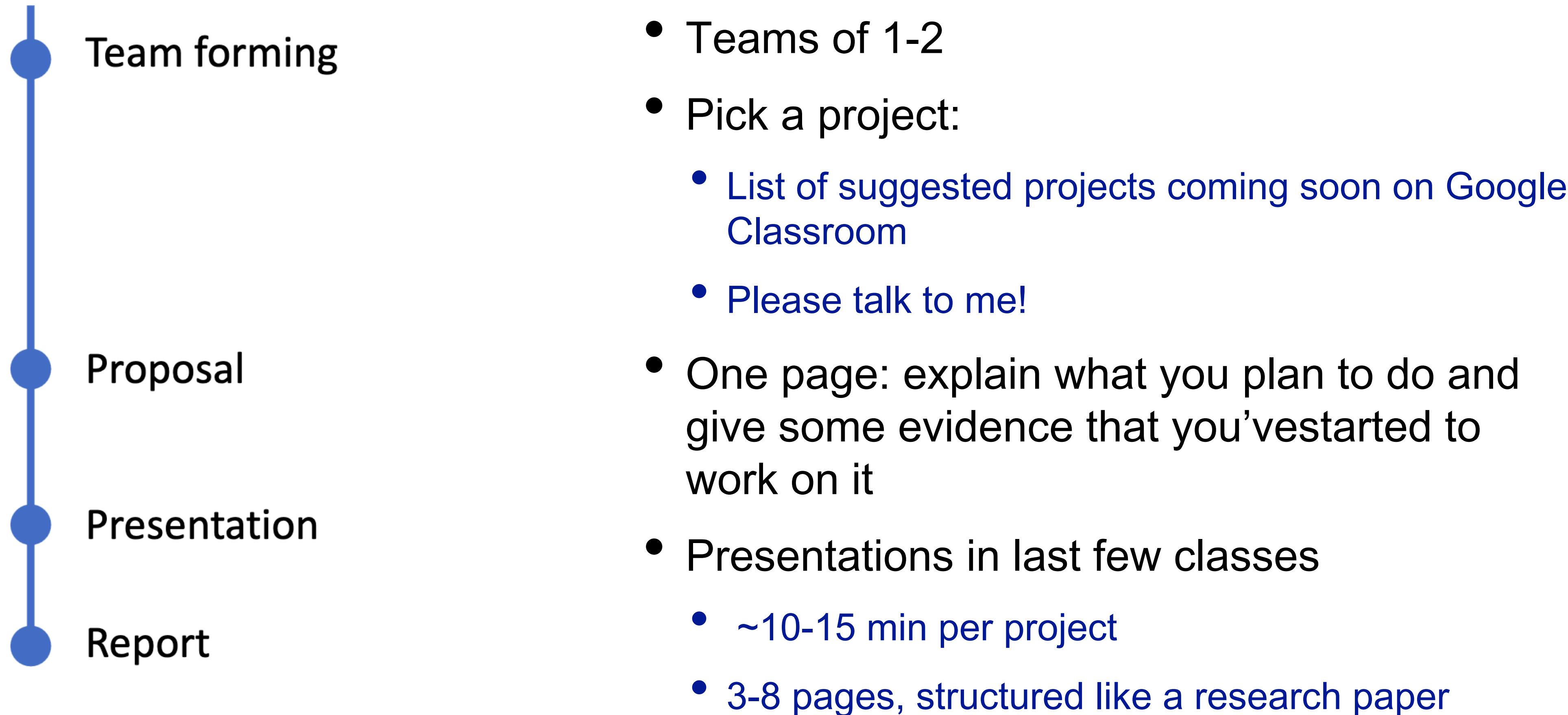
Paper Reviews

- Due on Wed of weeks 2 onwards, by the end of the day
- First review due next week
 - Posted on the Reading List at least a week before due date
- Reviews submitted via a Google Form: see course page
 - Link posted on Reading List (add this to the page)
- Review content: see course page
- Discussion:
 - before due date: discuss on Google Classroom
 - after due date: discuss in class

Project

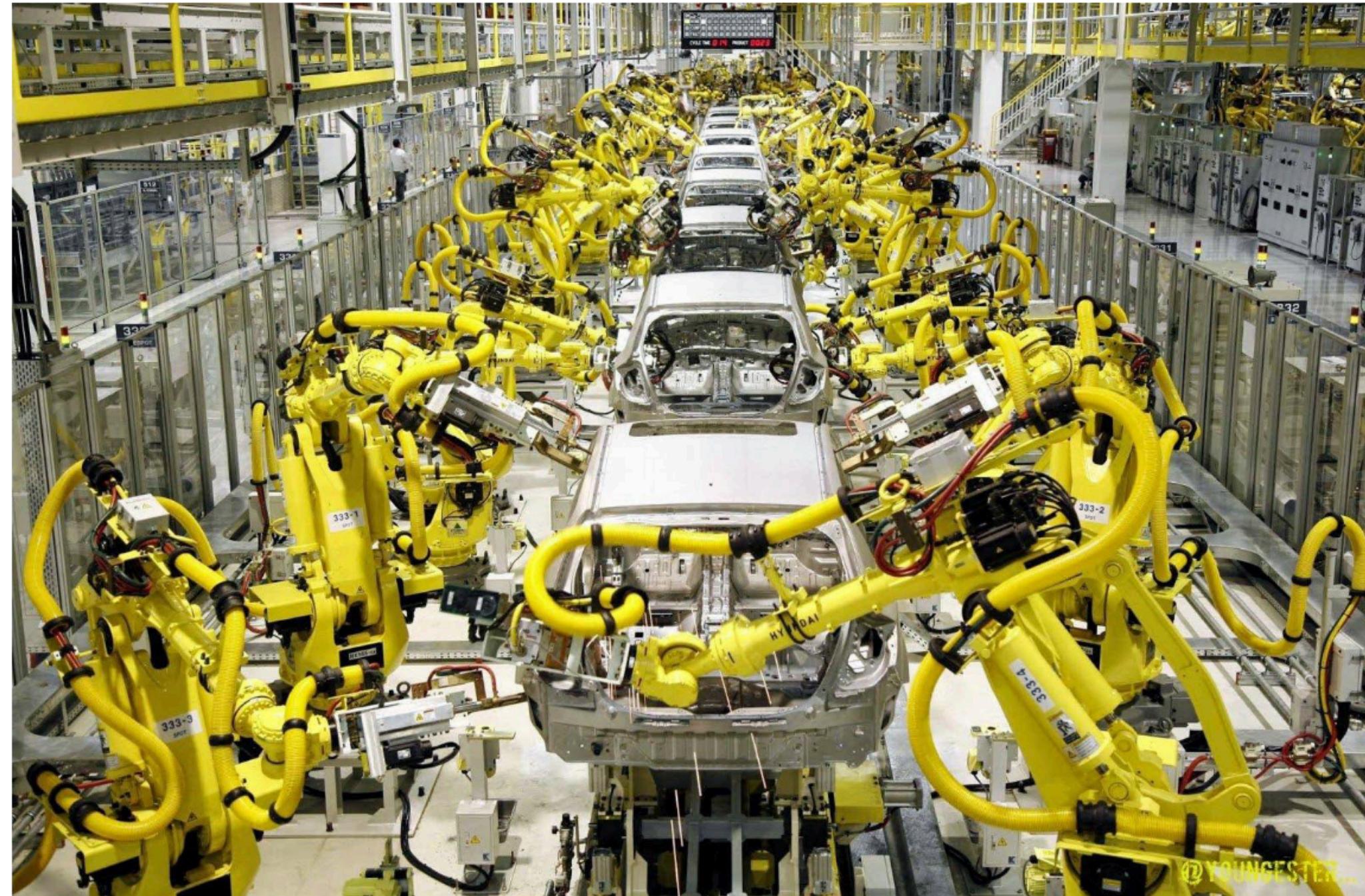
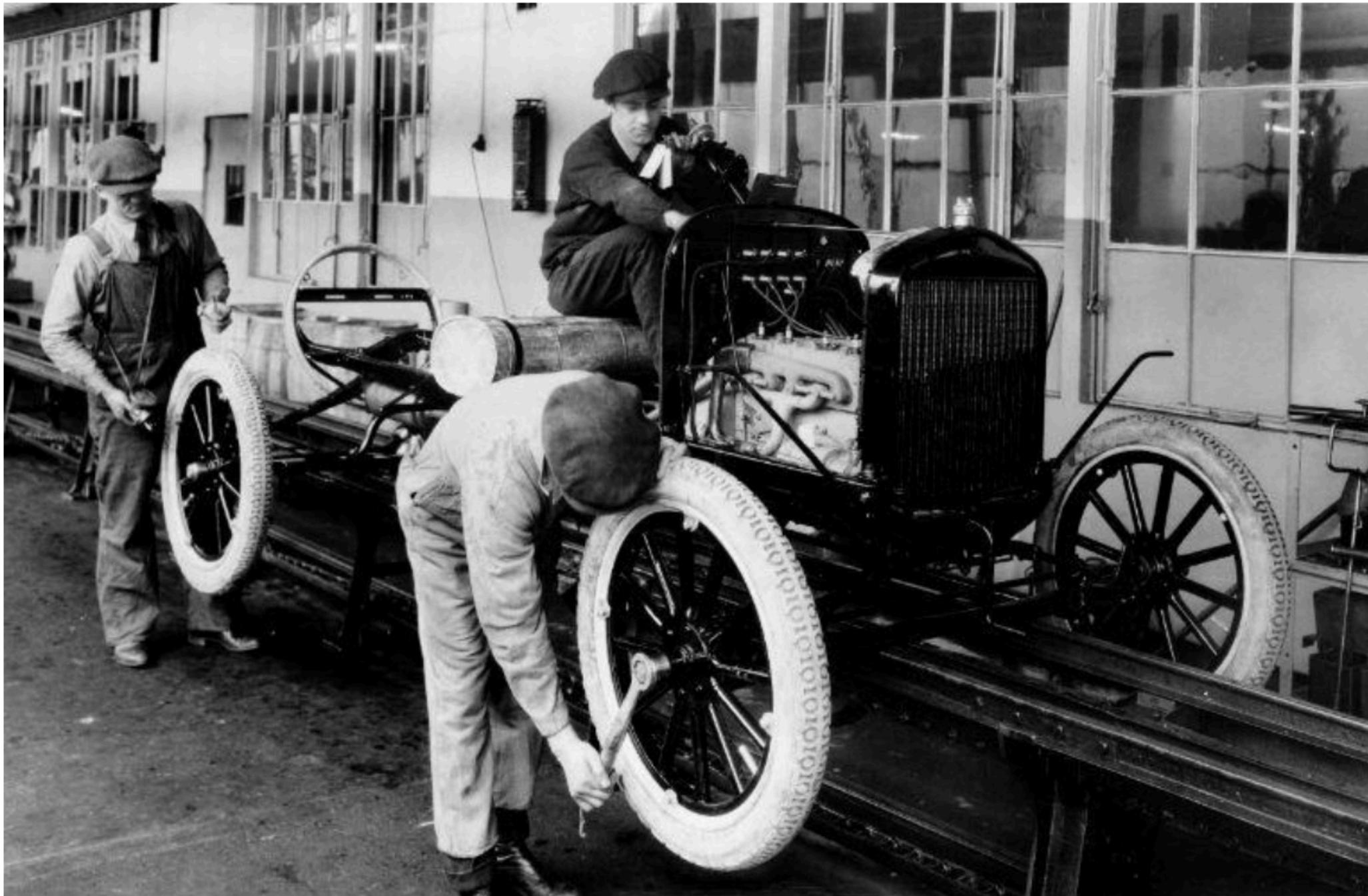
- Kinds of projects:
 - Re-implement techniques from a paper
 - Apply existing synthesis framework to a new domain
 - Extend/improve existing synthesis algorithm or tool
 - Develop a new synthesis algorithm or tool ...
- Judged in terms of
 - Quality of execution
 - Originality
 - Scope

Project



Lets begin the good stuff...

The goal: Automated Programming



A classic goal

The FORTRAN Automatic Coding System

J. W. BACKUS†, R. J. BEEBER†, S. BEST‡, R. GOLDBERG†, L. M. HAIBT†,
H. L. HERRICK†, R. A. NELSON†, D. SAYRE†, P. B. SHERIDAN†,
H. STERN†, I. ZILLER†, R. A. HUGHES§, AND R. NUTT||

INTRODUCTION

THE FORTRAN project was begun in the summer of 1954. Its purpose was to reduce by a large factor the task of preparing scientific problems for IBM's next large computer, the 704. If it were possible for the 704 to code problems for itself and produce as

system is now complete. It has two components: the FORTRAN language, in which programs are written, and the translator or executive routine for the 704 which effects the translation of FORTRAN language programs into 704 programs. Descriptions of the FORTRAN language and the translator form the principal

On the Synthesis of a Reactive Module

Amir Pnueli and Roni Rosner*

Department of Computer Science
The Weizmann Institute of Science
Rehovot 76100, Israel
amir@wisdom.bitnet, roni@wisdom.bitnet

Abstract

We consider the synthesis of a reactive module with input x and output y , specified by the linear temporal formula $\varphi(x, y)$. We show that there exists a program satisfying φ iff the branching time formula $(\forall x)(\exists y)A\varphi(x, y)$ is valid over all tree models. For the restricted case that all variables range over finite domains, the validity problem is decidable, and we present an algorithm for constructing the pro-

gram with input x and output y , specified by the formula $\varphi(x, y)$, is constructed as a by-product of proving the theorem $(\forall x)(\exists y)\varphi(x, y)$. The specification $\varphi(x, y)$ characterizes the expected relation between the input x presented to the program and the output y computed by the program. For example, the specification for a root extracting program may be presented by the formula $|x - y^2| < \epsilon$.

This approach, which may be called the *AE-*

IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, VOL. SE-5, NO. 4, JULY 1979

Synthesis: Dreams \Rightarrow Programs

ZOHAR MANNA AND RICHARD WALDINGER

techniques are presented for deriving programs from specifications. The specifications express the desired program without giving any hint of the algorithm. The basic approach is to transform the specification according to certain rules, until a satisfactory pro-

INTRODUCTION

IN RECENT years there has been increasing activity in the field of program verification. The goal of these efforts is to construct computer systems for determining whether a

What is Program Synthesis



- Automatically finding programs
 - from the underlying programming language/set of components.
 - satisfying the user intent, expressed using some constraints.

This sounds familiar: Compiler/Logic Prog/ML

- Compilers:
 - Fully specified High-Level code → Low-level machine rep.
 - Syntax-directed translation.
- Logic Programming
 - dream: express the requirements in a logical form.
 - generic algorithm for all problems.
- ML
 - Find a function/learned model, whose behavior closely matches the dataset.
 - the space of functions that the algorithm considers is very tightly prescribed
 - linear classifiers, decision trees and neural networks

Synthesis: discover how to perform the desired task. Some notion of Search

Synthesis: general algorithms for general classes of programs, that support recursion or other forms of iteration

```

append:
    push ebp
    mov ebp, esp
    push eax
    push ebx
    push len
    call malloc
    mov ebx, [ebp + 12]
    mov [eax + info], ebx
    mov dword [eax + next], 0
    mov ebx, [ebp + 8]
    cmp dword [ebx], 0
    je null_pointer
    mov ebx, [ebx]

next_element:
    cmp dword [ebx + next], 0
    je found_last
    mov ebx, [ebx + next]
    jmp next_element

found_last:
    push eax
    push addMes
    call puts
    add esp, 4
    pop eax
    mov [ebx + next], eax

go_out:
    pop ebx
    pop eax
    mov esp, ebp
    pop ebp
    ret 8

null_pointer:
    push eax
    push nullMes
    call puts
    add esp, 4
    pop eax
    mov [ebx], eax
    jmp go_out

```

```

void insert(node *xs, int x) {
    node *new;
    node *temp;
    node *prev;

    new = (node *)malloc(sizeof(node));
    if(new == NULL) {
        printf("Insufficient memory.");
        return;
    }
    new->val = x;
    new->next = NULL;
    if (xs == NULL) {
        xs = new;
    } else if(x < xs->val) {
        new->next = xs;
        xs = new;
    } else {
        prev = xs;
        temp = xs->next;
        while(temp != NULL && x > temp->val) {
            prev = temp;
            temp = temp->next;
        }
        if(temp == NULL) {
            prev->next = new;
        } else {
            new->next = temp;
            prev->next = new;
        }
    }
}

```

“Any sufficiently advanced compiler is indistinguishable from a synthesizer”

$$\begin{aligned}
 \text{insert } x \text{ []} &= [x] \\
 \text{insert } x \text{ (y:ys)} &= x:y:ys \\
 | \quad x \leq y &= x:y:ys \\
 | \quad \text{otherwise} &= y:(\text{insert } x \text{ ys})
 \end{aligned}$$



Assembly

C

Haskell

modern program
synthesis

Synthesis: Example

Library/Language

library

sort: list -> list

reverse: list -> list

take: list -> int -> list

sum: list -> int

query

best_ksum: (l : list) -> (k : int) -> int

i/o examples

49	62	82	54	76
----	----	----	----	----

k=2

158

User Intent

best_ksum l k = sum (take (reverse (sort l)) k)

Modern Program Synthesis: FlashFill



[Gulwani 2011]

FlashFill : A Feature of Excel 2013

[Gulwani 2011]

The screenshot shows a Microsoft Excel 2013 window with the ribbon menu open. The 'Table Tools' tab is selected. In the 'Design' section of the ribbon, the 'Quick Layout' dropdown is open, showing options like 'Quick Fill', 'Auto Fill', 'Quick Layout', 'Apply', 'HiLight', 'CurrencyWidget', 'Undo', 'Commit', and 'AddressWidget'. The main worksheet area displays a table titled 'Table116' with 26 rows of data. The first row contains the header 'Column1'. The second row contains the value 'Ana Trujillo' in column A and the full address '357 21th Place SE, Redmond, WA, (757) 555-1634, 140-37-6064, 27171' split across columns B through F. Subsequent rows show various names and addresses, with the address components being automatically filled into multiple columns. The bottom status bar shows 'Ready' and various toolbars like 'ssn', 'FixTrunc2', 'FixTrunc3', 'bigbets', 'CustomerData', 'Dates2', 'Layout', 'Currency', 'Dates', and 'Abbreviations'.

1	Column1	A	B	C	D	E	F
2	Ana Trujillo	357 21th Place SE, Redmond, WA, (757) 555-1634, 140-37-6064, 27171	Redmond	WA	(757) 555-1634	140-37-6064	27171
3	Antonio Moreno	515 93th Lane , Renton, WA, (411) 555-2786, 562-87-3127, 28581					
4	Thomas Hardy	742 17th Street NE, Seattle, WA, (412) 555-5719, 921-29-4931, 24607					
5	Christina Berglund	475 22th Lane , Redmond, WA, (443) 555-6774, 844-35-6764, 30146					
6	Hanna Moos	785 45th Street NE, Puyallup, WA, (376) 555-2462, 515-68-1285, 29284					
7	Frédérique Citeaux	308 66th Place , Redmond, WA, (689) 555-2770, 552-23-2508, 21415					
8	Martin Sommer	887 86th Place , Kent, WA, (715) 555-5450, 870-91-9824, 21536					
9	Laurence Lebihan	944 13th Street NE, Redmond, WA, (620) 555-2361, 649-25-5312, 25252					
10	Elizabeth Lincoln	452 73th Lane NE, Renton, WA, (851) 555-4561, 425-97-6344, 22279					
11	Victoria Ashworth	463 16th Street , Renton, WA, (696) 555-6044, 690-29-7926, 22832					
12	Patricia Simpson	630 20th Street , Redmond, WA, (179) 555-3265, 389-78-3236, 24525					
13	Francisco Chang	683 49th Lane , Seattle, WA, (272) 555-7434, 665-18-6435, 29453					
14	Yang Wang	944 28th Lane , Redmond, WA, (151) 555-2272, 846-78-8452, 24388					
15	Pedro Afonso	411 70th Place , Kent, WA, (170) 555-2964, 774-35-2298, 29485					
16	Elizabeth Brown	971 20th Lane , Puyallup, WA, (373) 555-4134, 476-53-7164, 26417					
17	Sven Ottlieb	676 17th Lane NE, Redmond, WA, (828) 555-1593, 548-73-8633, 27440					
18	Janine Labrune	267 95th Place SE, Seattle, WA, (949) 555-1316, 350-27-8300, 28074					
19	Ann Devon	694 53th Place , Kent, WA, (194) 555-8124, 559-74-4016, 22367					
20	Roland Mendel	581 12th Street NW, Kent, WA, (103) 555-2146, 303-79-1328, 20518					
21	Aria Cruz	594 85th Lane , Renton, WA, (431) 555-1376, 329-93-9992, 21498					
22	Diego Roel	550 22th Lane , Renton, WA, (639) 555-6238, 918-34-5172, 25931					
23	Martine Rancé	688 93th Place NW, Kent, WA, (573) 555-3571, 695-94-3479, 22424					
24							
25							
26							

FlashFill : A Feature of Excel 2013

The screenshot shows a Microsoft Excel 2013 window with the ribbon menu open. The 'Table Tools' tab is selected. A dropdown menu titled 'Quick Fill' is open, showing options like 'Auto Fill' and 'Quick Layout'. The main worksheet area displays a table titled 'Table116' containing 23 rows of data. The first column is labeled 'Column1'. The data includes names and addresses from various locations in Washington State. The 'Table Tools' ribbon tab is highlighted in yellow, indicating it is active.

	A	B	C	D	E	F
1	Column1	Col 2	Col 3	Col 4	Col 5	Col 6
2	Ana Trujillo 357 21th Place SE,Redmond,WA,(757) 555-1634,140-37-6064,27171	Redmond	WA	(757) 555-1634	140-37-6064	27171
3	Antonio Moreno 515 93th Lane ,Renton,WA,(411) 555-2786,562-87-3127,28581	Renton	WA	(411) 555-2786	562-87-3127	28581
4	Thomas Hardy 742 17th Street NE,Seattle,WA,(412) 555-5719,921-29-4931,24607	Seattle	WA	(412) 555-5719	921-29-4931	24607
5	Christina Berglund 475 22th Lane ,Redmond,WA,(443) 555-6774,844-35-6764,30146	Redmond	WA	(443) 555-6774	844-35-6764	30146
6	Hanna Moos 785 45th Street NE,Puyallup,WA,(376) 555-2462,515-68-1285,29284	Puyallup	WA	(376) 555-2462	515-68-1285	29284
7	Frédérique Citeaux 308 66th Place ,Redmond,WA,(689) 555-2770,552-23-2508,21415	Redmond	WA	(689) 555-2770	552-23-2508	21415
8	Martin Sommer 887 86th Place ,Kent,WA,(715) 555-5450,870-91-9824,21536	Kent	WA	(715) 555-5450	870-91-9824	21536
9	Laurence Lebihan 944 13th Street NE,Redmond,WA,(620) 555-2361,649-25-5312,2525	Redmond	WA	(620) 555-2361	649-25-5312	2525
10	Elizabeth Lincoln 452 73th Lane NE,Renton,WA,(851) 555-4561,425-97-6344,22279	Renton	WA	(851) 555-4561	425-97-6344	22279
11	Victoria Ashworth 463 16th Street ,Renton,WA,(696) 555-6044,690-29-7926,22832	Renton	WA	(696) 555-6044	690-29-7926	22832
12	Patricia Simpson 630 20th Street ,Redmond,WA,(179) 555-3265,389-78-3236,24525	Redmond	WA	(179) 555-3265	389-78-3236	24525
13	Francisco Chang 683 49th Lane ,Seattle,WA,(272) 555-7434,665-18-6435,29453	Seattle	WA	(272) 555-7434	665-18-6435	29453
14	Yang Wang 944 28th Lane ,Redmond,WA,(151) 555-2272,846-78-8452,24388	Redmond	WA	(151) 555-2272	846-78-8452	24388
15	Pedro Afonso 411 70th Place ,Kent,WA,(170) 555-2964,774-35-2298,29485	Kent	WA	(170) 555-2964	774-35-2298	29485
16	Elizabeth Brown 971 20th Lane ,Puyallup,WA,(373) 555-4134,476-53-7164,26417	Puyallup	WA	(373) 555-4134	476-53-7164	26417
17	Sven Ottlieb 676 17th Lane NE,Redmond,WA,(828) 555-1593,548-73-8633,27440	Redmond	WA	(828) 555-1593	548-73-8633	27440
18	Janine Labrune 267 95th Place SE,Seattle,WA,(949) 555-1316,350-27-8300,28074	Seattle	WA	(949) 555-1316	350-27-8300	28074
19	Ann Devon 694 53th Place ,Kent,WA,(194) 555-8124,559-74-4016,22367	Kent	WA	(194) 555-8124	559-74-4016	22367
20	Roland Mendel 581 12th Street NW,Kent,WA,(103) 555-2146,303-79-1328,20518	Kent	WA	(103) 555-2146	303-79-1328	20518
21	Aria Cruz 594 85th Lane ,Renton,WA,(431) 555-1376,329-93-9992,21498	Renton	WA	(431) 555-1376	329-93-9992	21498
22	Diego Roel 550 22th Lane ,Renton,WA,(639) 555-6238,918-34-5172,25931	Renton	WA	(639) 555-6238	918-34-5172	25931
23	Martine Rancé 688 93th Place NW,Kent,WA,(573) 555-3571,695-94-3479,22424	Kent	WA	(573) 555-3571	695-94-3479	22424
24						
25						
26						

Under the hood

Automating String Processing in Spreadsheets Using Input-Output Examples

Sumit Gulwani

Microsoft Research, Redmond, WA, USA
sumitg@microsoft.com

Abstract

We describe the design of a string programming/expression language that supports restricted forms of regular expressions, conditionals and loops. The language is expressive enough to represent a wide variety of string manipulation tasks that end-users struggle with. We describe an algorithm based on several novel concepts for synthesizing a desired program in this language from input-output examples.

their task [9]. More significantly, programmers perform tedious and repetitive tasks such as transforming entities like names/phone-numbers/dates from one format to another, data cleansing, extracting data from several text files or web pages into a single document, etc. Spreadsheet systems like Excel allow users to write macros using a rich inbuilt library of string and numerical functions, or to write arbitrary scripts using a variety of programming languages like Visual Basic, or .Net. Since end-users are not proficient in programming, they find it too difficult to write desired macros or scripts.

Abstract
We describe the design of a string programming/expression language that supports restricted forms of regular expressions, conditionals and loops. The language is expressive enough to represent a wide variety of string manipulation tasks that end-users struggle with. We describe an algorithm based on several novel concepts for synthesizing a desired program in this language from input-output examples. The synthesis algorithm in this language is very efficient taking a fraction of a second for various benchmark examples. The synthesis algorithm is interactive and has fast convergence features: if can rank multiple solutions and it supports an active interaction model wherein the user is prompted to provide outputs on inputs that may have multiple computational interpretations.

Excel 2013's coolest new feature that should have been available years ago"

Email	Last Name
Airplane.Lady@lufthansa.com	Lady
Excel.Team@microsoft.com	Team
Rishabh.Singh@mit.edu	Singh
Vu.Le@ucdavis.edu	Le
Alex.Polozov@uw.edu	Polozov
Rico.Malvar@microsoft.com	Malvar
Ben.Zorn@microsoft.com	Zorn
Piali.Choudhury@microsoft.com	Choudhury
Dany.Rouhana@microsoft.com	Rouhana
Shobana.Balakrishnan@microsoft.com	Balakrishnan
Vasudev.Gulwani@gmail.com	Gulwani
Sumay.Gulwani@gmail.com	Gulwani
Mooly.Sagiv@acm.org	Sagiv

Automating String Processing in Spreadsheets Using Input-Output Examples

Sumit Gulwani
Microsoft Research, Redmond, WA, USA
sumitg@microsoft.com

their task [9]. More significantly, programming is still required to perform tedious and repetitive tasks such as transforming entities like names/phone-numbers/dates from one format to another, data cleansing, extracting data from several text files or web pages into a single document, etc. Spreadsheet systems like Microsoft Excel allow users to write macros using a rich inbuilt library of string and numerical functions, or to write arbitrary scripts using a variety of programming languages like Visual Basic, or .Net. Since end-users are not proficient in programming, they find it too difficult to write desired macros or scripts.

We have performed an extensive case study of spreadsheet forums and identified that string processing is one of the most common class of programming problems that end-users struggle with. This is not surprising given that languages like Perl,

FlashFill++: Scaling Programming by Example by Cutting to the Chase

JOSÉ CAMBRONERO*, Microsoft, USA
SUMIT GULWANI*, Microsoft, USA
VU LE*, Microsoft, USA

Major Idea is VSAs

Modern Program Synthesis: Sketch [Solar-Lezama 2013]

- Problem: isolate the least significant zero bit in a word
 - example: 0010 0101 → 0000 0010
- Easy to implement with a loop

```
int W = 32;  
  
bit[W] isolate0 (bit[W] x) {      // W: word size  
    bit[W] ret = 0;  
    for (int i = 0; i < W; i++)  
        if (!x[i]) { ret[i] = 1; return ret; }  
}
```

- Can this be done more efficiently with bit manipulation?
 - Trick: adding 1 to a string of ones turns the next zero to a 1
 - i.e. 000111 + 1 = 001000

Sketch: space of possible implementations

```
/**  
 * Generate the set of all bit-vector expressions  
 * involving +, &, xor and bitwise negation (~).  
 */  
  
Missing constants!  
  
generator bit[W] gen(bit[W] x){  
    if(??) return x;  
    if(??) return ??;  
    if(??) return ~gen(x);  
    if(??){  
        return { | gen(x) (+ | & | ^) gen(x) | };  
    }  
}
```

Sketch Idea : program space as a parametric program $P[c]$

Different values of c gives different program in the space

translate requirements on the behavior of the program $P[c]$ into constraints on the parameters c .

Any value of c that satisfies the constraints is guaranteed to lead to a program satisfying all the requirements.

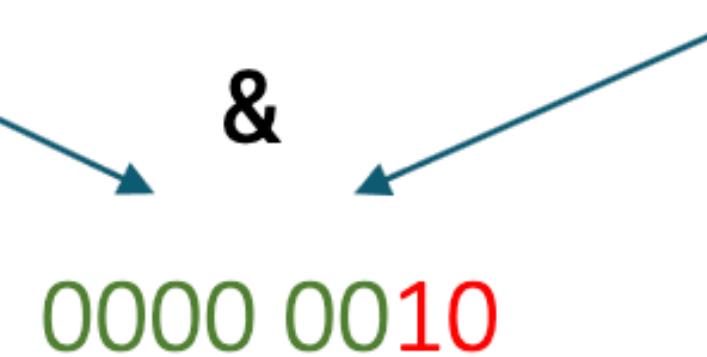
Sketch: synthesis goal

```
generator bit[W] gen(bit[W] x, int depth){  
    assert depth > 0;  
    if(??) return x;  
    if(??) return ??;  
    if(??) return ~gen(x, depth-1);  
    if(??){  
        return {| gen(x, depth-1) (+ | & | ^) gen(x, depth-1) |};  
    }  
}  
  
bit[W] isolate0fast (bit[W] x) implements isolate0 {  
    return gen(x, 3);  
}
```

Sketch: output

```
bit[W] isolate0fast (bit[W] x) {  
    return (~x) & (x + 1);  
}
```

$$\begin{array}{ll} \sim 0010\ 0101 & 0010\ 0101 + 1 \\ = \textcolor{red}{1101\ 1010} & = 0010\ 01\textcolor{red}{10} \end{array}$$



Modern Program Synthesis: Synquid

[Polikarpova et al. 2016]

Problem: intersection of sets represented as strictly sorted lists

- example: intersect [4, 8, 15, 16, 23, 42] [8, 16, 32, 64] → [8, 16]

Also: we want a guarantee that it's correct on all inputs!

Synquid: synthesis goal and components

Step 1: define synthesis goal as a *type*

```
intersect :: xs:List a    →      ys:List a    →  
                  List a
```



Step 2: define a set of components

- Which primitive operations is our function likely to use?
- Here: {Nil, Cons, <}

Synquid Output

intersection = \xs . \ys .	xs	ys	result
match xs with	[4, 8, 15, 16, 23, 42]	[8, 16, 32, 64]	
Nil -> xs			
Cons x xt ->	[8, 15, 16, 23, 42]	[8, 16, 32, 64]	[8]
match ys with	[15, 16, 23, 42]	[16, 32, 64]	
Nil -> ys			
Cons y yt ->	[16, 23, 42]	[16, 32, 64]	[8, 16]
if x < y	[23, 42]	[32, 64]	
then intersection xt ys			
else	[42]	[32, 64]	
if y < x	[42]	[64]	
then intersection xs yt			
else Cons x (intersection xt yt)	[]	[64]	

Modern Program Synthesis: GitHub Copilot

```
// find all images  
// and add a green border around them  
// and add class “githubCopilot” to them  
  
function go() {  
  
    var images = document.getElementsByTagName('img');  
    for (var i = 0; i < images.length; i++) {  
        if (images[i].className.indexOf('githubCopilot') == -1) {  
            images[i].className += ' githubCopilot';  
            images[i].style.border = '1px solid green';  
        }  
    }  
}
```

input

output

Other Program Synthesis Successes

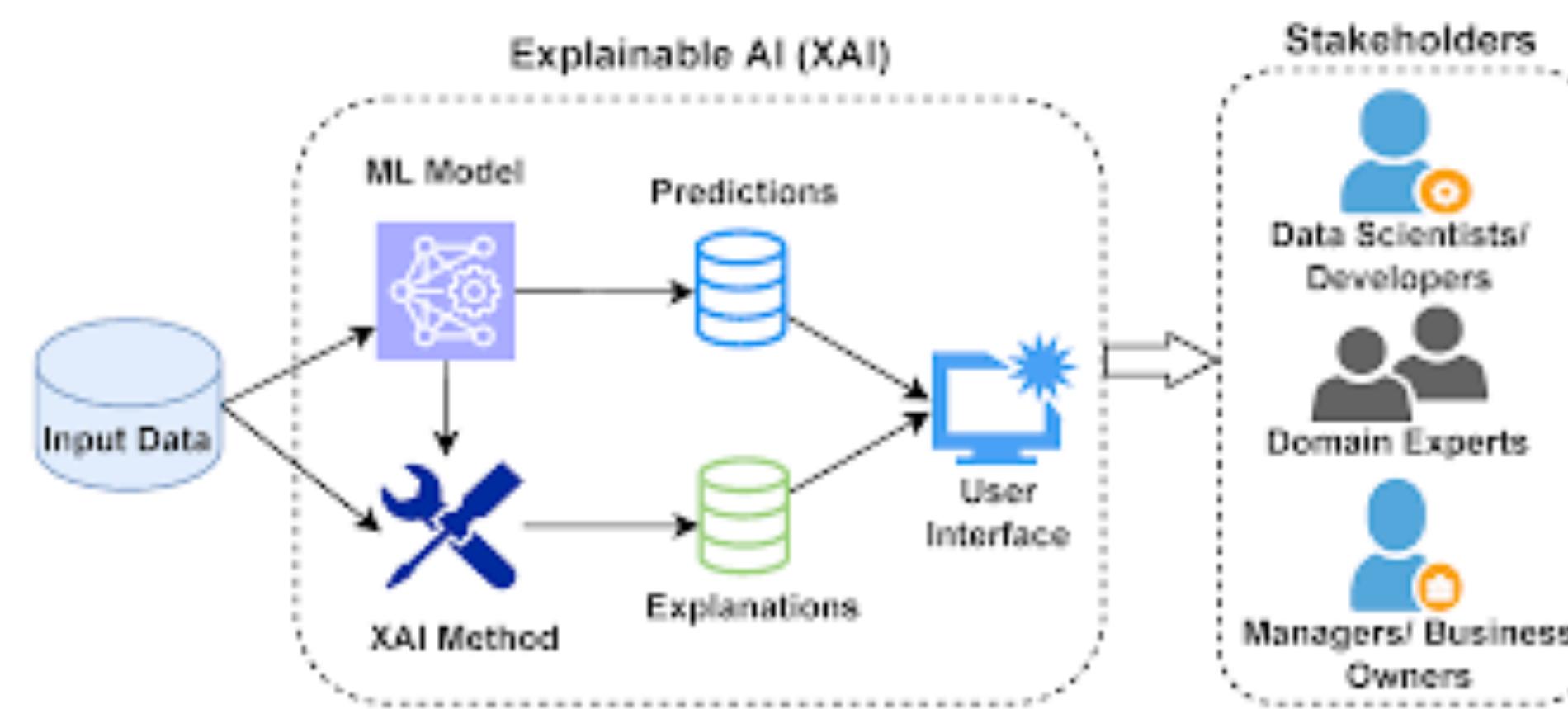
Robotics design and path planning



Data Migration

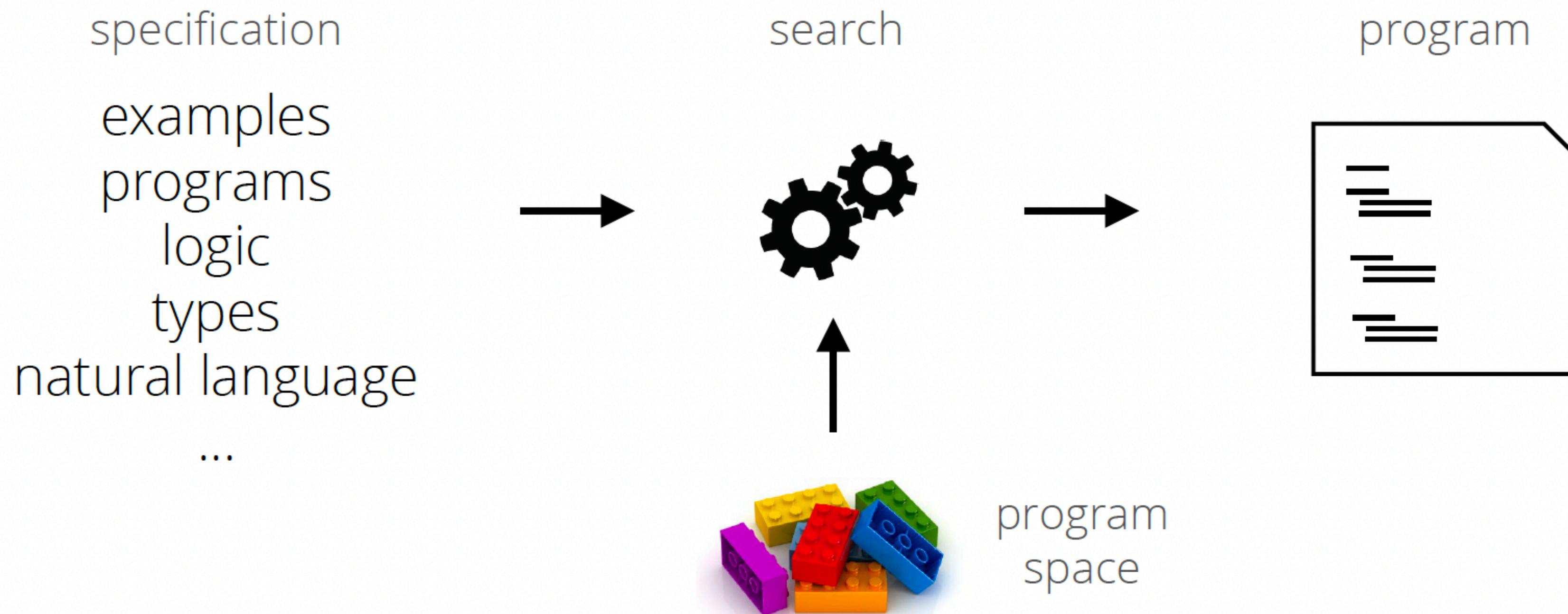


Explainable AI (XAI)



Clement, T. et. al. 2023

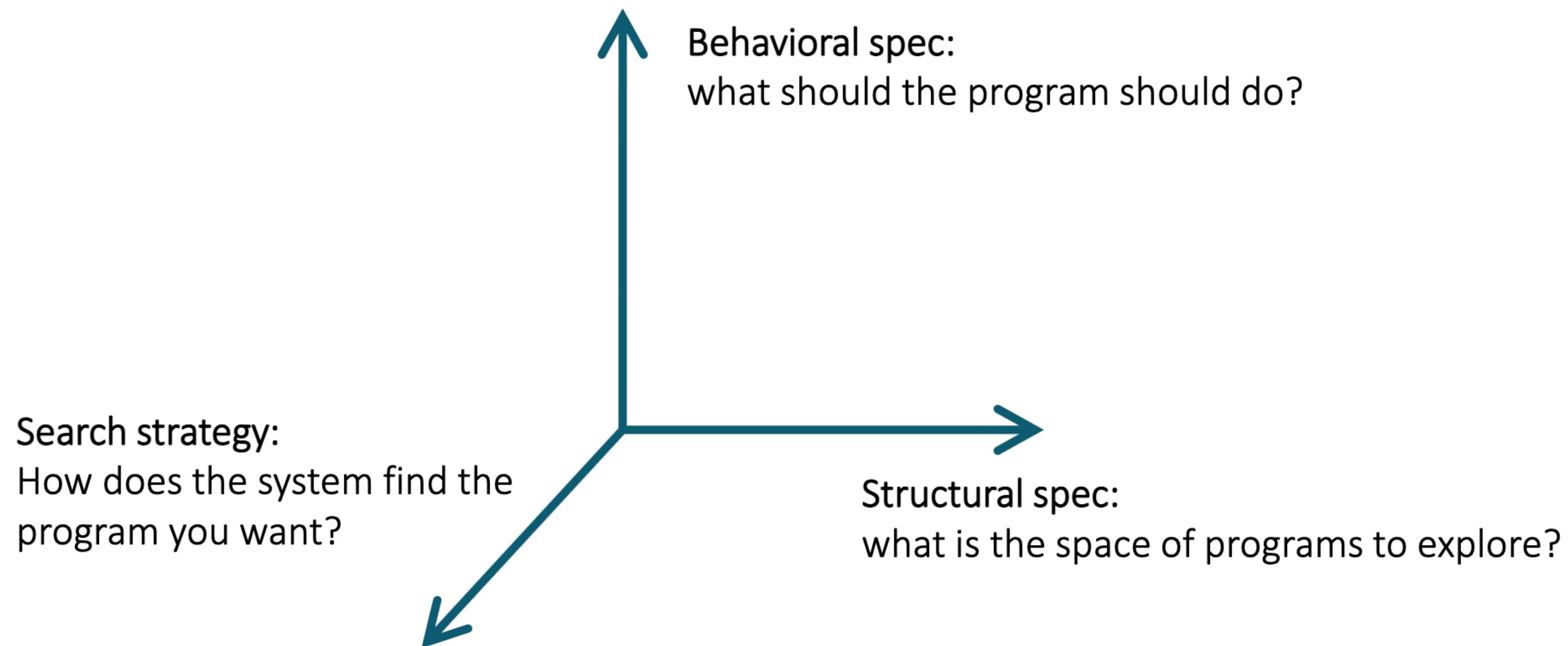
Program Synthesis



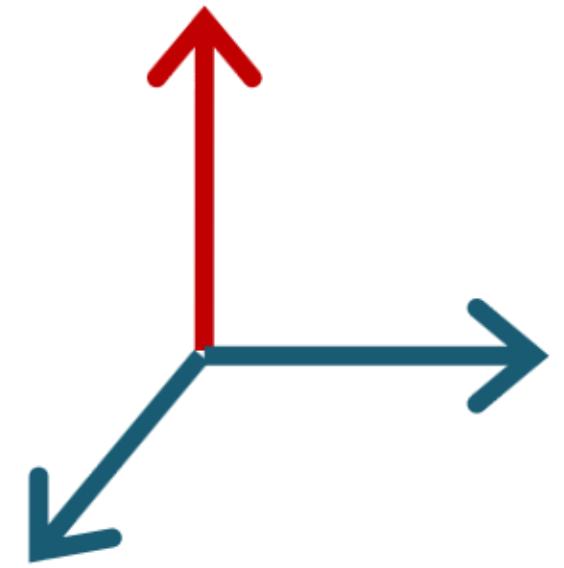
Notice the Duality with Program Verification

Dimensions in Program Synthesis

[Gulwani 2010]



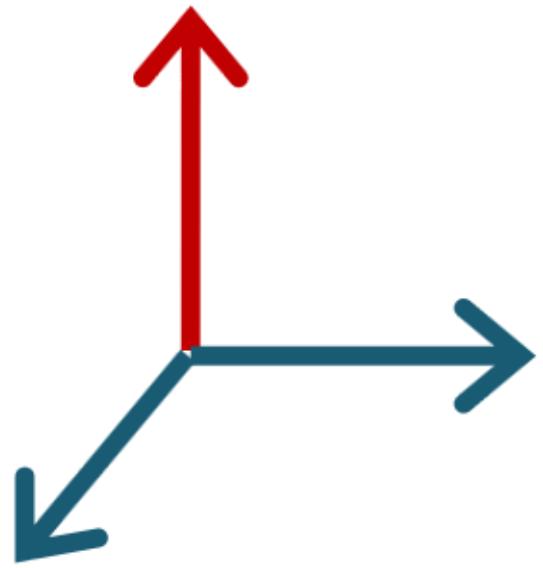
Behavioral Spec



- How do you tell the system what the program should do?
 - What is the input language / format?
 - What is the interaction model?
 - What happens when the intent is ambiguous?

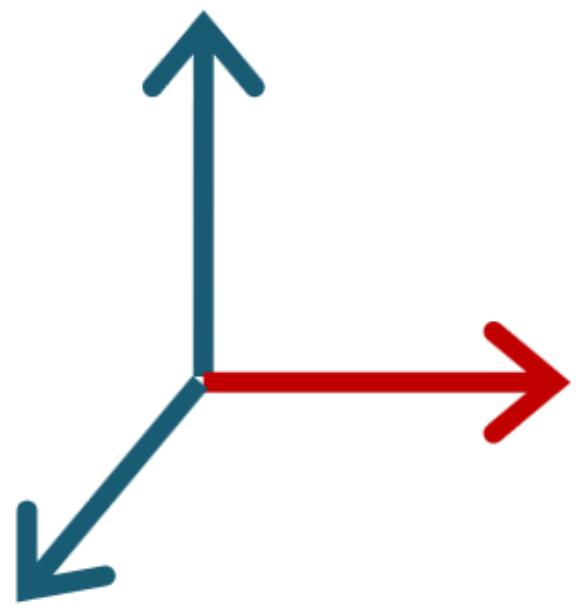
Q: What did the behavioral spec look like in FlashFill / Sketch / Synquid / Copilot?

Behavioral Spec: Examples



- Input/output examples
- Reference implementation
- Formal specifications (pre/post conditions, types, ...)
- Natural language
- Context

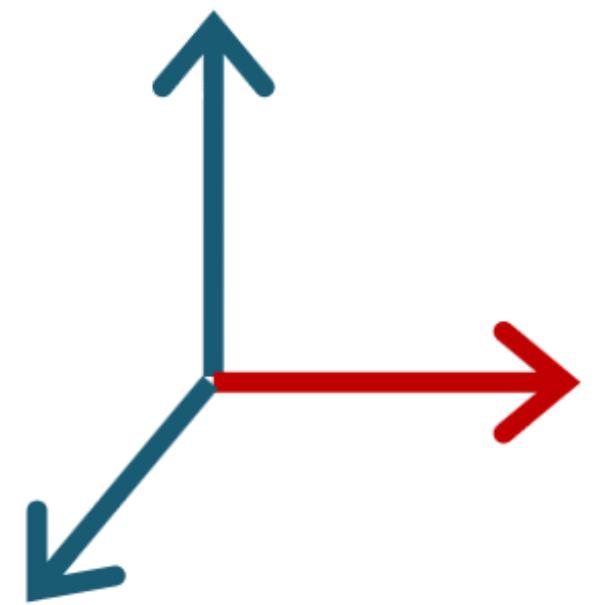
Structural Spec



- What is the space of programs to explore?
 - Large enough to contain interesting programs, yet small enough to exclude garbage and enable efficient search
 - Built-in or user-defined or learned from existing code?

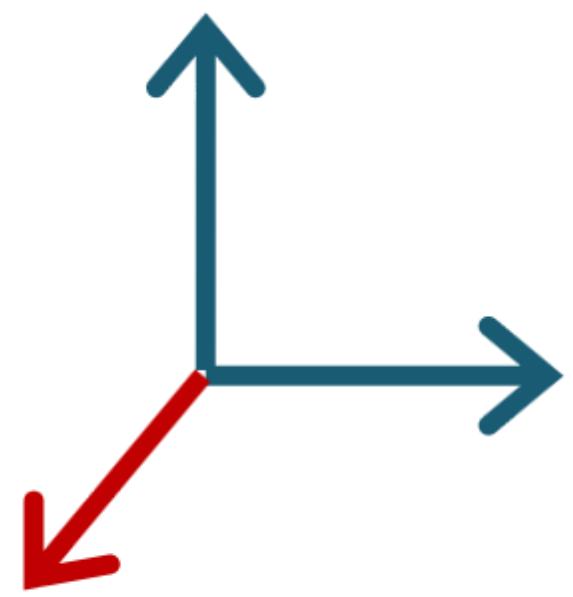
Q: What did the structural spec look like in FlashFill / Sketch / Synquid / Copilot?

Structural Spec: Examples



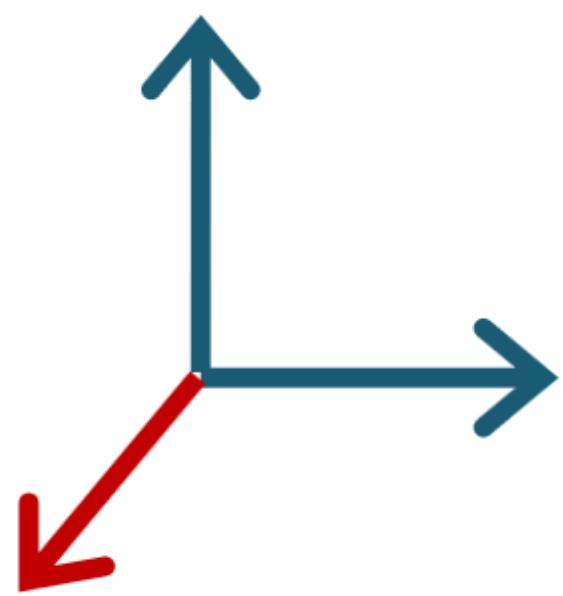
- Built-in DSLs — e.g. DSL for mathematical expressions/ Excel
- User-defined DSL (grammar)
- User-provided components. — Component-based synthesis —
- Languages with synthesis constructs
 - e.g. generators in Sketch
- Learned language model

Search Strategies



- Synthesis is search:
 - Find a program in the space defined by structural constraints that satisfies behavioral constraints
- Challenge: the space is astronomically large
 - The search algorithm is the heart of a synthesis technique
- How does the system find the program you want?
 - How does it know it's the program you want?
 - How can it leverage structural constraints to guide the search?
 - How can it leverage behavioral constraints to guide the search?

Search Strategies: Examples



- Enumerative (explicit) search
 - exhaustively enumerate all programs in the language in the order of increasing size
- Deductive
 - Top-down search with recursive reduction of problem to smaller ones.
- Stochastic and Statistical search
 - random exploration of the search space guided by a fitness function
- Representation-based search
 - use a data structure to represent a large set of programs
- Constraint-based search
 - translate to constraints and use a solve

Applications

- Data Wrangling
 - Transformations
 - Syntactic String Transformations

FirstName.Lastname@domain → Firstname Lastname

- Semantic Transformations

selling price =

$f(\text{Name}, \text{Selling date}, \text{MarkupRec}, \text{CostRec})$

- Splitting Table Transformations

- Extractions

- Layouts

A	B
1 Email	Column 2
2 Nancy.FreeHafer@fourthcoffee.com	nancy freehafer
3 Andrew.Cencici@northwindtraders.com	andrew cencici
4 Jan.Kotas@litwareinc.com	jan kotas
5 Mariya.Sergienko@gradicdesigninstitute.com	mariya sergienko
6 Steven.Thorpe@northwindtraders.com	steven thorpe
7 Michael.Neipper@northwindtraders.com	michael neipper
8 Robert.Zare@northwindtraders.com	robert zare
9 Laura.Giussani@adventure-works.com	laura giussani
10 Anne.HL@northwindtraders.com	anne hl
11 Alexander.David@contoso.com	alexander david
12 Kim.Shane@northwindtraders.com	kim shane
13 Manish.Chopra@northwindtraders.com	manish chopra
14 Gerwald.Oberleitner@northwindtraders.com	gerwald oberleitner
15 Amr.Zaki@northwindtraders.com	amr zaki
16 Yvonne.McKay@northwindtraders.com	yvonne mckay
17 Amanda.Pinto@northwindtraders.com	amanda pinto

Input v_1	Input v_2	Output
Stroller	10/12/2010	\$145.67+0.30*145.67
Bib	23/12/2010	\$3.56+0.45*3.56
Diapers	21/1/2011	\$21.45+0.35*21.45
Wipes	2/4/2009	\$5.12+0.40*5.12
Aspirator	23/2/2010	\$2.56+0.30*2.56

MarkupRec			CostRec		
Id	Name	Markup	Id	Date	Price
S30	Stroller	30%	S30	12/2010	\$145.67
B56	Bib	45%	S30	11/2010	\$142.38
D32	Diapers	35%	B56	12/2010	\$3.56
W98	Wipes	40%	D32	1/2011	\$21.45
A46	Aspirator	30%	W98	4/2009	\$5.12
...	A46	2/2010	\$2.56
		

More Applications (Many of these will be part of the Projects)

- Graphics Programming
- Code Repair
- Code Suggestions
- Synthesizing Error-prone, hard to write programs:
 - Distributed Programming : CRDTs
 - Concurrent Programs:
- Modeling of Systems
 - Probabilistic Programs

Structure of the course

- Module 1: Synthesis of Simple Programs
 - Easy to decide when a program is correct
 - Challenge: search in a large space
- Module 2: Synthesis of Complex Programs
 - Deciding when a program is correct can be hard
 - Search in a large space is still a problem
- Module 3: Advance Topics
 - Neural Synthesis, Neural + Symbolic Approaches, Synthesis for xAI
 - Synthesis + X ($X \in \{\text{Frameworks, Compilers, Network, Databases, etc.}\}$)

Reading Weeks 1 and 2

- Topic: Enumerative synthesis from examples
 - Paper: Alur, Radhakrishna, Udupa. Scaling Enumerative Program Synthesis via Divide and Conquer
 - Review due Wednesday
 - Link to PDF on the course web page
- Submit through Google Form (link will be in webpage/Classroom)
- Project:
 - Teams due in two weeks.
 - Submit through a Google Sheet (check email for invite and instructions)

Announcements

- Non-CS students
- Registration on ERP as well as Google Classroom
- Start looking for Projects
- Plagiarism Policy.
 - A bit different.