



Natural Language Programming

David Vadas

The University of Sydney

Supervisor: James Curran



About Me

- Undergraduate degree in computer science
- Honours work in NLP
 - POS tagging
 - Maximum Entropy



Overview

- What do I mean by Natural Language Programming?
- Background
- Components of the system
- Work so far
- Work to do

Natural Language Programming

- Write a program (code) in English
- Unrestricted syntax

Read in a number

Add 2 to the number

Print out the number

```
number = sys.stdin.readline()
```

```
number += 2
```

```
print number
```

Compare each number in the list to the min

If the number is smaller, print it out

```
for num in list:
```

```
    if num < min:
```

```
        print number
```

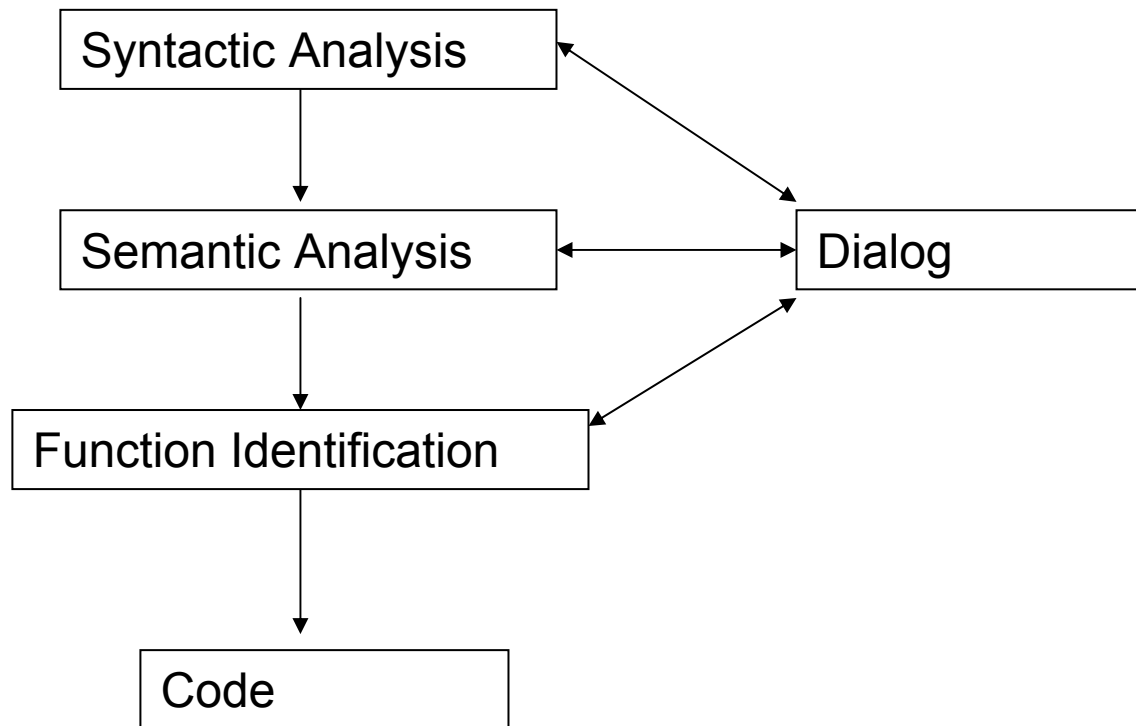
Why?

- Why is a natural language better than a programming language?
 - Easier to understand
 - Requires less technical knowledge to write
 - Especially suitable for people just learning to program
- An interesting challenge

Background

- SHRDLU (1972)
 - Interactive, dialog system
- NLIDBs (1980's)
 - Mostly domain specific
 - Dialog
- Pane and Myers (2000)
 - How do people express code?
- Hugo Liu (2003)
 - Creating a code skeleton

System Components



Syntactic Analysis

- Parsing with Combinatory Categorical Grammar (CCG)
- Small corpus of ~400 sentences
 - Imperative, very different to newspaper text
 - Manually annotated



Work to be done

- Semantic Analysis
- Function identification
 - Deciding on primitives
- Dialog
 - Resolve ambiguity
 - Ask for elaborations
 - Integration with other components

More work

- Anaphora
 - Very prevalent in this type of text
- Empirical evaluation
 - Parser
 - Finding functions and arguments
- User evaluation
 - 1st years and/or experienced programmers



Conclusion

- Unusual idea, interesting application
- Many different areas
- Still in early stages