Lecture 2: Formal Systems and Languages

CS150: Computer Science
University of Virginia
Computer Science

David Evans
http://www.cs.virginia.edu/evans

Menu

- Nuclear Weapons
- Questions from Lecture 1 Notes
- Survey Summary
- Formal Systems
  - *MIU*-system
- Languages
  - English
  - Scheme

Megabytes vs. Megatons

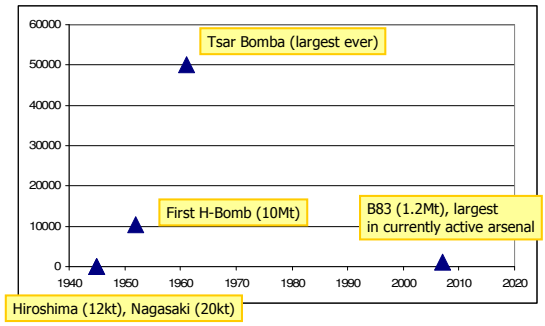
- Computing: 30,000,000 times increase in power since 1969
- Nuclear weapons?

Tsar Bomba 50 Megaton explosion, island in Arctic Sea, 1961

If Nuclear Weapons followed Moore's Law...

- 30M * 50 Megatons = 1.5 Teratons
- 1 Megaton TNT = $4.184 * 10^{15}$ Joules
- 1.5 Teratons TNT = $6.3 * 10^{21}$ Joules
- Energy from Sun to Earth
  = 4×10^{18} Joules/ Year
- One bomb today ~ all the energy to reach the Earth from the Sun since 400 AD

Actual Nuclear Weapons



If it takes 60 seconds to compute a photomosaic for Problem Set 1 today on a typical PC, estimate how long it will take CS150 students in 2010 to compute the same photomosaic? How long will it take in 2013?

```
> (/ (* (- 2010 2007) 12) 18)
2
> (/ 60 (* 2 2))
15
> (/ (* (- 2013 2007) 12) 18)
4
> (/ 60 (* 2 2 2 2))
15/4
> (exact->inexact (/ 60 (* 2 2 2 2)))
3.75
```

Difference in years * 12 = number of months
Number of months / 18 = number of doublings according to Moore's Law

60 seconds today, 2 doublings by 2010
15 seconds in 2010

Reality check: Moore's "law" is just an "observation". We'll see one reason later today why it won't continue forever.

60 seconds today, 4 doublings by 2013
3.75 seconds in 2013

Are there any non-recursive natural languages? What would happen to a society that spoke one?

Not for humans at least.
They would run out of original things to say.

Chimps and Dolphins are able to learn non-recursive "languages" (some linguists argue they are not really "languages"), but **only humans can learn recursive languages**.

Running out of Ideas

"Its all been said before."

Eventually true for a non-recursive language.

Never true for a recursive language.
There is always something original left to say!

Post Production Systems

Production Systems

- Set of symbols
  - *Primitives*
- Set of rules for manipulating symbols
  - Hofstadter: Rules of Production, Rules of Inference
  - Also: Rules of Combination

The MIU System

- Symbols: M, I, U
- Rules of Production:
  - **Rule I:** If you have a string ending in I, you can add a U at the end.
  - **Rule II:** Suppose you have Mx. Then you may add Mxx to your collection.
  - **Rule III:** If III occurs in one of the strings in your collection you may make a new string with U in place of III.
  - **Rule IV:** If UU occurs inside one of your strings, you can drop it.

MIU System Example

Start with MUI, produce MIU

Rules of Production:

Rule I: If you have a string ending in I, you can add a U at the end.

Rule II: Suppose you have Mx. Then you may add Mxx to your collection.

Rule III: If III occurs in one of the strings in your collection you may make a new string with U in place of III.

Rule IV: If UU occurs inside one of your strings, you can drop it.

Survey Summary

- 53 Responses
  - **63 are registered**
- Problem Set Partners
  - If you selected "Yes" for the question about wanting to be assigned a partner for PS1, you should have received an email from me telling you who your partner is
  - For PS2 everyone will be assigned a partner
  - For others, some you will choose, others you may be assigned

Very Diverse Class

- Years: 12 First, 15 Second, 18 Third, 7 Fourth+
- Majors:
  - 19 Computer Science
  - 11 Undecided
  - 7 Cognitive Science
  - 3 Economics, Math
  - 2 Psychology
  - 1 Anthropology, Architecture, Commerce, Foreign Affairs, Media Studies, Music, Philosophy, Systems Engineering

Survey Responses Continued

- Previous programming: 19 None, 32 Some
- Food: 28 Bodos, 11 Krispy Kreme, 10 pizza, 1 Korean Food, 1 Outback, 1 Paccino's, 1 Arch's, 1 Dunkin Donuts
- Topic: 18 Google Maps, 16 Facebook, 5 Second Life, 5 Java

See course website (by Monday) for my responses to questions and survey summary

Languages

What is a language?

Webster:

A ~~systematic~~ means of communicating ~~ideas or feelings~~ by the use of ~~conventionalized~~ signs, sounds, gestures, or marks having ~~understood~~ meanings.

Linguist's Definition

(Charles Yang)

A description of pairs (*S*, *M*), where *S* stands for sound, or any kind of surface forms, and *M* stands for meaning.

A theory of language must specify the properties of *S* and *M*, and how they are related.

Languages and Formal Systems

What is the difference between a formal system and a language?

With a language, the surface forms have **meaning**.

Caveat: computer scientists often use *language* to mean just a set of surface forms.

What are languages made of?

- **Primitives** (almost all languages have these)
  - The simplest surface forms with **meaning**
- **Means of Combination** (**all** languages have these)
  - Like Rules of Production for Formal Systems
  - Ways to make new surface forms from ones you already have
- **Means of Abstraction** (all **powerful** languages have these)
  - Ways to use simple surface forms to represent complicated ones

Does English have these?

- Primitives
  - ~~Words (?)~~
    - e.g., "antifloccipoccinihilipilification" – **not** ~~a primitive~~
  - Morphemes – smallest units of meaning
    - e.g., **anti-** ("opposite")
- Means of combination
  - e.g., *Sentence* ::= *Subject Verb Object*
  - Precise rules, but not the ones you learned in grammar school

Ending a sentence with a preposition is something up with which we will not put.
Winston Churchill

Does English have these?

- Means of abstraction
  - Pronouns: she, he, it, they, which, etc.
  - Confusing since they don't always mean the same thing, it depends on where they are used.

The "**these"** in the slide title is an abstraction for the three elements of language introduced 2 slides ago.

The "**they"** in the confusing sentence is an abstraction for pronouns.

How should we describe languages?

Backus Naur Form

symbol ::= *replacement*

We can replace *symbol* with *replacement*

A ::= *B* means anywhere you have an *A*, you can replace it with a *B*.

nonterminal – symbol that appears on left side of rule

terminals – symbol that **never** appears on the left side of a rule

BNF Example

Sentence ::= *NP Verb*

NP ::= *Noun*

Noun ::= **Dave**

Noun ::= **Scheme**

Verb ::= **rocks**

Verb ::= **sucks**

What are the *terminals*?

Dave, **Scheme**, **rocks**, **sucks**

How many different things can we express with this language?

4, but only 2 are true.

BNF Example

Sentence ::= *NP Verb*

NP ::= *Noun*

NP ::= *Noun* **and** *NP*

Noun ::= **Dave**

Noun ::= **Scheme**

Verb ::= **rocks**

Verb ::= **sucks**

How many different things can we express with this language?

Infinitely many!
Recursion is powerful.

Most Essential Scheme

Expr	::=	*PrimitiveExpr*
PrimitiveExpr	::=	***Number***
PrimitiveExpr	::=	**+** | ***** | **<=** | ...
Expr	::=	***Name***
Expr	::=	*ApplicationExpr*
ApplicationExpr	::=	**(***Expr MoreExprs***)**
MoreExprs	::=	
MoreExprs	::=	*Expr MoreExprs*

This is enough for everything you need to write for PS1

Charge

- Problem Set 1: due Monday
- Lab Hours: posted on website
  - Now and Sunday 4-5:30, 8-9:30
  - Take advantage of them!
  - If you can, follow us to lab now