# Automata and Formal Languages

# CS138, Fall 2007

Wim van Dam
Room 2151, Harold Frank Hall
vandam@cs.ucsb.edu
http://www.cs.ucsb.edu/~vandam/

# Current Enrollment

- At the moment we are at the maximum of 42 students.
- If you decide to drop the course, let us know sooner rather than later.

# Formalities

- Latest course news and slides can be found at http://vandamteaching.googlepages.com/cs138
- Do not visit the old CS138 sites
- The required reader should be available at The Alternative Copy Shop (6556 Pardall Rd. IV)
- Let me know how all of this works out

# Tentative CS138 Schedule

Tuesday 9:30–10:45: Class (Phelps 1508)

Wednesday afternoon: Homework due

Wednesday evening: Homework announcement

Thursday 9:30–10:45: Class (Phelps 1508)

Friday 12:00–12:50: Discussion (Phelps 1401)

Office hours are “to be announced”

You are expected to regularly check the CS138 web site for last minute information.

# Discussion Sessions

- TAs: Pegah Kamousi and Shravan Samindla
- The TAs will take turns: homework announced and discussed (by the TA) in week N, is due in week N+1 and will be returned in week N+2 (by the same TA).

# HW, Exams

**Homework:** Done on an individual basis.
In CS138 it is all about proofs: you have to able to convince others of your results. If your proof is sloppy or unnecessarily long, you will lose points.

**Examinations:** Much like homework but now with a time limit. Do no expect to get a 100% score.

The final grades are determined by "curving".

CS138 does not have projects
Examination schedule is to be announced.

# Rules of the Game

**Homework:** Collaboration is fine, but it should be you alone who writes up the answers. You should be able to explain your answers when questioned.

**If you are caught cheating you will get an "F".** UCSB's misconduct policies will be strictly enforced.

**Late homework policy:** Homework that is late will not be accepted and thus gets a score 0.

Grading homework or exams for a second time can lead to a decrease in your grade.

# Staying Up-To-Date

Frequently check the web site of the course at:
http://vandamteaching.googlepages.com/cs138

In case of exceptional situations:
email me as soon as possible.

When emailing me: make it clear that it is about CS138 and sign your message.

# Reader vs Slides vs Homework

- The reader is made up of several books, so the notation might change between the chapters, such is life.
- The slides in class will be posted on the CS138 site.
- "If it is not discussed in class, it will not be on the exams."
- The homework questions are the ultimate indicator of the material that you can expect for the examinations.

# This Week and the Next

- There will be no Discussion tomorrow
- The first homework assignment will be announced next Wednesday, October 3 (due one week later)

# What CS138 is about

# A Poll

What … do people frequently use

- operating system?
- editor?
- browser?
- programming language?
- search commands?

# Highlights in TCS

• Only in the 20th century did we get a true understanding of what it means to be able to solve a problem. This was done by defining the notion of a universal computer (cf. the Turing machine model of computation).
It was discovered that there are **uncomputable** problems that no computer ever will be able to solve.

• If we are capable of solving a problem in theory, **in practice** the problem might be **too hard** to solve within reasonable time and memory restrictions.

• The computability/complexity of a problem can **depend** on the computational **model** that you use to solve it.

# What CS138 is About

*"What does it mean when we say that we can compute the solution to a problem?"*

**Computability:**

Under which assumptions on our computer can we solve a given problem?

How complex does our computer need to be?

**Complexity:**

Assuming that we can solve a problem, how hard is it to solve? Think: memory and time requirements.

# CS138 Components

- **Automata and Languages**
  How to define computation
- **Regular Languages and Finite Automata**
  What can we compute with only finite memory?
- **Context Free Languages and Pushdown Automata**
  Computation with a stack memory
- **Turing Machines and Computability**
  General computation and its inherent limitations

# Automata: The Methods and the Madness

# Introduction

- Chapter 1: "Automata: The Methods and the Madness", from J.E. Hopcroft, R. Motwani and J.D. Ullman's *Introduction to Automata Theory, Languages and Computation,* 2001, Second Edition

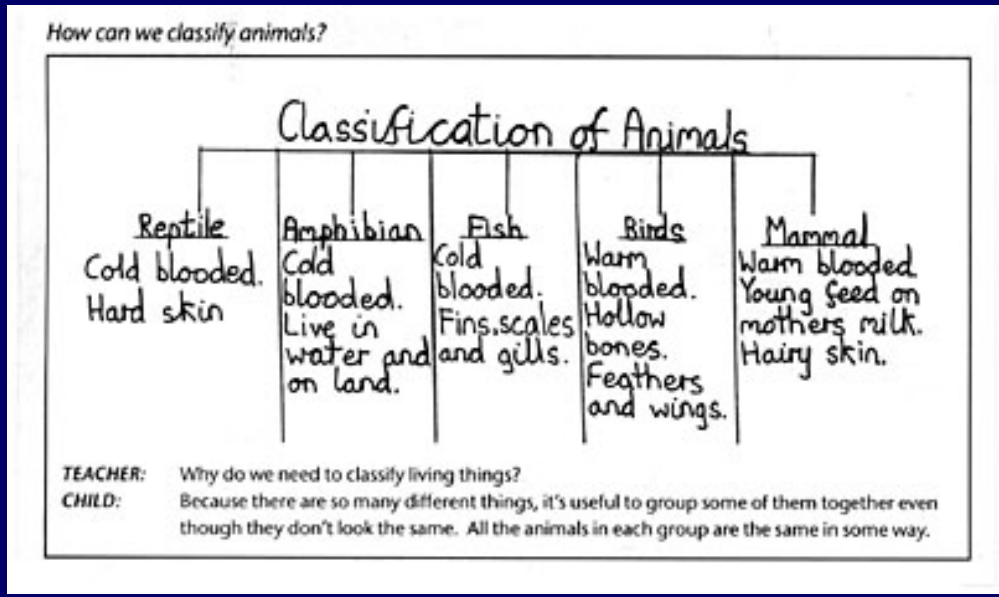
First:

- Why study automata theory?
- How to study automata theory?
- Standard mathematical ingredients that you should be familiar with like proof techniques and set theory

# Why Study Automata Theory?

- When dealing with computation it is an empirical important fact that certain problems are harder than others.
- Sometimes the hardness of a problem stems from the specific machine with which you try to solve it.
- Other times, the hardness is intrinsic to the problem itself.

*The crucial insight is that you can define large classes of problems and or machines that are natural and useful.*



# Classifying Problems

- In this course we will classify the different kinds of automata (= computing devices). By doing so we will automatically find a classification of problems as well.

*The complexity of a problem will be indicated by the kind of machine that is required to solve it.*

# CS138 Assumes that you are familiar with the following...

# Proof Techniques

- Throughout the course we will use several proof techniques to obtain our results. You should be familiar with (§1.2–4):
- Deductive proofs
- Proofs in set theory
- Proofs by contradiction
- Proofs by induction on integers
- Elementary graph theory
- Proofs by structural induction on strings, trees, et cetara.

# Standard Set Theory

- We will deal with a lot of sets and set theory in CS138. You should be familiar with concepts and notation like:

- Conditional: $A = \{ x \mid x \in \mathbb{N}, f(x)=0 \}$
- Union: $A \cup B$
- Intersection: $A \cap B$
- Complement: $A'$ or $A^c$ or $\bar{A}$.
- Cartesian Product: $A \times B$
- Power set: $\mathcal{P}(A)$

# Some Examples

$L_{<6} = \{ x \mid x \in \mathbb{N}, x<6 \}$
$L_{<6} \cap L_{prime} = \{2,3,5\}$

$\Sigma = \{0,1\}$
$\Sigma \times \Sigma = \{(0,0), (0,1), (1,0), (1,1)\}$

Formal notation:

- $A \cup B = \{ x \mid x \in A \text{ or } x \in B\}$
- $A \cap B = \{ x \mid x \in A \text{ and } x \in B\}$
- $A^c = \{ x \mid x \notin A\}$
- $A \times B = \{ (x,y) \mid x \in A \text{ and } y \in B\}$

# Power Sets

Formal: $\mathcal{P}(A) = \{ S \mid S \subseteq A \}$

Example: $A = \{x,y\}$
$\mathcal{P}(A) = \{ \{\} , \{x\} , \{y\} , \{x,y\} \}$

Note the different sizes:

$|\mathcal{P}(A)| = 2^{|A|}$

$|A \times A| = |A|^2$