So … you want to get a PhD … Now what?

Things you want to hear
Things you don't want to hear

Who am I you say?

- Doesn't really matter ☺
- What do I bring:
  - 5 (BS) + 2 (MS) + 3 (PhD) + 2.5 (post-doc) = 12.5 years in school … 25.5 if you count kinder/elementary/middle/high school
  - 9 years working for a major (the largest?) corporation in the world – ExxonMobil
  - 14 years as a faculty at RU (CBE and BME) and currently serving as the UGD in BME

Does research add any value to anything?
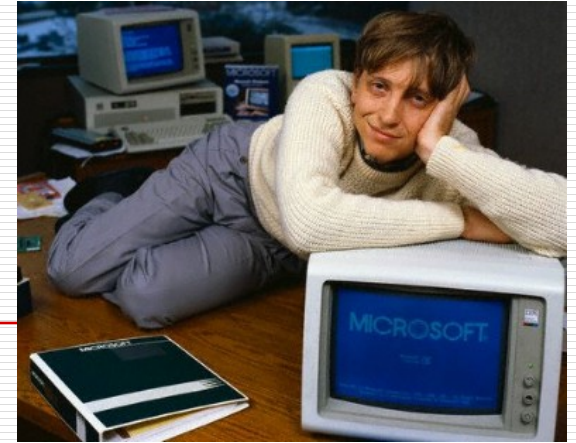
BOUNDS FOR SORTING BY PREFIX REVERSAL

William H. GATES
Microsoft, Albuquerque, New Mexico

Christos H. PAPADIMITRIOU*†
Department of Electrical Engineering, University of California, Berkeley, CA 94720, U.S.A.

For a permutation σ of the integers from 1 to n, let $f(\sigma)$ be the smallest number of prefix reversals that will transform σ to the identity permutation, and let $f(n)$ be the largest such $f(\sigma)$ for all σ in (the symmetric group) S_n. We show that $f(n) \leq (5n+5)/3$, and that $f(n) \geq 17n/16$ for n a multiple of 16. If, furthermore, each integer is required to participate in an even number of reversed prefixes, the corresponding function $g(n)$ is shown to obey $3n/2 - 1 \leq g(n) \leq 2n + 3$.

Does research add any value to anything?

The Anatomy of a Large-Scale Hypertextual Web Search Engine

Sergey Brin and Lawrence Page

Computer Science Department,
Stanford University, Stanford, CA 94305, USA
sergey@cs.stanford.edu and page@cs.stanford.edu

Abstract

In this paper, we present Google, a prototype of a large-scale search engine which makes heavy use of the structure present in hypertext. Google is designed to crawl and index the Web efficiently and produce much more satisfying search results than existing systems. The prototype with a full text and hyperlink database of at least 24 million pages is available at http://google.stanford.edu/ To engineer a search engine is a challenging task. Search engines index tens to hundreds of millions of web pages involving a comparable number of distinct terms. They answer tens of millions of queries every day. Despite the importance of large-scale search engines on the web, very little academic research has been done on them. Furthermore, due to rapid advance in technology and web proliferation, creating a web search engine today is very different from three years ago. This paper provides an in-depth description of our large-scale web search engine -- the first such detailed public description we know of to date. Apart from the problems of scaling traditional search techniques to data of this magnitude, there are new technical challenges involved with using the additional information present in hypertext to produce better search results. This paper addresses this question of how to build a practical large-scale system which can exploit the additional information present in hypertext. Also we look at the problem of how to effectively deal with uncontrolled hypertext collections where anyone can publish anything they want.

(12) **United States Patent**
Page

(54) **METHOD FOR NODE RANKING IN A LINKED DATABASE**

(75) Inventor: **Lawrence Page**, Stanford, CA (US)

(73) Assignee: **The Board of Trustees of the Leland Stanford Junior University**, Stanford, CA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **09/004,827**

(22) Filed: **Jan. 9, 1998**

Does research add any value to anything?

Larry (Lawrence) Page
Ph.D. Student
Computer Science Department
Stanford University

Member of Terry Winograd's Project on People, Computers, and Design.

Currently working on Google, a search engine for the Web. Papers and a demo are available off this page.

email: page@cs.stanford.edu
office: Gates 360
phone: (650) 330-0100
mail: Department of Computer Science
Gates Building Room #360
Stanford University
Stanford, CA 94305-214(

Sergey Brin's Home Page

Ph.D. student in Computer Science at Stanford - sergey@cs.stanford.edu

Research

Currently I am at Google.

In fall '98 I taught CS 349.

Data Mining

A major research interest is data mining and I run a meeting group here at Stanford. For more information take a look at the MIDAS home page or see the datamine maling list achive. Here are some recent publications:

- **Extracting Patterns and Relations from the World Wide Web**

  by Sergey Brin.
  We demonstrate a technique for extracting relations from the WWW based on the duality of patterns and relations. We experiment with it by extracting a relations of books. WebDB Workshop at EDBT '98 (postscript).

- **Dynamic Data Mining: A New Architecture for Data with High Dimensionality**

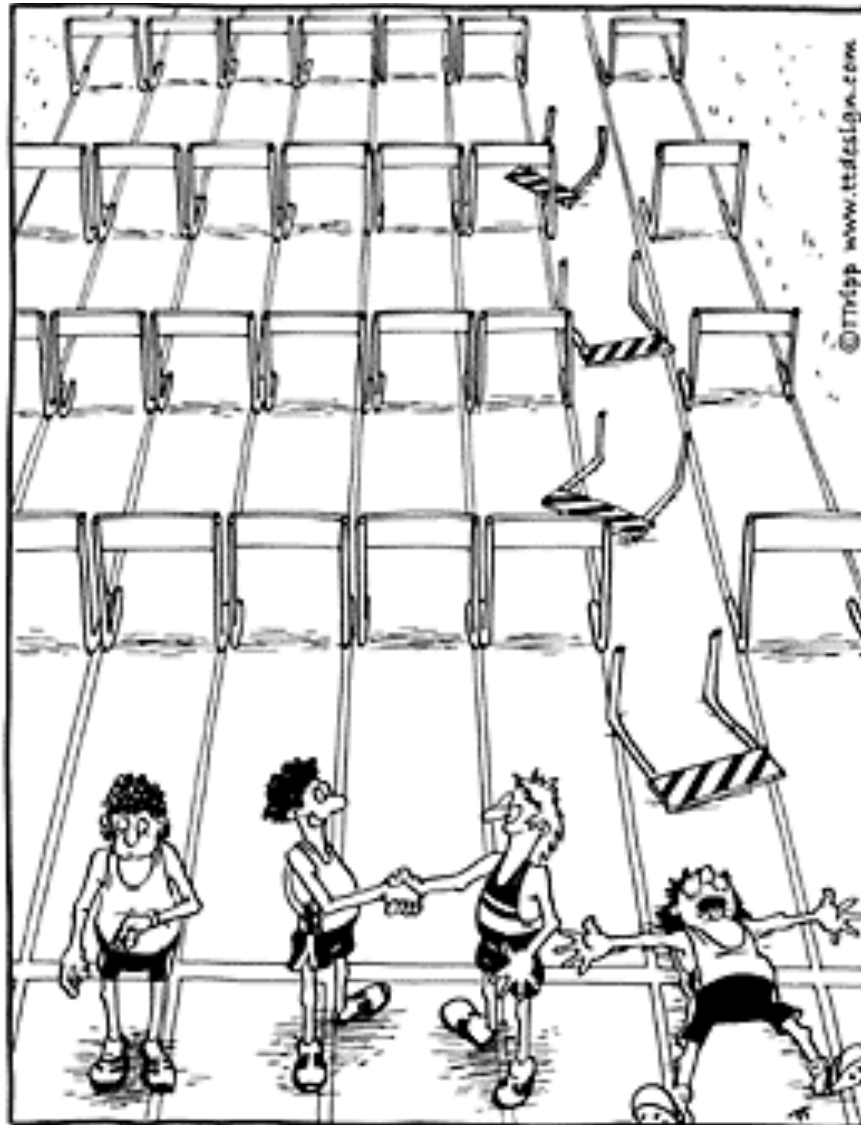# So … research is important because it teaches you how to think!

QUIZ …

- Question
  - What is the main difference between being an "undergraduate student" and "doing research"?

- Answer
  - It's all about learning how to ask your own questions and not look for answers to someone else's questions

The most serious hurdle to overcome:
Stop acting like a "student"

Exam vs Research

Exams

- You are told exactly what you are looking for

- You are given all the necessary data

- The answer exists, is unique, someone already knows it and you have to figure it out or verify it

Research

- You do not know what you are looking for

- You do not have all the data, if any

- The answer may not exist, may not be unique, no one knows whether it exists or how to find it

Interpret … don't just answer

- Develop your interpretive skills
- In research, unlike hwks/exams, we are not looking for "a" single answer which is the end of it
  - Usually it is hard to define an answer
- In research we are looking for understanding which will leads us to the next question

Data vs. Information vs. Knowledge:

- Data: Produce numbers, completely out of context: 100 and 5
- Information: Provide the basis for context: Bank savings account
  - $100 is the amount of money in the savings account, the principal
  - 5% is the factor used to calculate return on the principal every year, the interest rate
- Knowledge: Provide a pattern; how the numbers relate to each other within the context
  - If I put $100 in my savings account, and the bank pays 5% interest yearly, then at the end of one year the bank will compute the interest of $5 and add it to my principal and I will have $105 in the bank
  - Knowledge: What does this mean?
    - If I deposit more money in my account, I will earn more interest; if I withdraw money from my account, I will earn less interest
    - I will keep earning interest even if I do not deposit any more money
- Wisdom: Deduce overarching system principles: how does the pattern connect, interact, and evolve in the larger world
  - If the result of an action encourages more of the same action
- Leads you to the next questions! Manage your portfolio !

Qualities you have to develop

- Be responsible
- Be a team player
- Be proactive
- Be aggressive-minded
- Be prepared
- Be trustworthy
- Be persistent
- Be focused
- Be in charge
- Be aware of your field
- Be involved … You are researchers and not willing slaves

It's the process that is exciting!
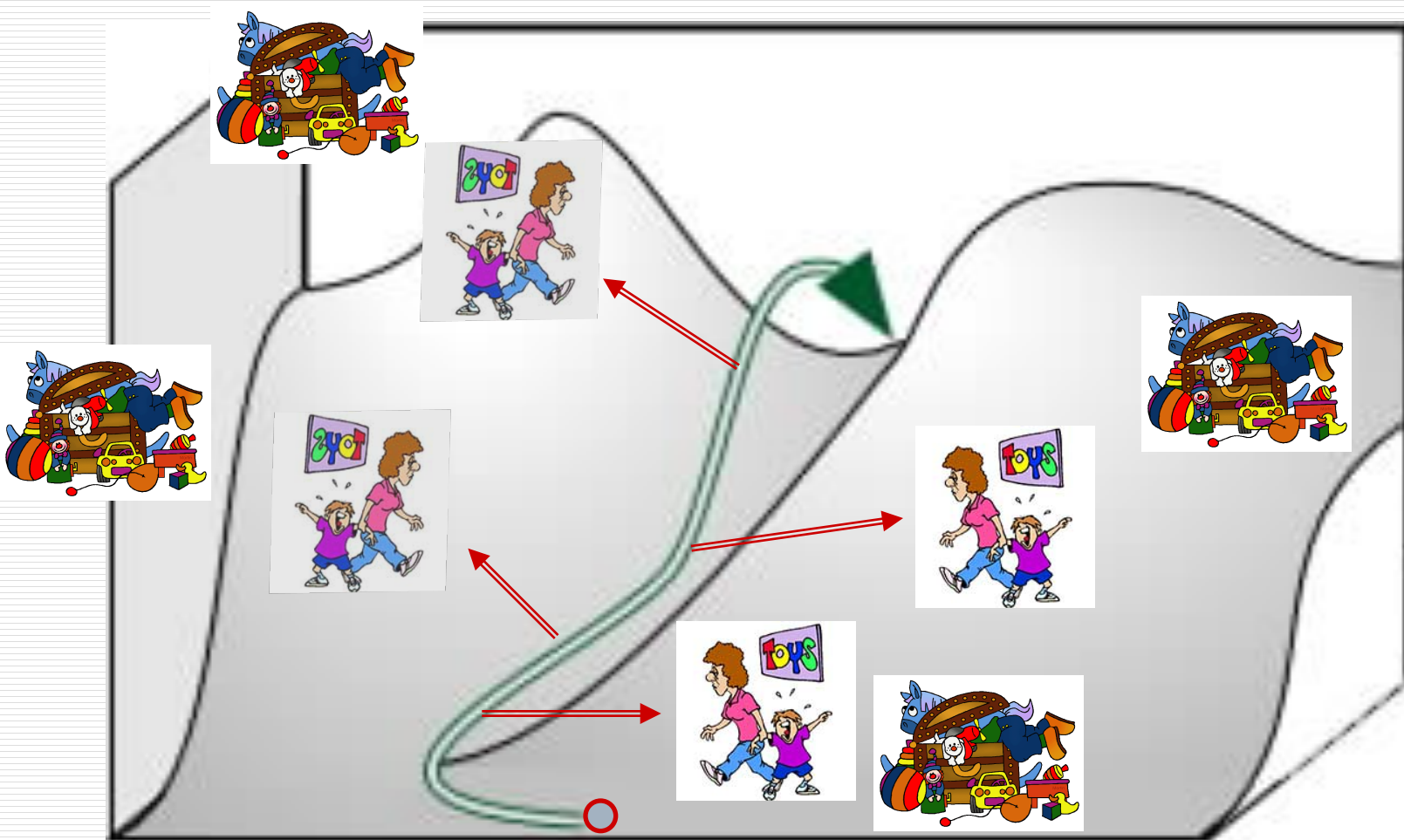
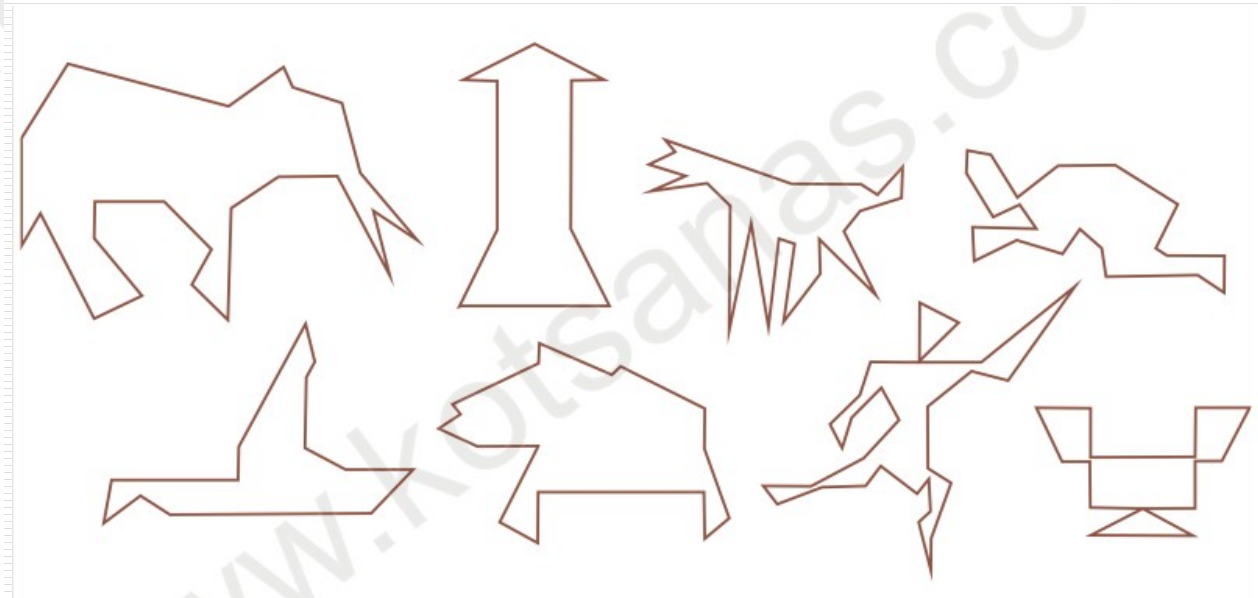# It's not about answering questions
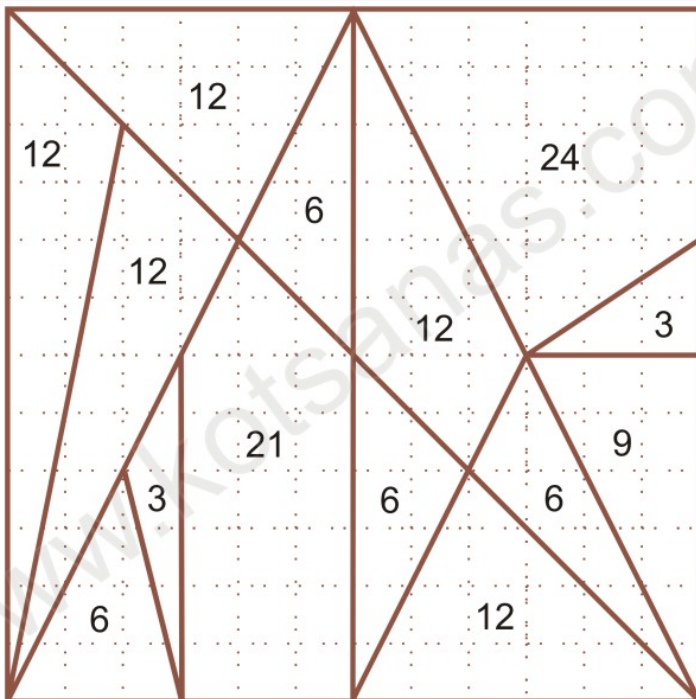# It's about asking questions

Project

Thesis
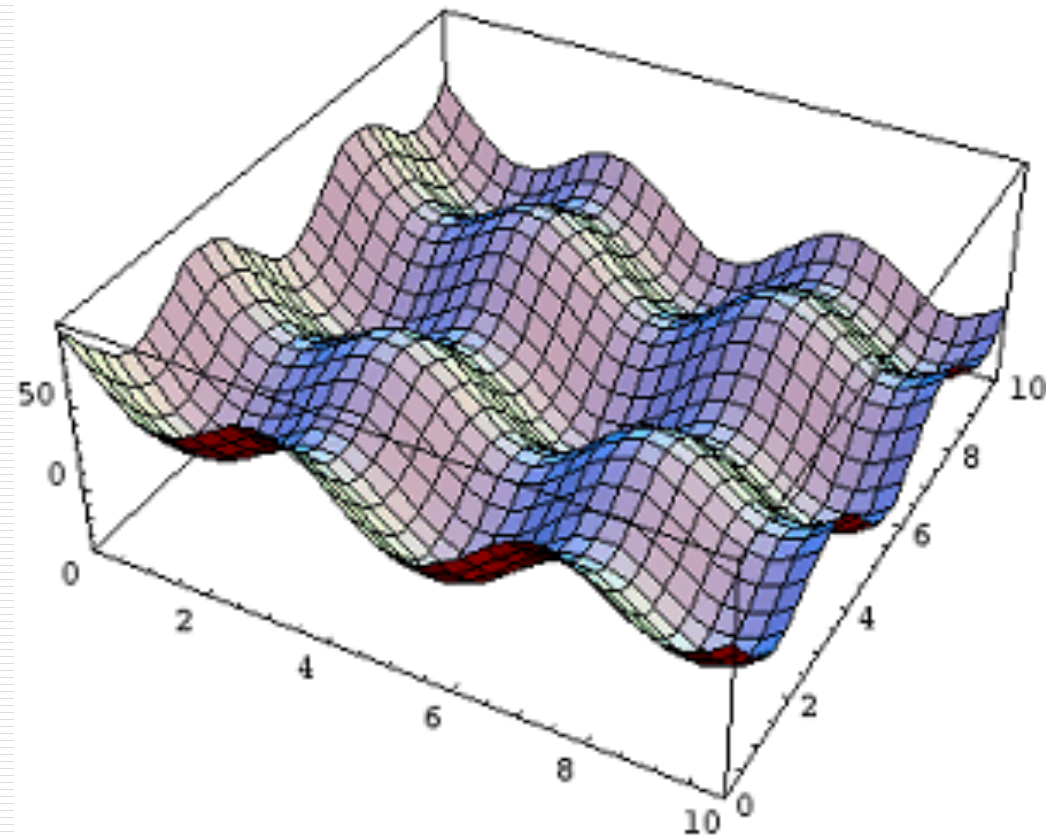
Imagine the big picture

- Research is like a jigsaw puzzle only you don't really know what you are trying to make and not sure where the pieces are

But the picture might change !

- The ingredients are the same but your target may evolve, so stay focused

Archimedes *Stomachion*

Learn how to ask questions

- How? WORK AND STUDY!!!
- Learn the area, be the "master of your domain"
- Follow the literature and the recent advances
  - There is nothing more embarrassing than re-inventing the wheel
- Identify the challenges and learn how to ask tough questions
- Your favorite question should be … so what?
- Realize that research is not like homework problems where the answer can be found in some book. You have to synthesize the approach
  - It is not magic! It's hard work!

You have to master the process and not just the project

- Projects evolve and often come and go
- What stays with you is your ability to analyze problems and synthesize answers
- You have to learn the process of learning
- There is no textbook to teach you that it is trial-and-error and hard work
- Dissect the problem and proceed to look at it piece by piece
- Learn to speak other peoples' language (and I do not mean Greek …)

The things I learned in graduate school

- Learn how to read, especially between the lines
- Learn how to ask questions
- Learn how to identify targets
- Learn how to analyze systems
- Learn how to synthesize systems
- Learn how to think
- These are the qualities future employers are looking for in future successful people
  - Otherwise they would be looking for lab technicians

Your work is an extension of yourself

- One of the most challenging tasks is to learn how to excite others and how to communicate with people from other fields
- No one has patience for nonsense. If you have something to say, say it! It has to be clear, consistent and meaningful
  - Otherwise, no one will ever pay attention to you. Why should they?
- If you don't' care about expressing yourself properly it is usually an indication that don't really care about what you do. So, why should I waste my time trying to understand what you are saying?

Your work is an extension of yourself

**Commodity
$0.60**

**Specialty
$2.00**

Standardized
No-Frills
No brand
Meets requirements
Competitive price
Low margins/profits
Push

Unique
Superior
Visibly Branded
Exceeds expectations
Premium price
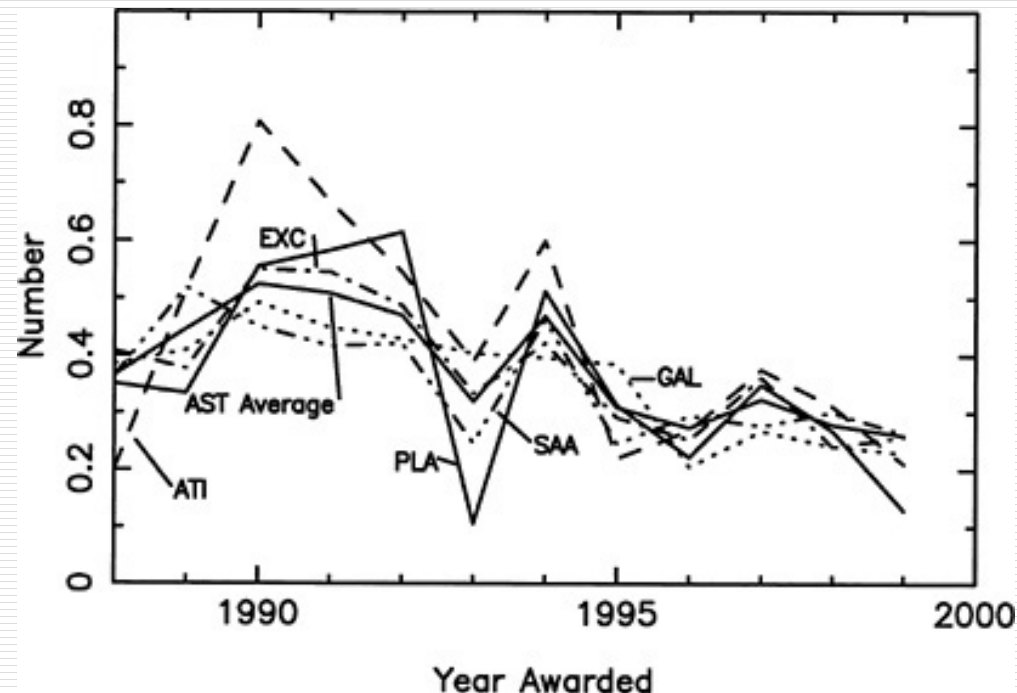High margins/profits
Pull

Take every task seriously

- When preparing a manuscript, assume your are working on the final version
- When preparing a talk, assume it's a job interview
- The way you present yourself and your work says much about who you are
- Deadlines are there for a reason
- Everybody has lucky breaks … You have to be ready when yours comes along
  - It's not about being lucky, it's about being well prepared

The world is self-similar

- We all have deadlines
- We all have bosses
- We all have constraints

- Realize that we all are part of a team and we all have responsibilities to each other

$$ is not everything, but ...

- No free lunch
- We all are resource-limited so learn to respect that
    - A PhD is not an all-expenses-paid vacation

Annual Cost of a student

Rutgers, The State University of New Jersey

Title:	who knows what
PI:	IPA
Spon:	Anyone

Facilities & Administrative Costs Rate & Base:

Description		*Enter Rate Code*
Sponsor Funding		**1**
	If rate code=9, enter rate percentages for Years 1-5:	0.00%
	If base code=3, enter amt. subject to F&A for Years 1-5:	$ -
	Base Amount	38,511.00
	If rate code=10, enter Custom FA:	-

Sub-code	Expense Classification (see comment)	**FY 2018**
	Cost Rate Base Percentages:	55.00%
	Enter Amounts in Whole $ only	**Year 1**
37100	Subcontract Expense over $25,000	0
83800	IRB or IACUC Fees	0
83700	**RU Gift Assessment**	0
	RU-Other (please identify)	0
H	**Total Other Non-MTDC**	15,000
I	**Total Direct Costs - TDC**	53,511
J	**Facilities & Administrative Costs**	21,181
	TOTAL PROJECT COSTS	74,692

Annual Cost of a student

Rutgers, The State University of New Jersey

Title:	who knows what
PI:	IPA
Spon:	Anyone

Facilities & Administrative Costs F

Description

Sponsor Funding

If rate code=9, enter rate perce

If base code=3, enter amt. sub

Base Amount

If rate code=10, enter Cus

Expense Classification (see

Cost Rate Base Percentage

Enter Amounts in Whole $ only

Sub-code		Year 1
37100		0
83800		0
83700		0
		0
		15,000
		53,511
J	Facilities & Administrative Costs	21,181
	TOTAL PROJECT COSTS	74,692

74,692

Rutgers, The State University of New Jersey

Title:	who knows what
PI:	IPA
Spon:	Anyone

Facilities & Administrative Costs Rate & Base: *Enter*

I	**Total Direct Costs - TDC**	53,511
J	**Facilities & Administrative Costs**	21,181
	TOTAL PROJECT COSTS	74,692

Annual Cost of a student

What a group meeting should look like

How much time should you spend in research

- Numbers in principle don't mean much, but there are limits …
  - *much in little … nothing in too little*

- If you work 40 hours a week get a job

- If you work 80 hours a week get a life

- 50-60 hours a week is what is realistic

Keep your curiosity alive!

- Research is a non-stop process
- It has to be consuming you and you have to be passionate
- I would strongly advise you to take advantage of the academic environment (courses, seminars, meetings) … don't overdo it …

- But … never loose your focus!

Perceived vs Actual usefulness of a student vs time
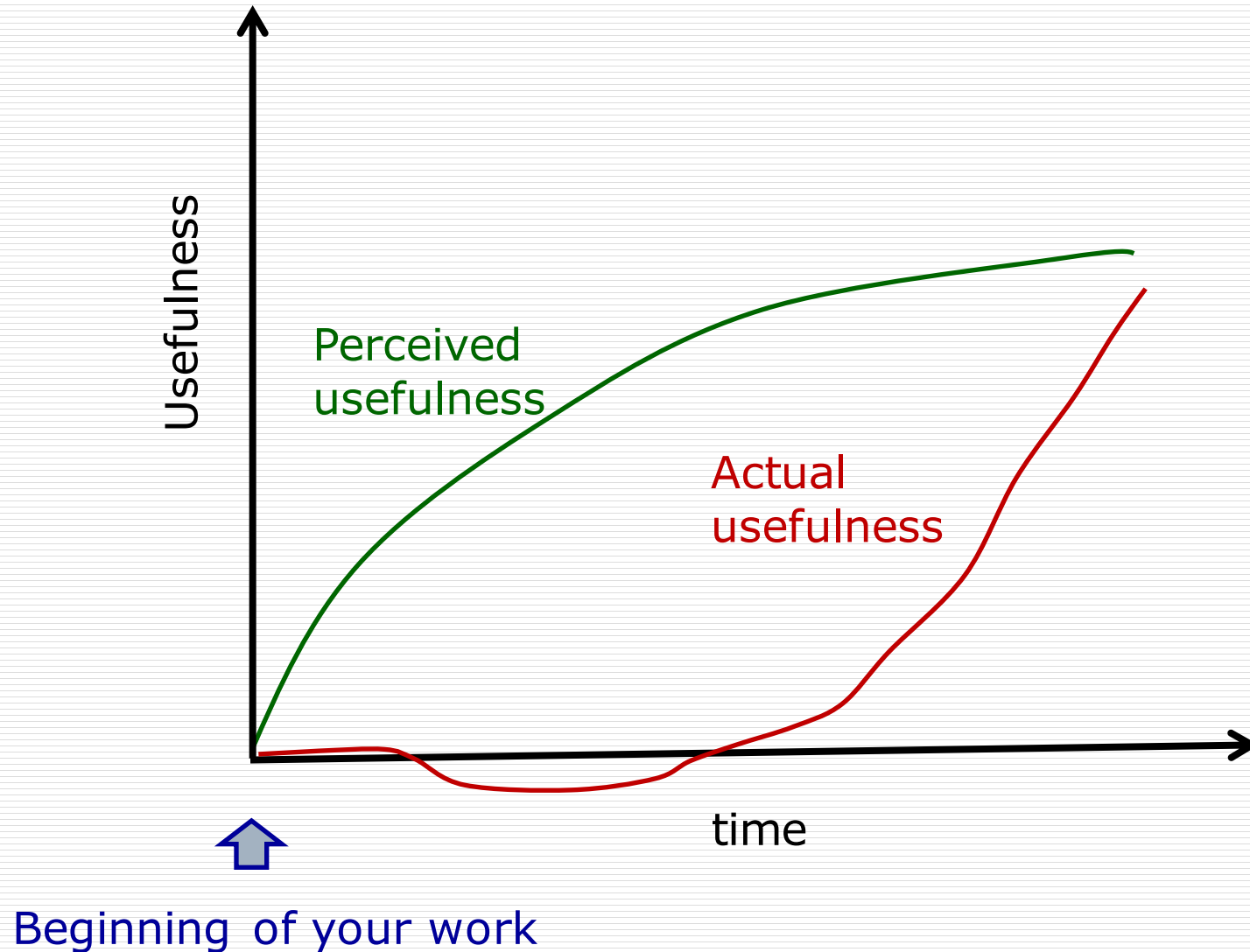
But ... don't rush!

- You have to know something before you can do anything

- Just because someone offered you a position in a lab and you can make cells light up, doesn't mean that you have accomplished, or learned, anything

- Take your time, do well in your courses and learn the fundamentals: Math, Phys, Chem, Biol

- The biggest mistake you can make is assume that you know what you are supposed to know! You don't, and that's why you are in college!
  - "all I know is I know nothing" (Socrates)
  - "You can't connect the dots looking forward; you can only connect them looking backwards" (Steve Jobs)

Finally …

I TAKE LIFE WITH A GRAIN OF SALT
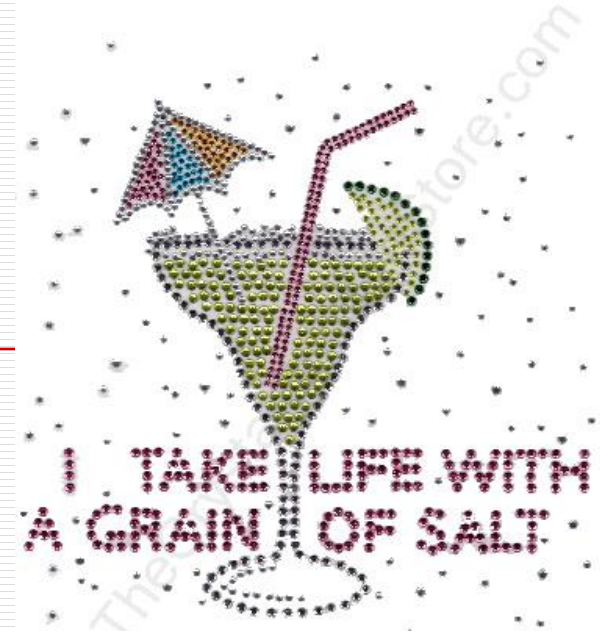
Lesson #1: Life is self-similar
… don't complain, work harder
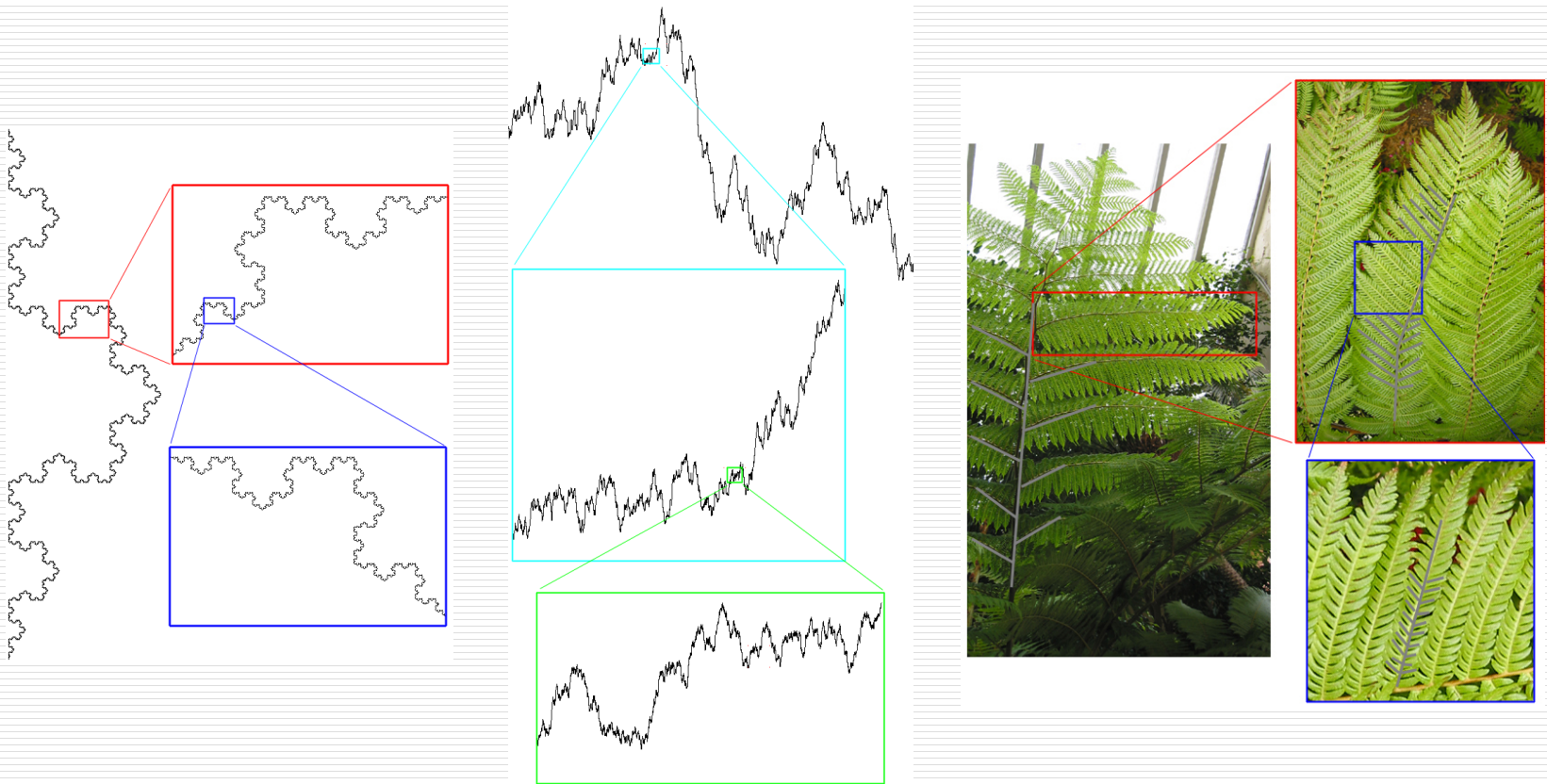
Lesson #2: Be proactive (proactive complainer)

- It's in your power to change things, but you have to be careful what can/cannot/should/should not change

- First, you have to **understand** the environment you are in, your skills, you advantages and disadvantages, your **limitations** and whose fault it is

- Our first reaction is to blame others because we do not like our job, our boss, our colleagues, Dr. Androulakis because he gave a lousy lecture etc

- You have to realize that sometimes it is your fault. The sooner you realize that, the faster you will readjust, the more you will succeed.

Lesson #3: Find a mentor and something you like

- Recognize what you would like to achieve and follow the footsteps of your role model
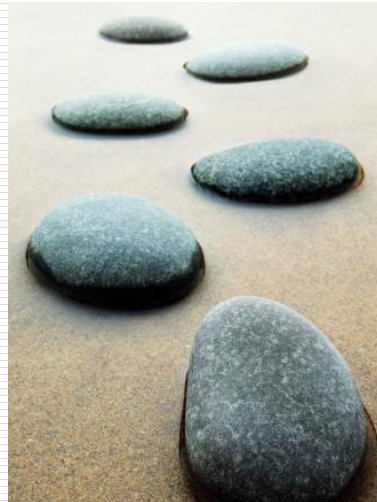
- Be careful!!!! Choose you role model wisely
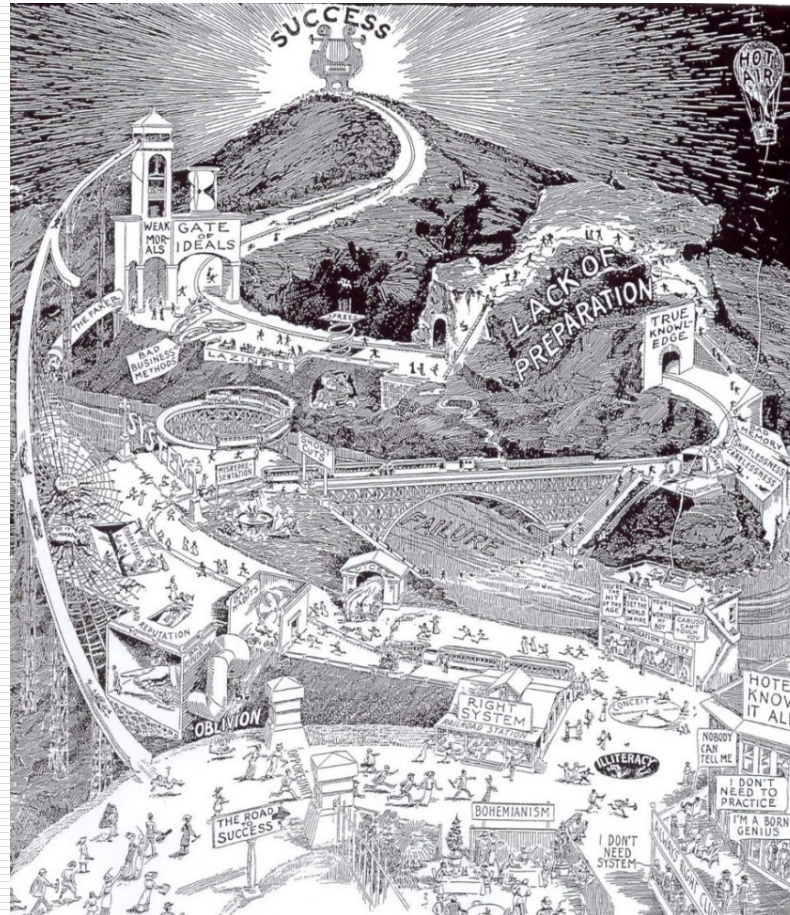- Define your goals and priorities clearly and work hard to make things happen

Lesson #4: Avoid the null space
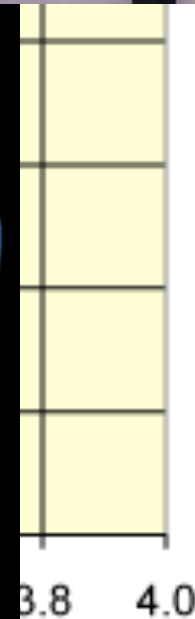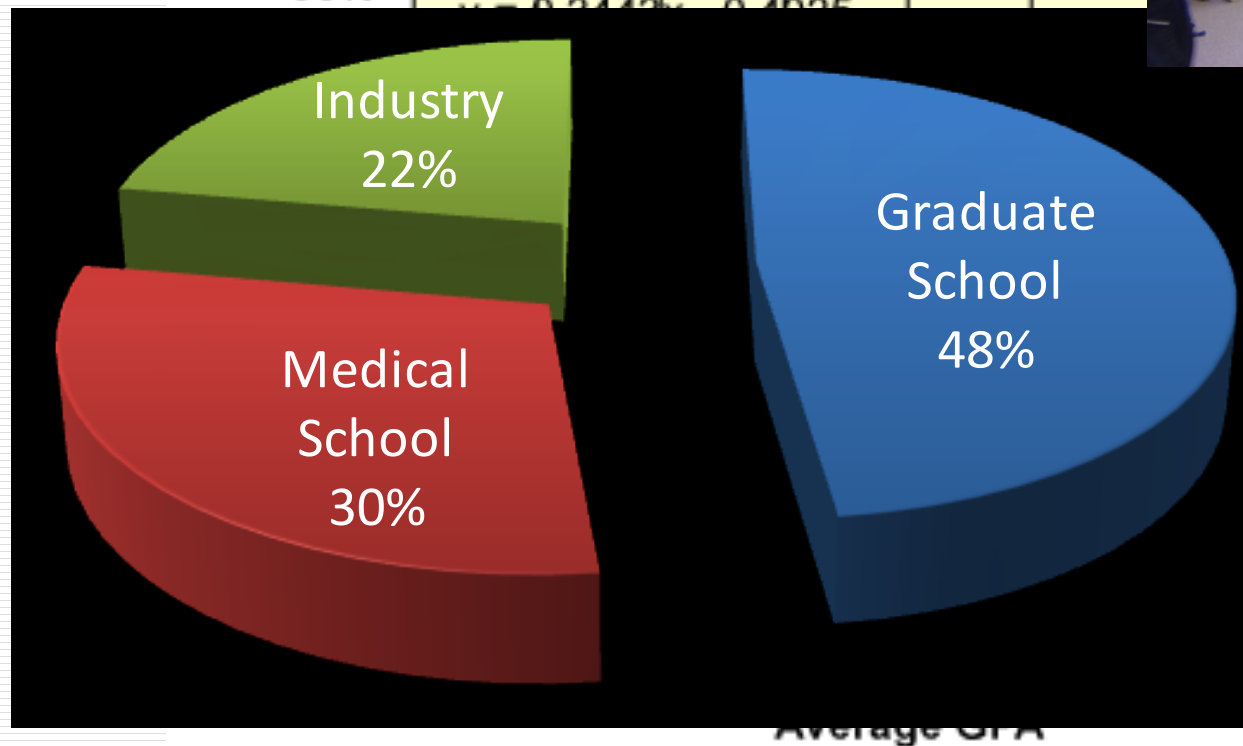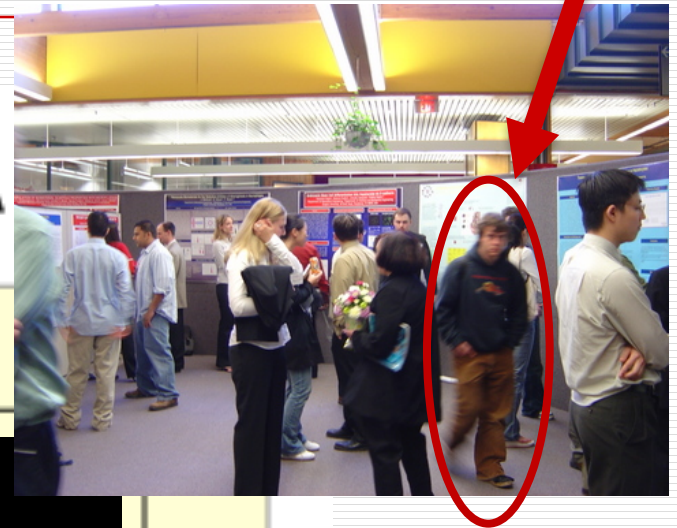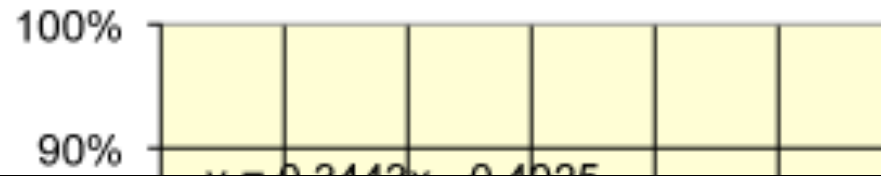my theory of life ☺
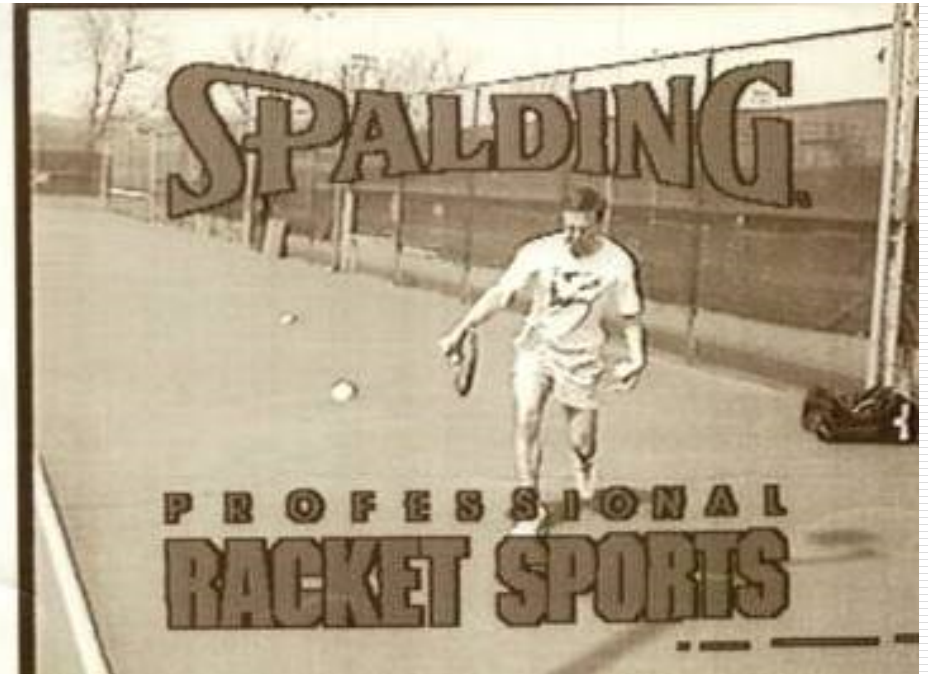
- Figure out S_2 (what you are really good at) and not S_1 (what you think you are good at)

Lesson #5: Do things that are fun, but realize that you have to work hard before you get wherever you are trying to go...
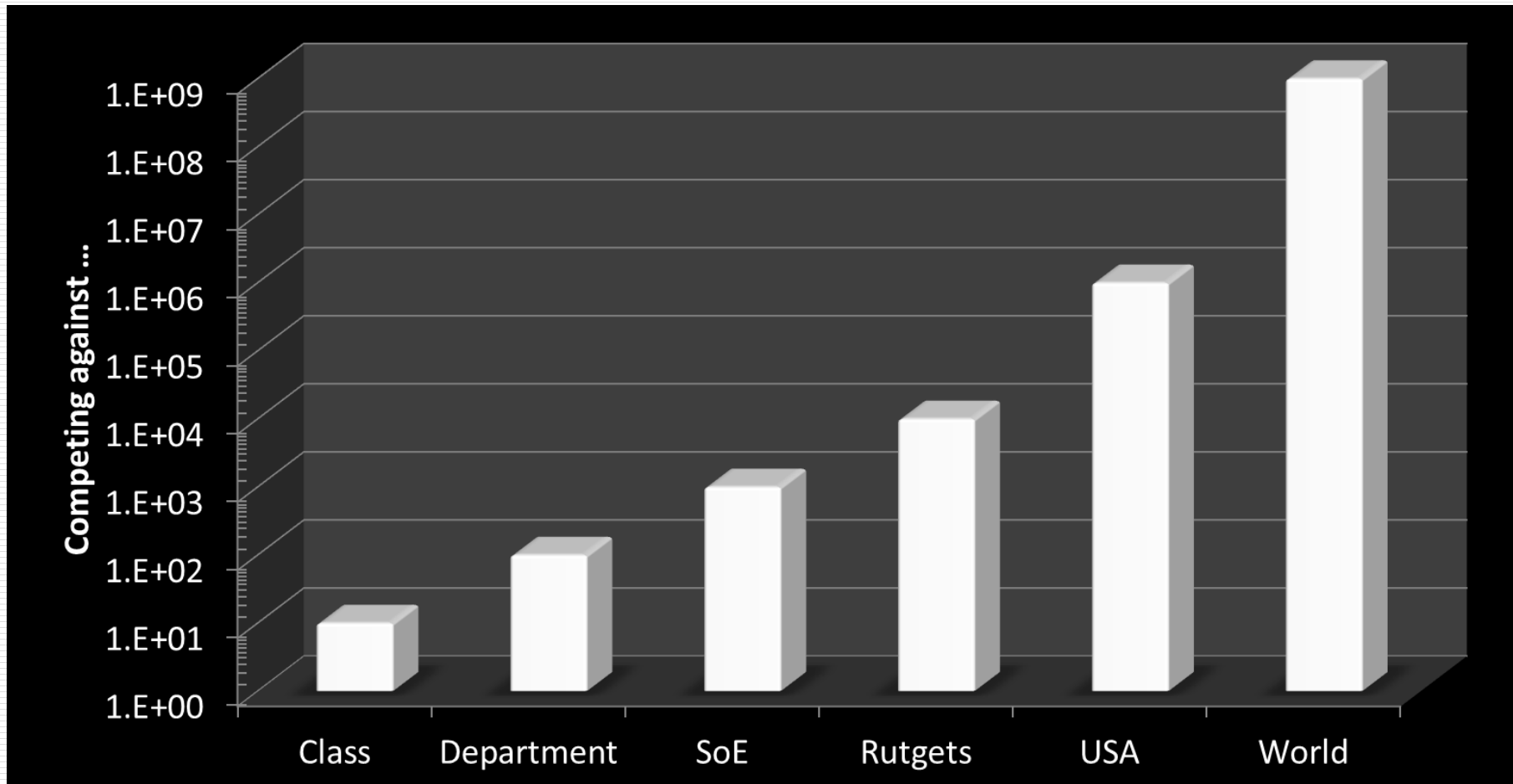
Lesson #6: Hard work pays off

Lesson #7: Never forget that you are as good as your most recent failure

You worked hard and you are at the top. A year from now, no one will remember unless you remain at the top!

Corollary for PhD advisors: they are only as good as their "worst" student ☺

Lesson #8: Keep things in perspective

Lesson #9: Be open minded

Nan-in, a Japanese master during the Meiji era (1868-1912), received a university professor who came to inquire about Zen.

Nan-in served tea. He poured his visitor's cup full, and then kept on pouring.

The professor watched the overflow until he no longer could restrain himself. "It is overfull. No more will go in!"

"Like this cup," Nan-in said, "you are full of your own opinions and speculations. How can I show you Zen unless you first empty your cup?"