The Impact of Technology in Future Employment & Education

Jaime Carbonell
Carnegie Mellon University
www.cs.cmu.edu/~jgc

Outline of Presentation

- **Trends in Technology**
  – big data, mobile, robotics, healthcare, services, …
- **A Whirlwind Visit to Macroeconomics**
  – Does Technology create or kill jobs?
- **Changes in the Employment Landscape**
  – Where will the jobs be?
- **Implications for Education**
  – What (e.g. STEM) and How (e.g. Learning by doing)

Trends in Technology

- Globalization (manufacturing, communication, …)
- Data Sciences (ubiquitous collection, analysis, …)
- Universal connectivity (internet of things)
- Mobility (smart phones, body media, …)
- Bio sciences (genes, proteins, microorganisms, …)
- Nano sciences (chemistry, drug delivery, …)
- 3D printing (mass customization, manufacturing, …)

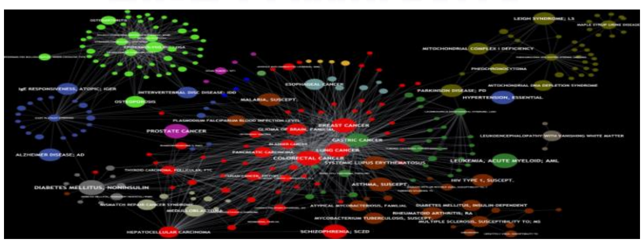
How Big is Big?

Dimensions of Big Data Analytics

LARGE-SCALE : TERABYTES → PETABYTES → EXOBYTES

Billions++ of entries:
Terabyes/Petabyes of data

HIGH-COMPLEXITY

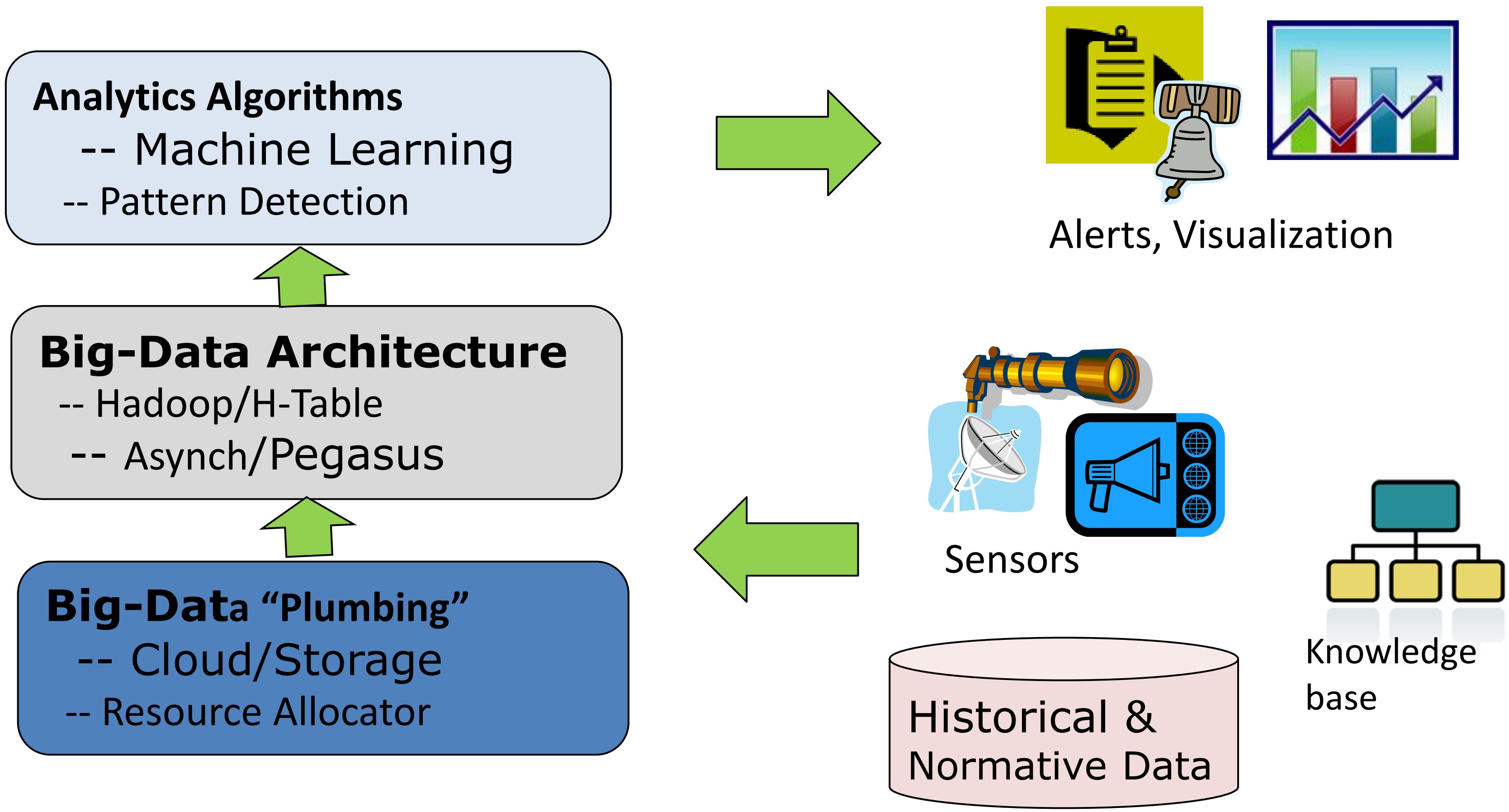


Trillions of potential relations among entries (graphs)

HIGH-DIMENSIONAL

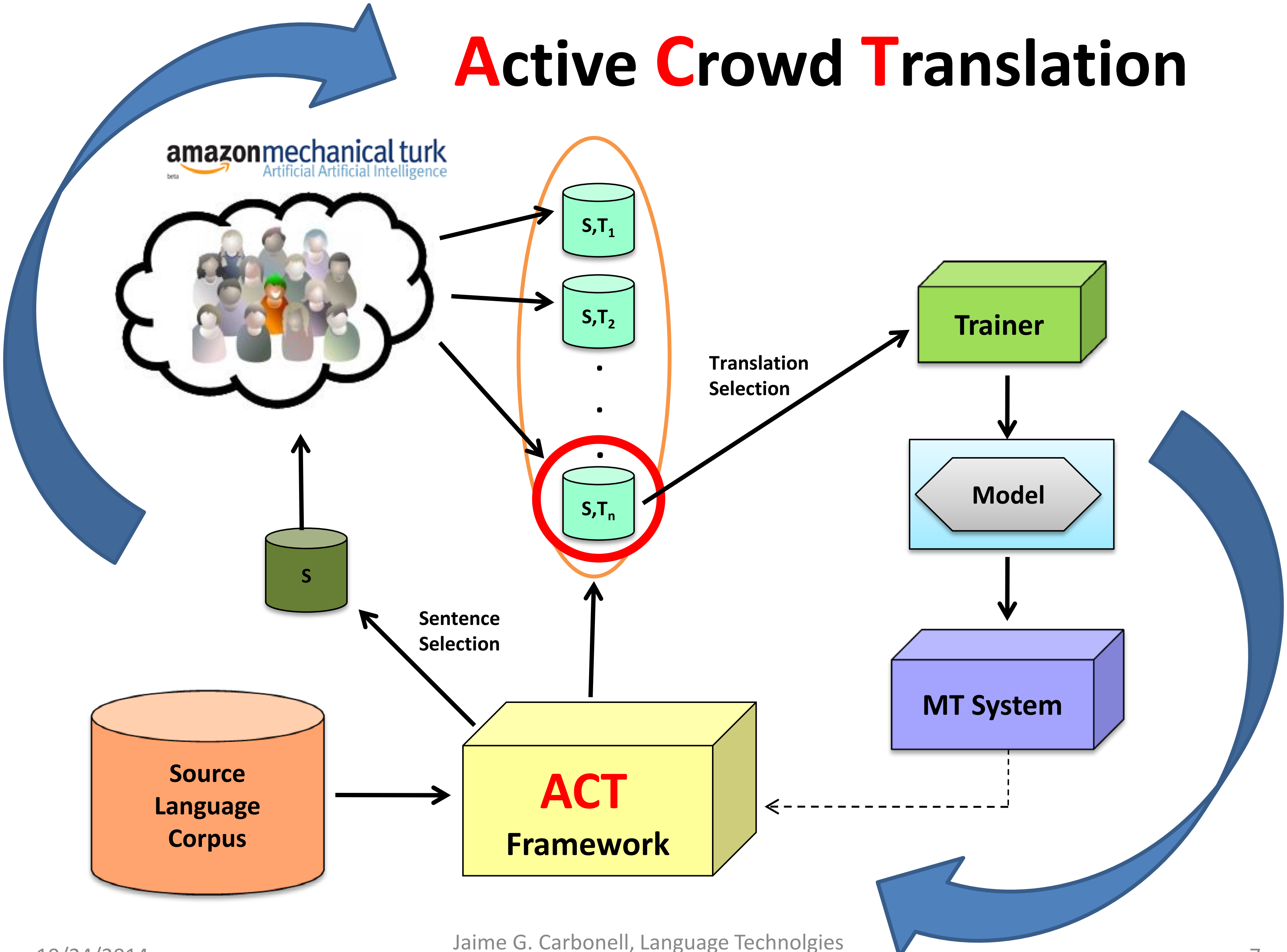
Millions of attributes per entry (but typically sparse encoding)

The Big-Data "Stack"




Roboceptionist

Active Crowd Translation



PROTEINS

(Borrowed from: Judith Klein-Seetharaman)

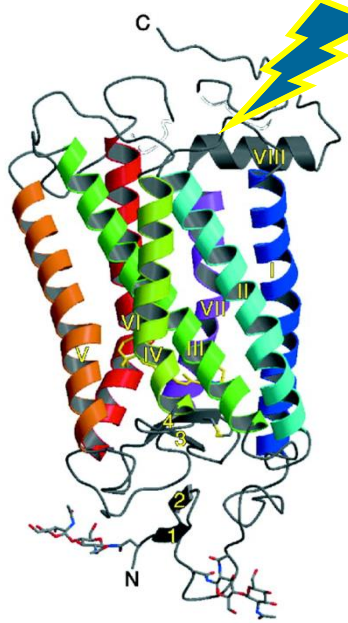
Sequence → Structure → Function

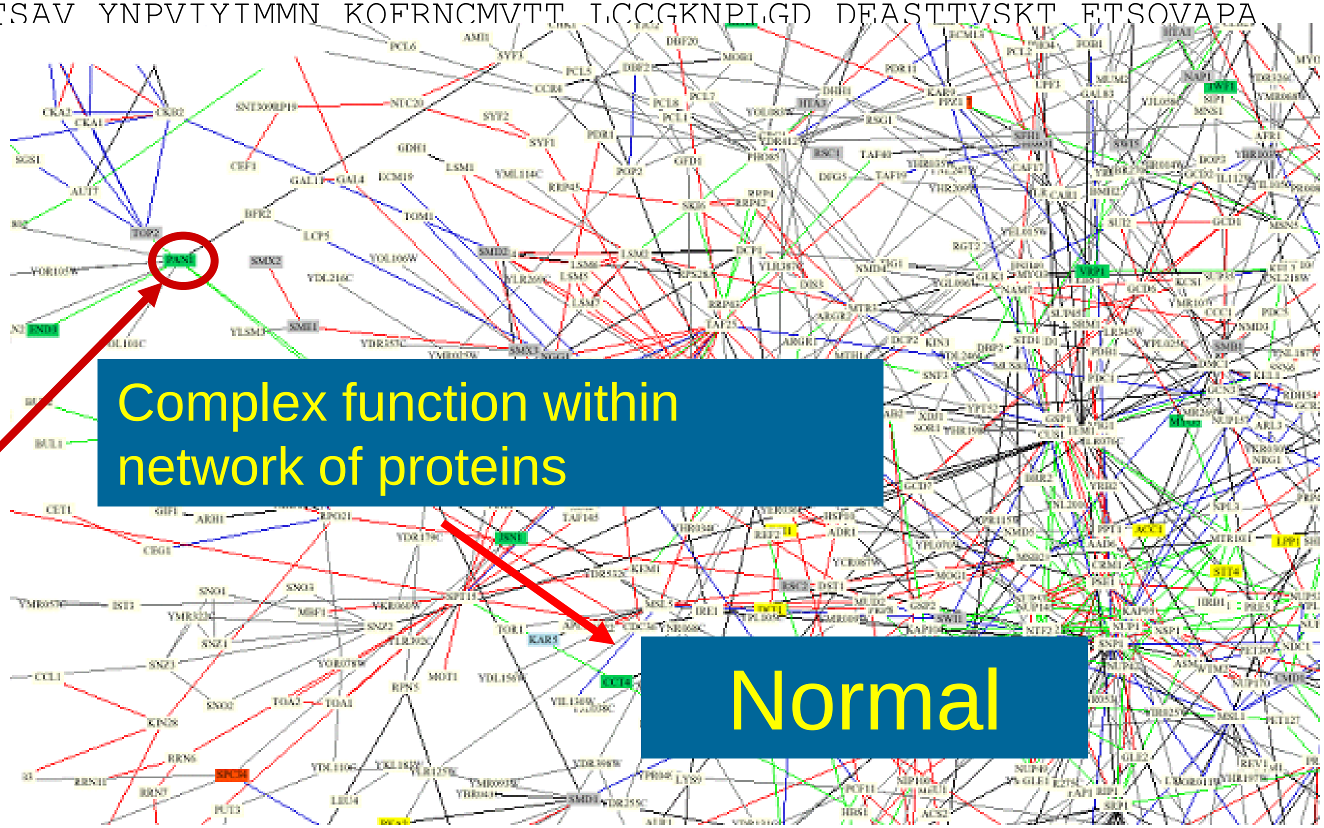
Primary Sequence

```
MNGTEGPNFY VPFSNKTGVV RSPFEAPQYY LAEPWQFSML AAYMFLLIML GFPINFLTLY VTVQHKKLRT
PLNYILLNLA VADLFMVFGG FTTTLYTSLH GYFVFGPTGC NLEGFFATLG GEIALWSLVV LAIERYVVVC
KPMSNFRFGE NHAIMGVAFT WVMALACAAP PLVGWSRYIP EGMQCSCGID YYTPHEETNN ESFVIYMFVV
HFIIPLIVIF FCYGQLVFTV KEAAAQQQES ATTQKAEKEV TRMVIIMVIA FLICWLPYAG VAFYIFTHQG
SDFGPIFMTI PAFFAKTSAV YNPVIYIMMN KQFRNCMVTT LCCGKNPLGD DEASTTVSKT ETSQVAPA
```

Folding

3D Structure





Complex function within network of proteins

Normal

PROTEINS

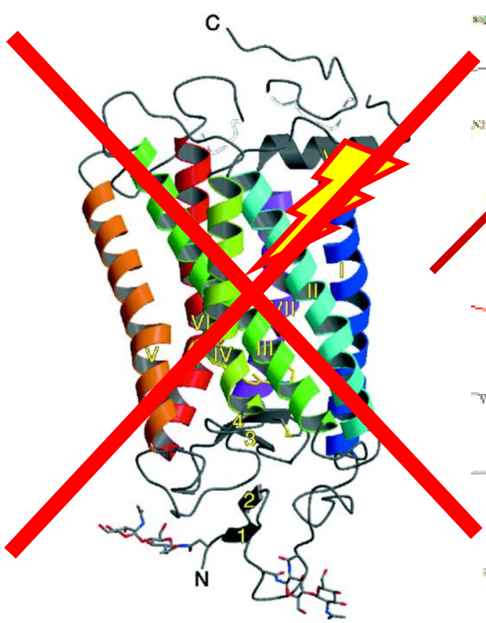
Sequence → Structure → Function

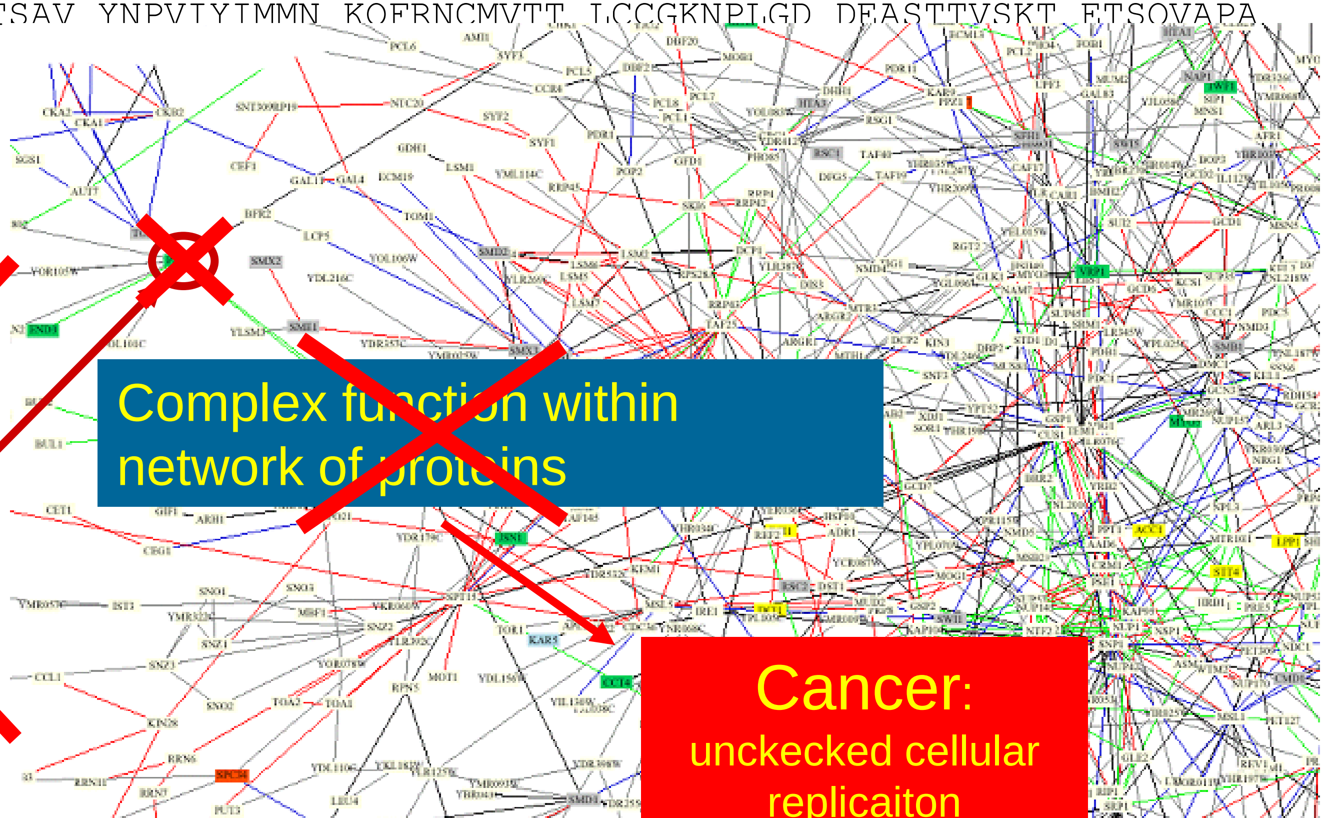
Primary Sequence

```
MNGTEGPNFY VPFSNKTGVV RSPFEAPQYY LAEPWQFSML AAYMFLLIML GFPINFLTLY VTVQHKKLRT
PLNYILLNLA VADLFMVFGG FTTTLYTSLH GYFVFGPTGC NLEGFFATLG GEIALWSLVV LAIERYVVVC
KPMSNFRFGE NHAIMGVAFT WVMALACAAP PLVGWSRYIP EGMQCSCGID YYTPHEETNN ESFVIYMFVV
HFIIPLIVIF FCYGQLVFTV KEAAAQQQES ATTQKAEKEV TRMVIIMVIA FLICWLPYAG VAFYIFTHQG
SDFGPIFMTI PAFFAKTSAV YNPVIYIMMN KQFRNCMVTT LCCGKNPLGD DEASTTVSKT ETSQVAPA
```

↓ Folding

3D Structure



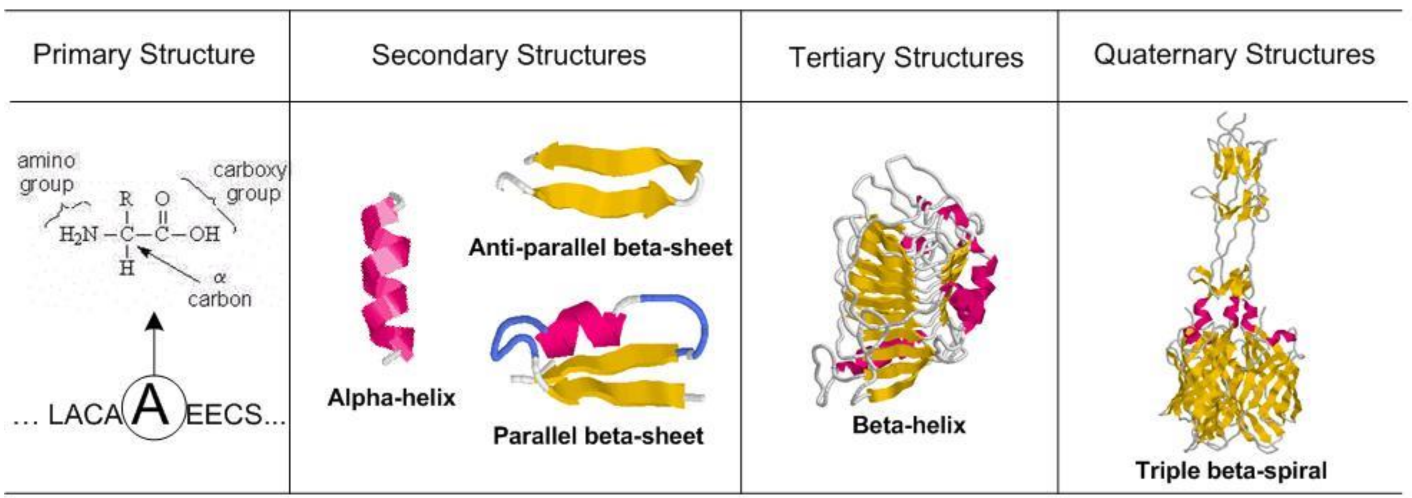


Complex function within network of proteins

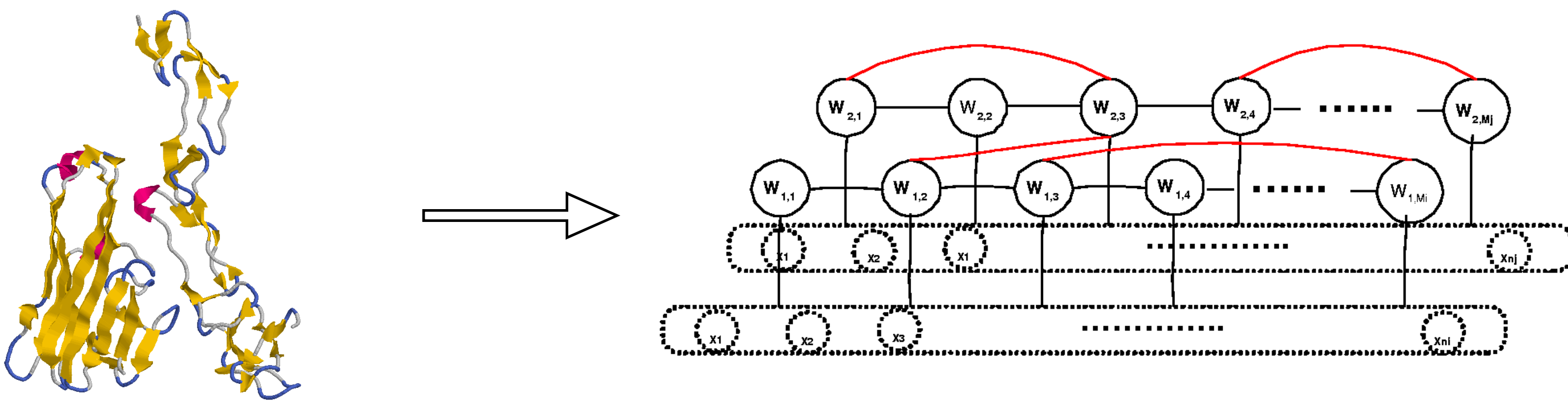
Cancer: unckecked cellular replicaiton

Predicting Protein Structures

- Protein Structure is a key determinant of protein function
- Crystalography to resolve protein structures experimentally in-vitro is very expensive, NMR can only resolve very-small proteins
- The gap between the known protein sequences and structures:
  - 3,023,461 sequences v.s. 36,247 resolved structures (1.2%)
  - Therefore we need to predict structures in-silico



Linked Segmentation CRF

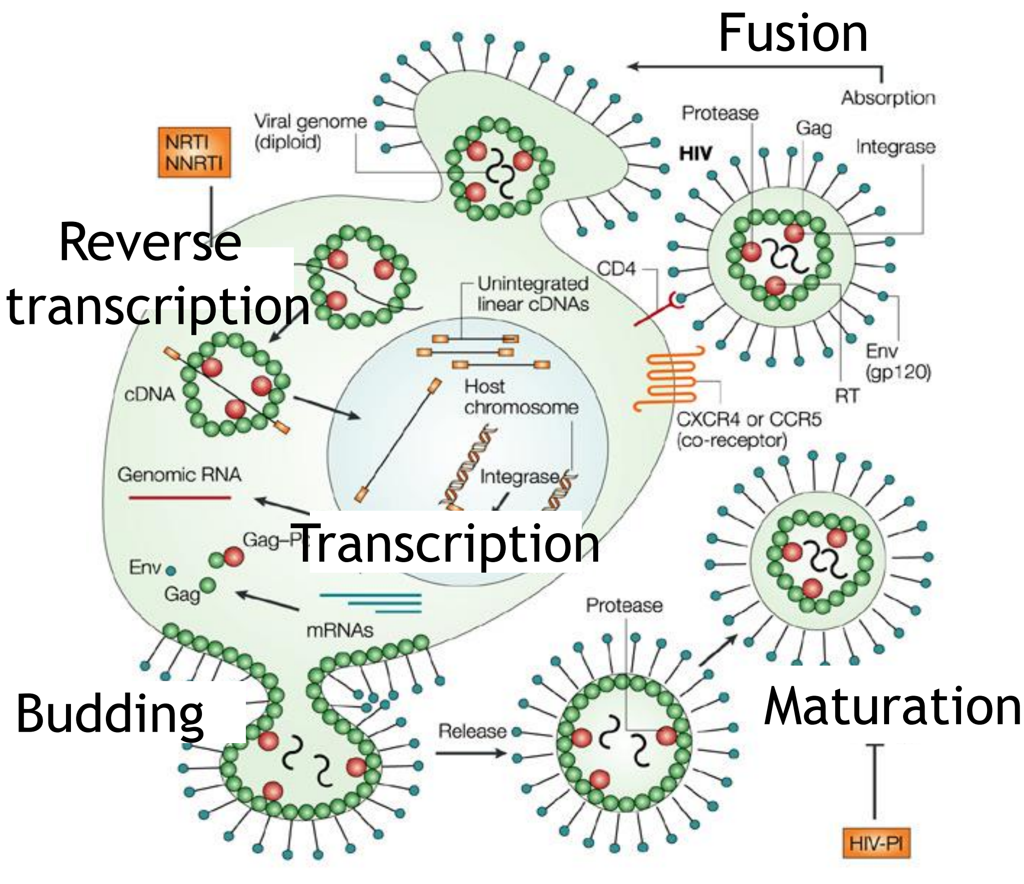


- Node: secondary structure elements and/or simple fold
- Edges: Local interactions and long-range *inter-chain* and *intra-chain* interactions
- L-SCRF: conditional probability of y given x is defined as

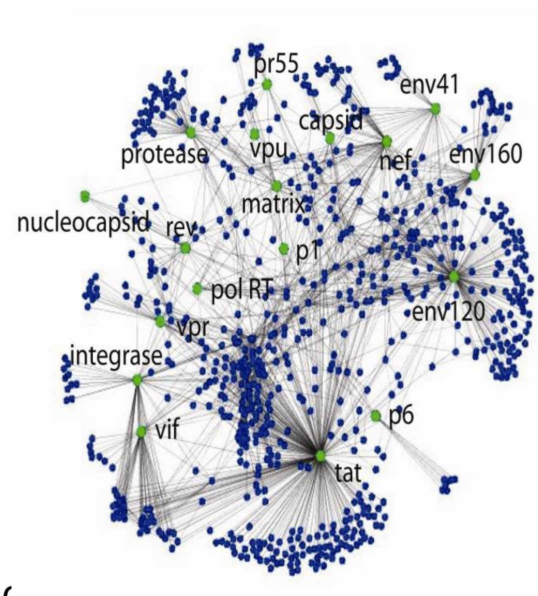
$$P(\mathbf{y}_1,\ldots,\mathbf{y}_R \mid \mathbf{x}_1,\ldots,\mathbf{x}_R) = \frac{1}{Z} \prod_{\mathbf{y}_{i,j} \in V_G} \exp\left(\sum_k \lambda_k f_k(\mathbf{x}_i, \mathbf{y}_{i,j})\right) \prod_{\langle \mathbf{y}_{i,j}, \mathbf{y}_{a,b} \rangle \in E_G} \exp\left(\sum_l \mu_l g_k(\mathbf{x}_i, \mathbf{x}_a, \mathbf{y}_{i,j}, y_{a,b})\right)$$

Joint Labels

HIV-1 + Human protein interactions



HIV-1 depends on the cellular machinery in every aspect of its life cycle.



Peterlin and Torono, *Nature Rev Immu* 2003.

Machine Learning

Data:

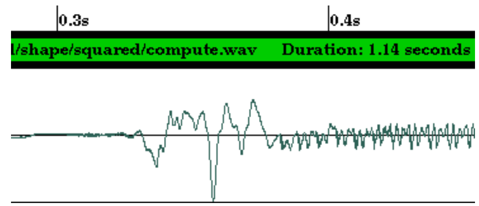
Patient103 time=1 →	Patient103 time=2 ... →	Patient103 time=n
Age: 23	Age: 23	Age: 23
FirstPregnancy: no	FirstPregnancy: no	FirstPregnancy: no
Anemia: no	Anemia: no	Anemia: no
Diabetes: no	Diabetes: YES	Diabetes: no
PreviousPrematureBirth: no	PreviousPrematureBirth: no	PreviousPrematureBirth: no
Ultrasound: ?	Ultrasound: abnormal	Ultrasound: ?
Elective C–Section: ?	Elective C–Section: no	Elective C–Section: no
Emergency C–Section: ?	Emergency C–Section: ?	**Emergency C–Section: Yes**
...	...	...

One of 18 learned rules:

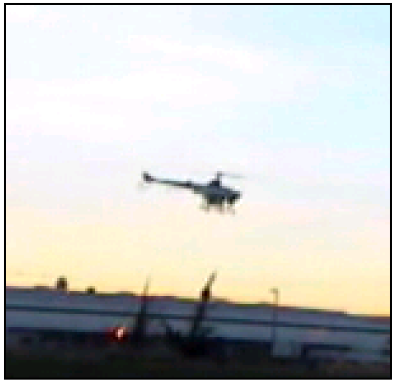
```
If   No previous vaginal delivery, and
     Abnormal 2nd Trimester Ultrasound, and
     Malpresentation at admission
Then Probability of Emergency C-Section is 0.6

 Over training data: 26/41 = .63,
 Over test data: 12/20 = .60
```

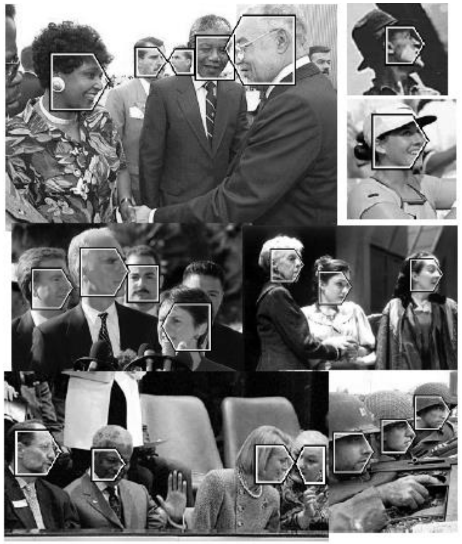
Data Mining



Speech Recognition



Automated Control learning



Object recognition

Extracting facts from text

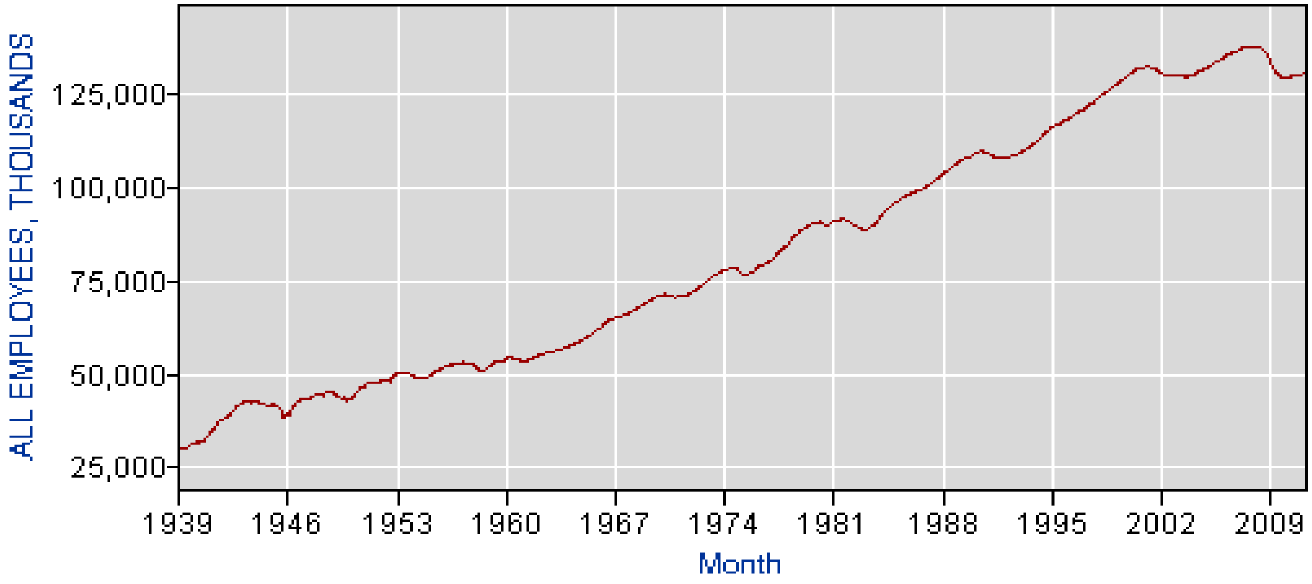
Peter H. van Oppen , Chairman of the Board & Chief Executive Officer
Mr. van Oppen has served as chairman of the board and chief executive officer of ADIC since its acquisition by Interpoint in 1994 and a director of ADIC since 1986. Until its acquisition by Crane Co. in October 1996, Mr. van Oppen served as chairman of the board of directors, president and chief executive officer of Interpoint . Prior to 1985, Mr. van Oppen worked as a consulting manager at Price Waterhouse LLP and at Bain & Company in Boston and London. He has additional experience in medical electronics and venture capital. Mr. van Oppen also serves as a director of Seattle FilmWorks Inc. and Spacelabs Medical, Inc.. He holds a B.A. from Whitman College and an M.B.A. from Harvard Business School, where he was a Baker Scholar.

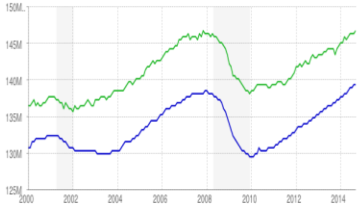
- Reinforcement learning
- Predictive modeling
- Pattern discovery
- Hidden Markov models
- Convex optimization
- Explanation-based learning
- ....

Macroeconomics of Disruptive Technologies

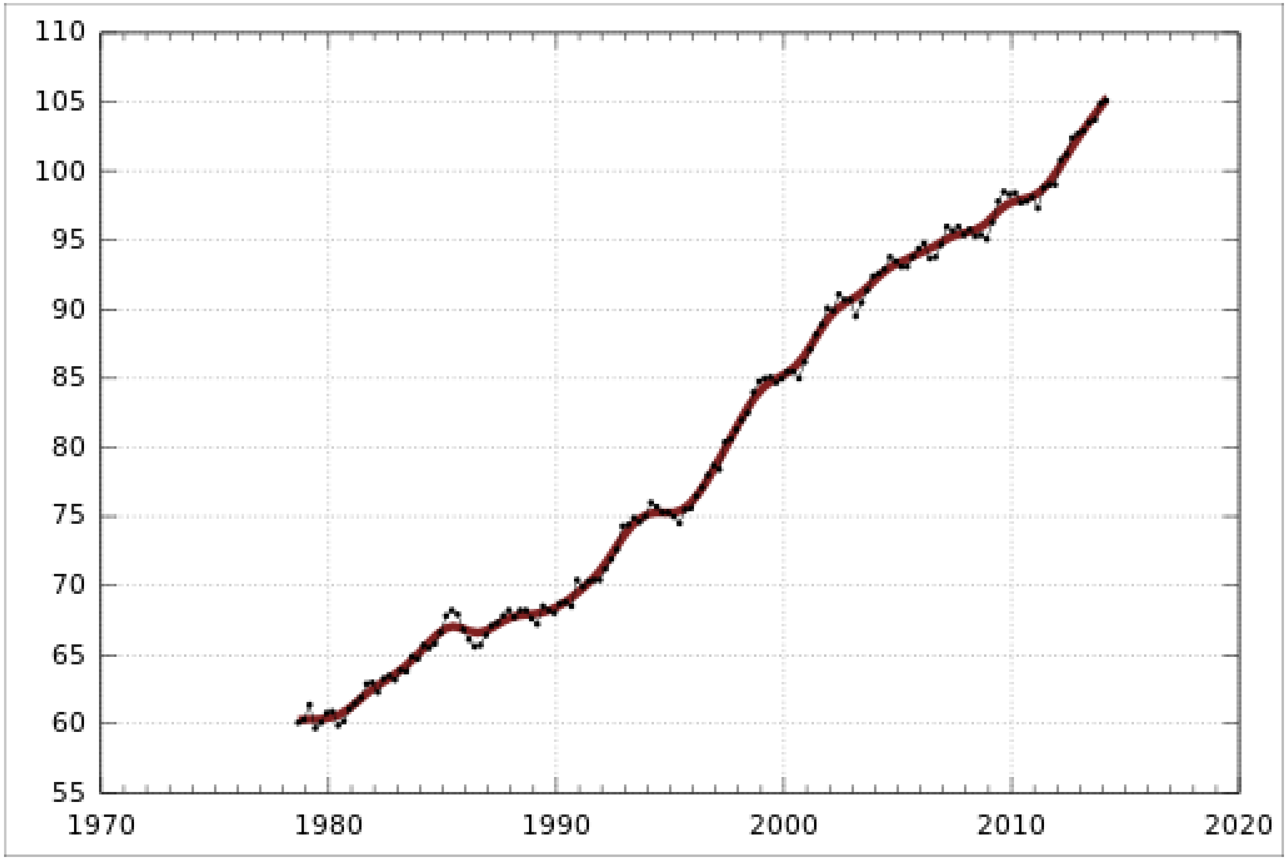
- Creates new industries (e.g. internet)
- Transforms others (e.g. x-rays, MRI, CAT)
- Eliminates some (e.g. robotic manufacturing)
- Contrary to (some) popular belief
  – Historically, new technology → more jobs
  – But different jobs, requiring more education
- Increased efficiency → more productivity

Total US Employment Since 1939 increased 500%

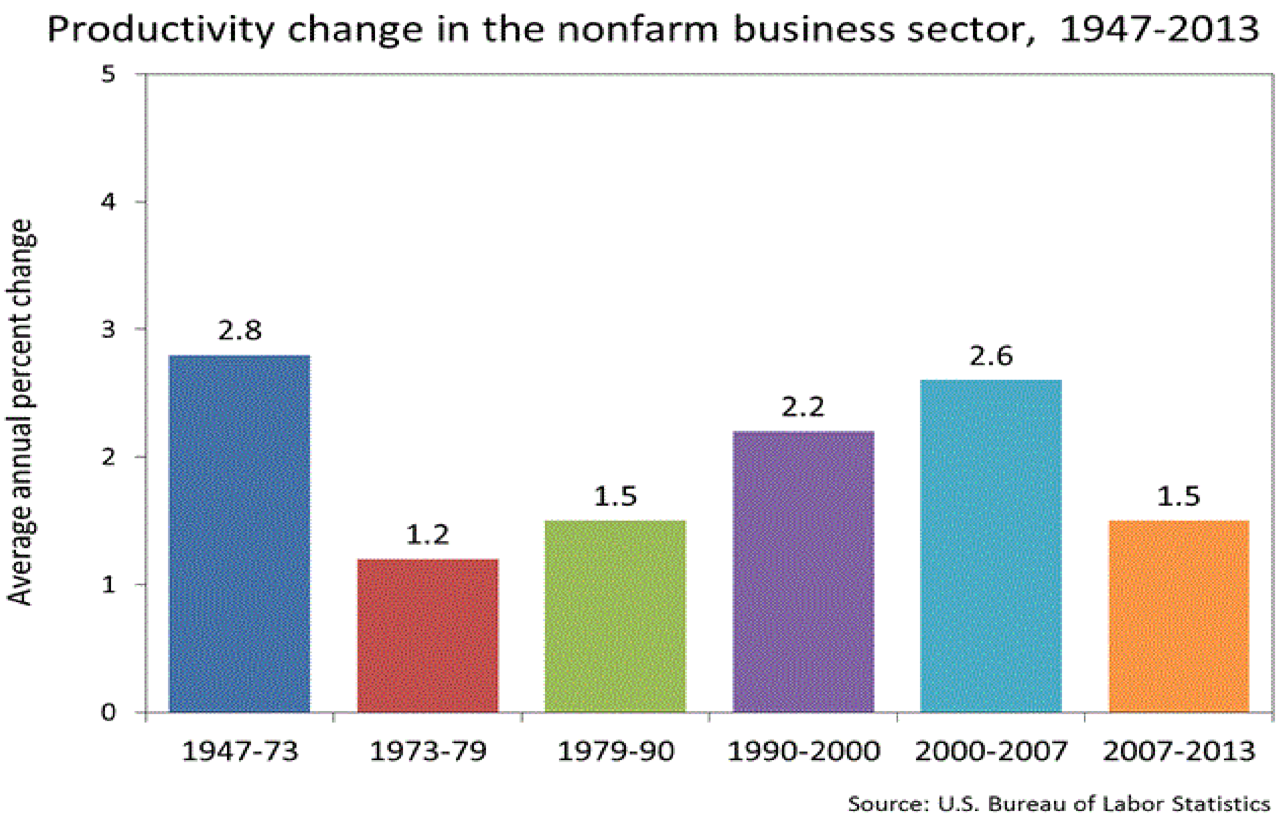




Australia: GDP per worker (measure of productivity)



US Productivity per worker Increase since 1947 > 400%



$$P' = Pe^{rt}$$

r=.023, t=66

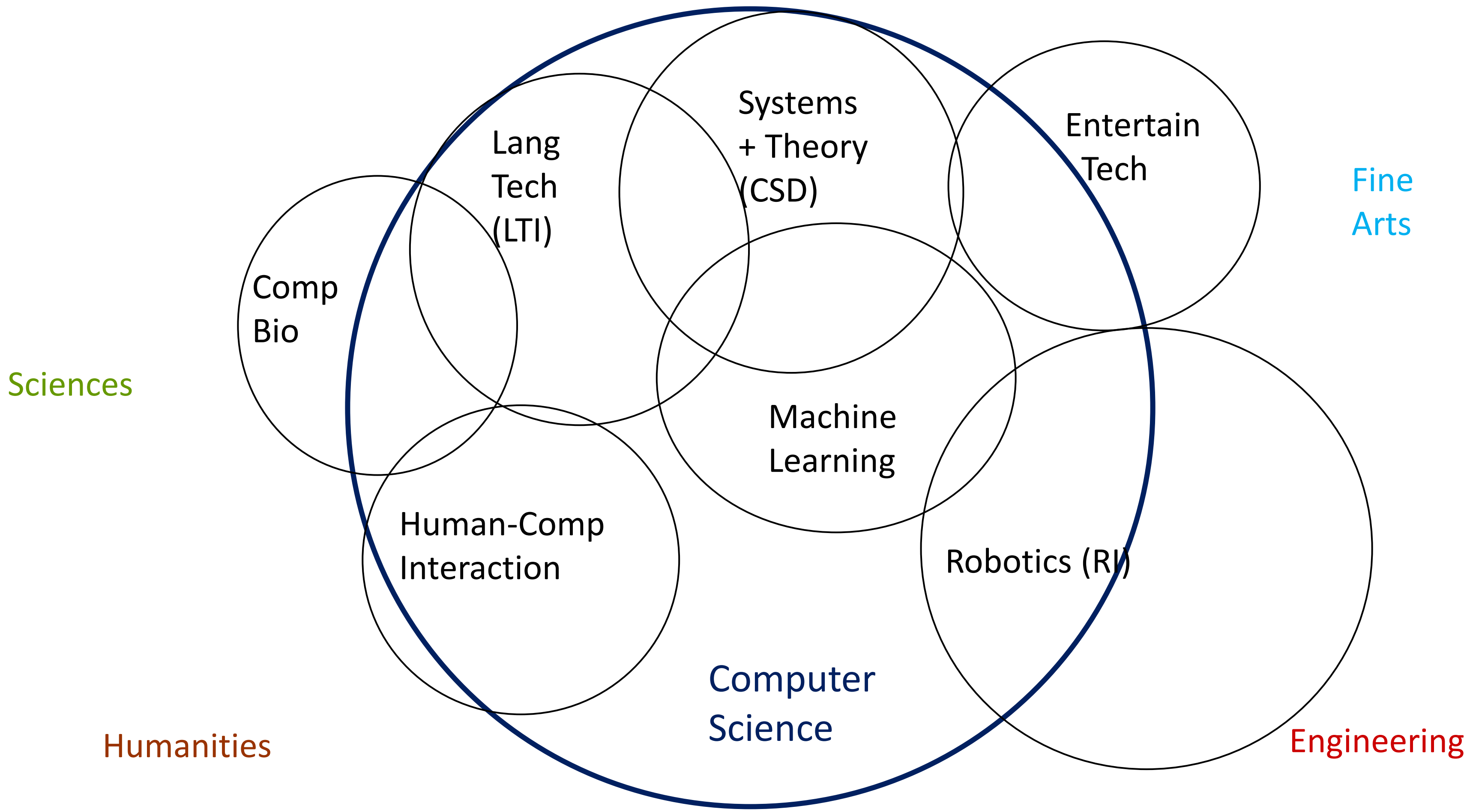
Changes to the Workplace

- Why are there more jobs, if fewer workers are needed due to increased efficiency/output?
  - Increase in quality of life (increase in demand)
  - Increase in complexity of products/services
  - Longevity (consumption after retirement)
- Education must forecast macro trends
  - Educate to jobs 10-20 years in future
  - Major burden on educators/schools/universities

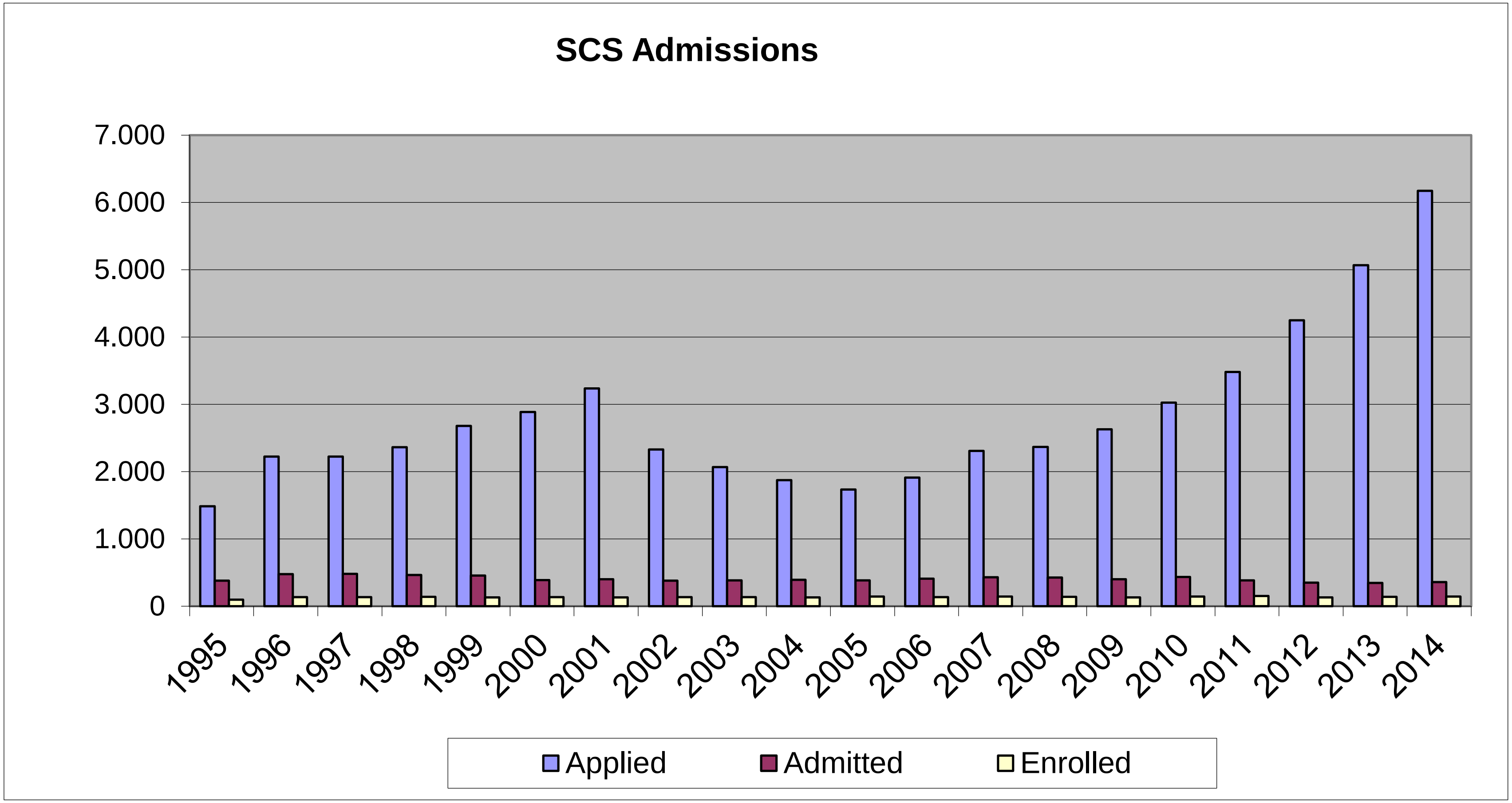
Example Areas of Job Growth

- Generally: STEM++
  - Science, Technology, Engineering, Mathematics
  - Healthcare, Tech services, Education, …
- Specific examples:
  - Mobile technologies (was Web/Internet)
  - Embedded smart computing
  - Medical (macro-molecular, complex devices)
  - Complex individualize manufacturing (3D printing)

School of Computer Science at CMU



Undergraduate Application Trends



- Fall ‘13 admissions: Avg SAT 776m 731r 736w

Technology for Education

- Intelligent Tutoring
  - Skill-based education (math, languages)
  - Example: Carnegie Speech: learn English
- MOOCs, Kahn Academy, …
  - Ingredient: lectures by best-in-brand
  - Do NOT replace interactive teaching, coaching,…
- Enabling Technologies
  - On-line forums, Social-media (e.g. WeSpeke)

Pronunciation Exercises

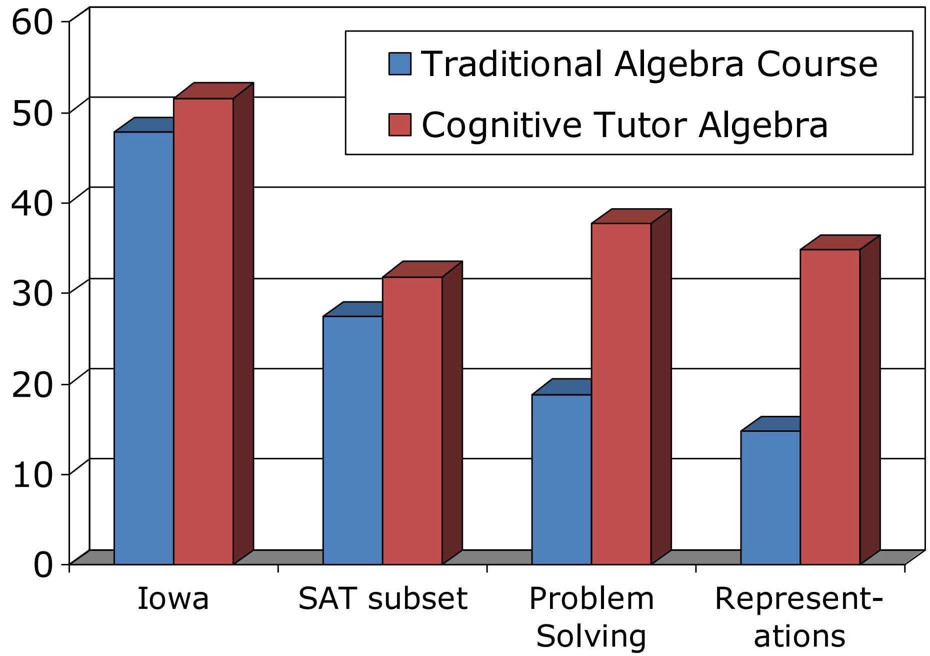
- Play model, record yourself, view graphical, audio, and text feedback

Impact of Cognitive Tutors Adaptive Learning

- Controlled, full year classroom experiments
- Replicated over 3 years in urban schools
  - Pittsburgh & Milwaukee
- Results:

  50-100% better on problem solving & representation use.

  15-25% better on standardized tests.



From Koedinger et al, CMU

Parting Thoughts on Education

- Learning-by-doing is key (Simon, Nobel Laureate)
  - "Tell me and I will forget, Show me and I might remember, Let me practice and I will master skills"
  - Low retention in both lectures & MOOCs
  - Learning must be interactive, goal driven, and focus on inherently interesting problems
  - Intelligent tutors & social-media can help a lot
- Learning takes place in one (or many) domains
  - Abstract problem solving is a myth
  - Students must be helped to generalize
  - Some modern education de-emphasizes content

Questions