PANEL DISCUSSION:
Achieving the Full Potential of Information Modeling and Processing

Moderator:

Ian Flood,

Panelists:

Dominique Thiebaut, Smith College, USA

Alexander Troussov, Russian Academy of Science, Russia

Ian Flood, University of Florida, USA

Challenges:

- Education:
  - Explosion in demand for computer science study programs
  - Demand not met by universities/colleges
  - Industry starting to bypass academia
- Need for very smart models/modeling systems:
  - Automated model development:
    - work for small incomplete datasets (make inferences to fill in the gaps) and knowledge bases;
    - required dataset size not geometrically dependent on the complexity of the problem;
    - noisy datasets and knowledge bases;
    - dynamic (problem changes with time), and
    - unstructured data sets (large or small):
      - contains no or little information on what may and may not be relevant,
      - no indication of how the problem may be broken down,
      - no prescribed format for the input information,
      - can automatically filter out or uncouple mixed streams of information….
  - Work for Complex Systems
  - Extensible:
    - extrapolate extend to new versions of the problem without having to rebuild the entire model,
  - Explainable:
    - to an appropriate level of abstraction (at least translucent, not black boxes)

Approaches:

- For education?
- For very smart models/modeling systems:
  - massive, richly structured models,
  - interface and internal structure learned automatically
  - internal structure analogous of the internal structure of the problem being learned (i.e. not just learn the mapping but also the structure of the problem):
  - inspiration from biological systems:
    - brain (the ultimate black box)
    - growth algorithms (for developing large, richly structured systems).

ICIW 2017, The Twelfth International Conference on Internet and Web Applications and Services.
June 25 - 29, 2017- Venice, Italy

Panel: Achieving Full Potential for Information Modeling and Processing

Alexander Troussov, Ph.D.

Mathematical Methods for Social Network Mining Laboratory,
The Russian Presidential Academy
of National Economy and Public Administration.

On Computer Science education

- More emphasis should be put on teaching Big Data analytics
  - Including engineering aspects of data integration
  - Analytic methods developed during the last decades
    Artificial Neutral Networks frameworks
    Natural Language Engineering
    Methods and applications of Network Science

On Information Modeling and Processing

- The use of artificial neural networks proved to be very successful in processing sensorial information
- However: artificial networks are just methods of Machine Learning:
  Garbage In – Garbage Out
- One of the leading figures in the development of Machine Learning methods is a Russian mathematician Mikhail Bongard. The algorithms he developed are widely used in Geophysics and other domains. His method was the most successful method in the recent history of presidential election in the USA forecasts.
  - In this seminal work: Bongard, M. M. (1970). Pattern Recognition. Rochelle Park, N.J.: Hayden Book Co., Spartan Books. (Original publication: Проблема Узнавания, Nauka Press, Moscow, 1967) wrote that

    the problem of Machine Learning is not only in finding the best algorithms (like hyperplanes in features space), but mostly in finding an adequate feature space. Once such space is found, the solution is already practically found.

On Information Modeling and Processing

- When machine learning (ML) gives recommendations of problems like facial recognition, the explanatory module could be redundant
- But in many other applications, recommendations without explanations are useless (and also offensive in my opinion)
- Knowledge based methods allow to build Explanatory Modules
- Some problems might be solved by one particular algorithm. Real life solution are usually build on dozens of methods. State of the art solutions, like IBM Watson Analytics employ hundreds of techniques
- In addition to the need for explanations, the results of machine learning should open possibilities for interpretation.

Machine Learning, Knowledge, Positivism and Interpretation

- Positivism helped to free Einstein from the notion that there is an absolute sense to a statement that two events are simultaneous; positivism also played an important part in the birth of modern quantum mechanics.
- Positivism has done as much harm as good. For instance, J. J. Thomson is widely credited for the discovery of the electron ☺, despite the fact that other scientists did the same experiments (these scientists were positivists ☹) .
- Hopefully, the humankind will not abandon the quest for understanding.
  The lecturer concurs with the position of Claus-Peter Rückemann expressed in his Keynote talk: Allow Knowledge to Prevail…"
- While Machine Learning is very useful, it is more an art than a science.
  The success hinges on the Knowledge
  - To build better models
  - To facilitate the explanation of the results
    (using hybrid method, where different techniques are used for the solution and for the explanation
    We forecast the weather by solving modified Navier-Stocks equation, but our trust in the results is based on the presentations of the military style maps with Red and Blue fronts and arrows indicated troops movements and charges)

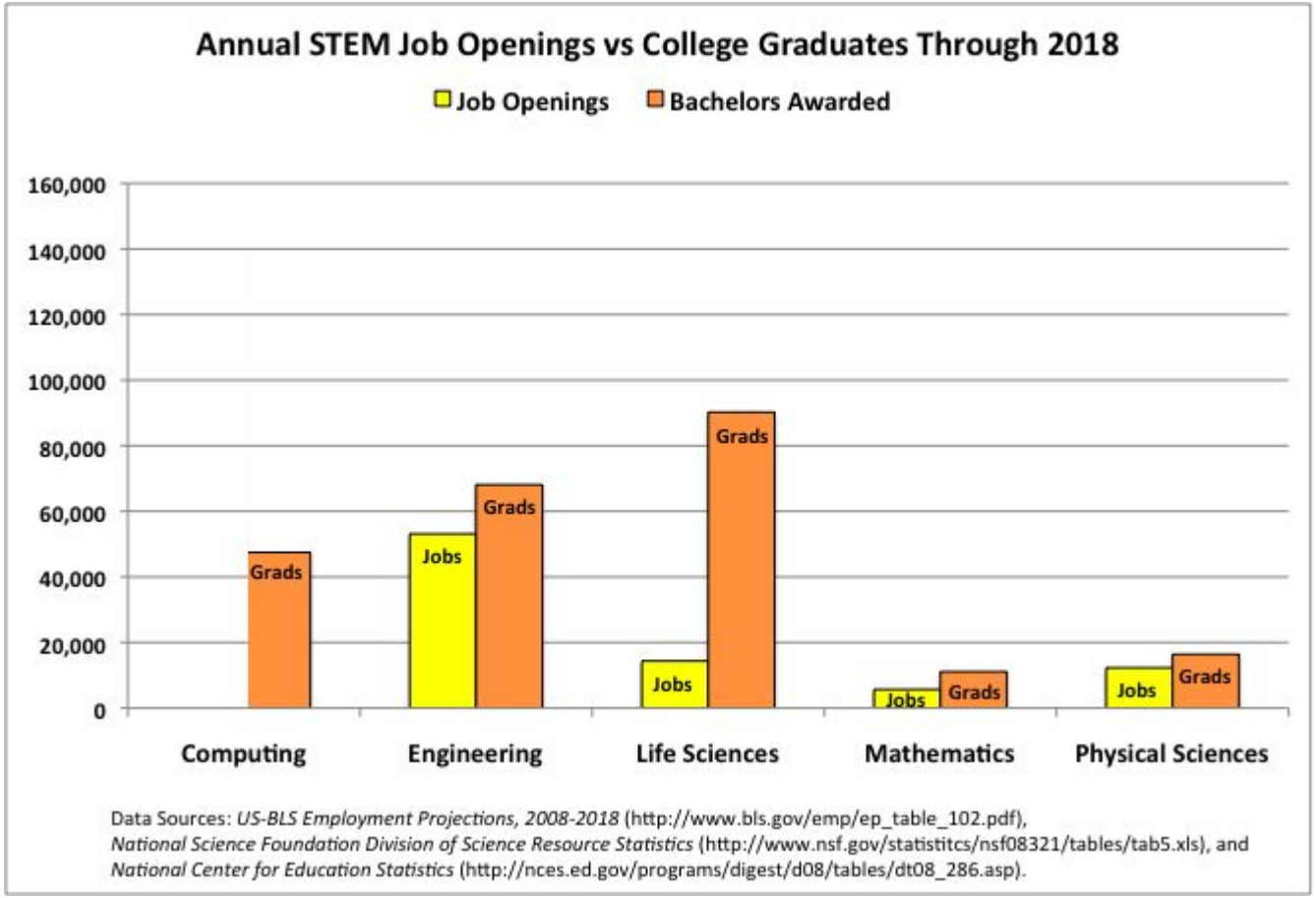
InfoComp 2017

Challenges Facing *Computer Science* Education

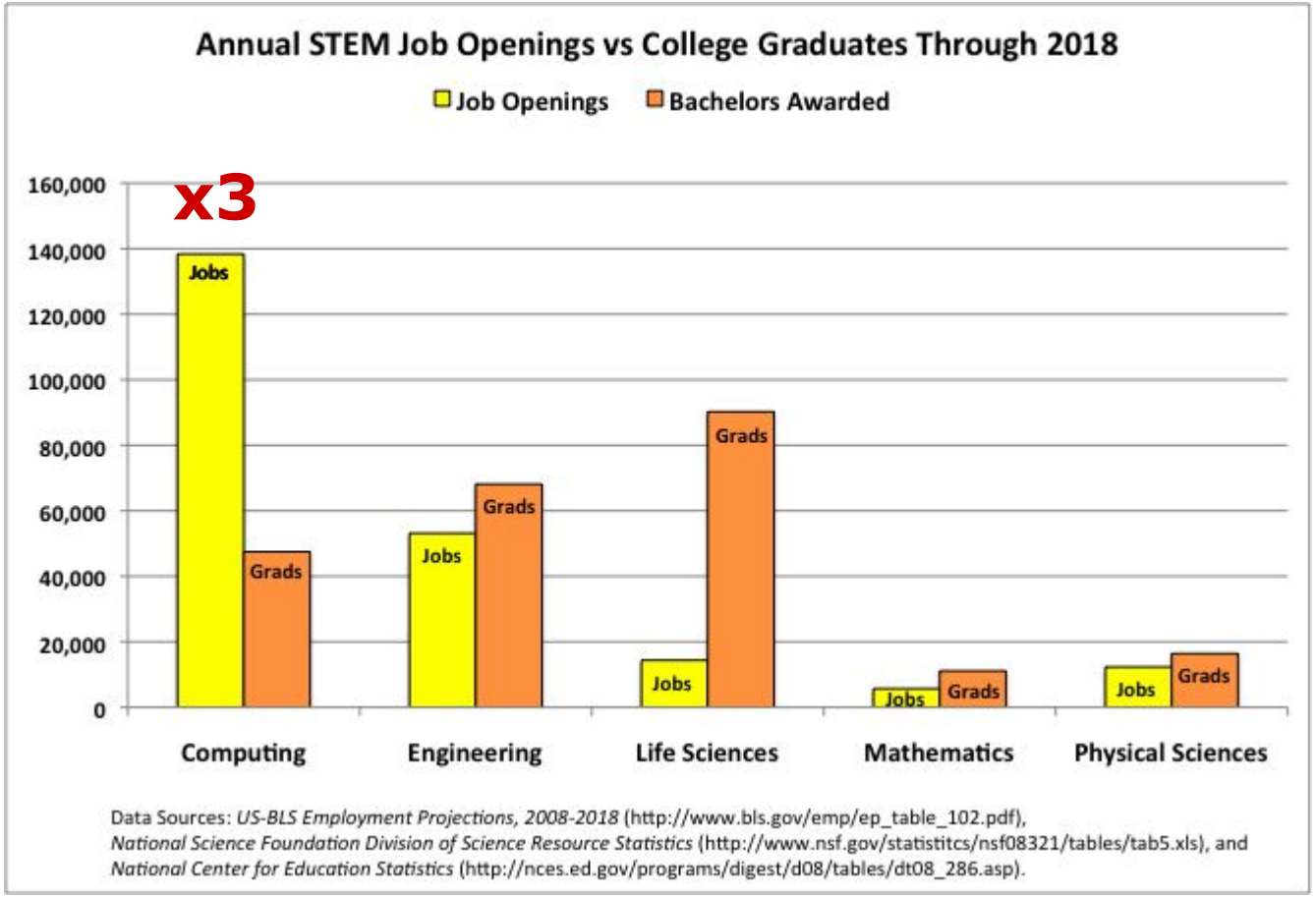
Dominique Thiebaut
Smith College
Northampton, Massachusetts, USA

[Photo credit: https://kindnessblogdotcom1.files.wordpress.com/2014/11/dominos-1.jpg]

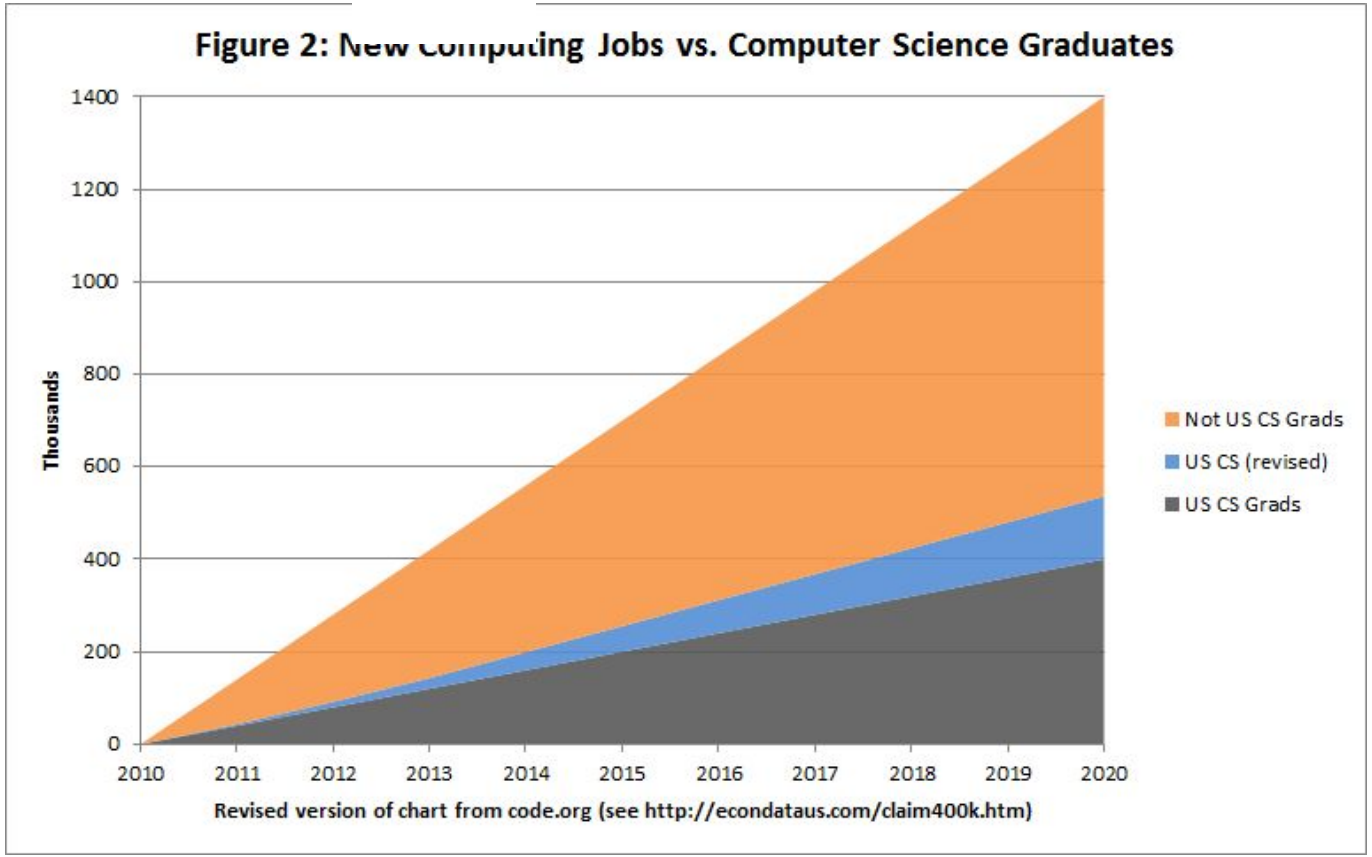
Google Online Education Council (2016-17)



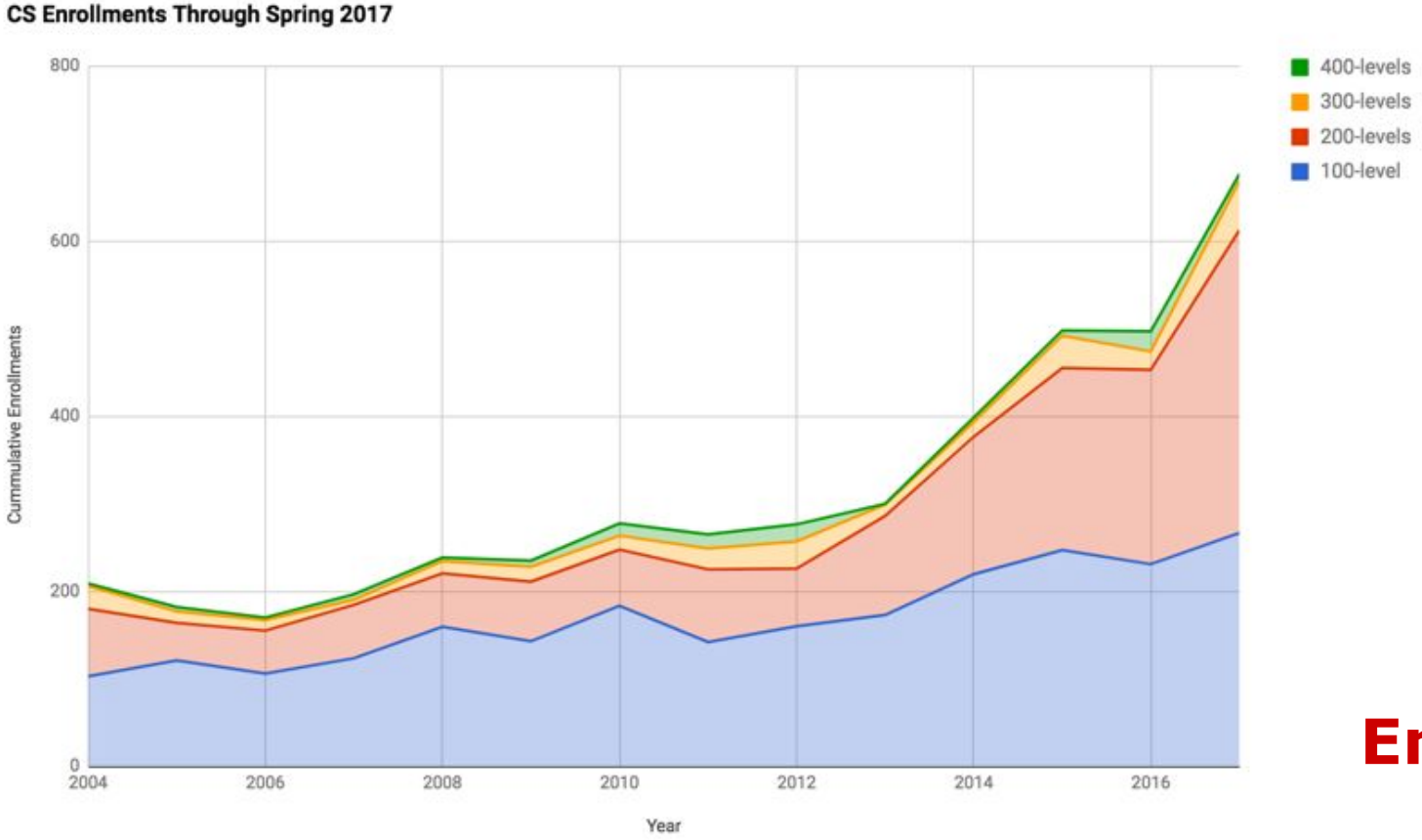
Jobs vs CS Graduates



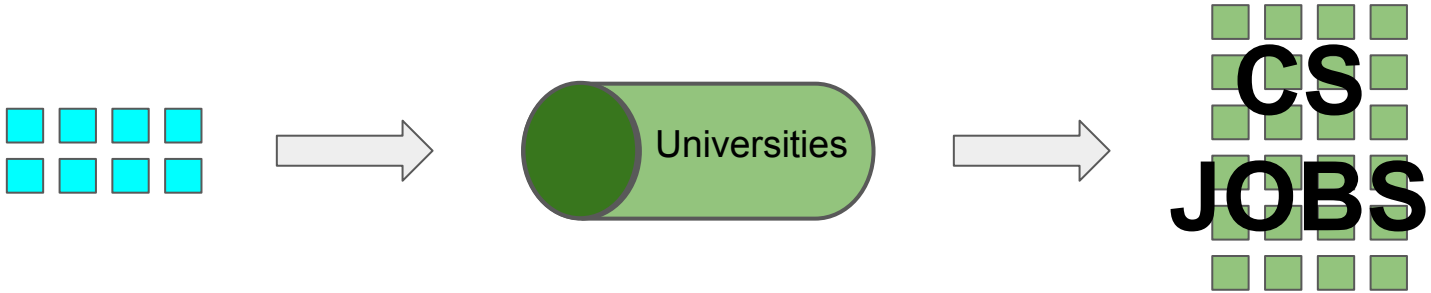
Jobs vs CS Graduates



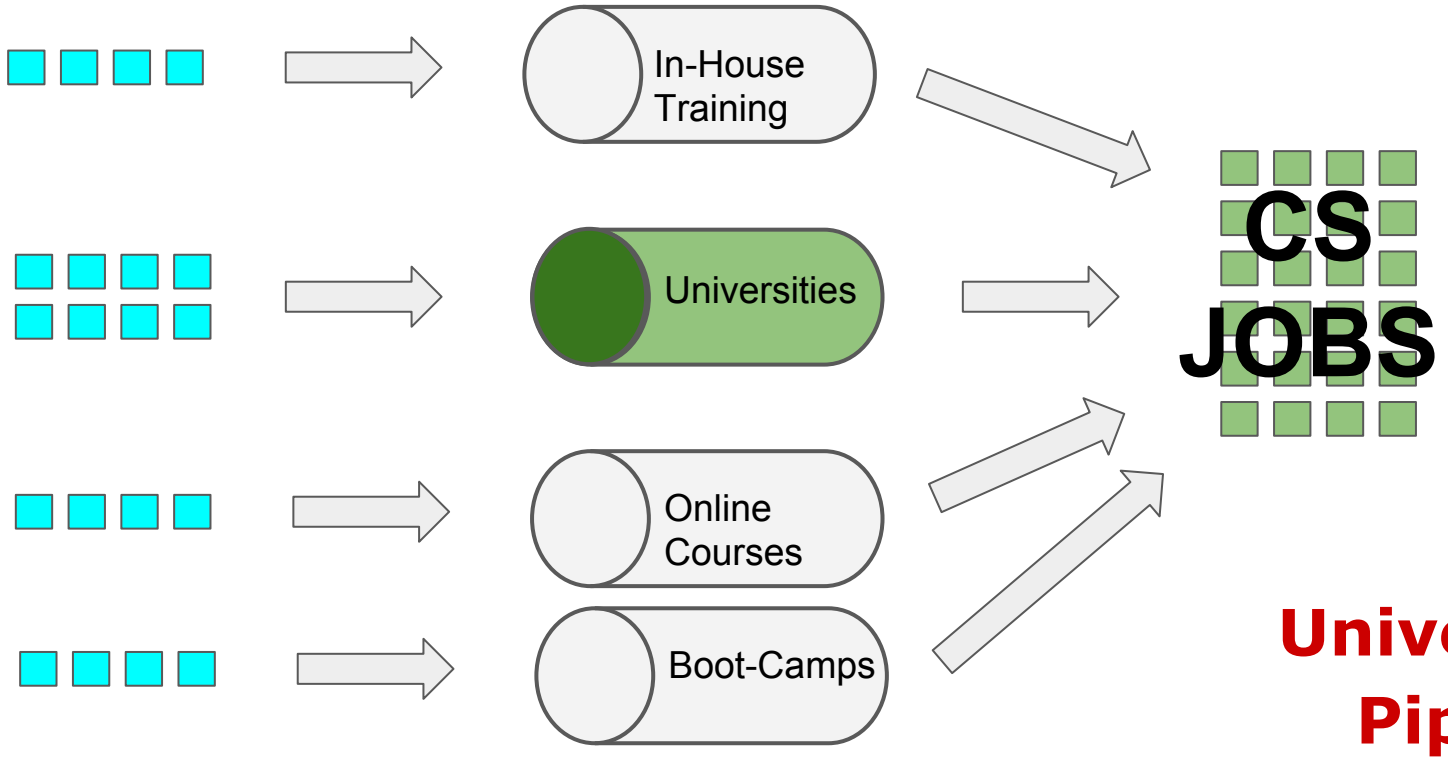
Jobs Growth Rate



CS Enrollments



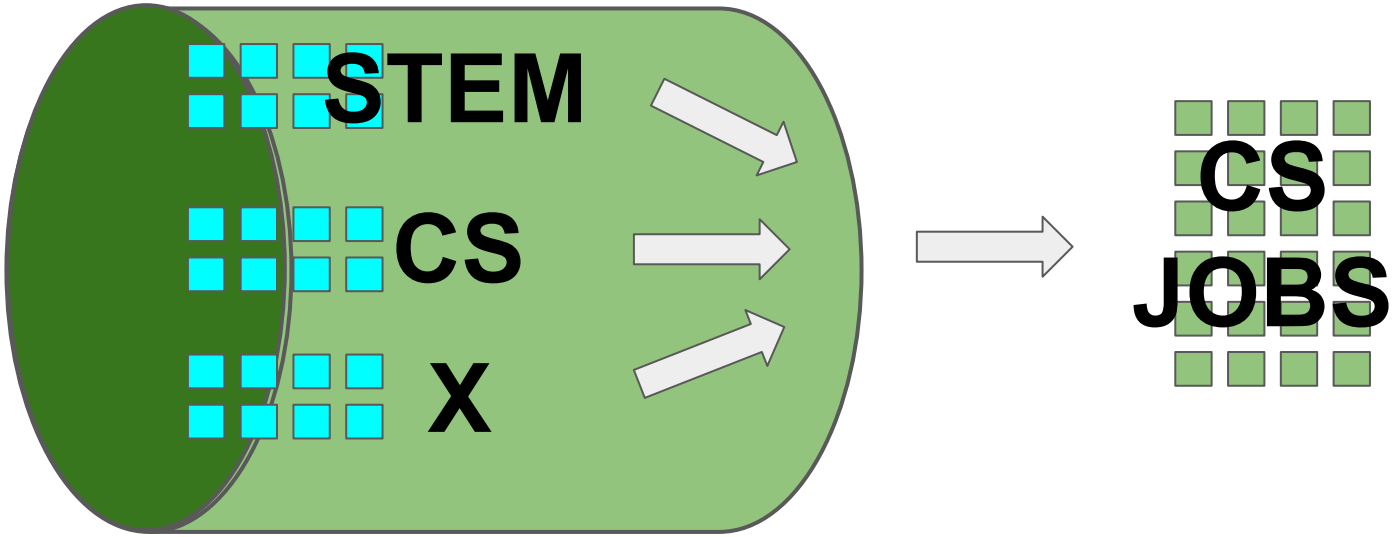
University Pipeline



University Pipeline

"Last month, Apple introduced a yearlong curriculum for high schools and community colleges to teach app design in Swift. Apple has also supported Code.org by hosting the group's popular Hour of Code events in its stores."

Education & Industry



Inside The Pipeline

Domino Effect

University pipeline unable to provide

- Universities bypassed
  - Fewer PhDs
    - Difficulty recruiting faculty
      - Industry - CS Dept. alliances
        - Talent bypasses universities

References

https://www.linkedin.com/pulse/computer-science-job-prospects-graduating-class-2016-barry-byers

http://sites.nationalacademies.org/CSTB/CurrentProjects/CSTB_173432

As CS Enrollments Grow, Are We Attracting Weaker Students?

Mehran Sahami, Chris Piech, Stanford U.

n recent years, enrollments in undergraduate computer science programs have seen tremendous growth nationally. Often accompanying such growth is a concern from faculty that the additional students choosing to pursue computing may not have the same aptitude for the subject as was seen in prior student populations. Thus such students may exhibit weaker performance in computing courses. To help address this question, we present a statistical analysis using mixture modeling of students' performance in an introductory programming class at Stanford University over an eight year period, during which enrollments in the course more than doubled. Importantly, in this setting many variables that would normally confound such a study are directly controlled for. We find that the distribution of student performance during this period, as reflected in their programming assignment scores, remains remarkably stable despite the large growth in enrollment. We then explain how the notion of having "more weak students" and the fact that the distribution of student ability is unchanged can readily co-exist and lead to misperceptions about the quality of incoming students during an enrollment boom.

TEALS: Teacher Professional Development Using Industry Volunteers

Nathaniel Granor Microsoft/TEALS New York, NY nathaniel@tealsk12.org
Leigh Ann DeLyser NYC Foundation for CS Education New York, NY leighann@csnyc.org
Kevin Wang Microsoft/TEALS Redmond, WA kevin@tealsk12.org

ABSTRACT Rising demand for high school computer science courses in the United States has created pressure to increase the number of computer science(CS) teachers in a short amount of time[2]. In this experience report, we present the TEALS program as a unique, high-touch, professional development model, pairing computing industry professionals with classroom teachers. By combining the relative strengths of the team (content and pedagogy) TEALS has been able to successfully train new CS teachers. We present the history of the TEALS program, the volunteer and teacher recruitment process, the volunteer training program, data from a study of the pedagogical content knowledge of the TEALS volunteers, and program growth and efficacy data. Additionally, we offer achievement of students on the AP CS A exam as an externally valid measurement of learning outcomes in TEALS classrooms.

Peter Norwig: Teach Yourself Programming in 10,000 hours

"Bad programming is easy. *Idiots* can learn it in *21 days*, even if they are *dummies*." Felleisen *et al, How to Design Programs*

Quincy Larson, FreeCodeCamp. https://www.youtube.com/watch?v=_FioceDs7JA

Emmanuel Schanzer “CS Education as engineering problem,” https://www.youtube.com/watch?v=turBxnXqIIs