Behind the Scenes of a University Education

John Hopcroft
Cornell University, faculty emeritus
SJTU, PKU, HUST

Introduction

- I am going to talk about many issues involving education:
  - How to be successful,
  - Having a good life,
  - What China is doing to improve education.
- Finally, I will talk a little about computer science.

Mission of a University

- To educate students for a good life.
- To help students discover what they really enjoy.
- Educate students to get a job they enjoy.
- Include humanity and social science courses in science and engineering programs.
- Missions of US and Chinese universities are different.

Applied and basic research

- The terms basic and applied do not refer to how fundamental the research is.
- They refer to why one is doing the research.
- Applied research is research done to solve a specific problem.
- Basic research is research that someone is curious about.

Basic Research Leads to New Ideas

- Applied research is unlikely to produce fundamentally new research directions since it is focused on solving a specific problem.

- Basic research goes off in all kinds of different directions and sometimes creates whole new areas leading to new industries and millions of jobs.

Missions of US and Chinese Universities are different

- The US has produced enough scientists and engineers for the nation’s needs.
- US universities do not engage in applied research.
- Applied research is done in national laboratories or company’s research divisions.

Apparent Contradiction

- The mission of top US universities does not include research although it looks like they focus on research.
- I will explain this apparent contradiction in a few minutes.

Chinese Universities

- China has not yet produced sufficient engineering and scientific talent for the needs of the nation.
- Chinese universities have two missions:
  - Educate the next generation of talent.
  - Help government and industries with applied research.

Why Hire Researchers

- If the mission of university is to educate, why hire researchers for faculty?
- Universities are hiring individuals for a career of 40 years.
- They want someone who is curious, and if a new direction appears in their field they are likely to explore it and update their course.

Hiring Evaluation

- Many recruiting committees look at the candidates research and ask how important is it.
- They should instead be asking, why did he pick the topics he has worked on.
- Will he continue to explore as he he gets older?

What are we teaching?

- What a student learns is far more important than the technical material taught.
- They learn how to think, solve problems, analyze issues, make decisions, predict consequences, distinguish truth from fiction, be creative, communicate clearly, to learn, to explore.
- When faculty are teaching the technical content, the students learn these softer skills.

Notion of a Limit

- A teacher ask me how to determine the area of a circle if the constant pi was not known.
- I divided the area of a circle into small squares and counted the number of square in the circle.

Notion of Limit Continued

- By making the squares smaller I realized I could get a better approximation.
- I learned that there are approximations and the intuitive concept of a limit.
- In college, I grasped that the mathematical definition of a limit formalized my intuitive notion of a limit.

One Gets What They Measure

- It is important to get correct metrics for measuring the quality of university programs.
- University presidents in China are focused on improving their institutions international ranking.
- These rankings are focused on research money and number and quality of research papers.
- A better metric is the quality of undergraduate teaching.

Evaluating Computer Science at 40 Universities

- I chaired an international board advising the premier on how to improve education in China.
- Our report was to be at most one page.
- We focused on undergraduate education.
- We recommended changing the metrics by which universities were evaluated from research funding and quality of papers to quality of teaching.

Evaluation Process

- Have a faculty member sit in on a lecture and observe:
  - Is the faculty member teaching the course engaging the students,
  - Is he excited and knowledgeable of the course material,
  - What fraction of the students are engaged.

Evaluation method is Subjective

- The MOE had many objective evaluation techniques.
- They were uncomfortable with a subjective technique.
- Issue solved by giving me permission to evaluate computer science teaching at the top forty universities.

Evaluation

- The evaluation now is in its fifth year.
- University presidents are informed if their institution's quality of teaching is in the top third, middle third, or bottom third.
- The project has had a major impact on university education in China.

Systemic Impact of Evaluation

- University presidents have five year terms.
- When completing their term they are assigned a new position depending on their performance.
- They changed their emphasis from research to quality of undergraduate teaching.

Do What You Enjoy

- Many individuals have asked me what my strategy was that led to my receiving the Turing Award.
- I had no strategy.
- If an exciting opportunity can along, I took it.
- If an opportunity was not exciting, I ignored it.

My Career

- Undergraduate work in Electrical Engineering at Seattle University.
- Graduate work in Electrical Engineering at Stanford.
- At Princeton, I was asked to teach a course in computer science.
- The content of that course was used around the world.

Discover What you Enjoy

- If as a child you liked to play with blocks and build things maybe you would enjoy a career in manufacturing.
- If there was a woods with a path and you explored where the path went maybe you would enjoy a career in research.
- If you enjoyed playing with other children maybe you would enjoy a career in education or medicine.

Project 101

- The Ministry of Education has created a project to improve computer science education at the top 33 universities in China.
- The project has two components:
  - Develop content for the important courses in computer science, and
  - Improve the quality of teaching the material.

Improving Content of Key Courses.

- The 33 universities selected 12 key courses.
- Each course was assigned to an individual.
- The assigned individual formed a team of 20 highly qualified faculty to develop material.
- This material should be made available to all 1,500 universities as soon as possible.

Portfolio

- Consists of 50-100 key concepts for a course.
- Two to three page write up for each concept.
- Should take only two to three weeks to develop for a course.
- A portfolio is fundamentally different from a book.

Comparison of Book and Portfolio.

- A book is developed by two or three coauthors and takes a year to complete.
- The result is the best content the two or three faculty can created.
- A portfolio is developed by 30 or 40 faculty and takes one or two weeks to create.
- The result is the best content that 30 or 40 faculty can create.

- Each faculty submits all the topics he feels are important.
- Union of topics submitted gives list of 50-100 topics.
- Each faculty is assigned approximately three topics to develop in two or three pages.
- Result is best content 30 or 40 faculty can create.

Comparison of Book and Portfolio

- If a book and a portfolio are placed on the internet and used by 1,500 universities, some faculty will point out additional material that should be included and some material that is not clear.
- Book unlikely to be updated.
- Portfolio can be updated in two or three weeks.
- Subjects are continuously evolving and need constant updates.

Course Content Project for 1,500 Universities.

- First step: each university designate an individual responsible for activity at their institution.
- Next step: these individuals develop list of disciplines to be included in project.
- Next, the universities that wanted a discipline developed, need to create a list of important courses for that discipline.

- Each course assigned to a qualified faculty who seeks 50 faculty teaching an existing version of course.
- These faculty submit list of intellectual ideas to include.
- Ideas are assigned to faculty to develop two or three pages.
- Each faculty develops only two or three topics.
- Process completed within a month and is of the quality 50 faculty can produce.

Improving Teaching

- When evaluating teaching one learns information that can help the professor teaching the course improve his teaching.
- Inform faculty member teaching the course of the interaction of the students to the lecture.
- The asumptions is that the discussion will lead to improved teaching.

Faculty Sitting in Lectures

- Proj 101 has faculty sit in lectures.
- Discuss with faculty member teaching the course how students interacted with material taught.
- The resulting discussion leads to improved teaching.

Evaluating is Different from Improving

- The two processes where a faculty member is sitting in on lectures are fundamentally different:
  - In evaluating, the faculty needs to be unbiased.
  - In improving teaching, the faculty member is working with the faculty member teaching the course and the two are a team.

Proj 101 Found Very Successful

- Extended to other disciplines.
  - Physics, Chemistry, Biology, Basic Medical Science, Traditional Chinese Medicine, Economics, Philosophy, etc.
- Important to educate all students:
  - In mathematics, the faculty in China may focus on high level courses for pure math majors.
  - Overwhelming number of students in math courses are not even math majors.

US Education

- Best strategy to improve US education is to focus on either:
  - The first three years of a child's life, or
  - Elementary school education.
- Education in the US is very distributed:
  - Each local community has its own school district,
  - Each district taxes its community, hires and funds teachers, controls what is taught.

The First Three Years of a Child's Life

- In the first three years the brain learns how to learn.
- A nurturing environment improves a child's life.
- Research falls into two categories:
  - Focus on impact of environment on child's life,
  - Focus on how the brain develops.

Research on Nurturing Environment

- Divide infants into two groups.
- For first groups create an intellectual and stable environment for first three year of child’s life.
- Leave second group alone for a comparison groups.

Nurturing Environment

- Thirty years later compare individuals in the two groups.
- Children in the first group had higher educational levels, had better jobs, and fewer had criminal records.
- The first three years had a significant statistical impact.

Researching Different Environments

- I talked with a researcher in Beijing who was studying the impact of different environments.
- I remarked that it must be difficult to do research where one needs to wait 30 years to see the outcome of an experiment.
- She remarked that was not how she was working.

Mice as a Proxy

- She was using mice as a proxy for humans.
- She divide 200 mice into two groups.
- First group she placed in cages with intellectually rich environmnets.
- Second group was placed in ordinary cages.
- After three weeks the two groups were combined.

Result

- After two years the mice were tested on mazes and other activities.
- The first group clearly outperformed the other mice confirming that the early environment has a major impact.
- Anyone can do this experiment.

What is a Good Intellectual Environment for a Mouse

- She set me an image of a cage with an intellectually rich environment.
- Each item was labeled as to what skill it developed.

Important Aspects of the Environment

- Two parent environment.
- Appropriate toys in year one.
- Age appropriate books in year two.

Elementary School

- Elementary school forms a child's view of the world.
- Their view once formed is difficult to change.
- They often live in a neighborhood that has the same view and confirms the view.
- Thus, what is taught in elementary school is important.

What is Currently Taught in Elementary School is not Correct

- Taught that democracy is the best form of government instead of there are may forms of government each having advantage and disadvantage.
- US history does not teach the true history, but a politically accepted version.

Improving Elementary Education

- Improving elementary education requires two components;
  - Determine what and how material should be taught,
  - Develop a strategy to get it well taught.

Determining the Material

- Approach professional organizations to get experts for each subject matter.
- Have historians develop US history as it actually happened, not as it is currently being taught.

Getting it taught

- Publish material in digital form on the internet where everyone would have access to it.
- Some politicians might not want it taught in schools but material would eventually filter down.

The World is Changing

- Civilization has undergone an agricultural revolution and an industrial revolution.
- Now it is undergoing an information revolution.
- Revolutions are occurring more frequently and happening faster.

The Agricultural Revolution

- Prior to the agricultural revolution, humans were simply hunter gathers.
- The agricultural revolution took thousands of years.
- It increased food supply and the total population.
- It created communities and small farms.

The industrial revolution

- Much quicker, only several hundred years.
- Created manufacturing jobs.
- Automated agriculture.
- Improve quality of life.

The information revolution

- The information revolution may only require one lifetime to complete.
- Automating many jobs and changing lifestyles.
- It may produce goods and services so only a small fraction of the population will work.
- Jobs will change from manufacturing to information processing.

The Future

- In the past, what made nations great was oil, gold, or agriculture.
- In the future, it will be talent.
- Nations that invest in education and make education one of their top priorities will become the important nations.

AI will be an important component of the information revolution.

- Developing AI will require nations to develop talent and that will require focusing on quality education.
- AI will be applied to agriculture, biology, medicine, manufacturing, and many other areas.

Research mathematically sophisticated

- Research has become mathematically sophisticated requiring highly educated talent.
- Nations need to focus on improving education.
- I will now focus on computer science development over the past 50 years and future directions.

Theoretical Computer Science

- Computer science started around 1964.
- At that time only computer programming was taught.
- I taught that was probably the first computer science course.
- There were no text books and I needed to develop the content.

Content of Course

- Finite automata
- Regular sets
- Nondeterminism
- Closure properties
- Context-free languages
- Turing machines, computability and undecidability

Finite Automata

- Input: strings of zeros and ones, Output: yes or no.
- Memory: finite set of states.
- A finite automaton accepts a set of strings that take the automaton to an accepting state.
- Sets accepted by a finite automaton are called regular.
- There exist sets not accepted by any finite automaton.

Nondeterminism

- Deterministic finite automaton: For each state and input symbol there is a unique move.
- Nondeterministic automaton: Several possible next move.
- A string is accepted if there exists a sequences of moves leading to an accepting state.
- Nondeterminism has become an important concept.

Regular Expressions

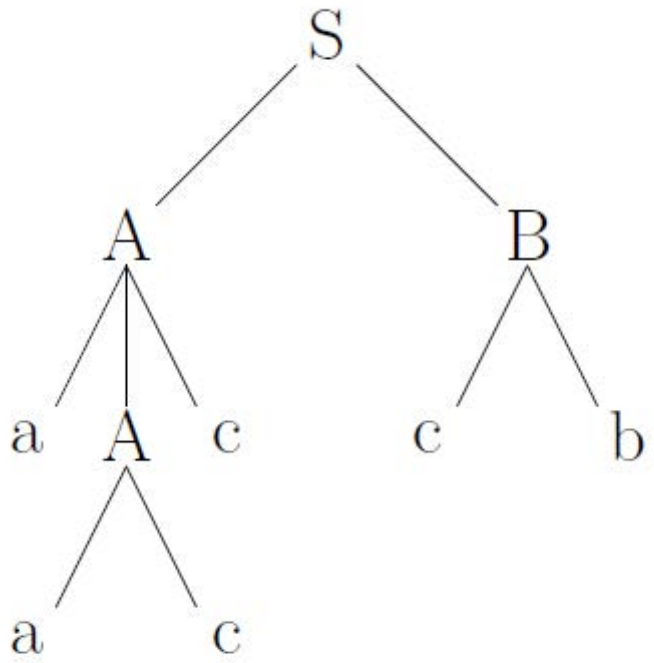
- Regular expressions denote sets described by union, concatenation, and closure of finite sets.
- Regular exressions denote precisely the sets accepted by finite automata.
- Showed how to construct a regular expression from a finite automaton and how to construct a finite automaton from a regular expression.

Context-Free Grammar

$S \rightarrow AB$

$A \rightarrow aCc|ac$

$B \rightarrow cBb|cb$



aacccb

ALGOL

- ALGOL syntax was defined by a context-free grammar.
- This lead to automatic creation of a compiler for ALGOL.
- Something was wrong.
- Two different compilers resulted in different answers to a problem.

Ambiguity

If A then if B then C else D

If A then (if B then C else D)

If A then (if B then C) else D

Undecidability

- Ambiguity for context-free grammars isundecidable.
- It was known that many math problems were undecidable.
- The course covered undecidability, computability, and Turing machines.

1970 algorithms

- Theoretical computer science expanded to include algorithms.
- Algorithms were measured by how fast their running times grew as the size of data increased.
- Design techniques that lead to good performance were developed.

P versus NP

- Could not find deterministic polynomial time algorithms for many problems that could be solved in nondeterministic polynomial time.

- Steve Cook provide that nondeterministic and deterministic polynomial time would be equivalent if there existed a deterministic polynomial time algorihm for a satisfiability problem.

Moore’s Law Ending

- Moore’s law stated that number of transistors on a chip doubled every two years.
- Focused changed from universal computers to special purpose computers.
- Graphical processing units that multiplied vectors extremely fast.

Graphical Processing Units

- GPU’s are not general-purpose computers, but computers that multiply vectors extremely fast.
- Found uses in areas other than graphics such as deep learning.
- Other areas where special purpose computers will be used includes quantum computing.

Quantum computing

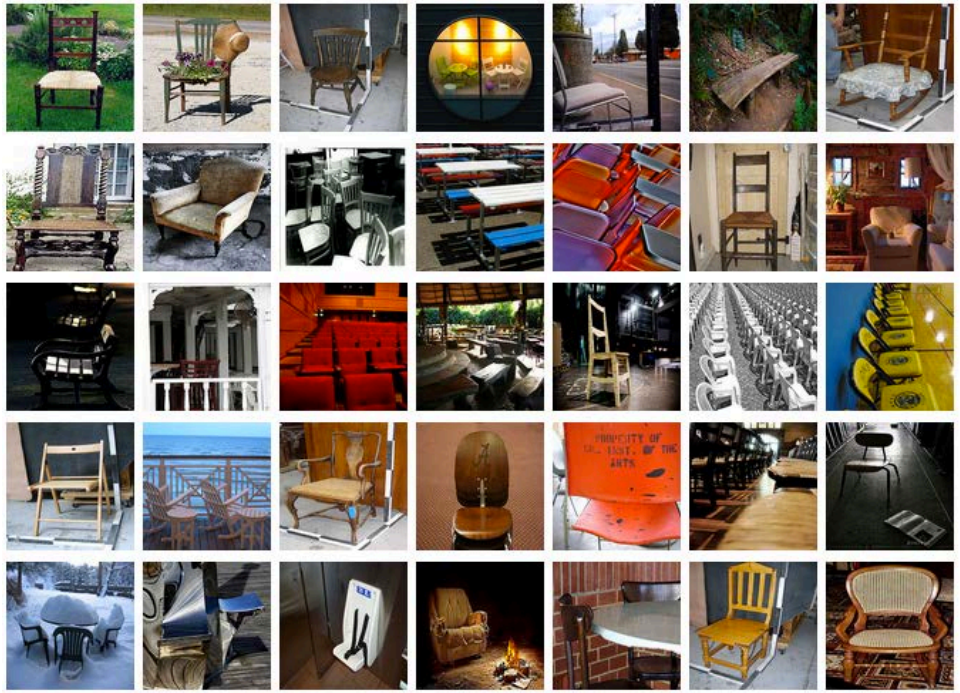
- Researchers in quantum computing realized that maintaining sufficiently many qubits for a general purpose computer will not occur in the foreseeable future.
- They focused on special purpose quantum computers such as a unit whose sole purpose is to factor integers.
- Proj 101 on quantum can be access at

  http:// 101.pku.edu.cn/tbzsd

A major advance in AI occurred in 2012 when deep learning was created.

h

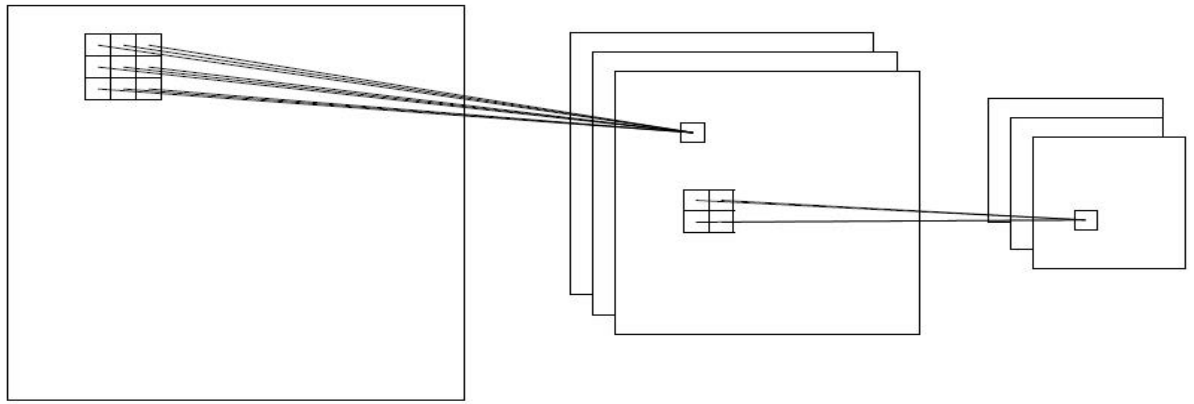
Image Net Competition



50,000 images, 1,000 categories

Advent of Deep Learning

- 2012 Image Net competition.
- AlexNet error 15%, previous rate 25%.
- Such an improvement lead researchers to apply the technology in many areas with significant success.
- 2015 ResNet, 1000 levels, error rate 3.6%. Training similar to stochastic gradient descent.
- Human with training has an error rate of 5%.


convolution
pooling

AlexNet

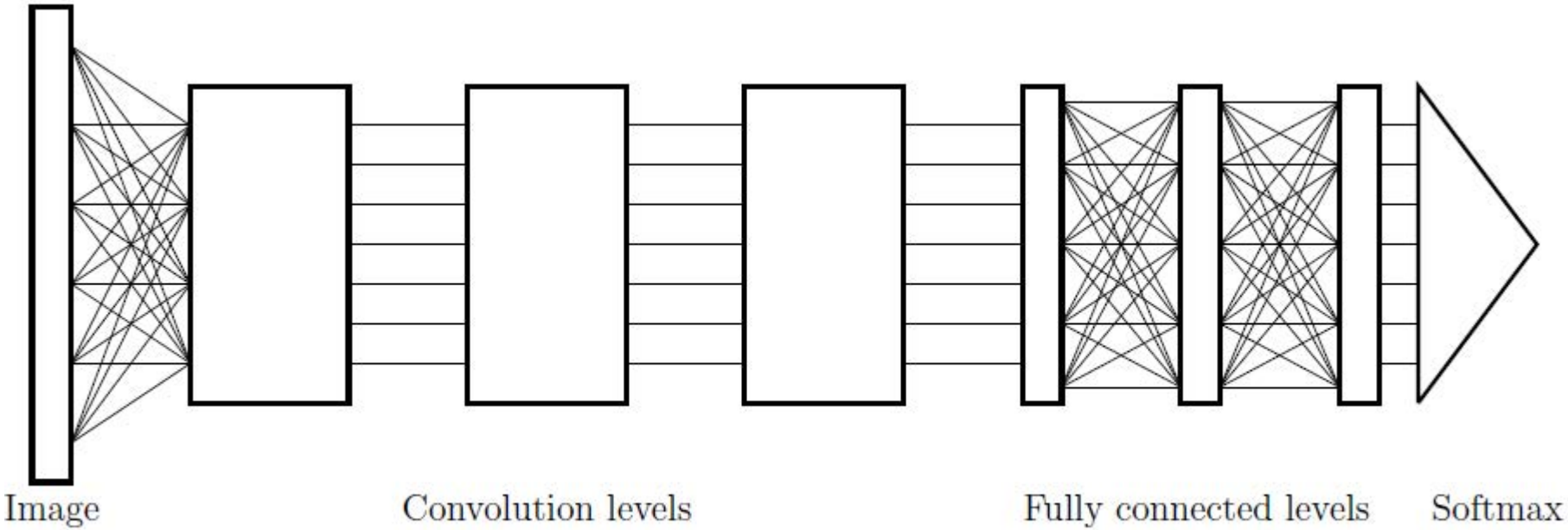
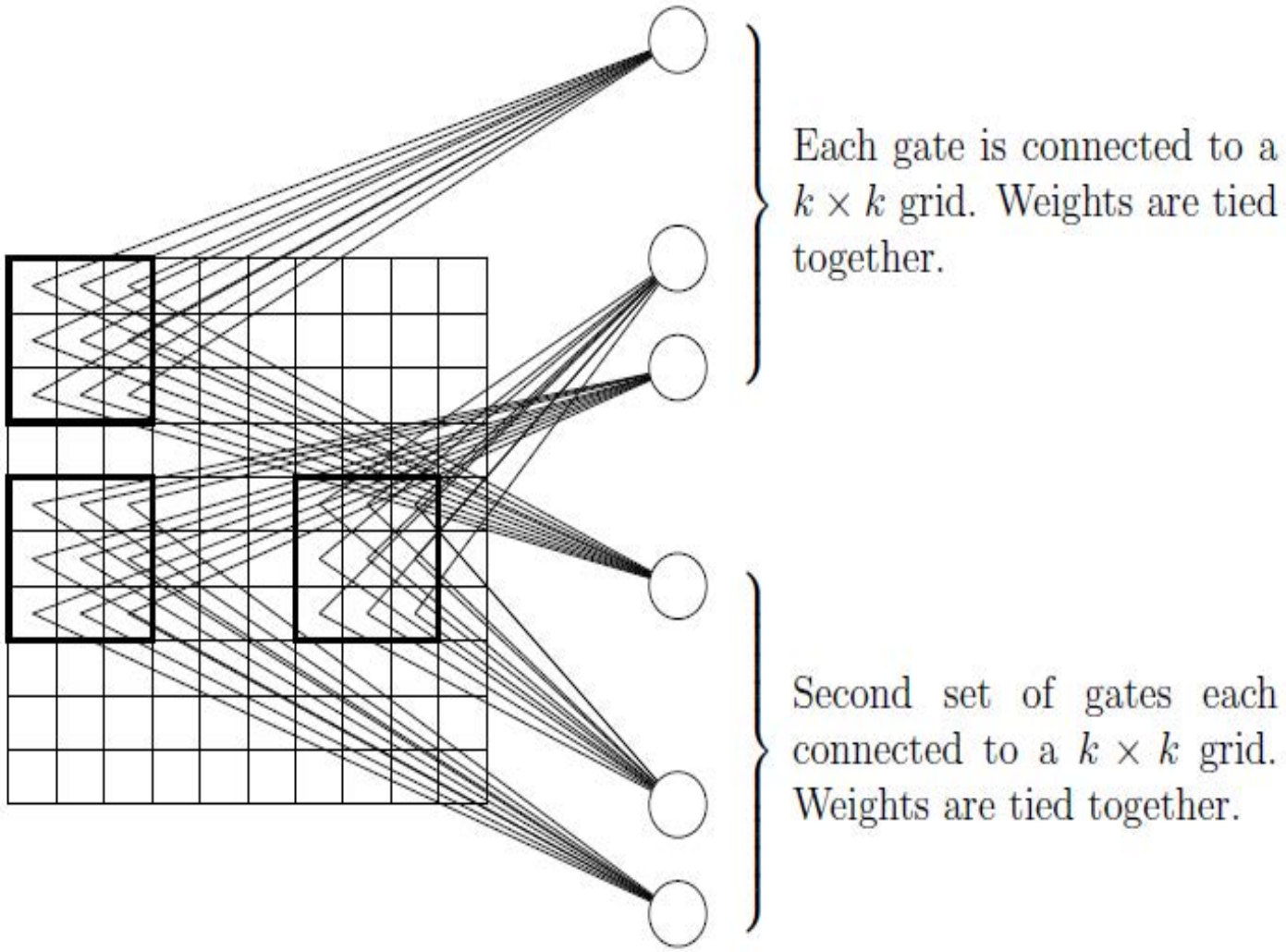


Image
Convolution levels
Fully connected levels
Softmax

AlexNe


Each gate is connected to a $k \times k$ grid. Weights are tied together.
Second set of gates each connected to a $k \times k$ grid. Weights are tied together.

Version of gradient descent

Gradient Descent

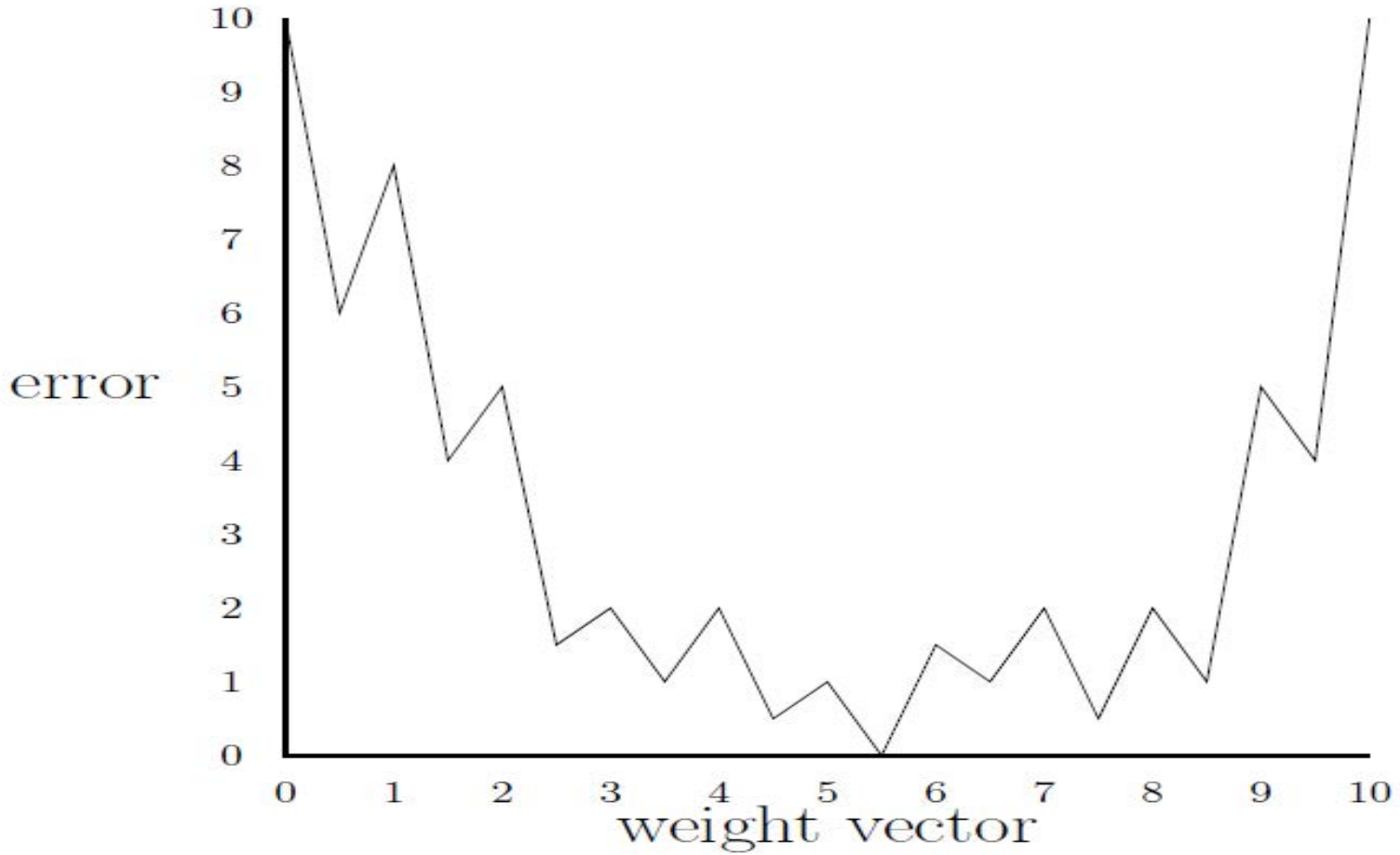
$$x \leftarrow x - \nabla f(x)$$

If f(x) is a sum of millions of terms, taking the derivative is time consuming

Stochastic gradient descent

Instead of taking the derivative of the sum, in each iteration replace the sum by the sum of a small random set of terms.

Gradient Descent Graph

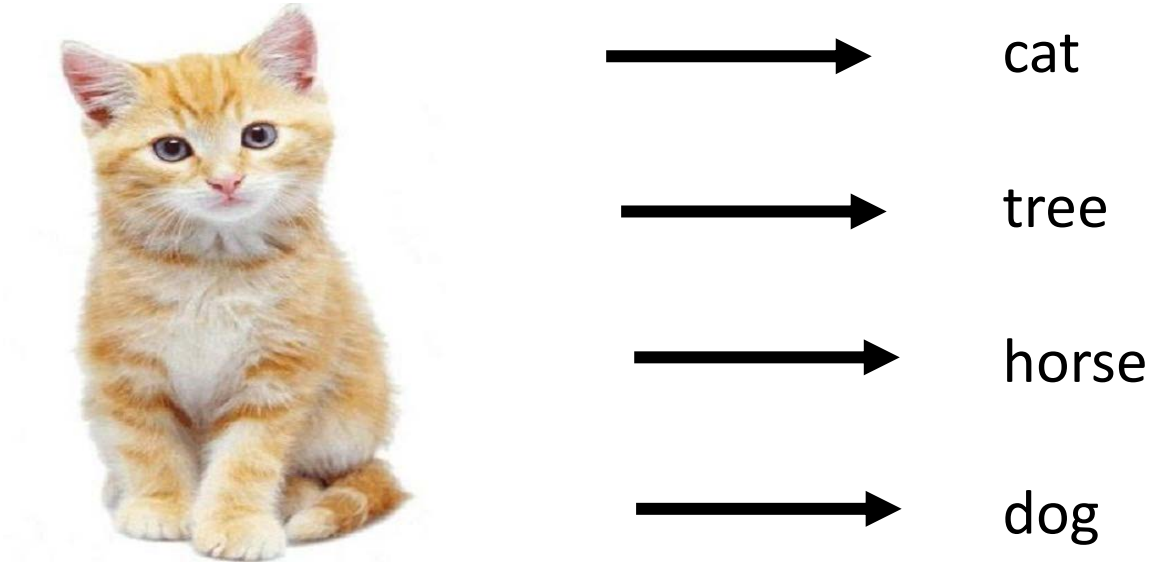


Fooling deep learning

Cat

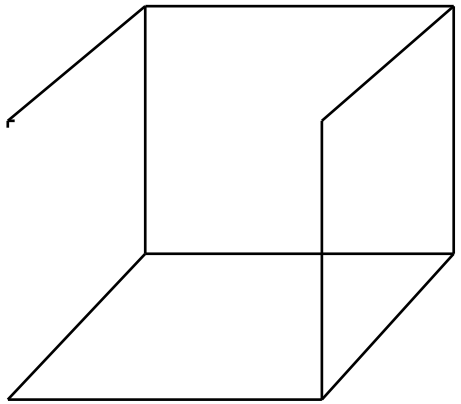
Automobile

A minor change to any image can change its classification to any arbitrary classification.

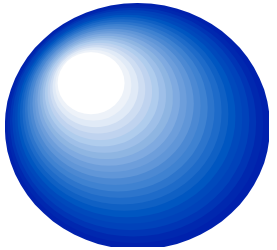


High Dimensional Space

Intuition from two and three dimensions is not valid for high dimensions.

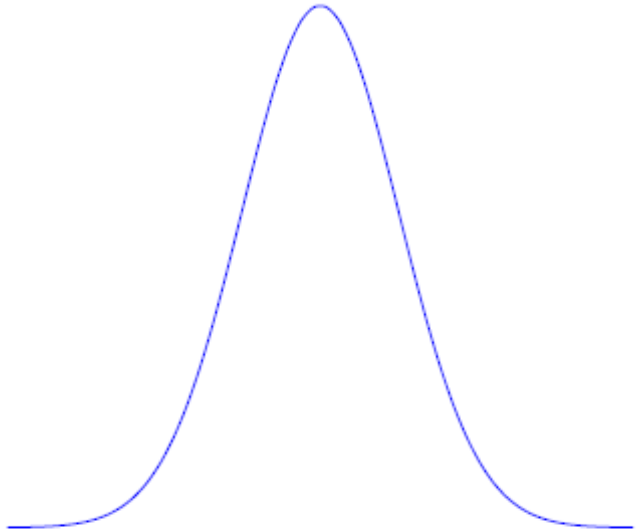


Volume of cube is one, independent of dimensions.

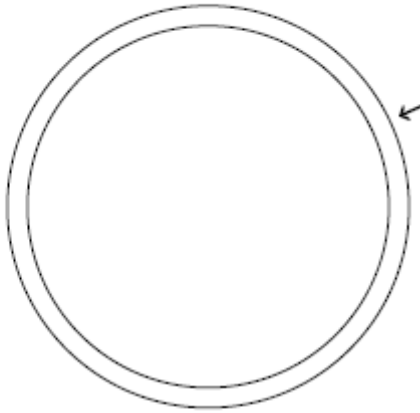


Volume of unit radius sphere goes to zero.

Data generated by Gaussian



data distribution
in 1-dim



data distribution
in high dimensions

Chat GPT

- Content of internet too large to store.
- Too large even if compressed losslessly.
- Use lossy compression.
  - Irrelevant material lost.
  - Essence of documents remains.

Chat GPT

- Learns structure of novels, research papers, obituaries, poems, recipes, and computer code.
- Can construct artificial novel or other item on subject you supply.
- Details may not be correct.

Chat GPT

- If internet content compressed in 2022, Chat GPT will not know about events in 2023.
- Need to continuously update compressed material efficiently.
- The area has many research topics.

Ending

This completes my remarks. I am happy to answer any questions you may have.