The Science of Information Meets the Liberal Arts

Sanjeev Kulkarni
Professor, Department of Electrical Engineering
Director, Keller Center
Center for Science of Information
kulkarni@princeton.edu

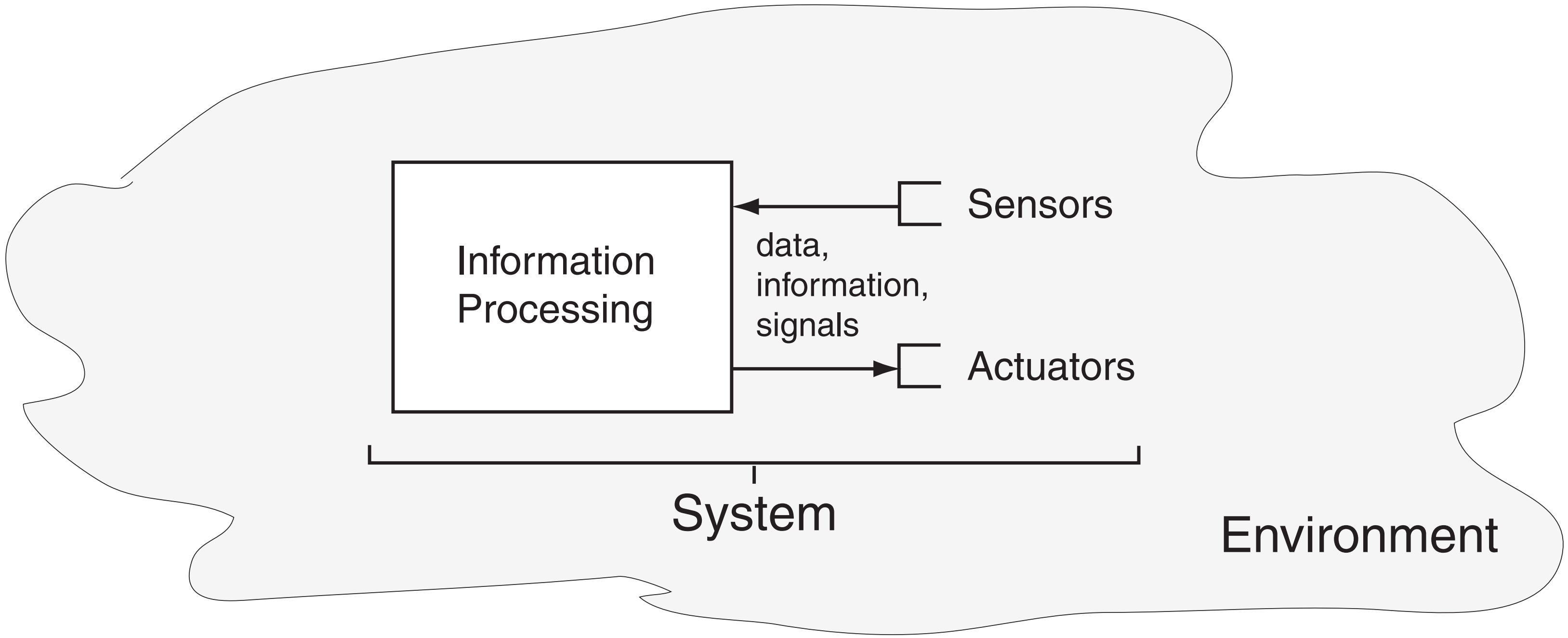
October 26, 2012

The Science of Information Meets the Liberal Arts

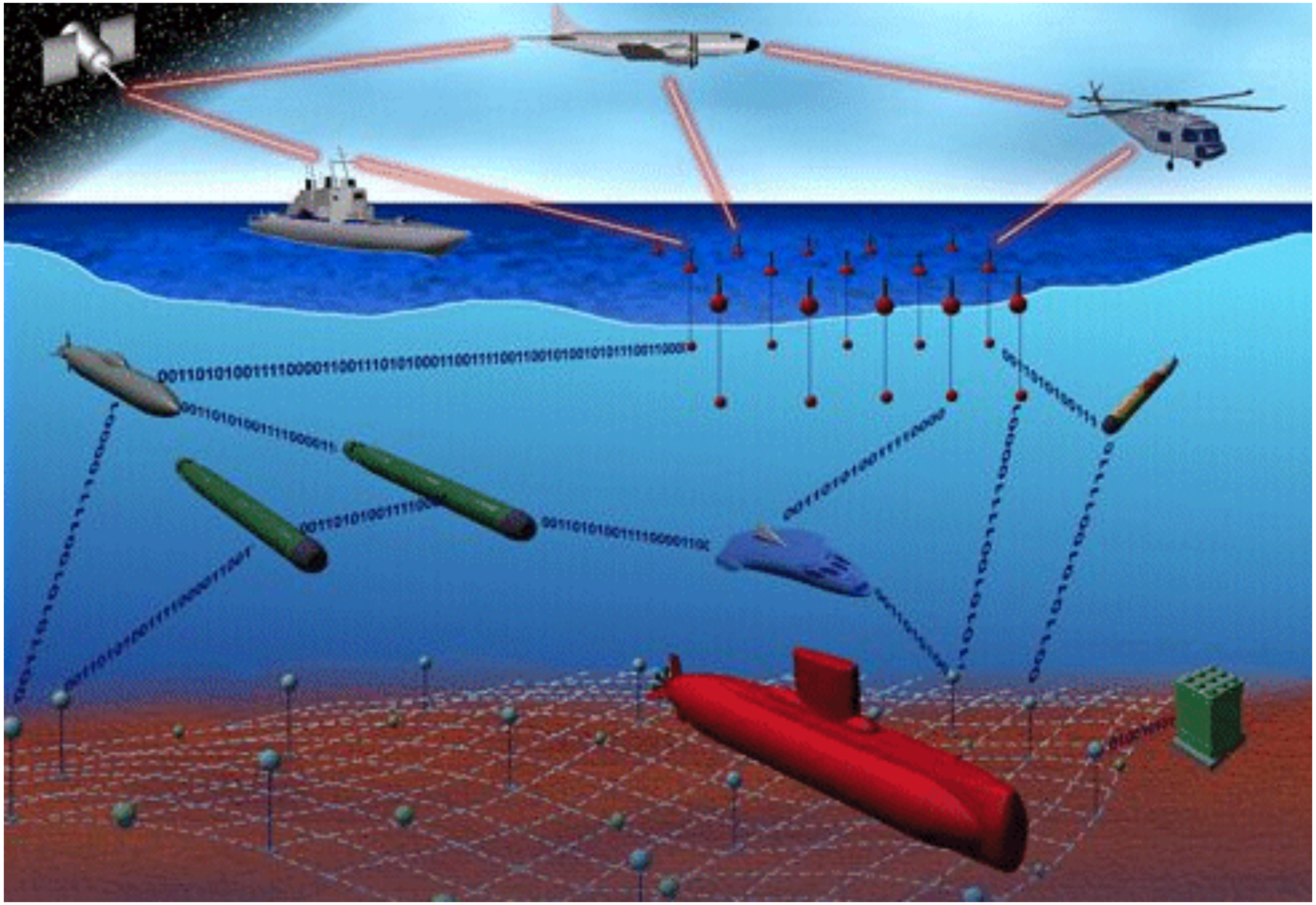
- A Broad View of the Science of Information
- Three Courses
    - Making Technical Material Accessible
      Introduction to Electrical Signals & Systems
    - Two Inherently Blended Fields
      Learning Theory and Epistemology
    - Technology in its Societal Context
      The Wireless Revolution
- Certificate in Information Technology and Society

A Broad View of the Science of Information

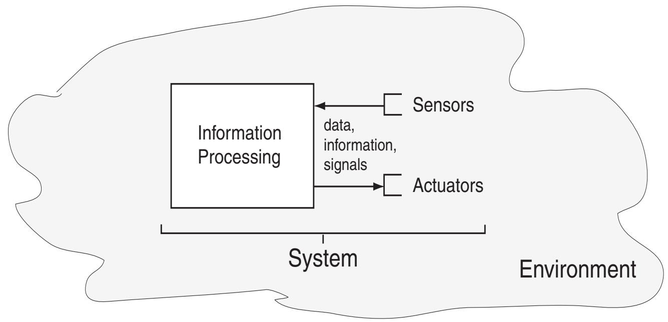
Information-based Paradigm for Designing Systems

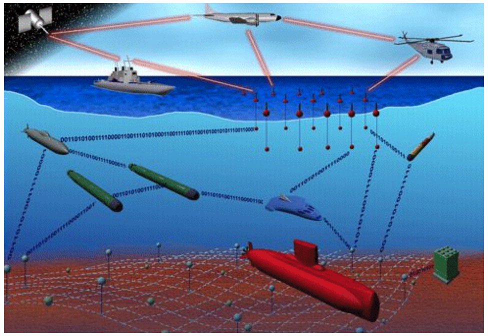


System may be Complex, Dynamic, Distributed



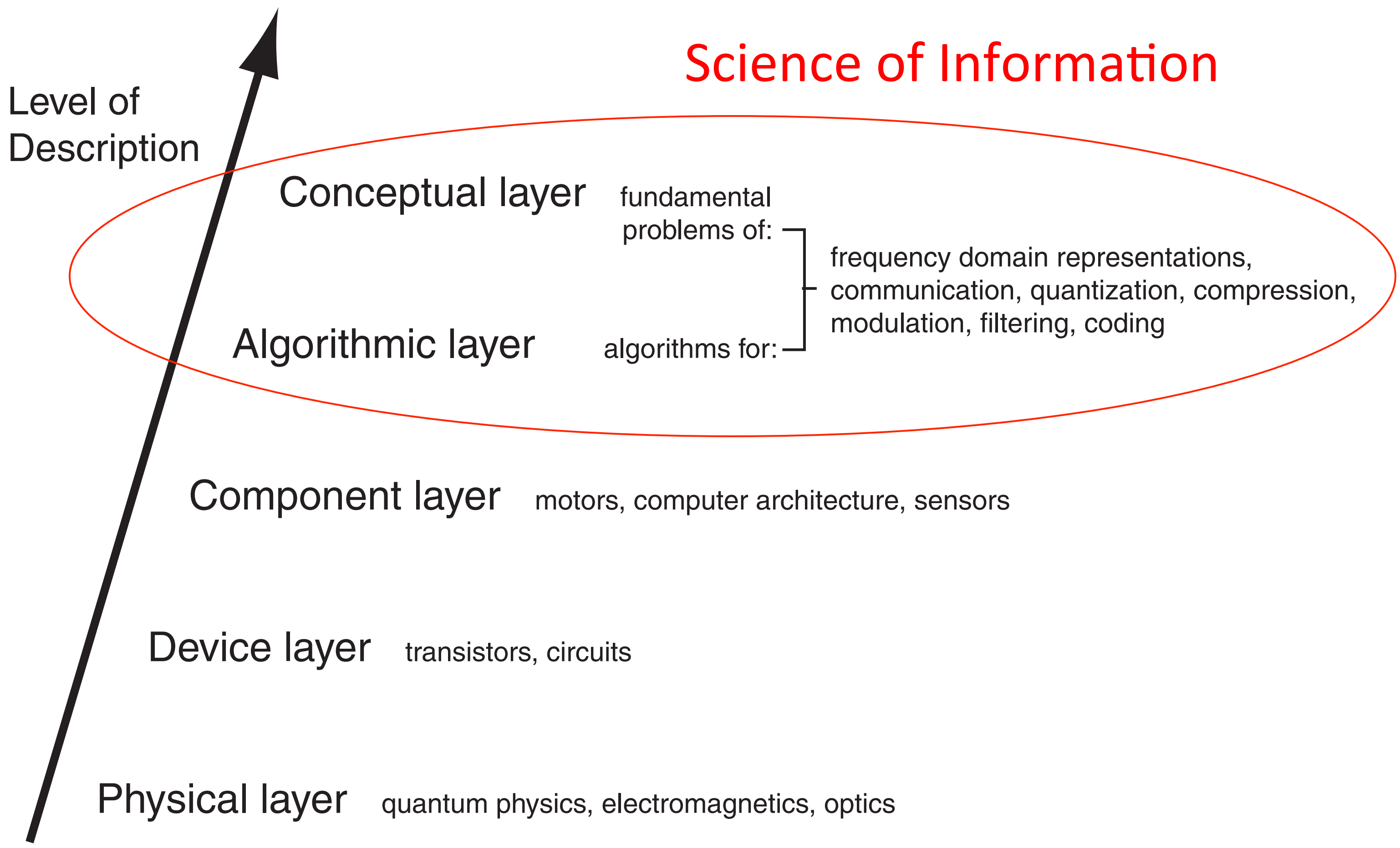
Some Information-Processing Tasks





- Sensing
- Sampling, Quantization (& D/A)
- Filtering
- Storage and Representation
- Search and retrieval
- Compression
- General purpose computation
- Communication, Data Transmission
- Error Detection/Correcton
- Cryptography
- Digital Rights Management
- Learning and Inference
- Control
- Actuation

Breakdown by Level of Description



Where and Why to Meet the Liberal Arts

- Science of information overlaps with liberal arts in many areas:
  - mathematics, statistics, psychology, philosophy, economics, politics, public policy, physics, biology, linguistics, etc.

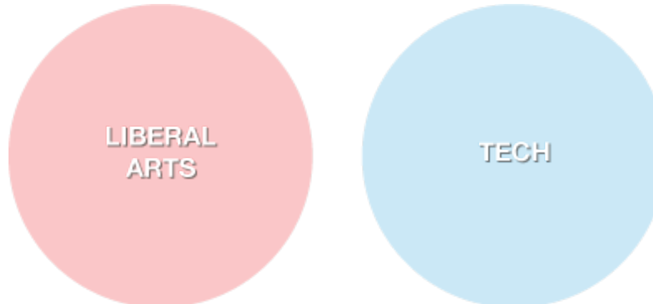
- All of our students use and are affected by information technology, and many will work in fields related to technology.

- A liberal arts education in the 21st century should include some basic understanding of technology (including information technology)

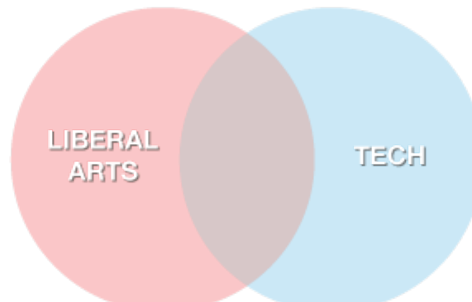
- It's all around us and it's interesting!

Liberal Arts Then/Now and How to Meet

Then



Now



How to teach at the intersection?

- Make technical subject matter accessible
- Teach material that inherently blends two or more fields
- Address technology in its broader societal context

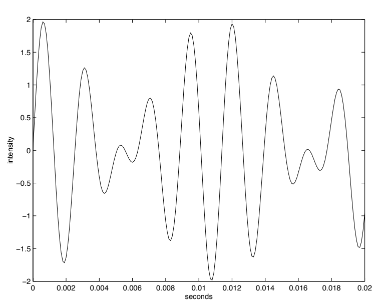
Making technical material accessible

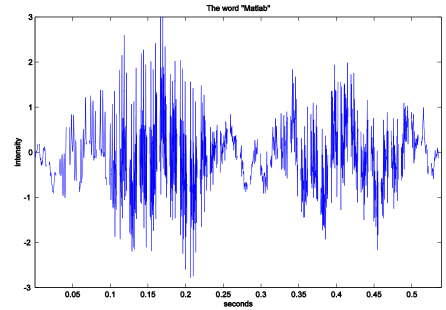
ELE 201 Introduction to Electrical Signals and Systems

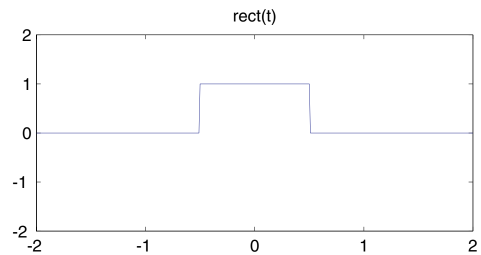
ELE 201 Introduction to Electrical Signals and Systems
Making technical material accessible

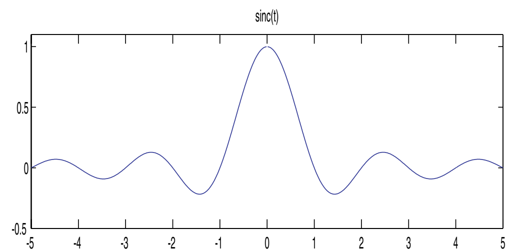
- Required core sophomore-level Electrical Engineering course.
- Open to all students with Calculus as only prerequisite. Also opened to qualified high-school students.
- Cover basics of signals, systems, and information theory.
- Has a lab component using Matlab working with audio and images.
- Now more than half of class is non-EE's, including many AB's and many undecided freshman.

Signals, Systems, Frequency Domain

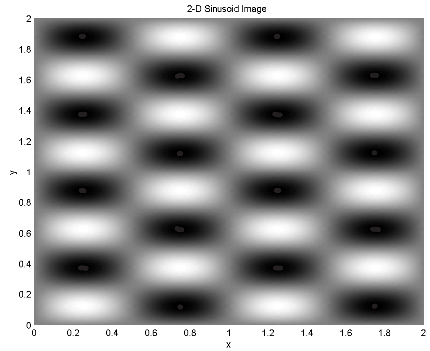


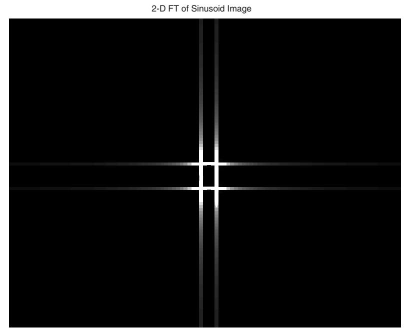






- What are signals?
- What are systems?
- Linear time-invariant systems
- δ-function, impulse response
- Convolution
- Frequency response
- Fourier transforms

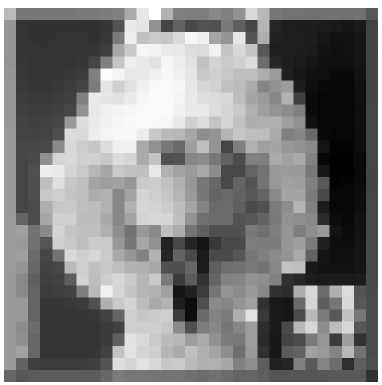




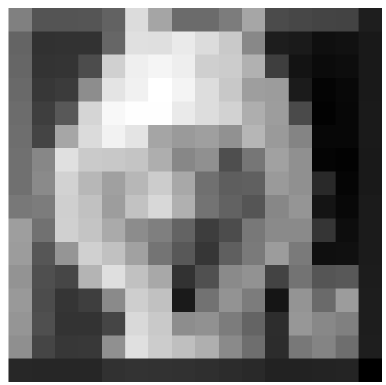
Sampling

No subsampling

4 x 4 blocks



8 x 8 blocks



16 x 16 blocks

- Bandlimited signals
- Sampling theorem
- Explaining sampling rate for digital audio and video

Quantization and Halftoning

256 levels

32 levels

16 levels

8 levels

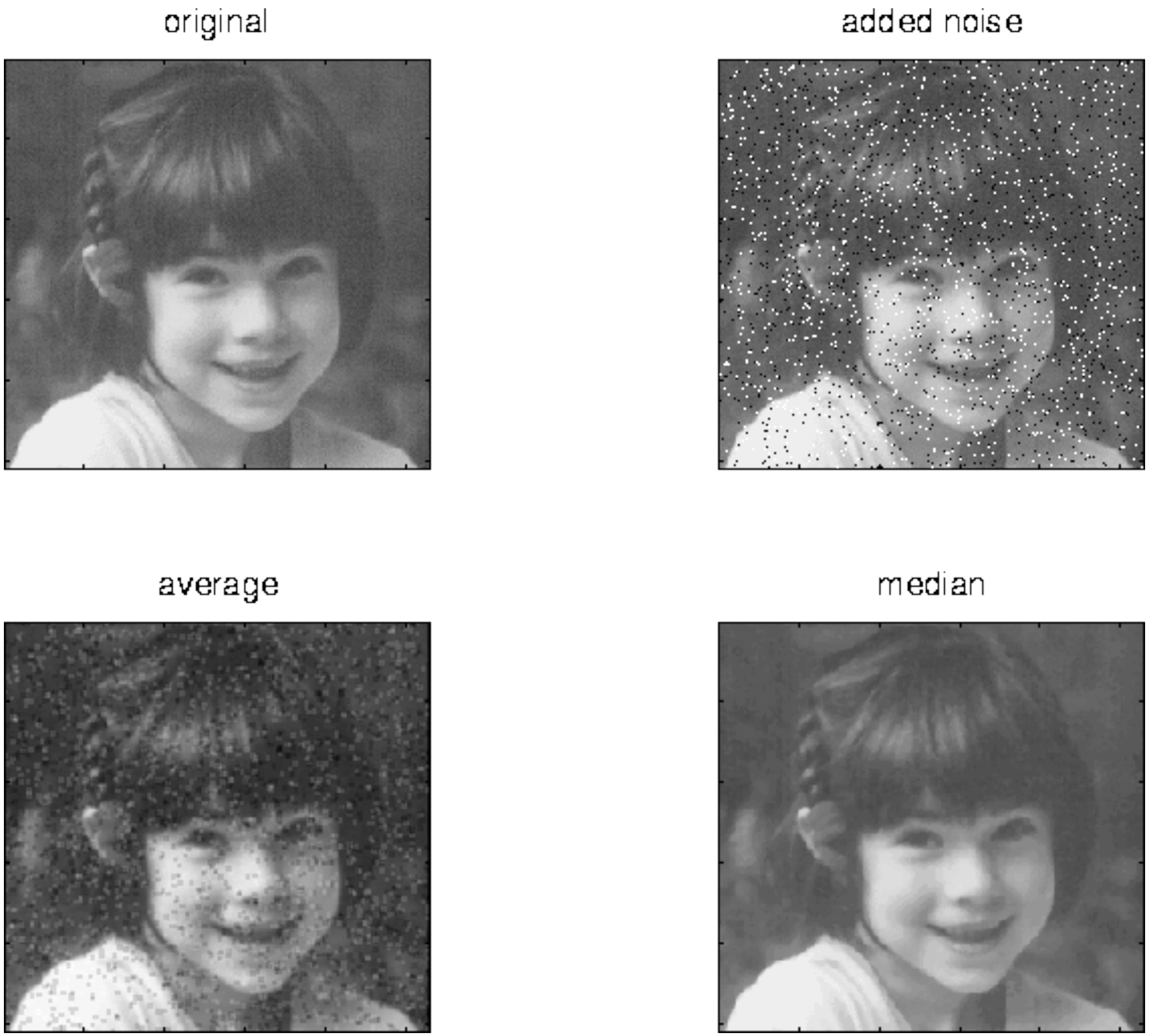
4 levels

2 levels

Original

Halftoned

Filtering



Data Compression

- Need for compression
  - Text: (1000 pages)(50 lines/page)(100 characters/line) = 5 MB
  - Audio: (44100 samples/sec)(16 bits/sample) = 88 kB/sec
  - Image: (512x512 pixels)(1 B/pixel) = 0.26 MB
  - Video: 30 frames/sec gives 7.86 MB/sec

- Storage and transmission both need compression.

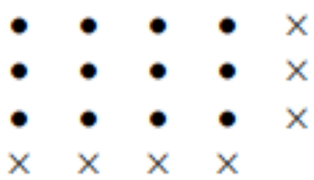
- Ability to compress based on exploiting redundancy. Fundamental limit based on inherent randomness (entropy).

- The more we know about the source, the better we can compress.

- Huffman coding, universal methods (zip), methods for specific types of data (JPEG, MPEG)

Error Detection and Correction

- Compression squeezes out redundancy
- To detect or correct errors, we add back *highly structured* redundancy
- Parity check bit for error detection: 0010110 → 0010110<u>1</u>
- More parity checks can allow correction:



1	0	1	1	1
0	1	0	1	0
1	0	0	0	1
0	1	1	0	

1	0	1	1	1	
0	1	1	1	0	F
1	0	0	0	1	
0	1	1	0		
		F			

- Also discuss better methods and fundamental limits

Some Comments

- Lab component (audio and images in Matlab) is popular.
- Blend of theory, hands-on, and real applications.
- Leads to substantive understanding.
- Demystifies technology.
- No exclusive domain for scientists/engineers.
- Biggest challenge is differing mathematical backgrounds.

Two inherently blended fields

ELE/PHI 218 Learning Theory and Epistemology

ELE/PHI 218 Learning Theory and Epistemology

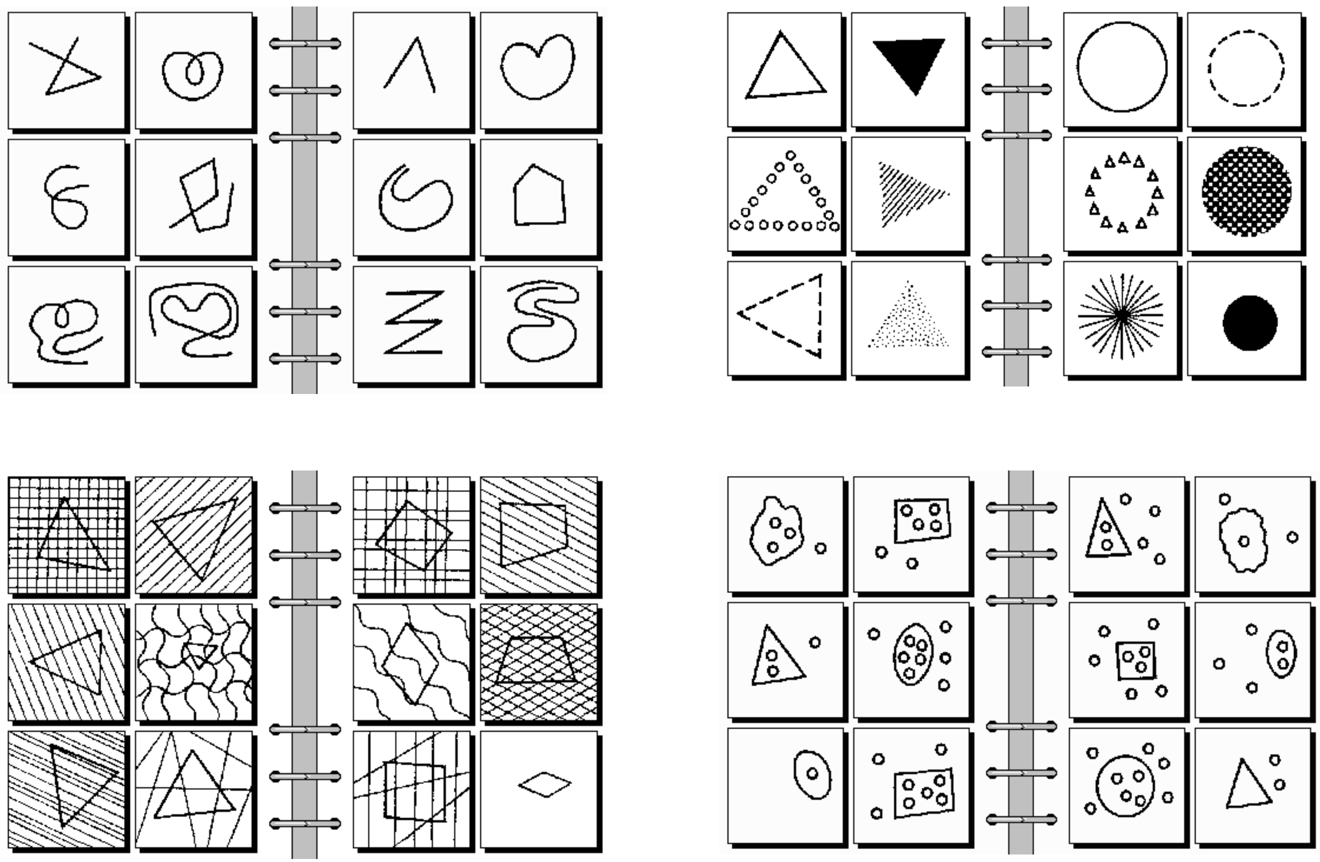
Two inherently blended fields

- Co-teach with Prof. Gil Harman in Philosophy.
- Students from wide range of departments and all levels (freshmen through seniors).
- Calculus is only prerequisite.

- Learning theory: Studies the fundamental limitations of learning (machine learning, pattern recognition). Are some learning/pattern recognition problems inherently hard? How can we design good algorithms?
- Epistemology: The branch of philosophy that deals with the nature and limitations of knowledge. What do we know and how do we know it?

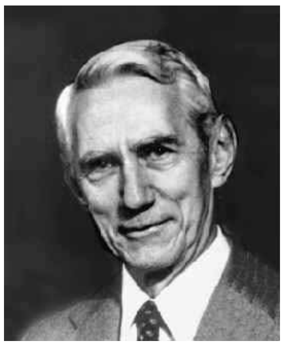
Pattern Recognition: Learning from Examples

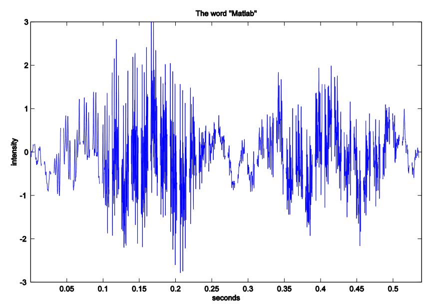
From M. Bongard, Pattern Recognition, 1970

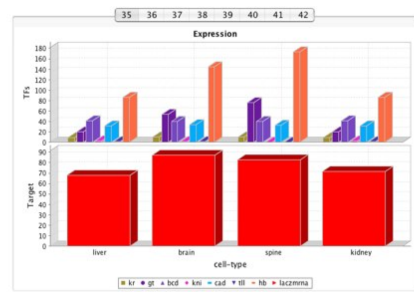


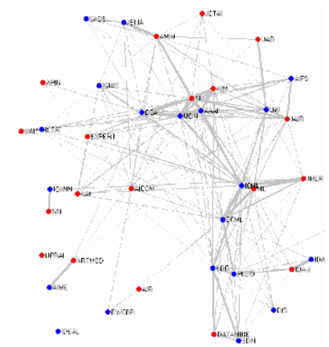
Machine Learning/Pattern Recognition

- Often don't know how to design good rules for classification or estimation.
- Learning can replace this knowledge, allow adaptation, and robustness to changing conditions.



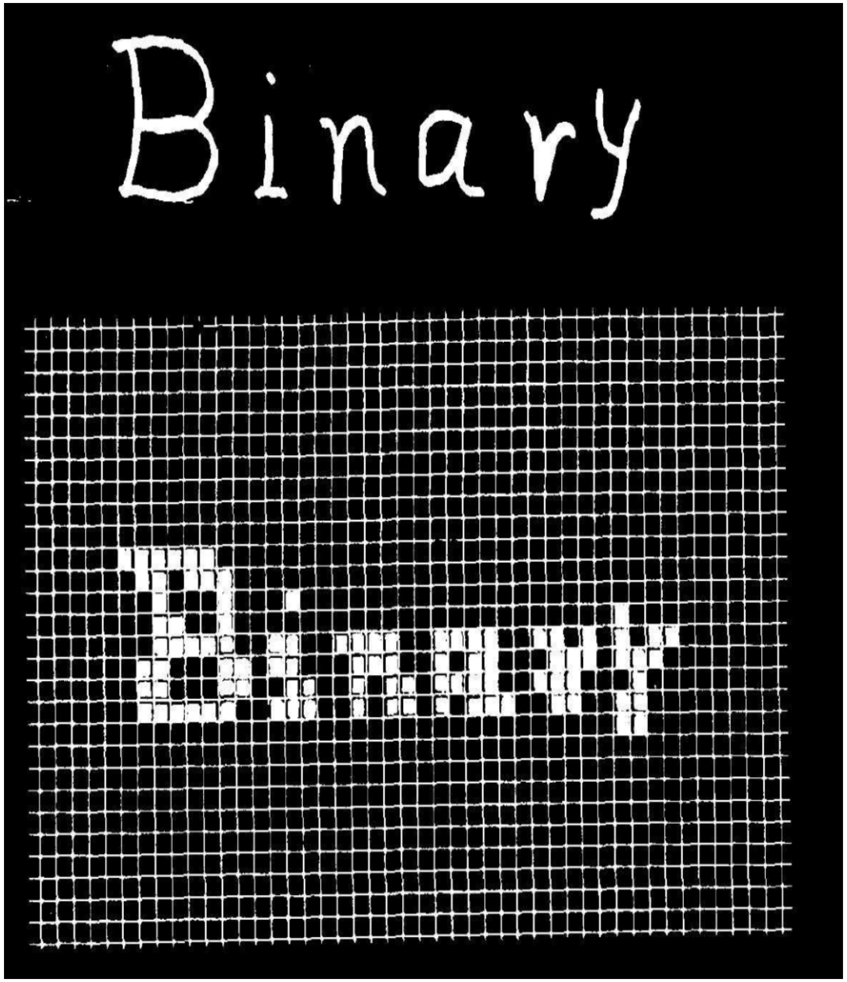






- Applications to recognition of images (faces, targets, etc.), speech, handwriting, medical diagnosis, spam, fraud, etc.
- Design effective algorithms
- Understand fundamental limits. What can be learned? What can't? Why?

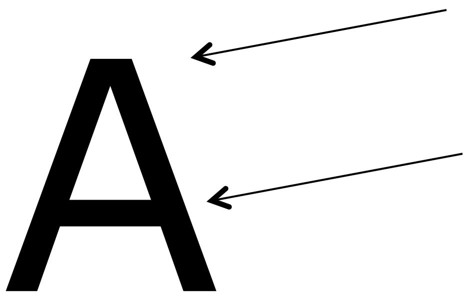
Example: Character Recognition



- Try to automatically recognize handwritten characters.
- Digitize characters to get a digital image.
- Segment into individual characters.
- Find features that distinguish each character.

Feature Extraction

- What are good features for recognizing characters?
- For example, what makes an “A” an “A”?



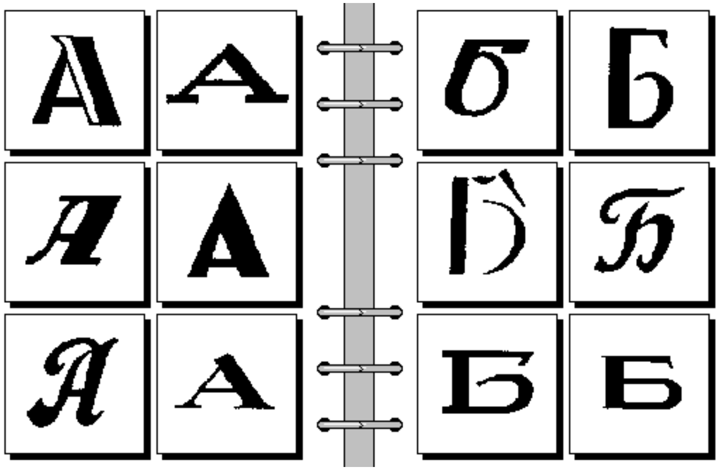
Angle at top

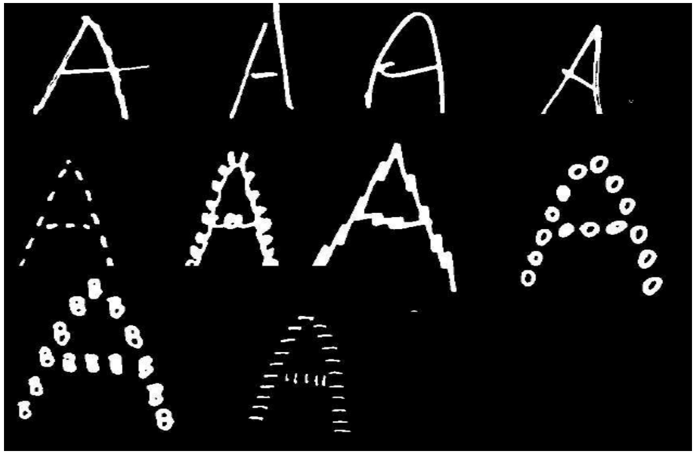
Horizontal line near the middle
that joins the two slanted lines

Anything else?

- Can we come up with good features for each letter and number?
- Even if can, how do we extract these features?

Problems With This Approach

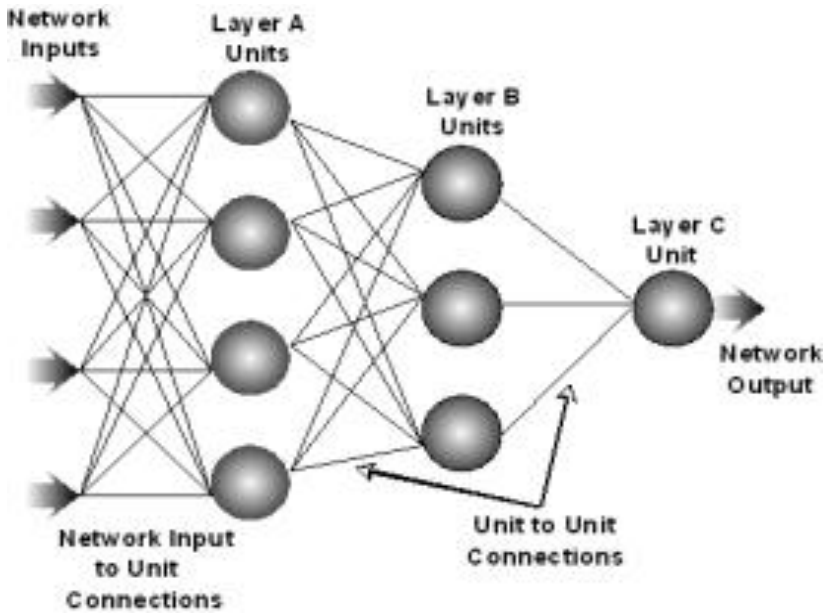




- Robust features are extremely difficult to identify and precisely define.
- And very difficult to extract.
- This is definitely not how humans learn!

A Different Approach

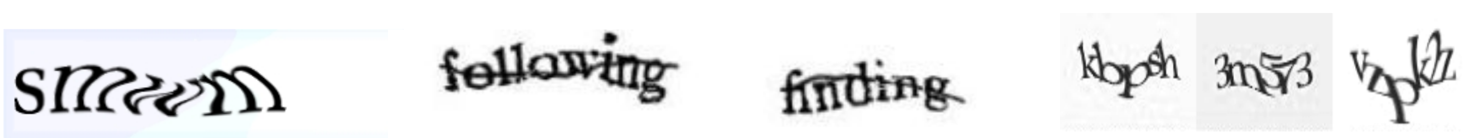
- Get lots of examples of A's, B's, etc.
- Use these training examples to come up with a rule.
- This is supervised learning.
- And this is closer to how humans learn.



- A number of learning techniques – e.g., neural networks, SVM's, boosting.
- Very successful in many applications.
- Still quite challenging:
  - Curse of dimensionality.
  - No Free Lunch theorems.
  - Understanding performance

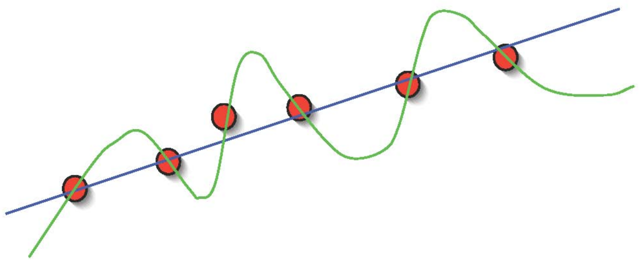
Exploiting the Limitations

- Captcha – a sort of “anti-Turing test”
- Tell humans and machines apart automatically
    - Prevent spam-bots from automatic email registration
    - Prevent vote-bots from disrupting on-line polls.



Connections to Many Other Fields and Many Fundamental Questions

- Mathematics, statistics, optimization.
- Neuroscience, cognitive science, psychology (brain, human learning, neural networks, etc.)
- Philosophy
  - The problem of induction
  - Role of simplicity, Occam's razor
  - Is the mind a computer?
  - Can a computer have a mind?
  - Can a computer be conscious? Be self-aware? Have intent? Feel?
  - If so, what are the ethical implications?



Some Comments

- Brings together a wide range of students from diverse backgrounds.
- Brings together two very different fields.
- Substantive in-class discussions
- Deep results from several fields
- Key ideas understandable
- Again, differing math backgrounds is biggest challenge

Technology in its societal context

ELE/EGR 391 The Wireless Revolution

ELE/EGR 391 The Wireless Revolution
Technology in its societal context

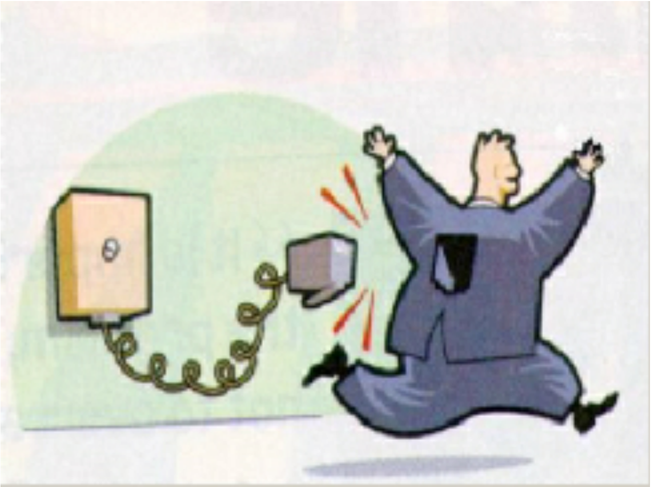
- Introduced by Prof. Vince Poor in 2001.
- No prerequisites. Not open to freshmen.
- Students from wide range of departments and sophomores through seniors.
- Open to engineers, but doesn't satisfy departmental requirement.
- Considering closing to Electrical Engineers

- First half: Cover basics of wireless technology.
- Second half: Guest lectures from academia, industry, government.

What is Wireless? Tetherless (Freedom)

- Wireless means communication by radio.
- Wireless typically implements only the last link between an end device (telephone, computer, etc.) and an access point to a network.
- Wireless usually involves significant wireline infrastructure (the "backbone").
- Wireless affords
  - mobility
  - portability
  - ease of connectivity

i.e., freedom.

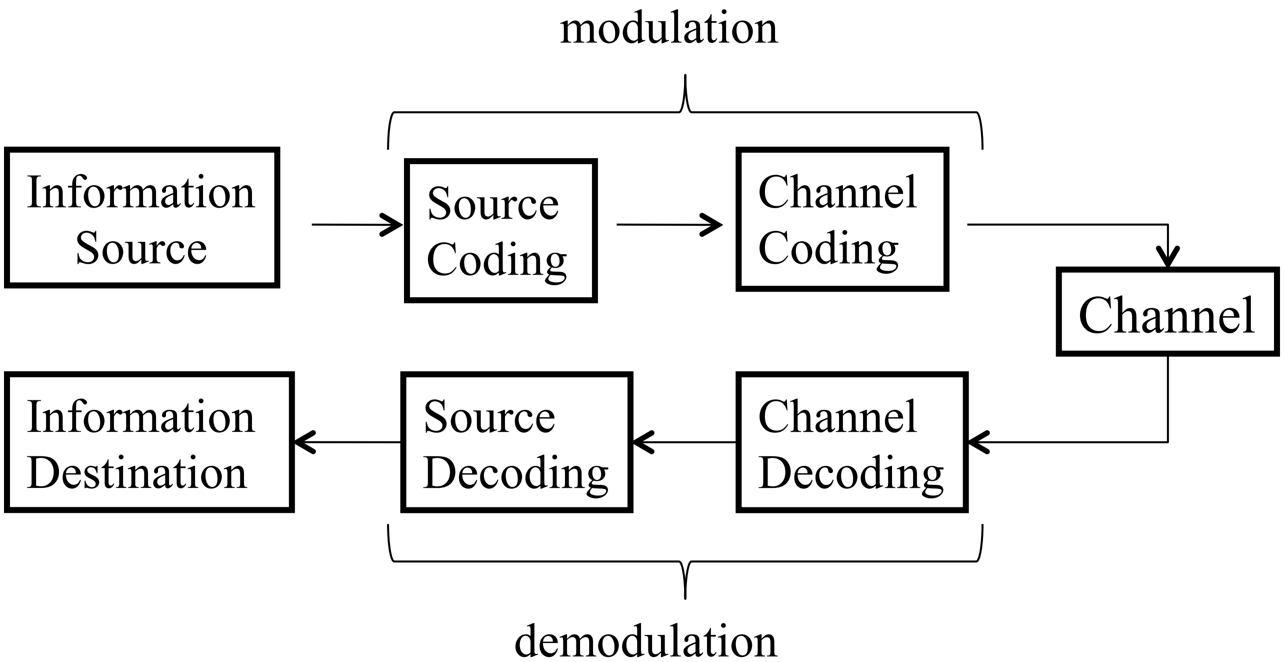


Wireless Challenges

Main Challenge: To provide the services of wireline systems, but with mobility.

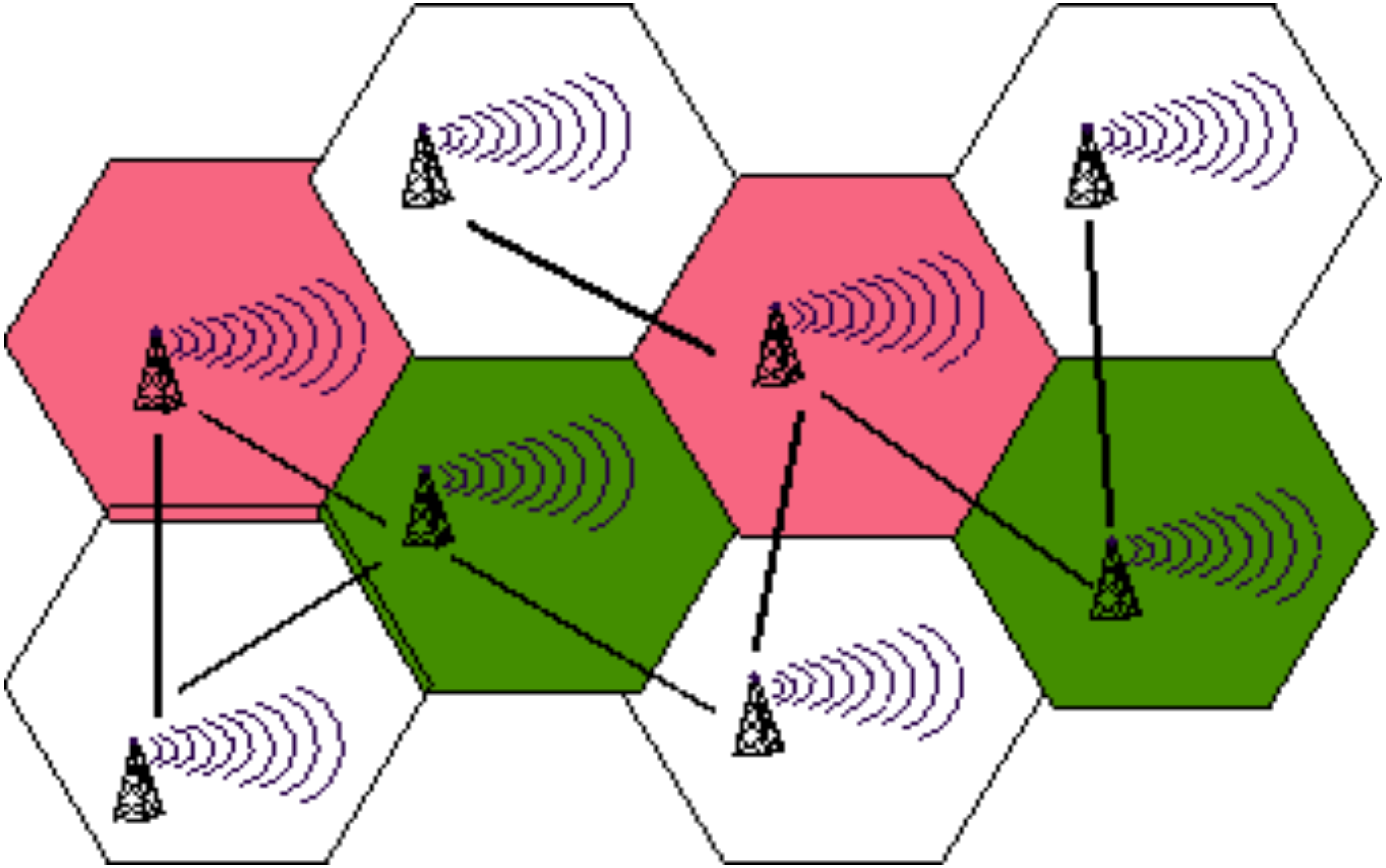
- High data rate (multimedia traffic)/greater capacity
- Networking (seamless connectivity)
- Resource allocation (quality of service - QoS)
- Manifold physical impairments
- Mobility (rapidly changing physical channel)
- Portability (battery life)
- Privacy/security (encryption)
- Global standardization (politics & $$$)

Point-to-Point Communication Model

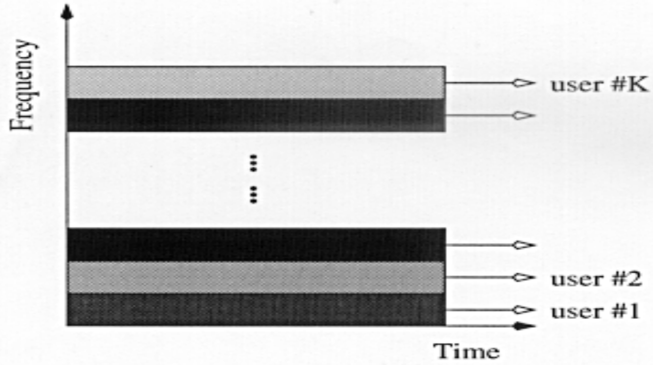


- Key ideas with multiple users
  - Cellular concept
  - Multiaccess techniques
  - Networks and protocols

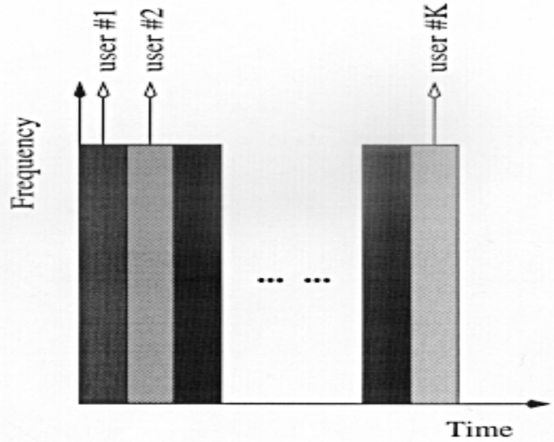
Cellular Telephony



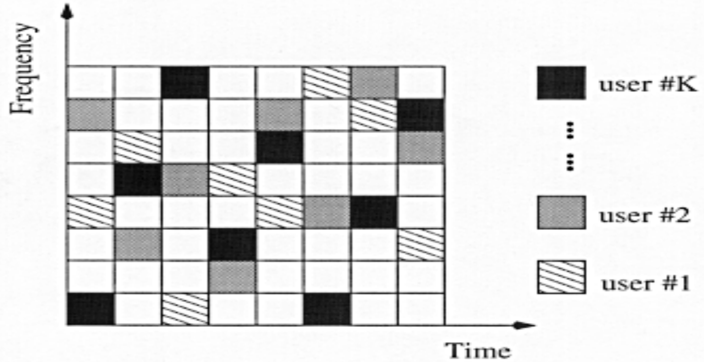
xDMA Summary



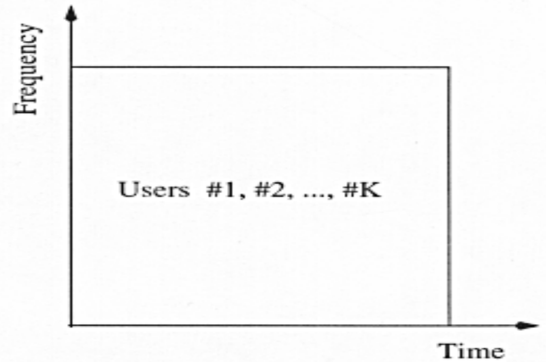
Frequency-Division Multiple-Access (FDMA)



Time-Division Multiple-Access (TDMA)



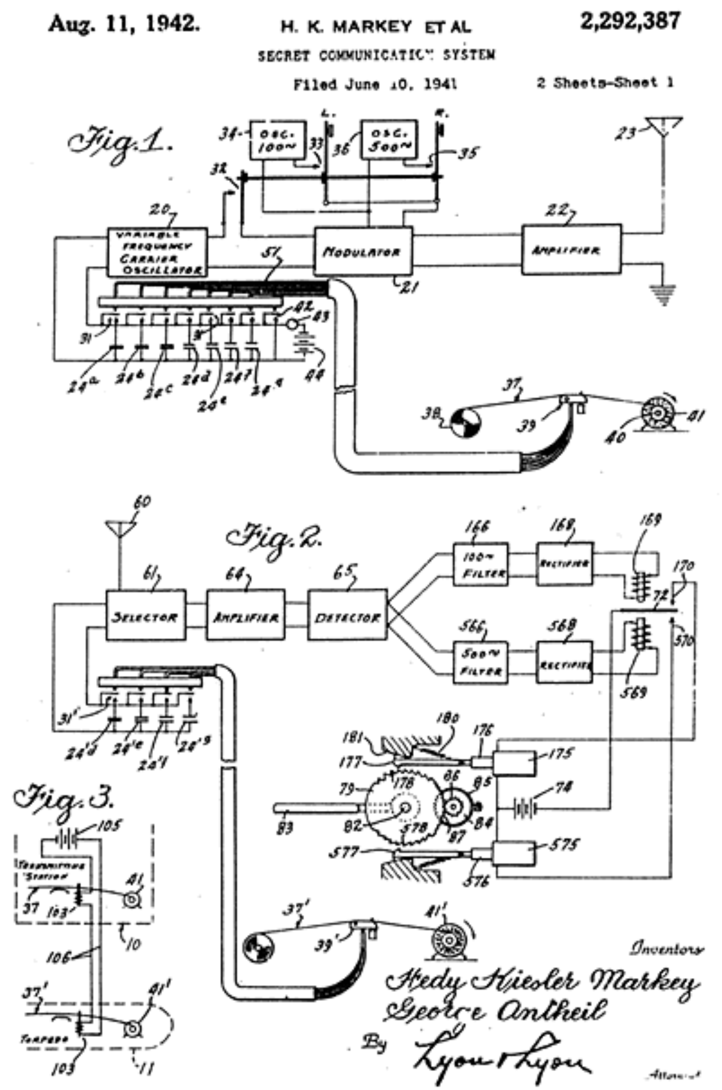
Frequency-Hopping Code-Division Multiple-Access (FH-CDMA)



Direct-Sequence Code-Divsion Multiple-Access (DS-CDMA)

Hedy Lamarr

Photo from *The Economist,* Jan. 25, 2000.



- Co-inventor of FH spread-spectrum.
- Invented in the context of torpedo guidance.

Packet Switching vs Circuit Switching

- In large data networks (e.g., the Internet), packets are switched through the network from source to destination by routers at the "nodes" of the network.

- This works like the postal system, where
  - the packets are like letters
  - the links are like postal routes and transportation routes between major cities
  - the nodes are like post offices
  - the end devices are like mailboxes

- Avoids need for end-to-end link.

Part II: Guest Lectures on Business, Regulatory, Social Issues, etc.

- Commercial enterprises/entrepreneurship.
- Wireless standards.
- Investment banking perspectives.
- Impact of regulatory policies/ role of the FCC in USA wireless development.
- Valuation and auctioning of the radio spectrum.
- Applications (e.g., environmental monitoring).
- Security and privacy in wireless networks.
- Social issues in wireless.
- Emerging techniques and the future of wireless.

For some, wireless is easy…

The wireless telegraph is not difficult to understand. The ordinary telegraph is like a very long cat. You pull the tail in New York, and it meows in Los Angeles.

\- *Albert Einstein*

For some, wireless is easy…

The wireless telegraph is not difficult to understand. The ordinary telegraph is like a very long cat. You pull the tail in New York, and it meows in Los Angeles. The wireless telegraph is the same, only without the cat.

\- *Albert Einstein*

Some Observations

- P/D/F-only levels playing field and promotes exploration outside of comfort zone.
- Understand revolutionary advance.
- Guest lectures are a big hit.
  - Appreciate broader impacts of technology.
  - Meet leaders in variety of areas.
  - Engage alumni.
- Biggest challenge is lining up compelling guest speakers. Differing math backgrounds addressed by P/D/F.

A Program of Study

Certificate in Information Technology and Society

Certificate in Information Technology and Society
Program of Study

- Jointly sponsored by Keller Center and Center for Information Technology Policy

Requirements

- Core course: EGR/HIS/SOC 277 – Technology and Society
- Two technology courses
- Two societal courses
- A breadth course
- Independent work
- Presentation at annual symposium

Technology Courses

- COS 109/EGR 109 – Computers in Our World
- COS 126 – General Computer Science
- COS 432 – Information Security
- COS 445 – Networks, Economics and Computing
- COS 455/MOL 455 – Intro to Genomics and Computational MolBio
- COS 597D – Advanced Topics in CS – Info. Privacy Technologies
- ELE 201 – Introduction to Signals and Systems
- ELE 222a/b/EGR 222a/b – The Computing Age
- ELE 381/COS 381 – Networks: Friends, Money, and Bytes
- ELE 386/EGR 386 – Cyber Security
- ELE 391/EGR 391 – The Wireless Revolution
- FRS 125 Friending, Following and Finding
- ORF 401 – Electronic Commerce
- ORF 411 – Operations and Information Engineering

Societal Courses

- COS 448* – Innovating Across Technology, Business, & Markets
- COS 495/ART 495 – Modeling the Past – Tech & Excav. in Polis, Cyprus
- COS 586/WWS 586F* – Information Technology and Public Policy
- FRS 101* – Facebook: The Social Impact of Social Networks
- FRS 163 – Technology and Policy
- PSY 214 – Human Identity in the Age of Neurosci. and Info. Technology
- PSY 322/ORF 322 – Human Machine Interaction
- SOC 204 – Social Networks
- SOC 214 – Creativity, Innovation, and Society
- SOC 344 – Communications, Culture, and Society
- SOC 357* – Sociology of Technology
- SOC 409*/COS 409 – Critical Approaches to Human Comp. Interaction
- WWS 334 – Media and Public Policy (formerly WWS 309)
- WWS 351/SOC 353/COS 351 – Info. Technology and Public Policy
- WWS 571B/NES 584 – New Media & Social Movements

Breadth Course

- CBE 260/EGR 260 – Ethics and Technology: Eng. in the Real World
- CEE 102a/b/EGR 102a/b – Engineering in the Modern World
- ENV 360* – Biotech Plants and Animals
- MAE 228/EGR 228/CBE 228 – Energy Solutions for the Next Century
- MAE 244*/EGR 244 – Intro to Biomedical Innovation and Global Health
- MAE 445/EGR 445 – Entrepreneurial Engineering
- MOL 205 – Genes, Health, and Society

- EGR 491/ELE 491 – High-Tech Entrepreneurship
- EGR 492* – Radical Innovation in Global Markets
- EGR 495 – Special Topics in Entrepreneurship – The Lean LaunchPad
- HIS 292 – Science in the Modern World
- HIS 398 – Technologies and Their Societies: Historical Perspectives
- NES 266*/ENV 266 – Oil, Energy and The Middle East
- WWS 315 – Bioethics and Public Policy

Projects and Student Presentations

- “TUBE (Time dependent Usage based Broadband price Engineering)”
- “Adolescents and Online Bullying”
- “Contested Control: European Data Privacy Regulations and the Assertion of Jurisdiction over American Businesses”
- “Evading Government Censorship; the Labor Movement's Use of the Internet”

Summary

- Science of Information is extremely broad
- Is embedded throughout our world
- Some understanding of technology should be part of a liberal education
- Many ways to teach at the interface
- Science of Information is a particularly rich area for bringing together engineering, sciences, social sciences, and humanities

Thank You!