Information Theory

Master of Logic 2015/16

2nd Block Nov/Dec 2015

first class: Wednesday, 28 October, 2015 9:00

http://homepages.cwi.nl/~schaffne/courses/inftheory/2015/

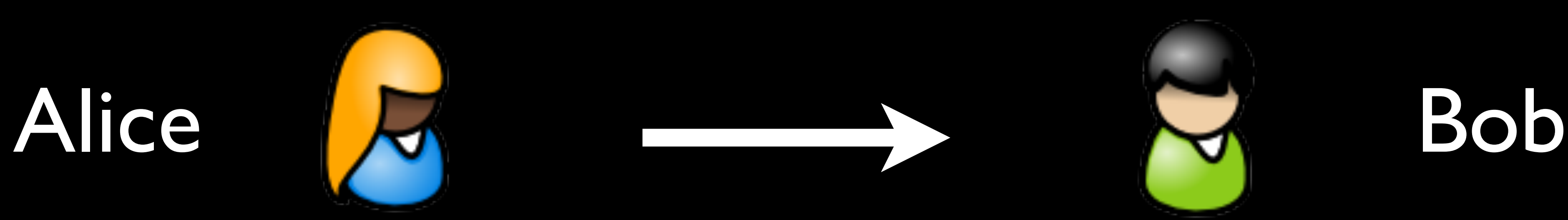
What is communication?

What is communication?

"The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point." - C.E. Shannon, 1948

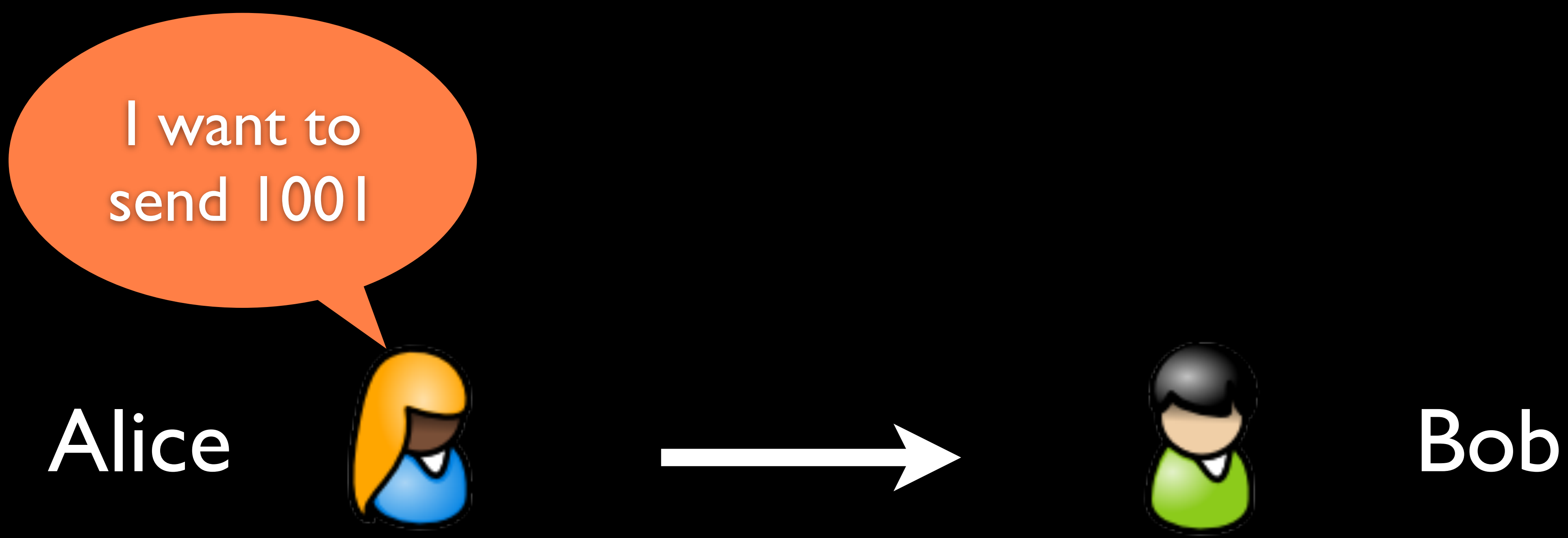
What is communication?

"The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point." - C.E. Shannon, 1948



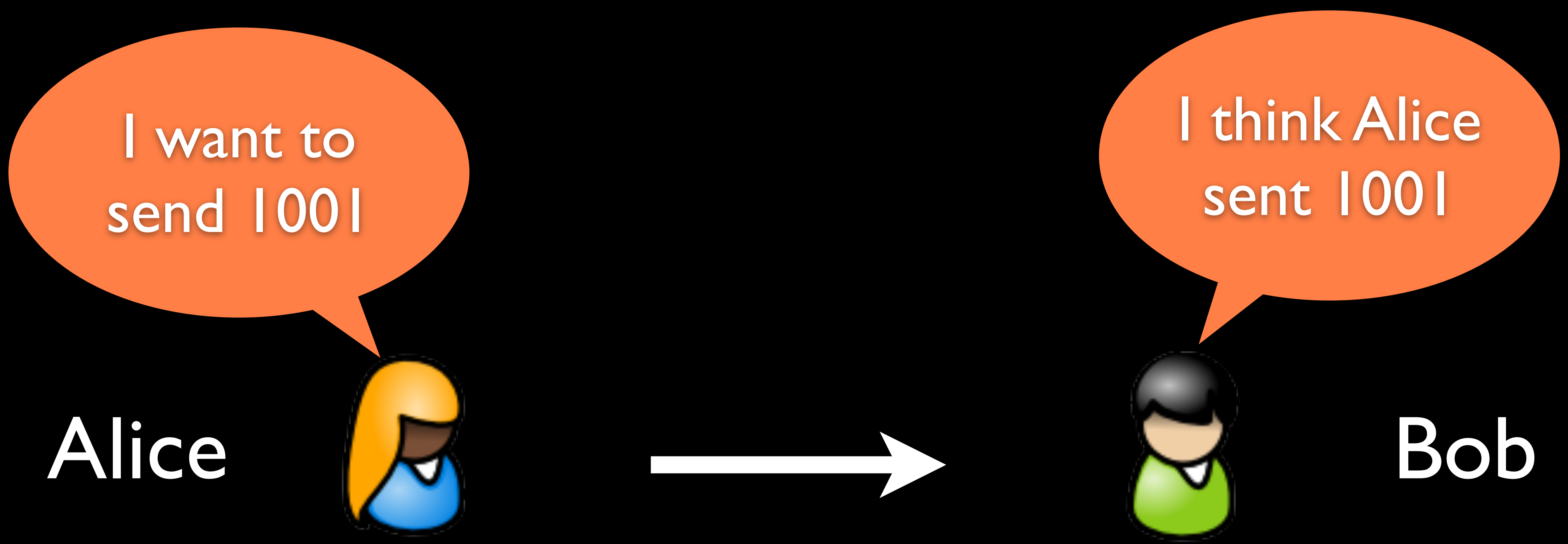
What is communication?

"The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point." - C.E. Shannon, 1948

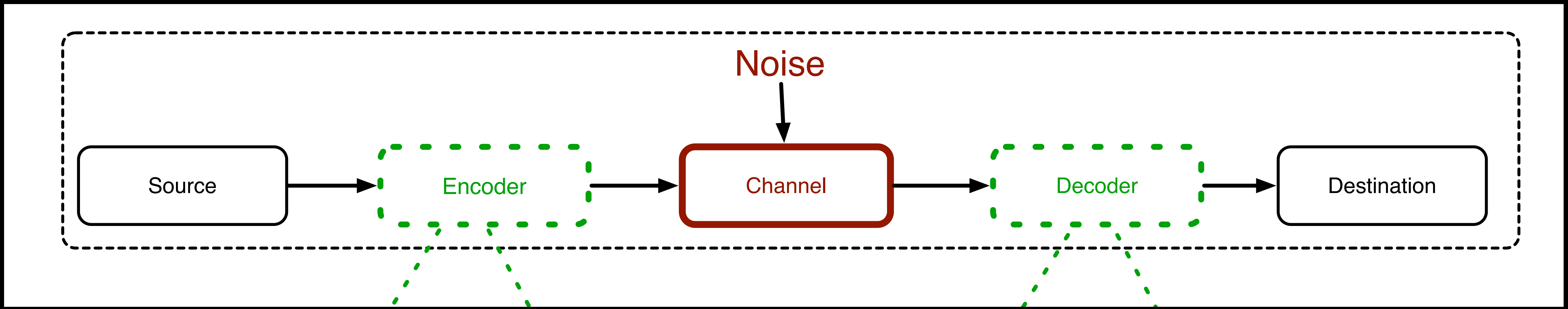


What is communication?

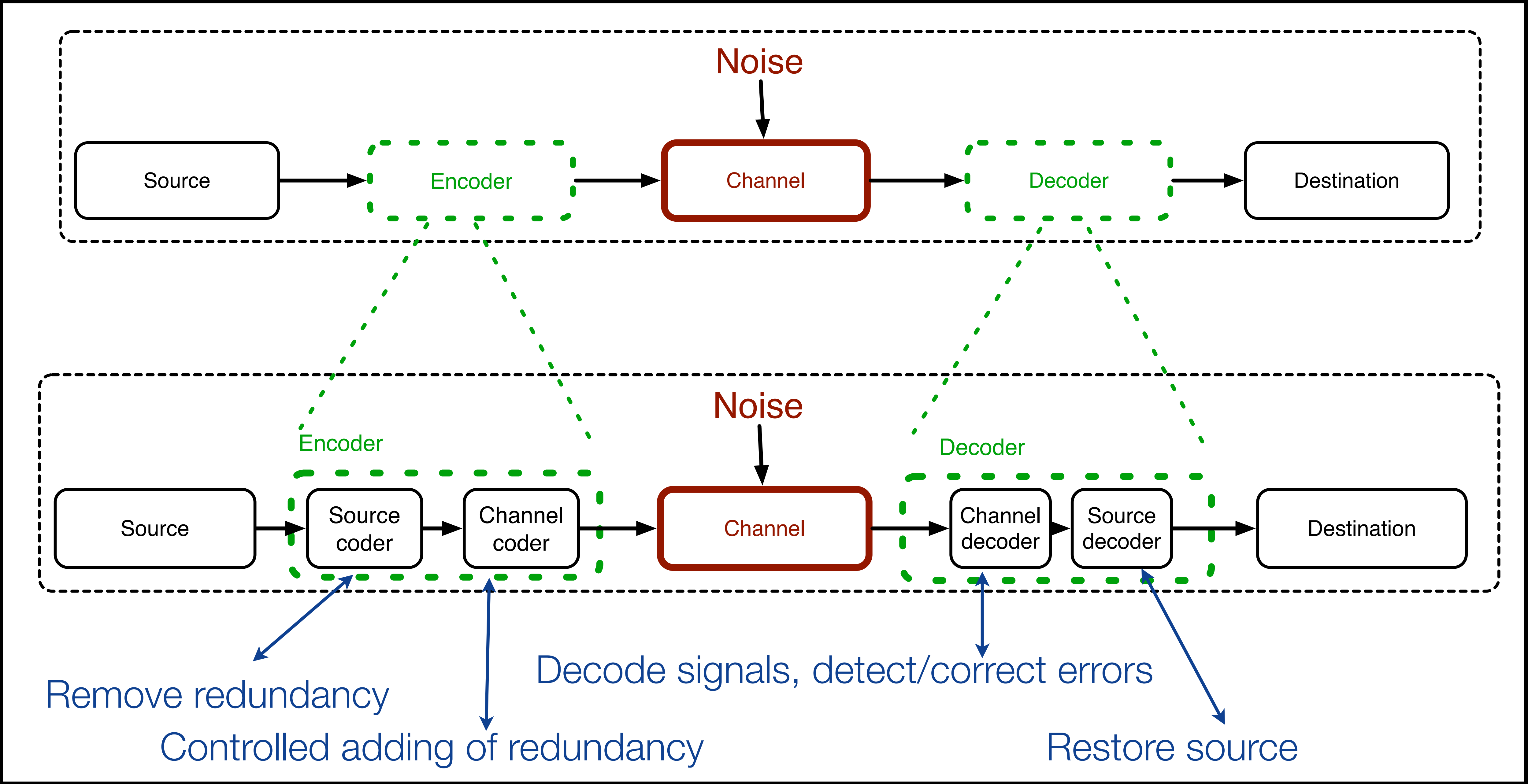
"The fundamental problem of communication is that of reproducing at one point either exactly or approximately a message selected at another point." - C.E. Shannon, 1948



Generic communication block diagram



Generic communication block diagram

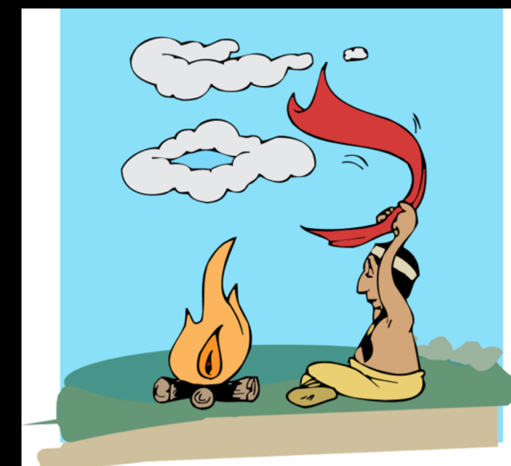


History of (wireless) communication

- Smoke signals

History of (wireless) communication

- Smoke signals
- 1861: Maxwell's equations

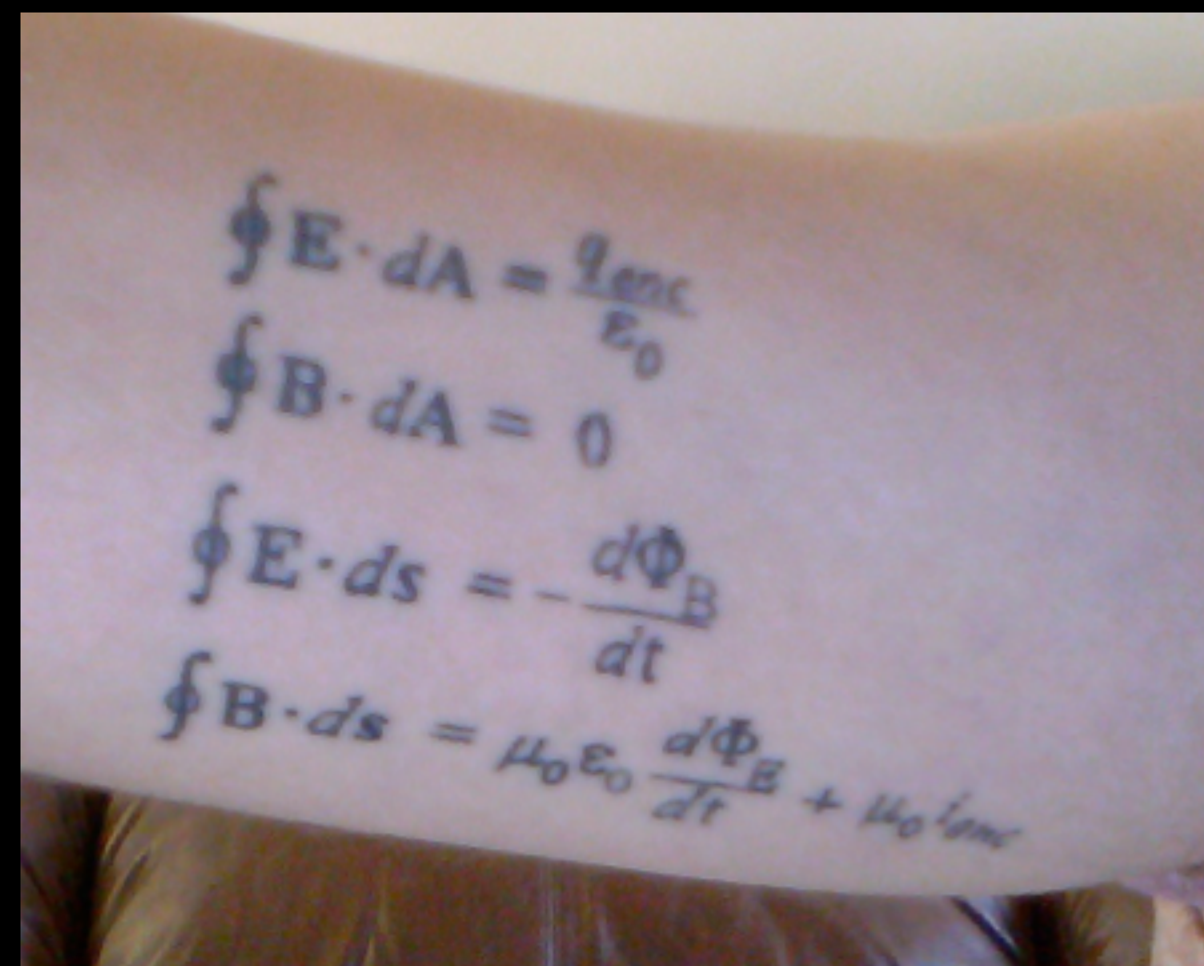


$$\oint \mathbf{E} \cdot d\mathbf{A} = \frac{q_{enc}}{\varepsilon_0}$$

$$\oint \mathbf{B} \cdot d\mathbf{A} = 0$$

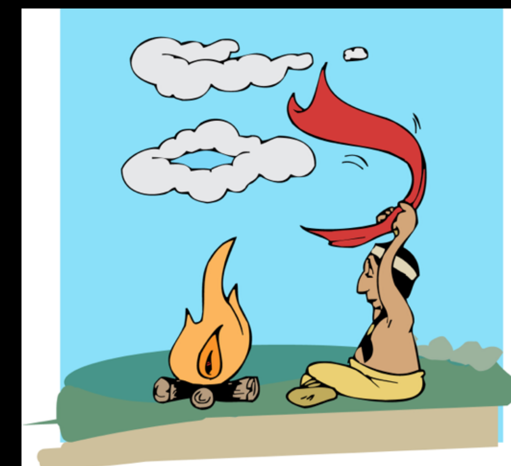
$$\oint \mathbf{E} \cdot d\mathbf{s} = -\frac{d\Phi_B}{dt}$$

$$\oint \mathbf{B} \cdot d\mathbf{s} = \mu_0 \varepsilon_0 \frac{d\Phi_E}{dt} + \mu_0 i_{enc}$$



History of (wireless) communication

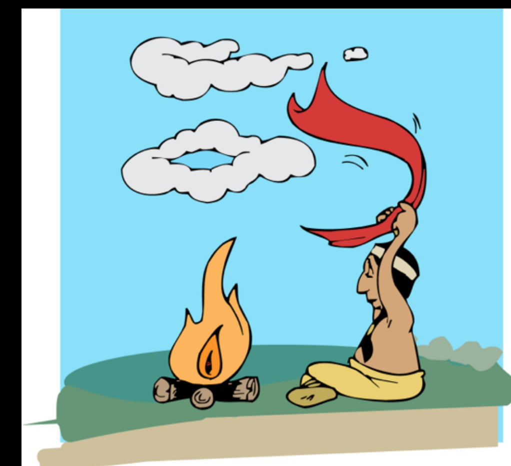
- Smoke signals
- 1861: Maxwell's equations
- 1900: Guglielmo Marconi demonstrates wireless telegraph



$$\oint \mathbf{E} \cdot d\mathbf{A} = \frac{q_{enc}}{\varepsilon_0}$$
$$\oint \mathbf{B} \cdot d\mathbf{A} = 0$$
$$\oint \mathbf{E} \cdot d\mathbf{s} = -\frac{d\Phi_B}{dt}$$
$$\oint \mathbf{B} \cdot d\mathbf{s} = \mu_0 \varepsilon_0 \frac{d\Phi_E}{dt} + \mu_0 i_{enc}$$

History of (wireless) communication

- Smoke signals
- 1861: Maxwell's equations
- 1900: Marconi demonstrates wireless telegraph
- 1920s: Edwin Howard Armstrong demonstrates FM radio



$$\oint \mathbf{E} \cdot d\mathbf{A} = \frac{q_{enc}}{\varepsilon_0}$$

$$\oint \mathbf{B} \cdot d\mathbf{A} = 0$$

$$\oint \mathbf{E} \cdot d\mathbf{s} = -\frac{d\Phi_B}{dt}$$

$$\oint \mathbf{B} \cdot d\mathbf{s} = \mu_0 \varepsilon_0 \frac{d\Phi_E}{dt} + \mu_0 i_{enc}$$

Big Open Questions

- mostly analog
- ad-hoc engineering, tailored to each application
- is there a general methodology for designing communication systems?
- can we communicate reliably in noise?
- how fast can we communicate?

Claude Elwood Shannon

1916 - 2001

- Father of Information Theory
- Graduate of MIT 1940:
  "An Algebra for Theoretical Genetics"
- 1941-1972: Scientist at Bell Labs
- 1958: Professor at MIT:

  When he returned to MIT in 1958, he continued to threaten corridor-walkers on his unicycle, sometimes augmenting the hazard by juggling. No one was ever sure whether these activities were part of some new breakthrough or whether he just found them amusing. He worked, for example, on a motorized pogo-stick, which he claimed would mean he could abandon the unicycle so feared by his colleagues ...
- juggling, unicycling, chess
- ultimate machine

History of (wireless) communication

- BITS !
- arguably, first to really define and use "bits"
- *"He's one of the great men of the century. Without him, none of the things we know today would exist. The whole digital revolution started with him."* -Neil Sloane, AT&T Fellow

The Bell System Technical Journal

Vol. XXVII *July, 1948* *No. 3*

A Mathematical Theory of Communication

By C. E. SHANNON

- Introduced a new field: Information Theory

What is communication?

What is information?

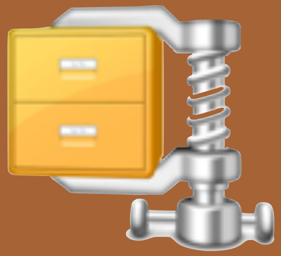
How much can we compress information?

How fast can we communicate?

Main Contributions of Inf Theory

Source coding

- source = random variable

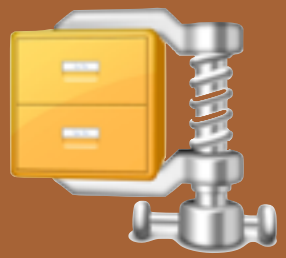


- ultimate data compression limit is the source's **entropy H**

Main Contributions of Inf Theory

Source coding

- source = random variable



- ultimate data compression limit is the source's **entropy H**

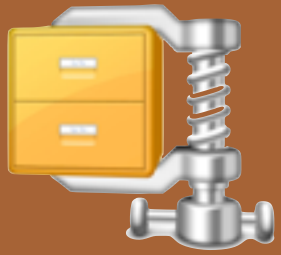
Channel coding

- channel = conditional distributions
- ultimate transmission rate is the **channel capacity C**

Main Contributions of Inf Theory

Source coding

- source = random variable



- ultimate data compression limit is the source's **entropy H**

Channel coding

- channel = conditional distributions
- ultimate transmission rate is the **channel capacity C**

Reliable communication possible ⇔ H < C

Reactions to This Theory

- Engineers in disbelief
- stuck in analogue world

How to approach the predicted limits?

Shannon says: can transmit at rates up to say 4Mbps over a certain channel without error. How to do it?

It Took 50 Years To Do It

How to approach the predicted limits?

review article by [Costello Forney 2006]

It Took 50 Years To Do It

- 50's: algebraic codes

How to approach the predicted limits?

review article by [Costello Forney 2006]

It Took 50 Years To Do It

- 50's: algebraic codes
- 60's 70's: convolutional codes

How to approach the predicted limits?

review article by [Costello Forney 2006]

It Took 50 Years To Do It

- 50‘s: algebraic codes
- 60‘s 70‘s: convolutional codes
- 80‘s: iterative codes (LDPC, turbo codes)

How to approach the predicted limits?

review article by [Costello Forney 2006]

It Took 50 Years To Do It

- 50‘s: algebraic codes
- 60‘s 70‘s: convolutional codes
- 80‘s: iterative codes (LDPC, turbo codes)
- 2009: polar codes

How to approach the predicted limits?

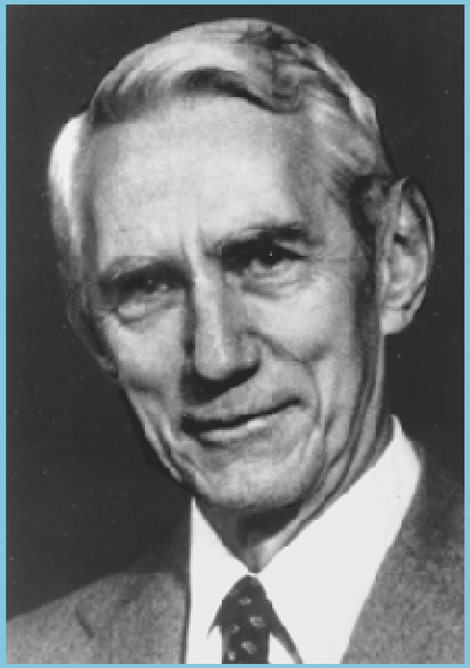
review article by [Costello Forney 2006]

It Took 50 Years To Do It

- 50‘s: algebraic codes
- 60‘s 70‘s: convolutional codes
- 80‘s: iterative codes (LDPC, turbo codes)
- 2009: polar codes

How to approach the predicted limits?

review article by [Costello Forney 2006]



Claude Shannon — Born on the planet Earth (Sol III) in the year 1916 A.D. Generally regarded as the father of the Information Age, he formulated the notion of channel capacity in 1948 A.D. Within **several decades**, mathematicians and engineers had devised practical ways to communicate reliably at data rates within 1% of the Shannon limit . . .

Encyclopedia Galactica, 166th ed.

Robert J. McEliece, Shannon Lecture 2004

Applications

- Communication Theory
- Computer Science (e.g. in cryptography)
- Physics (thermodynamics)
- Philosophy of Science (Occam's Razor)
- Economics (investments)
- Biology (genetics, bio-informatics)

Topics Overview

- Entropy and Mutual Information
- Entropy Diagrams
- Data Compression / Source Coding
- Perfectly Secure Encryption
- Error-Correction
- Zero-Error Information Theory
- Noisy-Channel Theorem
- Application to Machine Learning

Questions ?