COMP2610 — Information Theory

Lecture 1: Administration and Overview

Mark Reid and **Aditya Menon**

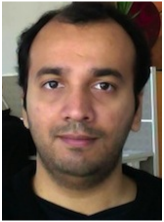
Research School of Computer Science
The Australian National University

THE AUSTRALIAN NATIONAL UNIVERSITY

July 22nd, 2014

Mark Reid

Marcus Hutter
(Guest Lecturer)



Aditya Menon

Welcome to COMP2610 - Information Theory

We will study the fundamental limits and potential of the *representation* and *transmission* of information.

- Mathematical Foundations
- Coding and Compression
- Communication
- Probabilistic Inference
- Kolmogorov Complexity (Guest Lectures)

Learning Outcomes

1. Understand fundamental concepts: probability, entropy, information content and their inter-relations
2. Understand principles of data compression
3. Compute entropy and mutual information
4. Solve simple probabilistic inference problems and understand their relation to information theory
5. Key theorems and inequalities
6. Basic concepts on communications over noisy channels

What Tools Will We Use?

- Elementary probability theory
  - "What's the probability of rolling an odd number using a fair die?"

What Tools Will We Use?

- Elementary probability theory
  - "What's the probability of rolling an odd number using a fair die?"
- Elementary linear algebra
  - "If $x = (1, 1, 0)$ and $y = (-2, 0, 1)$ what is $x \cdot y$ and $3x + 2y$?"

What Tools Will We Use?

- Elementary probability theory
  - "What's the probability of rolling an odd number using a fair die?"
- Elementary linear algebra
  - "If $x = (1, 1, 0)$ and $y = (-2, 0, 1)$ what is $x \cdot y$ and $3x + 2y$?"
- Basic programming skills
  - "Do you know your `for` loops from your `while` loops?"

What Tools Will We Use?

- Elementary probability theory
  - "What's the probability of rolling an odd number using a fair die?"
  - http://www.khanacademy.org/math/probability
- Elementary linear algebra
  - "If $x = (1, 1, 0)$ and $y = (-2, 0, 1)$ what is $x \cdot y$ and $3x + 2y$?"
  - http://www.khanacademy.org/math/linear-algebra
- Basic programming skills
  - "Do you know your `for` loops from your `while` loops?"

Outline

Course Overview

- Units: 6
- Lectures: 26×1 hour
- Tutorials: 5×2 hour
- Assignments: 3 (10%, 20%, 20% each)
- Final Exam (50%)
  - Unless otherwise stated, *everything* is examinable.

Lectures

- Attendance is compulsory
- 26×1 hour lectures
- Approximately half will be presented by Mark Reid and Aditya Menon
- Guest lectures at the end of the course by Marcus Hutter
- Mobile phones must be switched off or in silent mode during the duration of the lecture
- Laptops are allowed as long as they are used to take notes and/or to follow the lecture slides
- Participation is encouraged (and usually rewarded)
- Textbook reading, homework, additional reading, tutorials

Tutorials

- 5 tutorials, each 2 hours in duration
- Problem sets of 4-5 exercises will be provided for each tutorial
- These will review material covered in previous lectures
- You are meant to have tried the exercises beforehand

Note: *The course lecturers reserve the right to discount up to* 10% *of the marks for lack of attendance to tutorials.*

Lectures and Tutorials Schedule

Lectures

Lecture A : Tue 2:00 pm - 3:00 pm (CHEM T2 - Building 34)
Lecture B : Wed 2:00 pm - 3:00 pm (CHEM T2 - Building 34)

Tutorials
Only in Weeks 3, 5, 7, 11, 13

Tutorial A : Wed 4:00 pm - 6:00 pm (CSIT N101 - Building 108)
Tutorial B : Thu 2:00 pm - 4:00 pm (CSIT N101 - Building 108)
Tutorial C : Fri 2:00 pm - 4:00 pm (CSIT N108 - Building 108)

See Course info on Wattle for details and up-to-date postings:
`http://wattlecourses.anu.edu.au`

Note: *No food or drink is permitted in teaching venues.*

Assignments

- Assignment 1 (10%) : Available – 30th Jul. ; Due – 13th Aug.
- Assignment 2 (20%) : Available – 1st Sep. ; Due – 22nd Sep.
- Assignment 3 (20%) : Available – 30th Sep. ; Due – 17th Oct.
- Strong penalties on *plagiarism*. Please refer to the the corresponding ANU policies:
  `http://academichonesty.anu.edu.au/UniPolicy.html`

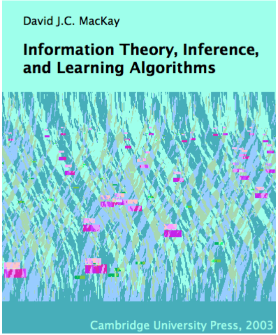
Assignments

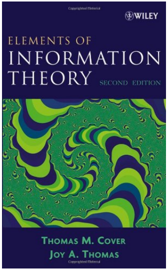
- Assignment 1 (10%) : Available – 30th Jul. ; Due – 13th Aug.
- Assignment 2 (20%) : Available – 1st Sep. ; Due – 22nd Sep.
- Assignment 3 (20%) : Available – 30th Sep. ; Due – 17th Oct.
- Strong penalties on *plagiarism*. Please refer to the the corresponding ANU policies:
  `http://academichonesty.anu.edu.au/UniPolicy.html`

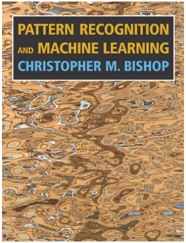
Note: In lieu of formal prerequisites, Assignment 1 will be testing how well you handle the level of mathematics we will be expecting in this course.

If you do not get a strong credit or above in Assignment 1, please see Mark or Aditya to discuss how to get up to speed or discontinue the course

TextBook



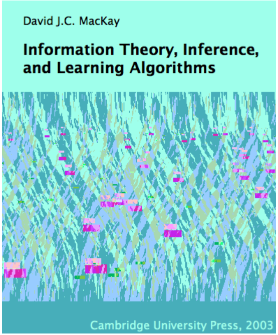


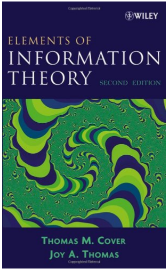


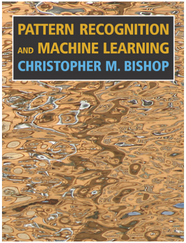
Mackay (ITILA, 2006) available online:
`http://www.inference.phy.cam.ac.uk/mackay/itila`

- Note copyright rules: e.g. copying the whole book onto paper is not permitted.
- We will follow a very different chapter order to that given in the book

TextBook







Mackay (ITILA, 2006) available online:
`http://www.inference.phy.cam.ac.uk/mackay/itila`

- Note copyright rules: e.g. copying the whole book onto paper is not permitted.
- We will follow a very different chapter order to that given in the book

For an alternative take – David MacKay's Lectures:

`http://www.inference.phy.cam.ac.uk/itprnn_lectures/`

Consultation & Other Issues

Consultation:

- Best way to contact the course lectures is via **email**

  Mark : `mark.reid@anu.edu.au`

  Aditya : `aditya.menon@nicta.edu.au`

- If you **really** need to meet them in person, send an email request first
- Email response times may vary but consider 1 day as a fast reply and up to three days as a normal response time

Consultation & Other Issues

Consultation:

- Best way to contact the course lectures is via **email**

  Mark : mark.reid@anu.edu.au
  Aditya : aditya.menon@nicta.edu.au

- If you **really** need to meet them in person, send an email request first
- Email response times may vary but consider 1 day as a fast reply and up to three days as a normal response time

Grievance resolution:

- You can contact the course lecturers in the first instance
- Alternatively, contact the Dean of students for advice
- If unresolved, you can lodge a formal complaint: http://policies.anu.edu.au/procedures/student_complaint_resolution/procedure

What Is Information? (1)

According to a dictionary definition, information can mean

1. Facts provided or learned about something or someone:
   a vital piece of information.
2. What is conveyed or represented by a particular arrangement or sequence of things:
   genetically transmitted information.

What Is Information? (1)

According to a dictionary definition, information can mean

1. Facts provided or learned about something or someone:
   a vital piece of information.
2. What is conveyed or represented by a particular arrangement or sequence of things:
   genetically transmitted information.

In this course: information in the context of *communication*:

- Explicitly include uncertainty
- Shannon (1948): "Amount of unexpected data a message contains"
  - A theory of information transmission
  - Source, destination, transmitter, receiver

What is Information? (2)

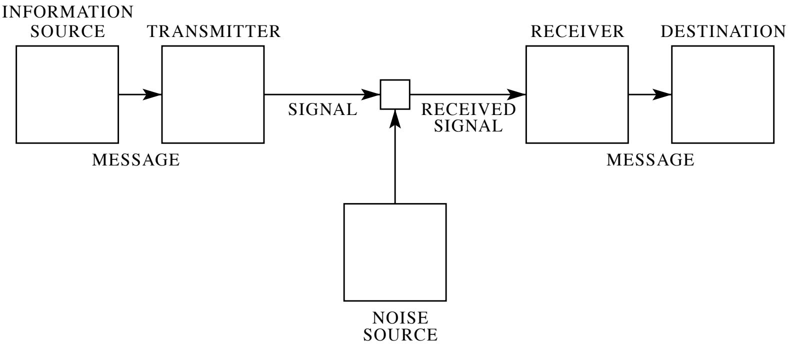


Fig. 1 — Schematic diagram of a general communication system.

From Shannon (1948)

What Is Information? (3)

Information is a message that is *uncertain* to receivers:

- If we receive something that we already knew with absolute certainty then it is non-informative.
- Uncertainty is crucial in measuring information content
- We will deal with uncertainty using probability theory

What Is Information? (3)

Information is a message that is *uncertain* to receivers:

- If we receive something that we already knew with absolute certainty then it is non-informative.
- Uncertainty is crucial in measuring information content
- We will deal with uncertainty using probability theory

Information Theory

Information theory is the study of the fundamental *limits* and *potential* of the *representation* and transmission of information.

Examples

Example 1: What Number Am I Thinking of?

- I have in mind a number that is between 1 and 20
- You are allowed to ask me one question at a time
- I can only answer yes/no
- Your goal is to figure out the number as quickly as possible
- What strategy would you follow?

Example 1: What Number Am I Thinking of?

- I have in mind a number that is between 1 and 20
- You are allowed to ask me one question at a time
- I can only answer yes/no
- Your goal is to figure out the number as quickly as possible
- What strategy would you follow?

Your strategy + my answers = a code for each number

Some variants:

- What if you knew I was twice as likely to pick numbers more than 10?
- What if you knew I never chose prime numbers?
- What if you knew I only ever chose one of 7 or 13?

$1000 Hidden in one of 16 cases.

- All equally likely to contain the prize

Example 2: How Much Is Information Worth?

Simplified Version of "Deal or No Deal"

\$1000 Hidden in one of 16 cases.

- All equally likely to contain the prize

How much would you pay to know:

1. Exactly which case contains the money?
2. Whether the case holding the money is numbered less than 8?
3. ... is less than 12?
4. Which range out of 0–3, 4–7, 8–11, or 12–15 the money case is in?

Example 2: How Much Is Information Worth?

Simplified Version of "Deal or No Deal"

$1000 Hidden in one of 16 cases.

- All equally likely to contain the prize

How much would you pay to know:

1. Exactly which case contains the money?
2. Whether the case holding the money is numbered less than 8?
3. ... is less than 12?
4. Which range out of 0–3, 4–7, 8–11, or 12–15 the money case is in?

Key Question:

- Can we use these ideas to *quantify* information?

Example 3: Redundancy and Compression

Cn y rd ths sntnc wtht ny vwls?

Example 3: Redundancy and Compression

Cn y rd ths sntnc wtht ny vwls?

Can you read this sentence without any vowels?

Written English (and other languages) has much redundancy:

- Approximately 1 bit of information per letter
- Naively there should be almost 5 bits per letter

(For the moment think of "bit" as "number of yes/no questions")

Example 3: Redundancy and Compression

Cn y rd ths sntnc wtht ny vwls?

Can you read this sentence without any vowels?

Written English (and other languages) has much redundancy:

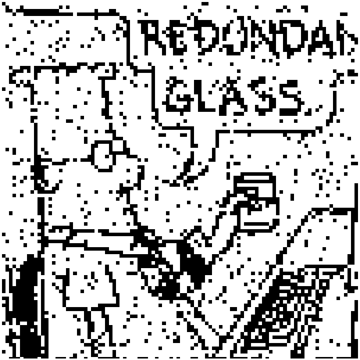
- Approximately 1 bit of information per letter
- Naively there should be almost 5 bits per letter

(For the moment think of “bit” as “number of yes/no questions”)

Key Question:

- How much redundancy can we *safely* remove?
  (Note: “rd” could be “read”, “red”, “road”, etc.)

Hmauns hvae the aitliby to cerroct for eorrrs in txet and iegmas.



Key Question:

- How much noise is it possible to correct for and how?

A Summary of the History of Information Theory

1920s : Nyquist & Hartley at Bell Labs

1940 : Turing and Good at Bletchley Park (WWII)

1942 : Hedy Lamarr and George Antheil

1948 : Claude Shannon: "A Mathematical Theory of Communication"

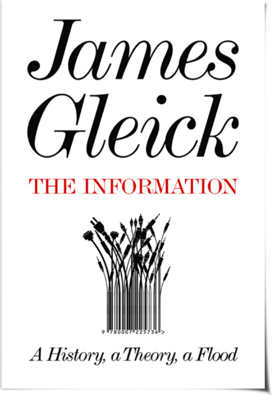
1951 : Huffman Coding

1958 : Peter Elias: "Two Famous Papers"

1970 : "Coding is Dead"

1970- : Revival with advent of digital computing
CDs, DVDs, MP3s, Digital TV, Mobiles, Internet, Deep-space comms (Voyager), ...

More on the History of Info. Theory



&

Information Theory and the Digital Age
by Aftab, Cheung, Kim, Thakkar, and Yeddanapudi.
http://web.mit.edu/6.933/www/Fall2001/Shannon2.pdf

Brief Overview of Course

- How can we quantify information?
  - Basic Definitions and Key Concepts
  - Probability, Entropy & Information

Brief Overview of Course

- How can we quantify information?
  - Basic Definitions and Key Concepts
  - Probability, Entropy & Information
- How can we make good guesses?
  - Probabilistic Inference
  - Bayes Theorem

Brief Overview of Course

- How can we quantify information?
  - Basic Definitions and Key Concepts
  - Probability, Entropy & Information
- How can we make good guesses?
  - Probabilistic Inference
  - Bayes Theorem
- How much redundancy can we safely remove?
  - Compression
  - Source Coding Theorem, Kraft Inequality
  - Block, Huffman, and Lempev-Ziv Coding

Brief Overview of Course

- How can we quantify information?
  - Basic Definitions and Key Concepts
  - Probability, Entropy & Information
- How can we make good guesses?
  - Probabilistic Inference
  - Bayes Theorem
- How much redundancy can we safely remove?
  - Compression
  - Source Coding Theorem, Kraft Inequality
  - Block, Huffman, and Lempev-Ziv Coding
- How much noise can we correct and how?
  - Noisy-Channel Coding
  - Repetition Codes, Hamming Codes

Brief Overview of Course

- How can we quantify information?
  - Basic Definitions and Key Concepts
  - Probability, Entropy & Information
- How can we make good guesses?
  - Probabilistic Inference
  - Bayes Theorem
- How much redundancy can we safely remove?
  - Compression
  - Source Coding Theorem, Kraft Inequality
  - Block, Huffman, and Lempev-Ziv Coding
- How much noise can we correct and how?
  - Noisy-Channel Coding
  - Repetition Codes, Hamming Codes
- What is randomness?
  - Kolmogorov Complexity
  - Algorithmic Information Theory

Brief Overview of Course

- How can we quantify information? [Aditya]
  - Basic Definitions and Key Concepts
  - Probability, Entropy & Information
- How can we make good guesses? [Aditya]
  - Probabilistic Inference
  - Bayes Theorem
- How much redundancy can we safely remove? [Mark]
  - Compression
  - Source Coding Theorem, Kraft Inequality
  - Block, Huffman, and Lempev-Ziv Coding
- How much noise can we correct and how? [Mark]
  - Noisy-Channel Coding
  - Repetition Codes, Hamming Codes
- What is randomness? [Marcus]
  - Kolmogorov Complexity
  - Algorithmic Information Theory