Information: Broader Perspectives

Deepak Kumar
Bryn Mawr College
First, a quick review...
The information went data way
The General Definition of Information

Information = Data + Meaning

General Definition of a Datum

\[ D_d = \text{def} \ x \text{ being distinct from } y \]

where the \( x \) and \( y \) are two un-interpreted variables and the domain is left open to further interpretation.

Common/Useful Representations of Data

• Analog

• Discrete
  (digital)

• Binary
  (y’all know what that is...)

Types of Data

• **Primary**
  (raw data in a database, or a table, ...)

• **Secondary**
  (converse of primary data, stuff that is missing...)

• **Metadata**
  (data about data, location, format, copyright,...)

• **Operational**
  (data about the operation of a data system...)

• **Derivative**
  (data extracted from other data, as in mining...)

The General Definition of Information

Definition: $\sigma$ is in instance of information, understood as semantic content, if and only if:

1) $\sigma$ consists of $n$ data, for $n \geq 1$;

2) The data are well-formed;

3) The well-formed data are meaningful.

Information: A Taxonomy

Three Types of Information

• **Syntactic Information**
  Related to symbols from which messages are formed, and to their interrelations (structural)

• **Semantic Information**
  related to the meaning of messages...

• **Pragmatic Information**
  Related to the usage and effect of messages

Example

I. John was brought to the railway station by taxi.

II. The taxi brought John to the railway station.

III. There is a traffic jam on highway A3, between Nuremberg and Munich in Germany.

IV. There is a traffic jam on highway A3 in Germany.
Example

I. John was brought to the railway station by taxi.
II. The taxi brought John to the railway station.

Syntactically different, but semantically and pragmatically identical.

Same meaning and equally informative.

Example

I. John was brought to the railway station by taxi.
II. The taxi brought John to the railway station.
III. There is a traffic jam on highway A3, between Nuremberg and Munich in Germany.
IV. There is a traffic jam on highway A3 in Germany.

Differ in syntax.

Also semantics: III gives more precise information than IV.

Pragmatic aspect of information depends on context.

Three Types of Information

• **Syntactic Information**
  Related to symbols from which *messages* are formed, and to their interrelations (structural)

• **Semantic Information**
  related to the meaning of *messages*...

• **Pragmatic Information**
  Related to the usage and effect of *messages*

Three Types of Information

• Syntactic Information
  Related to symbols from which messages are formed, and to their interrelations (structural)

• Semantic Information
  Related to the meaning of messages

• Pragmatic Information
  Related to the usage and effect of messages

Two Traditions...

• Concerned primarily with semantic & pragmatic aspects of information.

• Concerned with syntactic aspects of information:
  – How do you measure syntactic information?
  – Limits on the amount of information which can be transmitted?
  – Limits on compression of information?
  – How to build information processing systems approaching these limits?
Classical Information Theory

• Concerned with syntactic aspects of information:
  – How do you measure syntactic information?
  – Limits on the amount of information which can be transmitted?
  – Limits on compression of information?
  – How to build information processing systems approaching these limits?
Classical Information Theory

• Concerned with syntactic aspects of information:
  – How do you measure syntactic information?
    Answer: Entropy, $H$.
  – Limits on the amount of information which can be transmitted?
    Answer: Channel Capacity, $C$
  – Limits on compression of information?
    Answer: Entropy, $H$.
  – How to build information processing systems approaching these limits?
    Answer: Fo’ real!

Relationship with Other Fields

Floridi’s Roadmap

• **Mathematical Information**
  Information Theory

• **Semantic Information**

• **Physical Information**
  Universe is fundamentally composed of data, instead of matter or energy, with material objects as a complex secondary manifestation.

• **Biological Information**
  Genetic Information, Neural Information, Computational Neuroscience, ...

• **Economic Information**
  Commoditization of information, value of information, game theory, ...

• **Ethics of Information**

Five Fundamentals of Information

• **Entropy**
  Information Theory, bits, bandwidth, codes, ...

• **Economics**
  Strategies for value: how information is produced, priced, and distributed, ...

• **Encryption**
  Secure transmission, digital signatures, digital cash, ...

• **Extraction**
  Data organization, storage, extraction, etc...

• **Emission**
  Frequency, modulation, radio, TV, phones, networks, ...

Towards a Science of Information

• Information Theory needs to meet new challenges of current applications in biology, communication, knowledge extraction, economics, ...

• Understand new aspects of information in structure, time, space, and semantics.

• PLUS...dynamic information, limited resources, complexity, representation-invariant information, and cooperation & dependency
Some Challenges...

• **Structure**
  Measures are needed for quantifying information embodied in structures (e.g., information in material structures, nanostructures, biomolecules, gene regulatory networks, protein networks, social networks, financial transactions).

• **Time**
  Delay (e.g., information arriving late may be useless or has less value).

• **Semantics**
  Is there a way to account for the meaning or semantics of information?

• **Knowledge Extraction**
  Data driven science focuses on extracting information from data. How much information can actually be extracted from a given data repository? How?
Science of Information

Center for Science of Information (soihub.org->Research Overview), 2012.
## Emerging Frontiers of Sol

<table>
<thead>
<tr>
<th>Information Theory &amp; Communications</th>
<th>Biology</th>
<th>Knowledge Extraction</th>
<th>Economics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional</td>
<td>Quantum</td>
<td>Molecular</td>
<td>Neuroscience</td>
</tr>
</tbody>
</table>

- **Fundamentals**
- **Structure**
- **Time**
- **Space**
- **Semantics**
Our Sol Roadmap

• **Foundations of Information Theory**
  Entropy, codes, compression, channels, ...

• **Communications**
  Voice, data, ...

• **Biological Information**
  Bioinformatics, proteomics, epidemiology, neuroscience, ...

• **Information Extraction**
  Big data, storage, processing, IR, indexing, search engines, visualization, ...

• **Economic Information**
  Dynamic economic theory, behavior of continuously optimizing agents in markets, ...

• **Quantum Information**
How “dirty” do we want to get?

- Information Theory needs to meet new challenges of current applications in biology, communication, knowledge extraction, economics, ...
- Understand new aspects of information in structure, time, space, and semantics.
- PLUS...dynamic information, limited resources, complexity, representation-invariant information, and cooperation & dependency

- **Foundations of Information Theory**
  Entropy, codes, compression, channels, ...
- **Communications**
  Voice, data, ...
- **Biological Information**
  Bioinformatics, proteomics, epidemiology, neuroscience,...
- **Information Extraction**
  Big data, storage, processing, IR, indexing, search engines, visualization,...
- **Economic Information**
  Dynamic economic theory, behavior of continuously optimizing agents in markets, ...
- **Quantum Information**

- **Mathematical Information**
  Information Theory
- **Semantic Information**
- **Physical Information**
  Universe is fundamentally composed of data, instead of matter or energy, with material objects as a complex secondary manifestation.
- **Biological Information**
  Genetic Information, Neural Information, Computational Neuroscience, ...
- **Economic Information**
  Commoditization of information, value of information, game theory, ...
- **Ethics of Information**

- **Entropy**
  Information Theory, bits, bandwidth, codes,...
- **Economics**
  Strategies for value: how information is produced, priced, and distributed, ...
- **Encryption**
  Secure transmission, digital signatures, digital cash, ...
- **Extraction**
  Data organization, storage, extraction, etc...
- **Emission**
  Frequency, modulation, radio, TV, phones, networks, ...
Our Sol Roadmap

- **Foundations of Information Theory**
  Entropy, codes, compression, channels, ...

- **Communications**
  Voice, data, ...

- **Biological Information**
  Bioinformatics, proteomics, epidemiology, neuroscience,...

- **Information Extraction**
  Big data, storage, processing, IR, indexing, search engines, visualization,...

- **Economic Information**
  Dynamic economic theory, behavior of continuously optimizing agents in markets, ...

- **Quantum Information**
Next,

Fundamentals of Information Theory