



WHAT IS INFORMATION?

Jerzy Konorski¹, Wojciech Szpankowski²

¹Gdansk University of Technology, Poland, ²Purdue University, USA



==== Motivation ====

Intuition: *data* (collection of interpretable symbols) may or may not carry *information*.

Despite the...

- advances in **information** technology,
- abundance of **information** systems and services,
- much trumpeted advent of **information** society, or even the **Information** Age (cf. Web 2.0),

...the common i-buzzword remains undefined in precise terms.




This stands in the way of a systematic study: how much **information** (not data) are **information** systems able/supposed to carry?

Shannon's Statistical Information

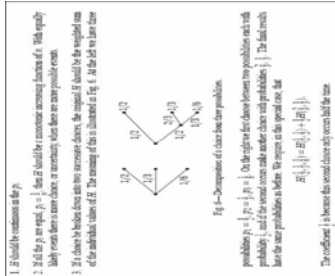


= reduction of source entropy at the output of a source-recipient *channel*.

source  recipient

Hence, "[The] semantic aspects of communication are irrelevant to the engineering problem." (Shannon 1948)

Skewed probability distribution of source symbols, high channel error rate may bring statistical information down to zero.



Technologie Informacyjne, Gdańsk, 22.5.2007

Formalizing Information



...brings into the picture:

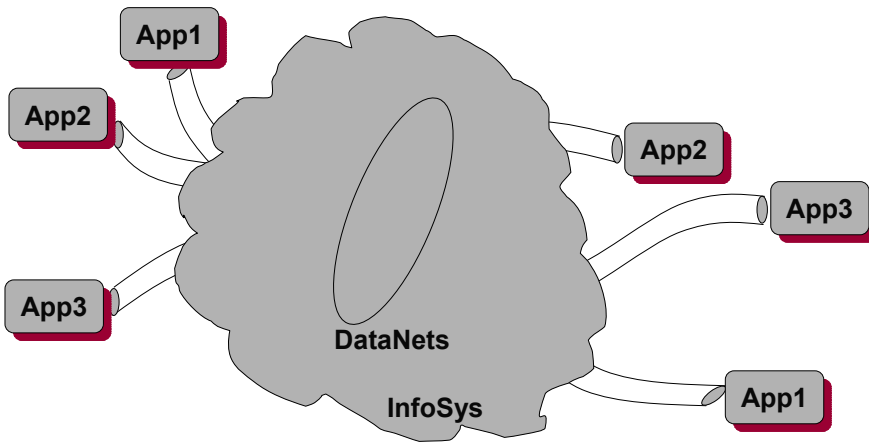
- *context* of data (even at a 50% error rate a transmitted math textbook might still be recognized as such),
- in particular, *timing* of data (consider a notice stating train departure time served a recipient after that time),
- recipient's *objective* (consider the same notice served a recipient not going anywhere),
- recipient's *protocol* i.e., rules of conduct (a duplicate of the same notice may still be of informational value if protocol requires confirmation).

Distribution of data can too play a role: secret-sharing decryption keys carry no information until brought together at one location and time.

Leads to a generalized (application-level?) notion of channel.

Technologie Informacyjne, Gdańsk, 22.5.2007

==== Vision ====



==== Data and Information ====



A starting point of some early IT textbooks:

A piece of data carries information if it helps its recipient achieve some objective.

Our viewpoint:

A piece of data carries information if it affects a recipient's objective under a given protocol and within a given context.

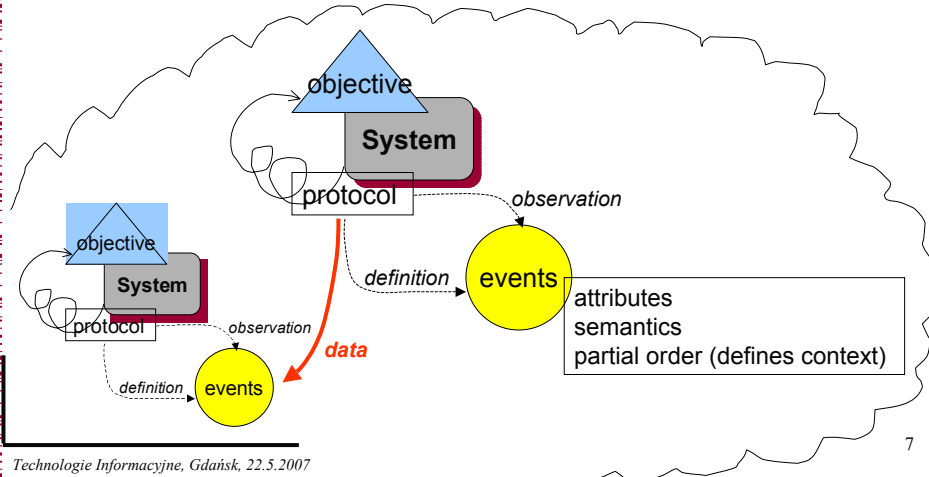
Little formal apparatus as yet to quantitatively account for all its facets.

Recent attempts at quantifying the "value" of information reduced to static decision problems, context- and protocol-free.

Event-Driven Paradigm



Can we formally define the *amount* of information and maximum transferable amount of information – *capacity* – without specification of data semantics?
One possibility: adopt the *event-driven paradigm*.



Event-Driven Paradigm



- well-established among CS community
- discrete and timeless in nature, yet allows for characterization of systems evolving in continuous time
- able to formalize such intuitions as causality and consistency of local views without specifying the semantics of events
- it is *events* that may or may not carry information; event = reception of data, but also event = clock tick etc.

Objective Functional



An *objective* functional maps a system's protocol R and the current context $C = (E1, E2, \dots)$ – sequence of events so far – into any space with a defined point order.

$objective(R, C) = \text{extent to which a stated objective has been achieved}$

Examples



[Train Departure]

E = train departure notice increases $objective(R, C)$ iff:

- (1) there is no E in C or R requires confirmation, and
- (2) clock ticks in C imply catching the train still possible.

[Decimal Representation]

Learn the number π through R = compute successive decimal digits using a finite-speed processor.

Here, $(E1, E2, \dots)$ = successive digits, $objective(R, C)$ increases monotonically in C and asymptotically stabilizes.

Examples



[Connectivity among Mobile Terminals]

Each MT can physically communicate only within its transmission range, so source-to-destination relay paths must be set up. The more time is allowed, the more connection discovery events occur as MTs move around and encounter other MTs within range.

If throughput is the objective then $objective(R, C)$ need not increase in C – new connections are discovered, but previously discovered ones disappear.

For $objective(R, C)$ to increase in C , quite unorthodox R is needed: restrict paths to two-hop, trade buffer space for bandwidth!

Examples



[Herding and Anti-Herding]

Previous example might suggest that $objective(R, C)$ increases in C provided that R is somehow "rational."

Enter herding effects: a rational Bayesian contemplating an action joins the majority of observed rational individuals. Finds very quickly that further observations bring no information on the benefits of the action.

Perhaps, then, $objective(R, C)$ is at least nondecreasing in C ? Consider a growing number of viewpoints contributed to a Web 2.0 community, which at some point may prevent a broad consensus.



Examples

[Noncooperative Behavior]

Objectives of different systems may be in conflict (e.g, DoS or selfish attacks on communication networks).

Consider two data sources contending for a multiple access channel. Various R may then calibrate the setting:

- from **cooperative** ($objective(R, C)$ increases in $C = \text{total data sent}$)...
- to **selfish** ($objective(R, C)$ increases in $C = \text{own data sent}$)...
- to **malicious** ($objective(R, C)$ decreases in $C = \text{data sent by the other source}$).



Amount of Information

...carried by event E in context C under protocol R:

$$info_{R,C}(E) = distance[objective(R, C), objective(R, C + E)]$$

where $distance[·, ·]$ = difference between two points according to the defined point order in the space of objectives.

Note:

- Definition allows negative information.
- Interesting notion: **nonconfoundable** R preclude negative information regardless of C (imagine a smart Web user always able to remove conflicting data from the context and proceed monotonically towards the objective). Whether such R exist is a tantalizing open problem.



Channel Capacity

R constrained by system's architecture. C constrained by the nature of source.

In the spirit of Shannon,

$$\text{channel_capacity} = \max_{\text{feasible } R, C} \max_{E \in C} \text{info}_{R, C}(E)$$

Interestingly, definition allows book chapter summaries channel to be more capacious than book content channel...



Example: Entry Deterrence

		Entrant	
		enter	not enter
Incumbent	Premium	$3 - K, -1$	$5 - K, 0$
	Standard	$2, 1$	$3, 0$

K = surcharge for Incumbent using Premium

- Entrant has only estimate K' of K ,
- Incumbent knows K and K' , can communicate K to Entrant

Game of incomplete information.

At a Nash equilibrium, E = communication of K is *worthwhile* and *credible* iff $K < 1 \leq K'$. Selfishness reduces channel capacity by 25%.



==== Conclusion ====

Design of information systems calls for notions of application-level information, capacity.

Approach successful if quantitative analysis of $info_{R,C}(E)$ possible with little or no use of event semantics.

Preliminary thoughts promise interesting results...