School of Computer Science University of Birmingham UK

What is information? Meaning? Semantic content?

Aaron Sloman

This file is available in HTML
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/whats-information.html
Also available in PDF format
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/whats-information.pdf
(Warning: sometimes the pdf file does not include the latest updates.)

Last updated: 2 May 2009
Now very different from the original version posted on 20th Sept 2006.
[Needs to be re-formatted]

CONTENTS

Introduction

What is 'information'? What is an information-user? What is involved in understanding something as expressing a meaning or referring to something? Is there a distinction between things that merely manipulate symbolic structures and things that manipulate them while understanding them and while using the manipulation to derive new information? In how many different ways do organisms acquire, store, extract, derive, combine, analyse, manipulate, transform, interpret, transmit, and use information? How many of these are, or could be, replicated in non-biological machines? Is 'information' as important a concept for science as 'matter' and 'energy', or is it just a term that is bandied about by undisciplined thinkers and popularists? Is it reasonable to think of the universe as containing matter, energy and information, with interdependencies between all three?

Can "information" be defined, or is it an unanalysable concept, implicitly defined by the structures of the powerful theories that use the concept (like most deep scientific concepts)? Is information something that should be *measurable* as energy and mass are, or is it more like the structure of an object, which which needs to be *described* not measured (e.g. the structure of this sentence, the structure of a molecule, the structure of an organism)?

If some of the information in a report is false can what is true be measured in bits? How come I can give you information without reducing the information I have? How come I can tell you something that gives you information I did not have?

What follows is an attempt to sum up what I think many scientists and engineers in many disciplines, and also historians, journalists, and lay people, are talking about when they talk about information, as they increasingly do, even though they don't realise precisely what they are doing.

> For example the idea pervades this excellent book about infant development:
> Eleanor J. Gibson, Anne D. Pick, (2000)
> *An Ecological Approach to Perceptual Learning and Development,*
> Oxford University Press, New York,

Unfortunately, I keep finding people who reveal new confusions about the notion of 'information' and are impervious to the points made below (some also made on the psyche-d discussion list: http://www.archive.org/details/PSYCHE-D

So I worry that the word may be a source of so much confusion that we should try to find a better one. I fear that is impossible now. "Meaning" is just as bad, or worse.

This is a topic on which I have written many things at different times including the following, but this is my first attempt to bring all the ideas together. The problem of explaining what information is includes the problem of how information can be processed in *virtual machines*, natural or artificial.

> NB: in this context, the word "virtual" does not imply "unreal", as explained in this

PDF presentation:
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#wpe08
Virtual Machines in Philosophy, Engineering & Biology (presented at WPE 2008)

These are some of the other papers and presentations on this topic:

- http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#4
'What enables a machine to understand?', in *Proceedings 9th International Joint Conference on AI,* pp 995-1001, Los Angeles, 1985.
- http://www.cs.bham.ac.uk/research/projects/cogaff/81-95.html#5
'Reference without causal links', in *Proceedings 7th European Conference on Artificial Intelligence,* Brighton, July 1986. Re-printed in J.B.H. du Boulay, D.Hogg, L.Steels (eds) *Advances in Artificial Intelligence - II* North Holland, 369-381, 1987.
- http://www.cs.bham.ac.uk/research/projects/cogaff/misc/information.text
A message about information processing posted to Psyche-D in 1998
- http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#inf
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#wpe08
PDF slide presentations on information processing in virtual machines
- http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#models
Why symbol-grounding (concept empiricism) is both impossible and unnecessary, and why theory-tethering is more powerful anyway.
(Latest version of PDF slide presentation on varieties of meaning, and why symbol grounding is neither necessary nor desirable. Includes an introduction to model-based semantics and symbol tethering.)
- http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0604
'Ontology extension' in evolution and in development, in animals and machines.
- http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#glang
What evolved first: Languages for communicating, or languages for thinking (Generalised Languages: GLs)?

```
This document is based on my answer to a question I was asked about
'semantic information' on the MINDMECHANISMS discussion list.

My original answer, posted on 20 Sep 2006, is available online at
   http://www.jiscmail.ac.uk/cgi-bin/webadmin?A2=ind0609&L=mindmechanisms&T=0&P=1717
as part of a thread with subject 'Analysis of conscious functions'.

This document is a much revised, expanded version of the above.

I expect to continue to revise/improve this. Criticisms welcome.
=======================================================================
```

Why I don't use the phrase 'semantic information'

```
Information is semantic (as I use the word 'information'), and anything
semantic is information content, as I use the world 'semantic', so
semantic information is just information, just as young youths are just
youths.

I sometimes quote other people who use the two-word phrase, and I have
often referred to the collection of pioneering papers edited by Marvin
Minsky in 1968 in a book called 'Semantic information processing', which
```

I think was wrongly named.

Occasionally I may contrast syntactic information with semantic information where the former is about the structure of something that conveys information (e.g. the syntactic information that this sentence includes a parenthesis, which is semantic information about syntactic structure) whereas the semantic information would be about what the content is, namely information about my habits.

Why I don't talk about 'information' in Shannon's sense
(Modified: 12 Apr 2009)

There is of course, another, more recent use of the word 'information' in the context of Shannon's so-called 'information theory'. But that does not refer to what is normally meant by 'information' (and what I mean by 'information'), since Shannon's information is a purely syntactic property of something like a bit-string, or other structure that might be transmitted from a sender to a receiver using a mechanism with a fixed repertoire of possible messages. For instance if a communication channel can carry N bits then each string transmitted makes a selection from 2**N possibilities. The larger N is the more alternative possibilities are excluded by each string actually received. In that sense longer strings carry more 'information'.

Having that sort of information does not, for example, allow something to be true or false, or to contradict or imply something else in the ordinary senses of 'contradict' or 'imply'. The ordinary concept of information includes control or imperative information, including information about what to do in order to bake a cake or knit a sweater. Imperative information is not true or false, but a particular process can be said to follow or not follow the instructions. Since bit patterns are used as instructions in computers, this could be taken to imply that Shannon's concept includes at least imperative (control) information. That would be a mistake: the Shannon measure indicates how many different instructions can be accommodated in fixed bit length, but says nothing about what particular action or process is specified by the instruction: that depends not only on the content of the bit-string but also also what is in the interpreter. In a computer, the interpreter of an instruction encoded as a bit-string could be hard-wired in the electronic circuits, or stored in changeable ``firmware'', or in an algorithm itself expressed as a procedure encoded in bit-strings.

In biological organisms genetic information about how to construct a particular organism by constructing a very complex collection of self-organising components is encoded in molecular sequences whose interpretation as instructions depends on complex chemical machinery assembled in a preceding organism, to kick-start the interpretation process, which builds additional components that continue the assembly partly influenced by the genetic information and partly by various aspects of the environment.

The importance of cascaded development of cognitive mechanisms influenced by the environment is discussed in
http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0609(PDF)
'Natural and artificial meta-configured altricial information-processing systems'
Jackie Chappell and Aaron Sloman
International Journal of Unconventional Computing 2,3, 2007, pp. 211--239

Though essential in other contexts, Shannon's concept is not what we need in talking about an animal or robot that acquires and uses information *about* various things (the environment, its own thinking, other agents, etc.)

NOTE

It is sometimes claimed that in Shannon's sense 'information' refers to physical properties of physical objects, structures, mechanisms. But that is a mistake. For example, it is possible to have structures in virtual machines that operate as bit-strings and are used for communication between machines, or for virtual memory systems.

That leaves the challenge of defining what we mean by 'information' in the semantic sense that involves being used to *refer* to something, being taken by a user to be *about* something, and not merely to having some syntactic or geometric form.

Why information need not be true: 'Information' *vs*. 'information content
(Added 3 Aug 2008)

This section presents and criticises a viewpoint that I think expresses unhelpful definitional dogmatism.

Some people, for example the philosopher Fred Dretske, claim that what we ordinarily mean by 'information' in the semantic sense being discussed here is something that is *true*. The claim is that you cannot really have information that is false. False information, on that view can be compared with the decoy ducks used by hunters. The decoys are not really ducks though some real ducks may be deceived into treating the decoys as real -- to their cost! Likewise, argues Dretske, false information is not really information, even though some people can be deceived into treating it as information. It is claimed that truth is what makes information valuable, therefore anything false would be of no value.

See Dretske's contribution to
Floridi, L. (Ed.). (2008). *Philosophy of Computing and Information: 5 Questions.* Copenhagen, Denmark: Automatic Press / VIP.
http://www.amazon.com/Philosophy-Computing-Information-5-Questions/dp/8792130097

Whatever the merits of this terminology may be in the context of conventional philosophical debates, the restriction of 'information' to what is true is such a useless encumbrance that it would force scientists and robot designers (and philosophers like me) to invent a new word or phrase that had the same meaning as 'information' but without truth being implied. For example, a phrase something like 'information content' might be used to refer to the kind of thing that is common to my belief that the noise outside my window is caused by a lawn-mower, and my belief that the noise in the next room is caused by a vacuum cleaner, when the second belief is true while first belief is false because the noise outside comes from a hedge trimmer.

> The philosopher R.M. Hare introduced the labels 'Phrastic' and 'Neustic' to distinguish the semantic content of an utterance and the speech act being performed regarding that content, e.g. asserting it, denying it, enquiring about its truth value, commanding that it be made true, etc. The concept of information being discussed here is close to Hare's notion of a 'Phrastic', except that semantic content here is not restricted to what can be expressed in a linguistic form: maps, models, diagrams and other things can encode such information.
> See R.M. Hare *The Language of Morals*, 1952, OUP.

When I say that humans, other animals and robots, acquire, manipulate, interpret, combine, analyse, store, use, communicate, and share information, this claim applies equally to false information and to true information, or to what could laboriously be referred to as the 'information content' or 'the potential information content' that can occur in false as well as true beliefs, expectations, explanations, and percepts, and moreover, can also occur in questions, goals, desires, fears, imaginings, hypotheses, where it is not known whether the information content is true, i.e. corresponds to how things are.

So in constructing the question

> "Is that noise outside caused by a lawnmower?"

I may use the same concepts and the same modes of composition of information as I use in formulating true beliefs like:

> "Lawnmowers are used to cut grass"
> "Lawnmowers often make a noise"
> "Lawnmowers are available in different sizes"

as well as many questions, plans, goals, requests, etc. involving lawnmowers. Dretske may find only true propositions valuable, whereas most people find all sorts of additional structures containing information very useful. Even false beliefs can be useful, because by acting on them you may learn that they are false, why they are false, and gain additional information. That's how science proceeds and I suspect much of the learning of young children depends heavily on their ability to construct information contents that are false. There's also the usefulness of false excuses, white lies, etc., which I shall not discuss.

Anyhow for the purposes of this discussion note, and more generally for cognitive science, neuroscience, biology, AI (including robotics) and many varieties of engineering, it is very important not to restrict the notion of 'information' to what is true, or even to whole propositions that are capable of being true or false. There are information fragments of many kinds that can be combined in many ways, some of which involve constructing propositions. But that's only one case: others may be used in controlling, questioning, desiring, planning, hypothesising, telling stories, etc. The uses of information in *control* are probably what evolved first in biological organisms, including, for example, microbes.

Forms of representation: information-bearers
(Added 3 Aug 2008)

It should not be assumed that anything that uses information expresses it in something like sentences, algebraic expressions, logical expressions (i.e. Fregean forms of representation that use application of function to arguments as the only way to combine information items to form larger information items. For example, some information may be expressed in the level of activation of some internal or external sensing device, some in patterns of activation of many units, some in geometrical or topological structures analogous to images or maps, some in chemical compounds, and many more. Exactly how many different forms exist in which information can be encoded, and what their costs and benefits are, is an interesting question that will not be discussed further here.

> See, for example, this discussion of alternatives to logical representations in 1
>
> "Interactions between philosophy and {AI}: The role of intuition and non-logical
> Proc 2nd IJCAI, 1971, pp. 209--226,
> http://www.cs.bham.ac.uk/research/cogaff/04.html#200407,
>
> and this more recent presentation:
> http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#glang
> What evolved first: Languages for communicating, or languages for thinking?

The investigation of the space of possible forms of representation and their tradeoffs is a major long term research project. Much of this paper is neutral as regards the form in which information is encoded.

Is 'information' definable?

After many years of thinking about this, I concluded that 'information' in the sense that is being discussed here is indefinable, like 'mass', 'energy' and other deep concepts used in important scientific theories.

That is to say, there is no informative way of writing down an explicit definition of the form 'X is Y' if X is such a concept.

All you'll end up with is something circular, or potentially circular, when you expand it, e.g. 'information is meaning', 'information is semantic content', 'information is what something is about'.

But that does not mean either that the word is meaningless or that we cannot say anything useful about it.

The same is true of 'energy'. It is sometimes defined in terms of 'work', but that eventually leads in circles.

Moreover, the specific things that might be said about what energy is change over time as we learn more about it. Newton knew about some forms of energy, but anything he might have written down about it probably would not have done full justice to what we now know about energy, e.g. that mass and energy are interconvertible, that there is chemical energy, electromagnetic energy, etc. Deepening theoretical knowledge gradually extends and deepens the concepts we use in expressing that knowledge.

> See also L.J. Cohen, (1962),
> *The Diversity of Meaning,* Methuen, London,

The same could happen to the concept of information.

Concepts implicitly defined by theories using them

So how do we (and physicists) manage to understand the word 'energy'? (Or the concept 'energy'?)

Answer: by making use of the concept in a rich, deep, widely applicable theory (or collection of theories) in which many things can be said about energy, e.g. that in any bounded portion of the universe there is a scalar (one-dimensional), discontinuously variable amount of it, that its totality is conserved, that it can be transmitted in various ways, that it can be stored in various forms, that it can be dissipated, that it flows from objects of higher to objects of lower temperatures, that it can produce forces that cause things to move or change their shape, etc. etc.
[Most of that would have to be made more precise for a physics text book.]

And as science progresses and we learn more things about energy the concept becomes deeper and more complex.

Of course, if nothing useful can be done with the theory, if it doesn't explain better than most alternative theories available, a variety of observed facts, and if it cannot be used in making predictions, or selecting courses of action to achieve practical goals, then the theory may not have content referring to our world, and the understanding of concepts implicitly defined by it will be limited to reference within the world postulated by the theory.

Note added 29 Dec 2006
Because a concept can be defined implicitly by its role in a powerful theory, and therefore some symbols referring to such concepts get much of their meaning from their *structural relations* with other symbols in the theory (including relations of derivability between formulae including those symbols) it follows that not all meaning has to come from experience of instances, as implied by the theory of concept empiricism.
Concept empiricism is a very old philosophical idea, refuted around 1780 by Immanuel Kant, and later by philosophers of science in the 20th century thinking about theoretical concepts like 'electron', 'gene', 'neutrino', 'electromagnetic field.'

The theory of concept empiricism was reinvented near the end of the century by Stevan Harnad and labelled 'symbol grounding theory'. This theory is highly plausible to people who are not properly educated in philosophy, so it has spread widely among AI theorists and cognitive scientists.

For a while I used the label 'symbol attachment', then later 'symbol tethering' (suggested by Jackie Chappell) for the alternative theory outlined above, that meaning can come to a large extent from a concept's role in a theory. I.e. meaning can be *primarily* determined by structural relations within a manipulable theory, independently of any causal links with the reality referred to. Residual ambiguity of reference can be reduced by 'symbol tethering', i.e. attaching *the whole theory* to means of performing observations, measurements, experiments and predictions.

There are also several relevant presentations here:
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/
Especially:
http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#models
Why symbol-grounding is both impossible and unnecessary, and why theory-tethering is more powerful anyway.
(Introduction to key ideas of semantic models, implicit definitions and symbol tethering through theory tethering.)

Also:
http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0603
(Discussion of the role of sensorimotor contingencies in cognition)
http://www.cs.bham.ac.uk/research/projects/cogaff/misc/nature-nurture-cube.html
Requirements for going beyond sensorimotor contingencies

to representing what's out there
(Learning to see a set of moving lines as a rotating cube.)
http://www.cs.bham.ac.uk/research/projects/cosy/papers/#pr0604
'Ontology extension' in evolution and in development, in
animals and machines.

For more on Concept Empiricism, see:
a recent paper by a philosopher attacking concept empiricism (PDF)
Concept Empiricism: A Methodological Criticism
(to appear in Cognition)
by Edouard Machery, University of Pittsburgh
Department of History and Philosophy of Science

For a recent defence of Concept Empiricism see
The Return of Concept Empiricism (PDF)
[Penultimate draft of chapter in H. Cohen and C. Leferbvre (Eds.)
Categorization and Cognitive Science, Elsevier (forthcoming).
by Jesse J. Prinz
Department of Philosophy, University of North Carolina at Chapel Hill

What was said above about 'energy' applies also to 'information':

An implicitly defined notion of 'information'

We understand the word 'information' insofar as we use it in a rich, deep, and widely applicable theory (or collection of theories) in which many things can be said about information, e.g. that it is not conserved (I can give you information without losing any), that instead of always having a *scalar* value, items of information have a *structure* (e.g. there are replaceable parts of an item of information such that if those parts are replaced the information changes but not necessarily the structure), that it can be transmitted in various ways, that it can vary both discontinuously (e.g. adding an adjective or a parenthetical phrase to a sentence, like this) or continuously (e.g. visually obtained information about a moving physical object), that it can be stored in various forms, that it can influence processes of reasoning and decision making, that it can be extracted from other information, that it can be combined with other information to form new information, that it can be expressed in different syntactic forms, that it can be more or less precise, that it can express a question, an instruction, putative matter of fact, and in the latter case it can be true or false, known by X, unknown by Y, while Z is uncertain about it, etc. etc. Some items of information allow infinitely many distinct items of information to be derived from them. (E.g. Peano's axioms for arithmetic, in combination with predicate logic.) Physically finite, even quite small, objects can therefore have infinite information content.

(Like brains and computers, or rather the systems containing them.)

Note that although there is not necessarily any useful scalar concept of 'amount' of information there is a partial ordering of containment. Thus one piece of information I1 may contain all the information in I2, and not vice versa. In that case we can say that I1 contains more information. But not every partial ordering determines a linear ordering, let alone a scalar measure.

Even the partial ordering may be relative to an information user. That's because, giving information I1 to an individual A, may allow A to derive I2, whereas another individual B may not be able to derive I2, because the derivation depends on additional in formation, besides that in I1.

Life and information

Every living thing processes information insofar as it uses (internal or external) sensors to detect states of itself or the environment and uses the results of that detection process either immediately or after further information processing to select from a behavioural repertoire, where the behaviour may be externally visible physical behaviour or new information processing. In the process of using information it also uses up stored energy, so that it also needs to use information to acquire more energy. (And there has to be an obvious inequality there.) There are huge variations between different ways in which information is used by organisms, including plants, single celled organisms, and everything else. Only a tiny subset have fully deliberative information processing competence, as defined here:

http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0604

A basic law of information and energy

I suspect it is one of the basic laws of the universe that operations in which the information content of some bounded system increases require energy to be used in the physical machine in which the information processing is implemented. But I suspect that what I have just stated is a special case of something more general.

Quantum computations are reversible and may be an exception to that, but I don't understand much about quantum computation and my initial conclusion is that such reversible computations must be incapable of deriving any new information; they perhaps merely produce syntactic rearrangements between informationally equivalent structures. But perhaps that's just my ignorance.

In any case, there are informationally equivalent (i.e. mutually derivable) rearrangements of information-bearing structures where one arrangement is more useful for certain purposes than others.

Information processing in virtual machines

Because possible operations on information are much more complex and far more varied than operations on matter and energy, e.g. insofar as information can have circular content relations (A is part of B, and B is part of A, which is impossible for physical structures), engineers discovered, as evolution had discovered much earlier, that relatively unfettered information processing requires use of a *virtual* machine rather than a physical machine, like using cog-wheels to do addition and multiplication.
A short tutorial introduction to this notion of a virtual machine, and an indication of some of the variety of possible virtual machines, can be found in this presentation (given in Bielefeld, October 2007):

> http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#bielefeld
> 'Why robot designers need to be philosophers'

> Some of the ideas are presented more clearly and in more depth here
> http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#wpe08
> 'Virtual Machines in Philosophy, Engineering and Biology'
> (presented at WPE 2008)

What I have written so far does not come near exhausting our current notion of information.

We also need to point out that whereas energy and physical structures simply exist, whether used or not, information is only information for a type of information-user. Thus a structure S refers to X or contains information about X for a user of X, U. The very same physical structure can contain different information, or refer to something different for another user U'.

Potential information content for a user

The information in S can be *potentially* usable by U even though U has never encountered S or anything with similar information content.

[Two points added: 28 Dec 2006] That's obviously true when U encounters a new sentence, diagram or picture for the first time. Even before U encountered the new item, it was potentially usable as an information bearer.

In some case the potential cannot be realised without U first learning a new language, or notation, or even a new theory within which the information has a place.

You cannot understand the information that is potentially available to others in your environment if you have not yet acquired all the concepts involved in the information. For example, it is likely that a new born human infant does not have the concept of a metal, i.e.

that is not part of its ontology. So it is incapable of acquiring the information that it is holding something made of metal even if a doting parent says "you are holding a metal object". The information-processing mechanisms (forms of representation, algorithms, architectures) that are required to think of things in the environment as made of different kinds of stuff, of which metals are a subset, take several years to develop in humans.

> **[Added 29 Mar 2008]**
> A recently completed paper on vision is now available here:
> http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0801
>
> > Architectural and representational requirements for seeing processes and affordances.
>
> This makes the point that whereas most people discuss affordances (following J.J.Gibson) as being concerned with what *actions* an agent can and cannot perform, we need to extend that idea of 'affordance' to include what information an agent can and cannot get in a particular situation. Thus we need to talk about both *action affordances* and *epistemic affordances.*
>
> Moreover, just as an agent will typically not make use of the majority of action possibilities available, it will also typically not make use of the majority of kinds of information that can be acquired in the environment. But nevertheless the affordances exist: the information is available, and the individual *has the potential* to make use of it.
>
> This is one of many ways in which states of an information-processing system (typically a virtual-machine, or a mind) are generally not just constituted by what is actually occurring in the system but by what would or could occur if various things were slightly different in various ways. In that sense, what it is for something to mean X rather than Y to an individual is intimately bound up with the causal powers of X. What those causal powers are is not just a matter of free interpretation, as some people suggest. It is a matter of which counterfactual conditional statements, or which causal statements, about that system are true and which false.
>
> Analysing what *that* means, however, is a hard philosophical problem, on which opinions differ.

Potential information content for a TYPE of user
(Added 20 Feb 2009)
It is possible for information to be potentially available for a TYPE of user even if NO instances of that type exist. For example, long before humans evolved there were things happening on earth that

could have been observed by human-like users using the visual apparatus and conceptual apparatus that humans have. But at the time there were no such observers, and perhaps nothing else existed on the planet that was capable of acquiring, manipulating, or using the information, e.g. information about about the patterns of behaviours of some of the animals on earth at the time.

There may also be things going on whose detection and description would require organisms or machines with a combination of capabilities, including perceptual and representational capabilities and an information-processing architecture, that are possible in principle, but have never existed in any organism or machine and never will -- since not everything that is possible has actual instances. Of course, I cannot give examples, since everything I can present is necessarily capable of being thought about by at least one human. Weaker, but still compelling, evidence is simply the fact that the set of things humans are capable of thinking of changes over time as humans acquire more sophisticated concepts, forms of representation and forms of reasoning, as clearly happens in mathematics, physics, and the other sciences. There are thoughts considered by current scientists and engineers that are beyond the semantic competences of any three year old child, or any adult human living 3000 years ago. If the earth had been destroyed three thousand years ago, that might have relegated such thoughts to the realm of possible information contents for types of individual that never existed, but could have. (This needs more discussion.)

Information content for a user determined partly by context

It is also possible for an information-bearing structure S to express different information, X, X', X", for the same user U in different contexts, e.g. because S includes an explicit indexical element (e.g. 'this', 'here', 'you', 'now', or non-local variables in a computer program).

Another factor that makes it possible for U to take a structure S to express different meanings in different contexts can be that S includes a component whose semantic role is to express a *higher order* function which generates semantic content from the context.

E.g. consider:

'He ran after the smallest pony'.

Which pony is the smallest pony can change as new ponies arrive or depart. More subtly what counts as a tall, big, heavy, or thin something or other can vary according to the range of heights, sizes, weights, thicknesses of examples in the current environment and in some cases may depend on why you are looking for something tall, big, heavy , etc., as explained in this discussion paper:

http://www.cs.bham.ac.uk/research/projects/cosy/papers/#dp0605
Spatial prepositions as higher order functions:
And implications of Grice's theory for evolution of language

There are many more examples in natural language that lead to incorrect diagnosis of words as vague or ambiguous, when they actually express precise higher order functions, applied to sometimes implicit arguments, e.g. 'big', 'tall', 'efficient', 'heap'.

> Note added 28 Dec 2006:
> This idea is developed in the context of Grice's theory of communication, with implications for the evolution of language, in the above web page. Examples include spatial prepositions and other constructs, which are analysed as having a semantics involving higher order functions some of whose arguments are non-linguistic.

A more complex example is:

'A motor mower is needed to mow a meadow'

which is true only if there's an implicit background assumption about constraints on desirable amounts of effort or time, size of meadow, etc.

So a person who utters that to a companion when they are standing in a very large meadow might be saying something true, whereas in a different context, where there are lots of willing helpers, several unpowered lawnmowers available, and the meadow under consideration is not much larger than a typical back lawn, the utterance would be taken to say something different, which is false, even if the utterances themselves are physically indistinguishable.

Moreover, where they are standing does not necessarily determine what sort of meadow is being referred to. E.g. they may have been talking about some remote very large or very small meadow.

There are other examples where what is said and understood, even by the same person, varies from one culture to another, e.g. the use of the word 'married', or 'rich'.

Visual information is highly context dependent

There are lots of structures in perceptual systems that change what information they represent because of the context. E.g. if what is on your retina is unchanged after you turn your head 90 degrees in a room, the visual information will be taken to be about a different wall, which may have the same wallpaper.
Many examples can be found in

Alain Berthoz, *The Brain's sense of movement,*
Harvard University Press, 2000,

Added: 10 Oct 2006

The importance of the role of extra-linguistic context in linguistic communication is developed in connection with indexicals, spatial prepositions, and Gricean semantics, into a theory of linguistic communications as using higher order functions some of whose arguments have to be extracted from non-linguistic sources by creative problem-solving. This has implications for language learning and the evolution of language, as discussed here.

Information content shared between users

Less obviously, it is sometimes possible for X to mean the same thing to different users U and U', and it is also possible for two users who never use the same information bearers (e.g. they talk different languages) to acquire and use the same information.

(This is why relativistic theories of truth are false: it cannot be true for me that my house has burned down but not true for my neighbour.)

Misguided definitions

What it means for S to mean X for U cannot be given any simple definition (though often people try to do that by saying U uses S to 'stand for' or 'stand in for' X, which is clearly either circular, because it merely repeats what is being defined, or else false if taken literally because there are all sorts of things you can do with something that you would never do with a thing that refers to it and vice versa. You can eat food, but not the word 'food', and even if you eat a piece of paper on which 'food' is written that is irrelevant to your use of the word to refer to food.

Information is normally used for quite different purposes from the purposes for which what is referred to in that information is used.

So the notion of standing in for is the wrong notion to use to explain information content. It's a very bad metaphor, even though its use is very common. We can make more progress by considering ways in which information can be *used*. If I give you the information that wet whether is approaching, you cannot use the information to take a shower. But you can use it to decide to take an umbrella when you go out, or, if you are a farmer you may use it as a reason for accelerating harvesting. The same information can be used by you or someone else in different ways in different contexts and the relationship between information content and information use is not a simple one.

The world is NOT the best representation of itself
(Section added 24 Feb 2009)

Another widely accepted but erroneous claim, related to confusing representing with standing in for, is the claim that "the world is its own best representation", or "the best representation of the environment is the environment".

Herbert Simon pointed out long ago, in *The Sciences of the Artificial* (1969) that sometimes the changes made to the environment while performing a task can serve as reminders or triggers regarding what has to be done next, giving examples from insect behaviours. The use of stigmergy, e.g. leaving tracks or pheromone trails or other indications of travel, which can later be used by other individuals, shows how sometimes changes made to the environment can be useful as means of sharing information with others. Similarly if you cannot be sure whether a chair will fit through a doorway you can try pushing it through, and if it is too large you will fail. Or if you cannot tell whether it could be made to fit by rotating it you can try rotating to see whether there is an orientation that allows it to fit through the space available.

But to go from the fact that more or less intelligent agents can use the environment as a store of information or as a source of information or as part of a mechanism for reasoning or inferring to the slogan that the world, or any part of it, is always, or even in those cases **the best** representation of itself is an error,

1. because it omits the role of the information-processing in the agent making use of the environment

2. because it sometimes is better to have specific instructions, or a map, or a blue-print or some other information structure that decomposes information in a usable way, than to have to use the portion of the world represented, as anyone learning to play the violin will notice if all the teacher does is play and say "just do what I do".

The error has various components, but the main component is a failure to acknowledge that information about X is something different from X and that the reasons for wanting or using information about X are different from the reasons for wanting or using X. E.g. you may wish to use information about X in order to ensure that you never get anywhere near X if X is something dangerous or unpleasant, but you cannot normally use X itself for that purpose, or you may wish to use information about Xs to destroy Xs, but if you have to use an X itself as the bearer of that information, rather than some other information-bearer (i.e. representation) then when the job is done you will have lost the information about how to destroy another X, which may include information about precautions to be taken in advance of meeting the next one.

```
I don't know where the error first arose. I don't think Herbert
Simon drew the general conclusion from his example, as shown by all
his subsequent research on forms of representation suited to various
tasks.

In 1998 Hubert Dreyfus wrote a paper on Merleau-Ponty entitled
"Intelligence Without Representation", published in 2002, in Phenomenology
and the Cognitive Sciences 1:367-83, available online:
http://www.class.uh.edu/cogsci/dreyfus.html
in which he stated:
    "The idea of an intentional arc is meant to capture the idea
    that all past experience is projected back into the world. The
    best representation of the world is thus the world itself."

As far as I can make out he is merely talking about expert servo
control, e.g. the kind of visual servoing which I discussed in
    http://www.cs.bham.ac.uk/research/projects/cogaff/06.html#604
    Image Interpretation, The Way Ahead? (1983)
But as any roboticist knows, and his own discussion suggests, this
kind of continuous action using sensory feedback requires quite
sophisticated internal information processing, though possibly not
the kind the Dreyfus assumed was the only kind available for use in
AI, e.g. something like logical reasoning.

Rodney Brooks wrote a series of papers attacking symbolic AI of
which one had the same title, published in Artificial Intelligence
1991, vol 47, pp 139--159, available here
http://people.csail.mit.edu/brooks/papers/representation.pdf
Instead of claiming that the world is its own best representation,
he repeatedly emphasises the need to test working systems on the
real world and not only in simulation, a point that has some
validity but can be over-stressed. (If aircraft designers find it
useful to test their designs in simulation, why not robot designers?)
However, he does not like the term "representation" as a label for
information structures required by intelligent systems, writing:

    "We hypothesize (following Agre and Chapman) that much of even
    human level activity is similarly a reflection of the world
    through very simple mechanisms without detailed representations."
and
    "we believe representations are not necessary and appear only in
    the eye or mind of the observer."

I have a critique of that general viewpoint in
http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0804
    Some Requirements for Human-like Robots: Why the recent
    over-emphasis on embodiment has held up progress.
That paper mostly criticises Brook's 1990 paper "Elephants don't
play chess", available at
    http://people.csail.mit.edu/brooks/papers/elephants.pdf
```

in which he goes further:

> "The key observation is that the world is its own best model. It is always exactly up to date. It always contains every detail there is to be known. The trick is to sense it appropriately and often enough."

Of course, that's impossible when you are planning the construction of a skyscraper using a new design, or working out the best way to build a bridge across a chasm, or even working out the best way to cross a busy road, which you suspect has a pedestrian crossing out of sight around the bend.

Another complication: complexity in information-users: information using subsystems

An information-user can bave parts that are information users and there are many complications such as that a part can have and use some information that the whole would not be said to have. E.g. your immune system and your digestive system and various metabolic processes use information and take decisions of many kinds though we would not say that you have, use or know about the information.

Likewise there are different parts of our brains that evolved at different times that use different kinds of information (even information obtained via the same route, e.g. the retina or ear-drum, or haptic feedback). Some of them are evolutionarily old parts, shared with other species, some newer, and some unique to humans.

That's why much of philosophical, psychological, and social theorising is misguided: it treats humans as unitary information users -- and that includes Dennett's intentional stance and what Newell refers to as 'the Knowledge level'.

This just the beginning of an analysis

I suspect that what I've written here probably amounts to no more than a tenth (or less) of what needs to be said about information in order to present the theory in terms of which the notion of information is implicitly defined in our present state of knowledge.

A hundred years from now the theory may be very much more deep and complex, just as what we know now about information is very much more deep and complex than what we knew 60 years ago, partly because we have begun designing, implementing, testing and using so many new kinds of information processing machines.

However the information processing machines produced by evolution are still orders of magnitude more complex than any that we so far understand.

I doubt that anyone has yet produced a clear, complete and definitive list of facts about information that constitute an implicit definition of how we (the current scientific community well-educated in mathematics, logic, psychology, neuroscience, biology, computer science, linguistics, social science, artificial intelligence, physics, cosmology, ...) currently understand the word 'information'.

Related documents

I have a draft incomplete paper on vision which includes a section about the evolution of different modes of expressing and using information, e.g. sometimes using information only implicitly in a pattern of activation within part of an information-using system (like the pattern of activation in an input or output layer of a neural net), and sometimes explicitly by creating a re-usable enduring structure. The paper is here

> http://www.cs.bham.ac.uk/research/projects/cogaff/sloman-diag03.pdf
> What the brain's mind tells the mind's eye.

I suspect that since the vast majority of species are micro-organisms the vast majority of information-using organisms can use only implicitly represented information. The conditions for enduring re-usable information structures to evolve are probably very special. There are also many sub-divisions within the implicit and explicit information users.

I tried to list a subset of the facts defining 'information' in these slide presentations on information processing machines:

> http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#inf
>
> http://www.cs.bham.ac.uk/research/projects/cogaff/talks/#meanings

But it's still too shallow and incomplete.

[Added: 28 May 2007]
Some of these ideas are developed in an invited talk for ENF 2007, entitled 'Machines in the ghost', available here
> http://www.cs.bham.ac.uk/research/projects/cosy/papers/#tr0702

Some references (Added 28 May 2007)

> Luciano Floridi, at the University of Oxford, has written much on 'The philosophy of information', a phrase I think he coined. E.g.

- http://orgs.unt.edu/asis/POI.htm
  The Philosophy of Information (PoI) (2003)
- http://www.philosophyofinformation.net/blackwell/index.htm
  The Blackwell Guide to the Philosophy of Computing and Information (Oxford - New York: Blackwell, 2004).

  The Preface.
  Chapter five on "Information"
  The Introduction
  The glossary
  Web site for The Blackwell Guide

- Full list of Floridi's papers.
- Article on "semantic information" in Stanford Encyclopaedia of Philosophy.

Some references to the idea of 'symbol tethering' are listed above.

Maintained by Aaron Sloman
School of Computer Science
The University of Birmingham