'No Representation without Information Flow'
- Measuring Efficacy and Efficiency of Representation: An Information Theoretic Approach

JUNKANG FENG
Business College
Beijing Union University, Beijing China
Database Research Group
School of Computing, University of the West of Scotland, UK
junkang.feng@uws.ac.uk

YANG WANG
Database Research Group
School of Computing, University of the West of Scotland, UK
yang.wang@uws.ac.uk

Abstract: -Representation is a key concept for semiotics and for information systems. Stamper's framework may be seen outlining what is required for the efficacy of signs in standing for (i.e., representing) something else in an organization, among many others. We explore how the efficacy and the efficiency of representation may be measured, which seems overlooked in available literature of information systems and organizational semiotics. And we approach this problem from the perspective of what we call the 'information carrying' relation between the representation and the represented. We model the represented as an information source and the representation an information carrier, and the 'representing' relationship between them as that of 'information carrying'. That is, information is carried and therefore flows. These then are further modeled mathematically as random variables and random events, and a special relationship between random events. This approach enables us to reveal a necessary condition for the efficacy and the efficiency of representation, and to measure it. To this end we extend Dretske's semantic theory of information. The conviction that we put forward here is 'No representation without information flow', based upon which the efficacy and efficiency of a representation system may be measurable.

Key-Words: - Representation, Information systems, Information theory, Information content, Database design

1 Introduction

It would seem that the notion of representation is highly relevant to information systems and semiotics. The most fundamental concept of semiotics is 'sign', which is defined as 'a signal or token which stands for something else' (Stamper 1997). Semiotics is said to be a domain of investigation that explores the nature and function of signs as well as the systems and process underlying signification, expression, representation, and communication (Guild to Semiotics).

Moreover, representation seems also a key issue for information systems (IS). For example, Shanks claims that representation is at the core of the discipline of information systems (Shanks 1999)

and he considers issues of representation on aspects of quality from four semiotics levels, namely syntactic, semantic, and pragmatic levels and social world. At the semantic level, it is said that the goal is that the representation is complete and accurate at particular points in time (Lindland et al 1994, Shanks and Darke 1998). This means that the representation should capture all the meaning accurately.

Therefore representation is a key concept in organizational semiotics and in information systems. Many researchers in various fields have made serious endeavours to explore issues around 'representation' including 'modes of representation' (Weber 1997, Shimojima 1996, Barwise and Seligman 1997). Some researchers like Weber (1997) puts forward some

criteria based on notions of representations to make IS development sound, accurate and complete, but they are mainly pragmatically driven. Stamper's (1997) framework may be seen outlining what is requited for the efficacy of signs in standing for (i.e., representing) something else in an organization, among many others.

One thing seems to have been overlooked in the available literature of organizational semiotics and well-known literature of information systems, namely given something a to be represented by something else b, how can we explore whether b is able to represent a apart from the linguistic or the conventional meaning, if any, of b that is inter-subjectively decided by a community is a? And indeed, how does the linguistic or the conventional meaning of a sign/symbol come about in the first place?

We submit that what links a sign and what the sign stands for is information in the sense that 'what information a signal carries is what it is capable of 'telling' us, telling us truly, about another state of affairs' (Dretske 1981, p.44). We explore one condition for this efficacy and efficiency, which we wish to call the 'information carrying' condition between the representation and the represented. Note that in this paper we give a dual meaning to the word 'representation'. One refers to the phenomenon that one thing stands for something else. The other refers to a thing that stands for something else. We hope that which one of the two meanings is used within a particular context would be clear within the context.

We suggest looking at the problem of representation with an information theoretic framework, i.e., information creation at a source, and information carrying and transmission through representation. We draw on theory on information and information flow (Dretske 1981, Devlin 1991, Barwise and Seligman 1997) through which we put forward a quantifiable measure (a necessary condition) for the efficacy and the efficiency of a representation system, which is mainly concerned with the semantic level of semiotics. We define this condition as whether in information carrying using representation, an information carrier (i.e., a representation) is able to carry in full the information that it is supposed to. We investigate how different situations in it can happen. We quantifiably measure this 'information carrying' condition for representation.

Our approach is this. We model the represented as an information source and the representations an information carrier, and the 'representing' relationship between them as that of 'information carrying'. In other words, information is carried and therefore flows. We observe that normally what is represented can be seen as a random event, so is a representation. Furthermore the aforementioned 'information carrying' relationship is therefore seen as a special relationship between two random events. This approach enables us to reveal a necessary condition for the efficacy and the efficiency of representation, and to measure it. To this end we extend Dretske's (1981) semantic theory of information.

The rest of the paper is structured as follows. We review some work in the literature on the link between representation and information flow in section 2, which should give some background on the problem that we are looking at, namely how the phenomenon of 'representation' may be approached from the perspective of information flow. We then give a simple motivating example in section 3. We model and analyze the represented, the representations, and the 'information carrying' relationship, which results in information flow between them in sections 4, 5 and 6 respectively. We give the notion of 'information content' of a state of affairs in section 7 in order to further emphasize that representation is possible only because of informational relationship between states of affairs. In section 8, we outline some implications of our work to be reported here for information systems design. We give conclusions in section 9.

2 Representation Is Seen Liked with Information Flow

There are many angles from which to investigate representation. Representation is highly relevant to presentation of information and information flow. Jerry Fodor feels that they are so closely related as to justify the slogan 'No information without representation' (Barwise and Seligman, p.235). Such a link is shown in Devlin's Infon theory (1991), situation semantics of Barwsie and Perry (1983), Shimojima's conceptual framework on representations (1996) and Barwise and Seligman's representation system (1997). In addition, Wobcke (2000) and Cavedon (1995) discuss the link between reasoning and

information.

Situation theory, especially the notion of 'infon' (Devlin 1991) and inference based upon 'infons' (Barwise and Etchemendy 1990), models representation from an information carrying perspective. For situation theory, information flow is about how we know about a situation by knowing something else in another situation. Following Devlin (1991), basic 'item of information' is called 'infons', which are a set of objects with their associated relations. Infon may be true or false only in certain situation. Some common properties of situations are situation types. Their relations are captured by using constraints by virtue of 'informational connections/relationships' (Dretske 1981). Information flow takes place when an agent applies a *constraint* between two situation types. That is, if an individual situation of one situation type is held through 'anchoring' parameters in the situation type, then due to the constraint, an individual situation of the other situation type is also held as a consequence of the former, also through anchoring parameters. What is made true by the latter is said to be in the information content of what is made true of the former.

Barwise and Etchemendy (1990) envision the semantic analysis of representation. They assess representation and associated inferences beyond linguistics. Significantly, they discuss inference over information by using 'constraint infon algebras'. In their theory, information (or information content) is represented in terms of 'infons'. Infon algebra is put forward to perform inference. The models of infons are formulated by 'distributed lattice'. Heterogeneous presentations of information content are captured in an independent way while the 'entailment' relations are found in the context of situation theory. There are many interesting points. For example, they emphasize 'coherent' sets of basic infons, which is a set that carries genuine simultaneous information about a single situation if there is no reason to make the situations not to support the infons. This notion is highly relevant to the concept of 'normality' mentioned in Cavedon (1995). The view taken by Cavedon (ibid.) is that to perform inference, tokens behave 'normally' unless certain information prevents it not to. Also, Barwise and Etchemendy (ibid.) state that

'pseudo-complements' are the infons that are incompatible with the original set of infons. Infons are incompatible when, for example, $\sigma \wedge \bar{\sigma}$. This notion is important especially when inference takes place following 'deterministic constraints'. Conflicts are not permitted in the inference if infons are detected to be incompatible. It appears that there are implied considerations about 'background' information during the inference, for example, in Barwise and Etchemendy (1990, p.71), although they do not provide an explicit definition.

There is another interesting work in the literature. Shimojima (1996) constructs a conceptual framework to account for the differences between linguistic representation and diagrammatic representation. He identifies a number of phenomena, including *free ride*s, *over-specificity* and *self-consistency* that are unique to diagrammatic representation. He stresses the significant role of *structural constraint*s on representation. The notion of sound and accurate representation is also notable. This work is centered on representation of information, and the question of how representation can ever happen is not fully addressed. The framework captures constraints between the 'source' and 'target' in a representation system by defining 'constraint projections'. The mechanism was then formulated under the 'information flow channel' framework to be the 'representation system' by Barwise and Seligman (1997). Although he concerns the connections between situations, i.e., the signaling relation ' ⤳ ', in presenting information, Shimojima does not put emphasis on this issue in further discussions on 'constraint projections'. Moreover, despite that the source and target domains are defined based on states of affairs, there is not much consideration on formulating 'inter-winded' relations between them in the framework. Without underpinnings of interconnections (Barwise and Seligman (1997) use the notion of 'infomorphisms' to formalize them), undesired representation may happen. Shimojima (ibid.) discusses the impact of different modes of representation to the efficacy of representation, but does not emphasize the content of a representation. Hence, his framework does not address the validity of a representation-based inference that may be linked

with information flow.

Barwise and Seligman (1997) give a definition of representation system from three aspects, i.e., IF channel, token connections and constraints. That is, a representation system is formulated in terms of IF channel and its associated IF logic. The definition of representation system seems working well in cognition involved circumstances. There is still lack of consideration on the content of information flow. Only 'normal tokens' are able to support representations if there are constraints on the core that support the connections. It is very interesting to notice that a token connection being 'normal' is neither sufficient nor necessary for making a representation accurate.

We observe that aforementioned Jerry Fodor's (1982) claim 'No information without representation' is not quite right in that information is created due to reduction in uncertainty, which may or may not result in any representation of it. 'No information *flow* without representation' would make more sense. Furthermore, we are convinced that representation is possible only because of information flow. That is, only because one thing, say X, carries information that another thing, say Y, exists, can X represent Y. Moreover, the works that we have reviewed thus far do not provide us with a quantitative measure on representation, which makes the notion 'representation' in the context of information systems at least appear immature. Based upon our conviction '*No representation without information flow*', in this paper, we describe how the efficacy and efficiency of a representation system may be quantitatively measured.

3   A Motivating Example

We would need a representation system for many things we do. For example, supposing at a junction on High Street, we want to give indications to the traffic that travel from the East to the West that they can do one of the following things:

a) 'you may go straight on if the way is clear'
b) 'you may turn left if the way is clear'
c) 'you may turn right if the way is clear'
d) 'stop and wait behind the stop line on the

carriageway'

The represented (instructions to traffic)	The representations (traffic lights)
go straight	green
turn left	
turn right	
stop	red
	amber
	red and amber

Table 1 A representation system

Suppose that we design a traffic signal system for the above as follows:

As shown in table 1, for the above a), b) and c), we give a 'green' light, and d) we give either 'red' light, 'amber' light or 'red and amber' together.

This would not appear to be a good representation system. But we want to find out why exactly it is not good and whether we can generalize what we will find. We give our analysis in this paper and it is based upon the basic idea that a representation carries the information that the represented exists. A representation and what is to be represented by the representation in the most general sense are both states of affairs. Thus, the idea of 'information carrying' is a relationship between a signal and another state of affairs in that (we repeat) 'what information a signal carries is what it is capable of 'telling' us, telling us truly, about another state of affairs' (Dretske 1981).

Mathematical Equations must be numbered as follows: (1), (2), …, (99) and not (1.1), (1.2),…, (2.1), (2.2),… depending on your various Sections.

4   The Represented

Our approach starts with creating a mathematical model for this system. The 'indication to traffic' can be seen as a random variable having the abovementioned four possible values, namely those listed as a) to d), and each of 'the variable of 'indication to traffic' having a particular value' is therefore a random event. We observe that the elementary component of what is represented normally can be seen as a random event, and the whole set of the represented can be seen as a collection of random events that are results of some selection process under certain conditions, and each 'run' of a selection process results in the realization of one of the possible random events.

Following Shannon (1949) and Dretske (1981),

we can now view such an aforementioned selection process as an information source. This is because we take the view that information is created due to *reduction in uncertainty*, and the realization of one of the possible random events removes other possible outcomes of the selection process whereby uncertainty namely *n* (*n*>1) possible outcomes is reduced to 1. Put informally, information is taken as created by or associated with a state of affairs among a set of other states of affairs of a situation, the occurrence or realization of which reduces the uncertainty of the situation. Reduction of uncertainty due to the realization of a random event can be quantified precisely as long as the unlikeliness, i.e., the probability, of the random event can be worked out. As a result of this, we can quantify a selection process as well.

Given s_a being a state of affairs (a random event) among a few others at a selection process S, then

$$I(s_a) = -\log P(s_a),$$

where $P(s_a)$ is the probability of s_a is taken as the quantity of information created by s_a, which is another and convenient measurement of the 'unlikeliness' of s_a. This is termed the *surprisal* of s_a. The weighted mean of surprisals of all random events of *S*, denoted as I(S), and

$$I(S) = -\Sigma P(s_i)\log P(s_i), i = 1,\ldots,n$$

is called the *entropy* of S.

For our example of 'indication to traffic' at that particular junction of High Street, let s_a, s_b, s_c and s_d denote the four random events involving the aforementioned following four values:

a) 'you may go straight on if the way is clear'
b) 'you may turn left if the way is clear'
c) 'you may turn right if the way is clear'
d) 'stop and wait behind the stop line on the carriageway'

Let us suppose that the four random events are equally likely, then the probability of s_a $P(s_a)$ is 1/4, and so are $P(s_b)$, $P(s_c)$ and $P(s_d)$. We would have:

$$I(s_a) = -\log P(s_a) = \log 4 = 2 \text{ (bits)}$$
$$I(s_b) = -\log P(s_b) = \log 4 = 2 \text{ (bits)}$$
$$I(s_c) = -\log P(s_c) = \log 4 = 2 \text{ (bits)}$$
$$I(s_d) = -\log P(s_d) = \log 4 = 2 \text{ (bits)}$$

These are the surprisals, and the entropy would be

$$I(S) = -\Sigma P(s_i)*\log P(s_i) = 4*1/4*(\log 4) = 2 \text{ (bits).}$$

5 The Representations

We now look at the representations, namely
1) 'green' light,
2) 'red' light,
3) 'amber' light,
4) 'red and amber' together.

We consider these in isolation, i.e., disregarding how they link to the represented except their probabilities, which are affected (in this case almost determined) by those of the represented. We can also model them as a selection process, a random variable and four random events. That is to say, the representations can also be seen as an information source in that some reduction in uncertainty takes place. Even though the representations are not independent of the represented, reduction in uncertainty does take place never the less.

Assume that the 2), 3) and 4) above can occur with equal chances to respond to the represented d), namely 'stop and wait behind the stop line on the carriageway'. Moreover, we denote the above 1) to 4) with ra, rb, rc and rd respectively. We would have probabilities for them respectively: ¾, 1/12, 1/12, and 1/12. Then we would have the following surprisals and entropy for the representations R:

$$I(r_a) = -\log p(r_a) = \log 4/3 = 2 - \log 3 \text{ (bits)}$$
$$I(r_b) = -\log p(r_b) = \log 12 = \log 4 + \log 3 = 2 + \log 3 \text{ (bits)}$$
$$I(r_c) = -\log p(r_c) = \log 12 = \log 4 + \log 3 = 2 + \log 3 \text{ (bits)}$$
$$I(r_d) = -\log p(r_d) = \log 12 = \log 4 + \log 3 = 2 + \log 3 \text{ (bits)}$$

These are the surprisals, and the entropy would be

$$I(R) = -\Sigma P(r_j)*logP(r_j) = \tfrac{3}{4}*(2 - log3) + 3*1/12*(2 + log3) \text{ (bits)}.$$

6 The 'Information Carrying' Relationship Between the Represented and The Representation

As aforementioned, the represented can be seen as an information source. We now submit that the representations are information carriers because the representations can tell us truly something about the represented. Moreover, when something to be represented is not fully represented (we will define what is meant by 'full representation' shortly), there must be the case where some information created at the information source is not carried by an information carrier. Such information is termed *Equivocation*. This is one hand. On the other hand, it is not always the case that all information that is created at the carrier (seen as an information source in its own right) comes from the represented. The information created by the representations themselves that is not accounted for by that at the represented is called *Noise*. Whether there is a relationship of 'being represented and representing' between two sets of things (random events) and how well the representation is can be measured precisely by means of *Equivocation* and *Noise* as long as the probabilities including conditional probabilities of the relevant random events are available. We now show how these can be done by using our running example.

Equivocation
Equivocation is the *lost information* that is created at the represented (events) but not carried by the representations. Information is created due to reduction in uncertainty. Therefore, if we can work out what the bit of 'reduction in uncertainty' that is lost is, i.e., not carried/represented, then we would obtain the equivocation. Our approach to this problem goes like this. Suppose that the represented (note

that this is only a random event, which would be a result of some selection process) is s_i, say s_a, namely a) 'you may go straight on if the way is clear', and due to the way we have designed the representation system shown in section 2, the representation would be r_a, namely 1) 'green' light. The whole 'reduction in uncertainty' measured by I(s_a) due to the realization of s_a can be seen as composed of two parts. One is the elimination of s_d from the all four possible outcomes, which can be measured by I(s_a or s_b or s_c) = $-logP(s_a$ or s_b or s_c). The other is the elimination of s_b and s_c from the set $\{s_a, s_b, s_c\}$, which can be measured by $-logP(s_a/(s_a$ or s_b or $s_c))$, where P($s_a/(s_a$ or s_b or $s_c)$) denotes the probability of s_a under the condition (s_a or s_b or s_c), and it is the same as r_a. So $-logP(s_a/(s_a$ or s_b or $s_c)$) = $-logP(s_a/r_a)$. As r_a only captures (represents) the first part of the reduced uncertainty, what is lost is the second part, which is equivocation. Both of them can be calculated by using the notion of surprisal as already shown above. We will generalize these and then move on to discuss average equivocations for a representation system in the subsections that follow.

Equivocation for a particular representation and one of its corresponding represented events
Following the approach just described above, given the represented s_i and the representation r_j, the equivocation denoted by E$s_i(r_j)$ would be
$$-logP(s_i/r_j),$$

where P(s_i/r_j) is the probability of s_i under the condition r_j.

For our example concerning s_a and r_a, we would have

$$Es_a(r_a) = -logP(s_a/r_a) = log3 \text{ (bits)}.$$

That is to say, the represented s_a is not fully represented by r_a. What is represented is one of s_a, s_b and s_c, which is the first part of the reduced uncertainty, namely 'you may go if the way is clear'. This applies to s_b and s_c also.

In the same way, we get

$$Es_d(r_b) = -logP(s_d/r_b) = log1 = 0 \text{ (bit)}.$$

So is $Es_d(r_c)$ and $Es_d(r_d)$. That is to say, s_d is fully represented by any of r_b, r_c or r_d.

Note that Shannon did not concern himself with such problems, so none provided. Dretske (1981) we believe gets it wrong as he uses a weighted mean of equivocations (see the next subsection) in place of the equivocation for the specific representation and one of its n ($n \geq 1$) corresponding represented events.

Equivocation for a particular representation and all its corresponding represented events

For a particular representation, there could be more than one represented events corresponding to it. That is, one representation may be used to represent more than one state of affairs (event). For example, for r_a, either s_a, s_b and s_c corresponds to it at one time. For each pair of them, i.e., r_a and s_a, r_a and s_b, etc, equivocation may be different. It is desirable to find out on average how much information is lost as for as a particular representation is concerned. This is the *weighted mean of equivocations* for a particular representation with each represented that corresponds to it. It can be calculated by using the following formula:

$$E(r_j) = \Sigma p(s_i/r_j)*Es_i(r_j) = -\Sigma p(s_i/r_j)*\log p(s_i/r_j), i =1,\ldots,n$$

For example,

$$E(r_a)=-\Sigma P(s_i/r_a)*\log P(s_i/r_a)=-P(s_a/r_a)*\log P(s_a/r_a)+P(s_b/r_a)*\log P(s_b/r_a)+P(s_c/r_a)*\log P(s_c/r_a)=3*1/3*\log 3 = \log 3 \text{ (bits)}.$$

Similarly we have

$$E(r_b) = -P(s_d/r_b)*\log P(s_d/r_b) = 1*\log 1 = 0 \text{ (bit)}.$$

In the same way, $E(r_c)$ and $E(r_d)$ are both 0 bit.

Equivocation for a representation system as a whole

The whole set of representations, i.e., the representation system as a whole, normally would have more than one representation. We can calculate the overall average amount of lost information as a measure for looking at the efficacy of a representation system. This is the weighted mean of equivocations for each representation. The formula for it is

$$E(r) = \Sigma P(r_j) *E(r_j), j = 1,\ldots, n$$

For our example,

$$E(r)= P(r_a)*E(r_a)+P(r_b)*E(r_b)+P(r_c)*E(r_c)+ P(r_d)*E(r_d) = ¾* \log 3 + 0+ 0+ 0 = ¾\log 3 \text{ (bits)}.$$

This shows the average amount of lost information, i.e., the average of the parts of the reduction in uncertainty created at the information source that is not captured by the representations with this representation system. Given the entropy of the represented is 2 bits (see the end of section 3), considerable information, i.e., ¾log3 bits, is lost, which leads to what may be called 'under representation'. In addition to 'under representation', 'over representation' can also occur, which, we suggest, can be captured by using the notion of *noise* mentioned before.

Noise

Noise can be handled in a similar way to that for equivocation but in an opposite direction. Noise is the *unaccounted-for information* that is created at the representations but is not accounted for by the represented (events). In other words, noise is the (part) of the reduction in uncertainty at the representations that is not accounted for by that at the represented.

Suppose that the representation is r_b, namely 2) 'red' light, and due to the way we have designed the representation system shown in section 2, the represented that corresponds to r_b would be s_d, namely d) 'stop and wait behind the stop line on the carriageway'. The whole 'reduction in uncertainty' due to the realization of r_b can be seen as composed of two parts. One is the elimination of r_a from the all four possible outcomes. The other is the elimination of r_c and r_d from the set of r_b, r_c and r_d. The first part is accounted for by s_d, as it eliminates r_a, which is the only possible outcome it eliminates. Thus the second part of the reduction in uncertainty is entirely due to the representations' 'initiative', which is noise. Both of them can be calculated by using the notion of surprisal, but we will only

discuss the second part here as it is noise. As we handle noise is the same as that for equivocation, we will only give the conclusions below.

Noise for a particular represented event and one of representations that represents it

Given the represented s_i and the representation r_j, the noise denoted by $Nr_i(s_j)$ would be

$$-\log P(r_j/s_i),$$

where $P(r_j/s_i)$ is the probability of r_j under the condition s_i.

For our example, if we consider s_d and r_b, we would have

$$Nr_b(s_d) = -\log P(r_b/s_d) = \log 3 \text{ (bits).}$$

This also applies to r_c and r_d in relation to s_d. In the same way, we get

$$Nr_a(s_a) = -\log P(r_a/s_a) = \log 1 = 0 \text{ (bit).}$$

So is $Nr_a(s_b)$ and $Nr_a(s_d)$. That is to say, with r_a, there is no noise.

Again, as for equivocations, Shannon (1949) did not concern himself with such problems, so none for it was provided. Dretske (1981) uses a weighted mean of noises (see the next subsection) in the place of the noise for a particular represented event and a particular representation that represents it, which we believe is incorrect.

Noise for a particular represented event and all representations that represents it respectively

This is the weighted mean of noises for each represented and a representation that represents it. It can be calculated by using the following formula:

$$N(s_i) = \Sigma P(r_j/s_i)*Nr_j(s_i) = -\Sigma P(r_j/s_i)*\log P(r_j/s_i), j = 1,\ldots, n$$

For our example,

$$N(s_d) = -\Sigma P(r_j/s_d)*Nr_j(s_d) = -\{P(r_b/s_d)*\log P(r_b/s_d) + P(r_c/s_d)*\log P(r_c/s_d) + P(r_d/s_d)*\log P(r_d/s_d)\} = 3*1/3*\log 3 = \log 3 \text{ (bits).}$$

$$N(s_a) = -P(r_a/s_a)*\log P(r_a/s_a) = 1*\log 1 = 0 \text{ (bit).}$$

In the same way, $N(s_b)$ and $N(s_c)$ are both 0 bit.

Noise for a whole set of represented events

This is the weighted mean of noises for each represented (event). The formula for it is

$$N(s) = \Sigma P(s_i)*N(s_i), i = 1,\ldots, n$$

For our example,

$$N(s) = P(s_a)*N(s_a) + P(s_b)*N(s_b) + P(s_c)*N(s_c) + P(s_d)*N(s_d) = 0 + 0 + 0 + 1/4*\log 3 = 1/4*\log 3 \text{ (bits).}$$

This shows the average amount of the unaccounted-for information at the representations, i.e., the average of the parts of the reduction in uncertainty created at the information carriers that is not accounted for by the represented events within this representation system.

Summary of the analysis of the 'information carrying' relationship for our motivating example
We now summarize our analysis for the running example in the tables below:

The represented	Equivocation (bits)	Noise (bits)	The representations
s_a	log3	0	r_a
s_b	log3	0	
s_c	log3	0	
s_d	0	log3	r_b
	0	log3	r_c
	0	log3	r_d

Table 2 Equivocation and noise between each pair of individual representation and individual represented event

The representations	Average equivocation (bits)
r_a	log3
r_b	0
r_c	0
r_d	0
Overall average equivocation	¾log3

Table 3 Average equivocations

The represented	Average noises (bits)
s_a	0
s_b	0
s_c	0
s_d	log3
Overall average noise	1/4log3

Table 4 Average noises

To make our discussion more complete than it is now, we should cover two trivial cases, namely an event that is not represented by any representation, and a representation that does

not represent anything. To this end we extend our example by adding two unrealistic situations shown in the last two rows in the following table.

The represented (instructions to traffic)	The representations (traffic lights)
go straight	green
turn left	
turn right	
stop	red
	amber
	red and amber
reverse	
	blue

Table 5 The motivating example extended

The state of affairs 'reverse' is not represented by any representation, so all information created (or should be created) by it would be lost, and it can be measured by its surprisal –logP(reverse), which is the equivocation associated with it. Similarly, the representation 'blue light' does correspond to any state of affair to be represented, thus all information crested due to its occurrence is noise, which can be measured also by its surprisal -logP(blue light). It is easy to note that the inclusion of the two trivial cases would change the overall average equivocation and the overall noise of the system as a whole.

7 The Notion of 'Information Content' and How It Is Related To Representation

The notion of 'information-carrying' between states of affairs can be further formalized with the notion of the information content of a state of affairs. Let us consider the following list:

Example 1. That there is smoke carries the information that there is a fire.

Example 2. That he is awarded a grade 'A' for his Programming course contains the information that Jack Brown has gained 70% or above for that course.

These two examples show that states of affairs are sometimes linked due to regularities such as natural laws and norms of an organization, and because of which one can learn something about one thing from another.

The strongest case of all is where the former can potentially tell the observer truly the existence of the latter. Such a phenomenon is termed the information content of a state of affairs.

Dretske (1981, P.45) defines the notion of the 'information content' of a state of affairs as follows:

A state of affairs contains information about X to just that extent to which a suitably placed observer could learn something about X by consulting it.

Following Dretske, we take information as in the form of 'de re', rather than 'de dicto', that is, in the form of '*a*'s being *F* carries the information that *b* is *G*'. Dretske (ibid.) establishes the following definition:

Information Content: A signal r carries the information that s is F = The conditional probability of s's being F, given r *(*and k*)*, is 1 *(*but , given k alone, less than 1*)*.

In this definition, *k* stands for prior knowledge about information source *s*.

A state of affairs, say X, becomes a representation of another state of affairs, say Y, only because Y is in the information content of X. When such a link between X and Y is accepted by a community, it is taken for granted for the community that X represents Y and the community would no longer wonder how it became so in the first place.

8 Implications For Information Systems Design

The ideas that we have been developing here should help further our understanding of the

nature of information systems. That is, an information system and the real world domain that the information system refers to constitute a 'representation system' in which the former is made of representations (also called the 'source' by Barwise and Seligman (1997)), and the latter the represented (also called the 'target' by Barwise and Seligman (ibid.)). On this point, we would agree with Seta et al. (2006) in that conceptual modeling is an activity to build an idealized and simplified representation of selected semantics about some real-world domain.

There is always a question of efficacy and efficiency for an information system, which may be taken as a foundation for formalizing methods and techniques for information systems development. For example, in database design, we would therefore want to make sure that all states of affairs of the real world domain be represented as the very minimal requirement. That is, there is no *equivocation* involved. Ideally, the database should provide just sufficient number of states of affairs to carry the information that the states of affairs of the real world domain exist. That is, there is no *noise* involved. In terms of 'information content' of states of affairs, the optimal design is achieved when real world states of affairs are all in the information content of a set of states of affairs of the database and this set is minimal in that the removal of any construct of the database would result in al least one of the real world state of affair being unrepresented. It is highly desirable to reveal how all these may be achieved.

To this end, we notice that the 'orthogonal database design principle' put forward by Eessaar (2006) would seem relevant to identifing the states of affairs of the real world that need to be captured. Moreover, Wang and Feng 2007 may also be seen as part of the endeavor, and we shall report further work in due course.

9   Conclusions

Our fundamental conviction '*No representation without information flow*' has led us to explore how the efficacy and efficiency of a representation system may be quantitatively measured. On this point, we now come to the following conclusions:

The efficacy and the efficiency of a representation system can be approached from an 'information carrying' perspective and by

using concepts and techniques associated with such a perspective. This results in the efficacy being looked at in terms of whether an individual state of affairs is fully, partially, not, or over represented by some representation (which is also a state of affairs in the most general sense), This also results in the efficiency being looked at in terms of whether an individual representation represents any state of affairs that is supposed to be represented. Then based upon these, we can look at the efficacy and the efficiency of the system as a whole in terms of the statistical characteristics of the 'representing and represented' relationship. In more details:

For a state of affair s_i to be *fully represented* within a representation system, there is at least one representation r_j such that the equivocation in relation to s_i and r_j is 0 bit, and this is equivalent to the condition that the probability of s_i given r_j is 1. The situation with 'stop' and 'read' in out running example is such a case.

When there is more than one such r_j, then s_i is *over-represented*, such as the case with the state of affairs 'stop', and the representations 'red', 'amber' and 'red and amber' lights.

For a state of affair s_i to be *partially represented* by a representation r_j, the probability of s_i given r_j must be grater than that of s_i without r_j, but it is not 1, such as the case with 'go straight on' and 'green'.
If the probability of s_i given r_j is the same as that of s_i without r_j, then s_i is *not represented* by r_j, such as the case with 'stop' and 'green'.

If the overall equivocation of a representation system is greater than 0 bit, then at least one state of affairs that is to be represented is not actually fully represented by the representations within the representation system. That is, the system is not of a full efficacy.

An *entirely irrelevant representation* r_j is a state of affairs such that for all s_i to be represented the probability of r_j given s_i is the same as that of r_j without s_i, such as the case with 'blue' light. In such a case, all information associated with the representation is noise.
In a case of over-representation, there must be some noise, i.e., the noise is greater than 0 bit, such as the case with 'red', 'amber' and 'red and amber'.

If the overall noise of a representation system is greater than 0 bit, there must be either at least one *entirely irrelevant representation* or at least one case of *over-representation*, or both. In such a case, a representation system is not the most efficient (regardless whether it is of a full

efficacy).

A point worth further investigation in the future by extending the work presented here is the question raised in section 1, namely how the linguistic or the conventional meaning of a sign/symbol comes about in the first place. We believe that to establish such meaning, the 'information carrying' condition discussed here has to be met as well between a sign (i.e., a representation) and the meaning (i.e., the represented).

References:

[1]. Barwise, J. and Seligman, J. (1997) *Information Flow: the Logic of Distributed Systems*, Cambridge University Press, Cambridge.

[2]. Barwise, J. and Etchemendy, J. (1990). *Information, Infons, and Inference*. In R. Cooper, K. Mukai and J. Perry (Eds.). Situation Theory and Its Applications. Stanford, CA: CSLI Publications.

[3]. Cavedon, L. (1995). *A Channel Theoretic Approach to Conditional Reasoning*. PhD Thesis, Edinburgh University.

[4]. Devlin, K. (1991) *Logic and Information*, Cambridge.

[5]. Dretske, F. (1981) *Knowledge and the Flow of Information*, Basil Blackwell, Oxford (reprinted 1999).

[6]. Eessaar, Erki. (2006). Guidelines about usage of the complex data types in a database. WSEAS Transactions on Information Science and Applications, v 3, n 4, April, 2006, p 712-719.

[7]. Fodor, J. A. (1982). *Representations: Philosophical Essays on the Foundations of Cognitive Science*. Contemporary Psychology 27(6):419-421.

[8]. Lindland, O., Sindre, G. and Solvberg, A. (1994) Understanding Quality in Conceptual Modelling, IEEE Software (March) 42-49, Mingers, J. and Stowell, F. (ed.), McGraw-Hill, London, 267-283.

[9]. Semiotics, http://www.press.jhu.edu/books/hopkins_guide_to_literary_theory/semiotics.html

[10]. Seta, Kazuhisa; Koyama, Kazuya; Hayashi, Yusuke; Ikeda, Mitsuru. Building ontologies for conceptual model management. WSEAS Transactions on Information Science and Applications, v 3, n 3, March, 2006, p 546-553.

[11]. Shanks, G. and Darke, P. (1998) Understanding Metadata and Data Quality in a Data Warehouse, Australian Computer Journal (November).

[12]. Shanks, G., (1999) Semiotic Approach to Understanding Representation in Information Systems, Proceedings of the information Systems Foundations Workshop, Ontology, Semiotics and Practice.

[13]. Shannon, C. E. and Weaver, W. (1949). The Mathematical Theory of Communication. University of Illinois Press, Urbana.

[14]. Shimojima, A. (1996) On the Efficacy of Representation, Ph.D. Thesis. The Department of Philosophy, Indiana University.

[15]. Stamper, R. (1997) Organisational Semiotics, in Information Systems: An Emerging Discipline?,

[16]. Weber, R. (1997) Ontological Foundations of Information Systems, Coopers and Lybrand, Melbourne.

[17]. Wang, Y; Feng, J (2007). FCA assisted IF channel construction towards formulating conceptual data modeling. WSEAS Transactions on Systems, v 6, n 6, June, 2007, p 1159-1167.

[18]. Wobcke, W. (2000). An Information-based Theory of Conditionals. Notre Dame Journal of Formal Logic 41(2): 95-141.