Contexts, Oracles, and Relevance

Varol Akman and Mehmet Surav
Department of Computer Engineering and Information Science
Bilkent University, Bilkent, Ankara 06533, Turkey
Phone: +90 (312) 266 4133 (sec.), Fax: 266 4126
E-mail: {akman,surav}@bilkent.edu.tr

Abstract

We focus on how we should define the relevance of information to a context for information processing agents, such as oracles. We build our formalization of relevance upon works in pragmatics which refer to contextual information without giving any explicit representation of context. We use a formalization of context (due to us) in Situation Theory, and demonstrate its power in this task. We also discuss some computational aspects of this formalization.

1 Introduction

The idea of an oracle comes from the ancient Greece. In the mythical literature, this word had two uses: (i) the place where questions about the future were asked of the gods, and (ii) the priest (or priestess) who gave answers to such questions. In our conception, viz. Computational Situation Theory [38; 39], oracles should correspond to information servers for the objects they represent, and one should be able to pose queries concerning an object to its oracle[1]. For example, in the ultimate interpretation, the oracle related to one of the organizers of an international colloquium, e.g., *Oracle*(Korta), should be able to answer the question "Where did Korta go last Friday?" as "Korta went to Donostia" if it can access that information by some means and as "I don't know" if it has no useful information in this regard.

We are not the first researchers grappling with the idea of having an oracle. One of the earliest uses of oracles goes back to Turing machines. In the analysis of (arguably) the toughest problem in computer science, namely "Is $\mathcal{P} = \mathcal{NP}$?", oracles were heavily used. Interestingly, it turns out that relative to some oracle it can be shown that $\mathcal{P} = \mathcal{NP}$, while relative to some other oracle $\mathcal{P} \neq \mathcal{NP}$ [4]. In our conception, this has two implications: (i) a conclusion obtained via hypothetical oracles is not necessarily unique, and (ii) using oracles we may find answers to very difficult problems under the assumptions embedded in the oracle. We must probably note that the issue of oracle appears not only in the theory of computation but also in practical fields of computer science such as computer networks, client-server systems, etc.

The plan of this paper is as follows. In the next section we will review the emergence of oracles in Situation Theory, a contribution due to Devlin [14; 15]. We will then summarize the concept of relevance (in Pragmatics) in the light of Grice [21], and Sperber and Wilson [33; 34]. Next, we will try to interpret these works in the light of Situation Theory and within our formalization of context [36; 37]. Also in this section a critique of the approach of Devlin will be given. The paper will be concluded with a tentative list of observations[2].

2 Situation Theory Background

Situation Theory is a principled program to develop a unified theory of meaning and information content, and to apply that theory to specific areas of language, computation, and cognition.

One of the most notable motivations of the theory is to provide a *mathematical* theory of meaning. Barwise and Perry [8] claim that for an expression to have meaning, it should convey information. They develop a theory of situations and of meaning as a relation between situations. The theory provides a system of abstract objects that makes it possible to describe the meaning of both expressions and mental states in terms of the information they carry about the external world.

Two major concepts of Situation Theory are infons and situations. Infons are the basic informational units. They should be considered as discrete items of information. Infons are denoted as $\ll P, a_1, \ldots a_n, i \gg$ where P is an n-place relation, $a_1, \ldots a_n$ are objects appropri-

[1] The reader is warned that our notion of oracle is *not* precisely in the spirit of Devlin [15] in which oracles are defined as situations containing all of the information relevant to the object. Our approach considers oracles as information collecting agents.

[2] "What can we do when things are hard to describe? We start by sketching out the roughest shapes to serve as scaffolds for the rest; it doesn't matter very much if some of those forms turn out partially wrong. Next, draw details to give these skeletons more lifelike flesh. Last, in the final filling-in, discard whichever first ideas no longer fit." [29, p. 17]. These ideas of Minsky seem to us a particularly apt way of studying context. (Minsky, by the way, offers several excellent ideas on context in the cited reference.)

ate for the respective argument places of P, and i is the polarity (0 or 1). If $i = 1$ (resp. $i = 0$) then the informational item that $a_1, \ldots, a_n$ stand (resp. do not stand) in the relation P is denoted. Situations are first-class citizens of the theory, and are defined intensionally. A situation is considered to be a structured part of the reality that an agent (somehow) manages to pick out.

Since (real) situations are partially definable objects, abstract situations are proposed to be their counterparts which are more amenable to mathematical manipulation. An abstract situation is defined as a (possibly non-well-founded) set of infons. Given a real situation s, the set $\{\alpha \mid s \models \alpha\}$ is the corresponding abstract situation. Note that s *supports* α (denoted as $s \models \alpha$) means that α is an infon that is true of s.

One important paradigm behind Situation Theory is a *schema of individuation*, a way of carving the world into uniformities. The notions of individual, relation, spatial and temporal location, and further entities depend upon this schema of individuation. Thus, constituents of Situation Theory such as infons, constraints, and situations are determined by the agent's schema of individuation.

Object types are determined over some initial situation. Let s be a given situation. If $\dot{x}$ is a parameter and I is a set of infons (involving $\dot{x}$), then there is a type $[\dot{x}|s \models I]$. This is the type of all objects to which $\dot{x}$ may be anchored (see presently) in s, for which the conditions imposed by I obtain. We refer to this process of obtaining a type from a parameter $\dot{x}$, a situation s, and a set I of infons, as *type-abstraction*. Here, $\dot{x}$ is known as the *abstraction parameter* and s as the *grounding situation*. An infon which contains a parameter, such as $\dot{x}$, is called a *parametric infon*.

Related to parametric infons, there is a formal construct by which we can assign "values" to parameters, i.e., an *anchor*. Formally, an anchor for a set, A, of basic parameters is a function defined on A, which assigns to each parameter T_i in A an object of type T. For example, in

$$\ll \text{goes}, \dot{a}, \text{Boston}, \dot{t}, \dot{l}, 1 \gg,$$

if f anchors $\dot{a}$ to the individual "Sullivan," we write

$$f(\dot{a}) = \text{Sullivan}$$

to denote this anchoring.

The task of anchoring may be considered to be the theory's correspondent to humans' mental processing effort, and is, accordingly, one of the most important components of the theory. In Situation Theory, the flow of information is realized via *constraints*, and anchoring plays the major role in the working of constraints. We represent a constraint as

$$\ll \textit{involves}, S_0, S_1, 1 \gg$$

where S_0 and S_1 are situation-types. Cognitively, if this relation holds then it is a fact that if S_0 is realized (i.e., there is a real situation $s_0 : S_0$), then so is S_1 (i.e., there is a real situation $s_1 : S_1$). For example, with the following constraint c, we might represent the regularity "Smoke means fire":

$$S_0 = [\dot{s}|\dot{s} \models \ll \textit{smoke-present}, \dot{l}, \dot{t}, 1 \gg]$$

$$S_1 = [\dot{s}|\dot{s} \models \ll \textit{fire-present}, \dot{l}, \dot{t}, 1 \gg]$$

$$c = \ll \textit{involves}, S_0, S_1, 1 \gg$$

Here, in order to invoke the constraint, we have to use an anchoring which binds the parameters $\dot{l}$ and $\dot{t}$ to appropriate objects present in the grounding situation, i.e., we have to find a place and time at which there is smoke, and by binding that place and time to $\dot{l}$ and $\dot{t}$, respectively, we get a realization of the constraint[3].

If R is an n-place relation and $a_1, \ldots, a_m$ $(m \leq n)$ are objects appropriate for the argument places $i_1, \ldots, i_m$ of R, and if the filling of argument places $i_1, \ldots, i_m$ is sufficient to satisfy the minimality conditions for R, then for $i \in \{0, 1\}$, the object

$$\ll R, a_1, \ldots, a_m, i \gg$$

is a well-defined *infon*. Here, *minimality conditions* are (for a particular relation) the collection of conditions that determine which particular groups of argument roles need to be filled in order to produce an infon. If $m < n$, the infon is said to be *unsaturated*; if $m = n$ it is *saturated*.

More information on Situation Theory can be found in Barwise and Perry [8], Barwise [6], Barwise and Etchemendy [7], and Devlin [14; 16; 17] (the most up-to-date version of the theory).

3 Oracles as Situations

In this paper, we shall mainly deal with the philosophy behind the nature of information and the ways of gathering *relevant* information, rather than the purely mathematical and technical issues. Thus, one of the best places to find a good reference on the nature of information is Situation Theory. Furthermore, we have a definition of oracle in Situation Theory: Devlin [14, pp. 79–85] defines oracles *as* situations.

In his definition of oracle, Devlin uses the notion a *set of issues* to mean a collection of parametric infons which provides us with an information-theoretic framework for discussing the world or some part of it. By anchoring the parameters of an infon in this set, we obtain an item of information. Clearly, using different sets of issues, we can talk (and obtain information) about different aspects of the world. For example, if we are talking about Korta in the context of a colloquium, we may include stuff like his being the organizer of the colloquium, things happened to him around the time of the colloquium, or the personal characteristics of Korta in order to help one recognize him when he gives his talk, and so on. However, what happened to Korta when he was 5 years old is, probably, not included in the set of issues in this context (unless his talk is about his experiences in his early youth!).

Let Γ be used to denote a particular set of issues and Γ-*infon* to denote any infon that results from anchoring the parameters in an infon in Γ. Then, following is the basic definition of an oracle from Devlin [14, p. 79].

[3] We must caution the reader that we are slightly abusing the notation here. In fact, this particular constraint will be denoted by the expression $S_0 \Rightarrow S_1$ (and this is read as S_0 *involves* S_1), and represents a fact (i.e., a factual, parameter-free infon): $\ll \textit{involves}, S_0, S_1, 1 \gg$.

Definition 1: Oracle$_\Gamma(a)$.
Given an individual (animate or inanimate) or a situation a, the Γ-oracle *of a,* Oracle$_\Gamma(a)$*, is the situation comprising that part of the world and the entire 'body of knowledge' that, within the framework provided by Γ, concerns a.*

Alternatively, $Oracle_\Gamma(a)$ is the 'minimal' situation, s, such that $s \models \sigma$ for any factual, parameter-free Γ-infon, σ, that 'genuinely involves' a. (The scare quotes signal stuff that will be more thoroughly examined in the sequel.)

For example, $Oracle_\Gamma$(Korta) contains all the information related to Korta; i.e.,

$$\ll \text{organizer}, \text{Korta}, \text{ICCS–95}, 1 \gg$$
$$\ll \text{goes}, \text{Korta}, \text{Donostia}, \dot{l}, \dot{t}, 1 \gg$$
$$\ll \text{male}, \text{Korta}, 1 \gg$$
$$\ldots \quad \ldots \quad \ldots$$

Γ is, in some sense, a template of information which determines what portion of *Oracle*(Korta) will be considered. For example, *Oracle*(Korta) contains an enormous amount of information about the medical state of Korta, but if Γ does not discuss these issues, $Oracle_\Gamma$(Korta) should not include such medical information.

In general, the above definition is quite adequate for a philosophical discussion of oracle. However, it has some terms and phrases to be clarified. Devlin elucidates some of them but leaves some vague; these include *genuinely involves*, *minimal situation*, and *entire body of knowledge*. In addition to these terms, the definition of *set of issues* does not seem appropriate in some cases. When we describe our account of oracle (together with the relevance concept), we will offer a somewhat more explicit treatment for these terms.

4 Relevance Theory

One of the vaguest concepts in Devlin's definition is the *relevance* of information. In fact, the issue of relevance is a well-researched topic in various fields of computer science, including mathematical logic [2]. For example, in Information Retrieval (IR), the measurement of the relevance is a major line of research. There are various syntactical approaches to measure the relevance of a term to a document. Until recently, the only respectable methods were the statistical methods, which deal with the frequency of the occurrences of the term. However, lately psychological, epistemic, and semantical considerations are beginning to appear. For example, Park [31] studies the role of context and importance of the relevance to improve information retrieval techniques in public libraries. According to Park, the search criteria for any query should be set according to the users' criteria of relevance. Since different users exhibit different relevance criteria, the query formation is a dynamic task. In this view, context is user dependent and includes psychological and semantical aspects of the user and the search topic. On the other hand, Hearst [24] studies the issue from a Computer Science point of view. He implements a full text information retrieval system in which relevance of the documents is presented to the user in a graphical manner so that the user can see the changes on relevance from the beginning to the end of the text. In this view, the measurement of the relevance is left to the user's cognitive powers.

In this paper, we will base our arguments on Sperber and Wilson [34], in which relevance is mainly considered to be the psychological relevance of a proposition to a context. Their assumption is that people have intuitions of relevance, i.e., they can consistently distinguish relevant from irrelevant information. However, these intuitions are not very easy to elicit or use as evidence, since the ordinary language notion of relevance comes along with a fuzzy (variable) meaning. Moreover, intuitions of relevance are relative to contexts, and there is no way of controlling exactly which context someone will have in mind at a given moment. Despite these difficulties, Sperber and Wilson intend to invoke intuitions of relevance. According to them, a proposition is relevant to a context if it interacts in a certain way with the (context's) existing assumptions about the world, i.e., if it has some contextual effects in some context that are accessible. These contextual effects include: (i) *Contextual implication:* A new assumption can be used together with the existing rules in the context to generate new assumptions; (ii) *Strengthening an existing axiom:* A new assumption can strengthen some of the existing assumptions of the context; and (iii) *Contradicting or eliminating an existing assumption:* A new assumption may change or eliminate some of the existing assumptions of the context.

Here, context is a psychological construct which represents an individual's assumptions about the world at any given time and place, and is supposed to include

1. *Logical information:* The logical inference rules valid in the context that allow us to reason. According to Sperber and Wilson, these rules are deductive. In our view, these rules should also support non-monotonicity.
2. *Encyclopedic information:* Information about the objects, properties, and events that are instantiated in the context [23]. This information in general will help us to form the left-hand side of the inference rules of that context.
3. *Lexical information:* The lexical rules that allow us interpret the natural language utterances and sentences.

Sperber and Wilson talk about "degrees of relevance." Clearly, one piece of information may be more relevant to a particular context than another. In the comparison of the relevance of pieces of information they consider the mental processing effort, i.e., the length of the chain of reasoning and the amount of encyclopedic information involved, and so on. (In the description of our account of relevance, we will give an explicit measure for the processing effort.) In their discussion, the resulting definition of relevance is the following [34, p. 125].

Definition 2: The Relevance Maxim.

1. *An assumption is relevant in a context to the extent that its contextual effects in this context are large.*

2. *An assumption is irrelevant in a context to the extent that the effort required to process it in this context is large.*

The measurement of contextual effects and processing effort is a difficult task due to the problems of qualification of mental effects and effort. Sperber and Wilson say [34, p. 130]:

> The problems involved in measuring contextual effects and processing effort are, of course, by no means specific to relevance theory or to pragmatics. They affect psychology as a whole. However, for relevance theory these problems take on a more specific form. Within relevance theory, the problem is not so much to assess contextual effects and processing effort from the outside, but to describe how the mind assesses it own achievements and efforts from the inside, and decides as a result to pursue its efforts or relocate them in different directions.

In our treatment of relevance, we will try to take advantage of the specific form of the measurement problem, though the character of the measurement task is still blurry.

In the theory of communication developed upon the relevance maxim, context is assumed to be a dynamic mechanism which can be changed during the communication in order to maximize the relevance of the utterance. In this respect, context is left partially-defined in the approaches of Grice, and Sperber and Wilson.

In order to proceed on the computational side, first we need a clarification of the notion of context (the explicit representation of its contributors) so that we can do some computation (e.g., the computation of relevance) over them. Second, the notion of relevance of Sperber and Wilson should be extended to work together with the *partial* information in a context, since the assumption of explicitly knowing everything in the context is quite far from the reality.

In the following section we will try to incorporate the ideas of Sperber and Wilson into a situation-theoretic account and particularize the above issues.

5 Situations and Relevance

In our situation-theoretic account of oracle, we first (re)define what an oracle is, and clarify the critical terms in this definition.

Definition 3: Oracle$_\Gamma(a)$.
Oracle$_\Gamma(a)$ *is the collection of information from the set of available (i.e., existing reachable) situations which is relevant to an object a in the context Γ.*

This definition can be used with a two-phase procedure in an application. We first search over the existing situations (i.e., the situations available in our database) for the occurrence of a. This is done by searching over the infons making up the situations, and whenever an infon containing a, say σ, is found in a situation, say S, we measure the relevance of σ to Γ (using the relevance criteria we discuss in the remainder of this section which is mainly based on Sperber and Wilson's relevance-to-a-context), and then if σ is found to be relevant to Γ, include $< S, \sigma >$ in the oracle. Figure 1 depicts the view of an oracle while an agent B operates as described above.

One of the most important things in the above definition is context. The notion of context was first introduced to AI in a logicist framework by McCarthy in his 1971 Turing Award talk. (This talk was later published as [26].) After that introduction, research on the topic was quite silent until the late eighties. McCarthy published his recent ideas on context in [27; 28]. In that proposal, McCarthy offers three reasons for introducing formal notion of context:

1. The use of contexts allows simple axiomatizations;
2. Contexts allow us to use a specific vocabulary and information about a circumstance; and
3. We can build AI systems which are not stuck with the concepts they use at a given time, since they can transcend the context they are in.

McCarthy and his co-workers (Guha [22; 23] and Buvač [10; 12; 13; 11; 28]) use $\mathtt{ist}(c, p)$ to assert that proposition p is true in context c. In this view, contexts are abstract and first-class objects, and there are some relations and functions working over the contexts (e.g., *more general than* might be used as a relation between two contexts, whereas *specialize-time*(c, t) returns a context in which time is fixed to t). Lifting rules are used to transfer propositions between contexts.

Giunchiglia and his co-workers [20; 19; 9; 18] approach context from a deductive point of view and use natural deduction as the reasoning mechanism over (and inside) the contexts. They offer the notion of a *MultiContext (MC) System.* An MC system is defined to be a pair $< \{c_i\}_{i \in I}, \Delta >$, where $\{c_i\}_{i \in I}$ is the set of contexts and Δ is the set of *bridge rules*. In this respect, context is a triple $c_i =< L_i, A_i, \Delta_i >$ where L_i is the language of c_i, A_i is the set of axioms of c_i, and Δ_i is the set of inference rules that can be used only in c_i. Bridge rules are inference rules, similar to lifting rules, linking different contexts. They are of the form

$$\frac{< A, C_1 >}{< B, C_2 >}$$

allowing us to derive a formula B in context C_2 from a formula A in context C_1.

Other notable works on formalizing context are due to Akman and Tın [1], Attardi and Simi [3], and Shoham [32]. We have surveyed the previous works on context in another paper [36], and due to space limitations, will not go into further details.

We are planning to use the context due to our recent research [36; 37] which grants us logical and encyclopedic information represented in a situation type with parametric and parameter-free infons. One extension to this representation of context is the inclusion of saturated and unsaturated infons by which we can measure relevance in a more naturalistic manner.

Definition 4: Context.
A context Γ is a situation type which contains

1. Constraints: *These are the constructs of Situation Theory which correspond to logical rules (i.e., if-then rules) valid in the context.*

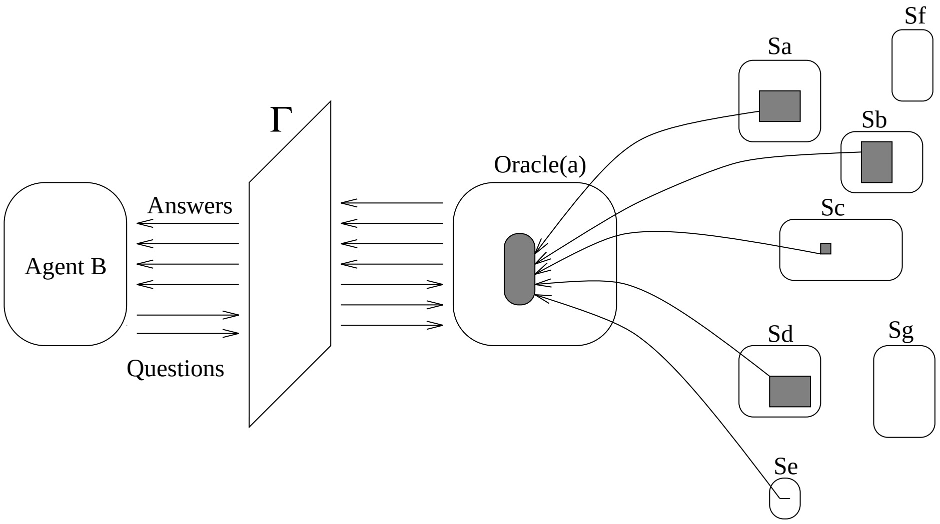

Figure 1: Using oracle(a) with context Γ.

2. Saturated and unsaturated infons: *These are the infons which correspond to the encyclopedic knowledge in Sperber and Wilson's conception of context. Saturated infons represent complete information whereas unsaturated infons represent partial knowledge.*

In the current state of this research, we are not including lexical rules in the context since natural language processing and understanding with computers is still a difficult (but promising) area of research. The inclusion of lexical rules might be waived in the cost of having a natural language-to-situation theory converter, since we consider Situation Theory powerful enough to work with information. Note that a full-fledged conversion from natural language to Situation Theory is, at the current state of the art, impossible. Still, using some shallow parsing methods we can do an incomplete conversion from natural language to Situation Theory. Since the measurement of relevance is still, in some sense, an intuitive task, we can afford the loss of lexical rules at the cost of a (hopefully) small potential error in the estimation of relevance.

In our estimation of relevance, we will try to measure the relevance of σ (cf. Definition 3 ff) to Γ. Here, we use the criterion proposed by Sperber and Wilson, namely, *maximum contextual effect and minimum processing effort*. In the following paragraphs, we interpret the effects of σ on Γ as contextual effects.

The first contextual effect is *contextual implication.* This is exactly the use of constraints available in Γ; if σ invokes any constraint we take it that σ creates a contextual implication in the context. The larger the number of contextual implications the larger will be their relevance.

The second contextual effect is that of *strengthening the existing knowledge.* This can be considered as saturation of unsaturated infons in Γ. In this case, the measurement criteria of relevance include: (a) The number of infons (partially) saturated by the inclusion of σ; (b) The amount of saturation in each saturated infon. In the computation of the similarity of σ and Γ, we can 'sum' the effects of σ on each infon of Γ to get a cumulative contextual effect of σ on Γ.

The last contextual effect is that of *contradicting some existing knowledge in context.* The contradiction of σ with some existing information in Γ is possible only if there is an infon in Γ similar to σ but has the opposite polarity. By similarity, we mean that the relation names of the infons are the same (or synonymous), the saturated fields of the infons are matching (except for the polarity), and there exists a consistent anchoring function which can anchor the parametric fields of these infons.

While considering these contextual effects, we must keep the processing effort as small as possible. In the measurement of processing effort we use the complexity of anchoring function, since an anchoring function might be thought to correspond to the mental operations of humans. As for the complexity of anchoring function we may utilize the size of the anchoring to achieve the desired contextual effects. Notice that this is rather reasonable. First, the mental operations of humans are, in general, very similar to what an anchoring function does: humans individuate relations and objects and fill the gaps in the relations with appropriate individuals, and reason over them. This is very similar (if not exactly the same) to the way an anchoring function operates. Second, from a computational point of view, the size of an anchoring function can easily be determined if the function is available. Needless to say, finding an appropriate anchoring function which creates the desired contextual effects might be quite difficult. (This issue might be considered, in some sense, to be analogous to the unification mechanism of Prolog.) Here, we will not cover the details of this task, but refer the interested reader to Computational Situation Theory [30; 39; 38].

In determining the relevance, a major problem is the measurement method. In our opinion there are three possibilities:

(i) *Relevance as a number in the interval* $[0,1]$*:* This method might be useful in computational applications of relevance. One such application area is IR. The use of certainty factors in assorted applications (with a probabilistic flavor) in AI may also be regarded as analogous to what we have in mind here.

(ii) *Relevance as a number in the set* $\{0,1\}$*:* The nature of relevance is quite inappropriate to make a 0-1 distinction, for we can always talk about degrees of relevance.

(iii) *Defining two (not necessarily disjoint) sets of relevant and not-so-relevant information:* In this view, the set of relevant items contains all 'potentially' relevant items, and the set of not-so-relevant items contains all 'potentially' irrelevant items of information. In some cases, both of these sets may be useful, and we can choose the appropriate one according to our optimistic or pessimistic tendencies.

In our measurement of relevance, we are planning to use the third method, i.e., we will define two sets for relevant and not-so-relevant information, respectively. Our definition criteria are the same with Sperber and Wilson's from an intuitive angle. Technically, we define the relevance of an infon σ to a context Γ.

Definition 5: Relevance.
An infon σ is relevant to a context Γ if has some contextual effects on Γ with a small size anchoring.

Definition 6: Irrelevance.
An infon σ is irrelevant to a context Γ if either it has no contextual effect on Γ or otherwise some contextual effects with a large size anchoring.

Note that, in these definitions, we are still leaving the meaning of two terms open, namely *large* and *small* anchoring, in order to be correct from a philosophical point of view. In a computer implementation of this approach, we can find an appropriate measure for these terms (either by an analysis of the domain or by some simulation results). To offer some insights of the method for various contextual effects, the following examples will be useful[4].

Example 1: Contextual Implications.
Let us consider a context, Γ, which contains the regularity "Birds fly." In Situation Theory, we represent this as

$$\begin{aligned} S_1 &= [\dot{s}|\dot{s} \models \ll \mathrm{bird}, \dot{x}, 1 \gg] \\ S_2 &= [\dot{s}|\dot{s} \models \ll \mathrm{flies}, \dot{x}, 1 \gg] \\ \Gamma &\models \ll \mathrm{involves}, S_1, S_2, B, 1 \gg \end{aligned}$$

where B is the background conditions to support the non-monotonicity [5]. The infon

$$\sigma = \ll \mathrm{bird}, \mathrm{Tweety}, 1 \gg$$

is relevant to Γ, since with the anchoring

$$f(\dot{x}) = \mathrm{Tweety}$$

we can conclude that "Tweety flies" using the constraint supported by Γ. The size of anchoring is 1 and thus the processing effort is minimal. Thus, σ is *relevant* to Γ because of the contextual implication and the small processing effort. □

Example 2: Contextual Implications.
Consider the following imaginary dialog between Sullivan and Korta:

(a) Sullivan: Did you watch the football game on TV5 last night?

(b) Korta: I always watch TV5.

In the context of the conversation, we must have the following information:

(c) A football game was shown on TV5 last night.

(d) If some event is shown on a visual device and someone watches that visual device then he/she also watches the event.

The encoded versions of (a)–(c) are respectively

$$\ll \mathrm{watch}_1, \mathrm{Korta}, \mathrm{the\_football\_game}, \dot{l}, \dot{t}, ? \gg \quad (1)$$
$$\ll \mathrm{watch}_2, \mathrm{Korta}, \mathrm{TV5}, 1 \gg \quad (2)$$
$$\ll \mathrm{shown}, \mathrm{the\_football\_game}, \mathrm{TV5}, \dot{l}, \dot{t}, 1 \gg \quad (3)$$

and the constraint in (d) is c:

$$\begin{aligned} S &= [\dot{s}|\dot{s} \models \ll \mathrm{watch}_2, \dot{p}, \dot{x}, 1 \gg \\ &\quad \wedge\ \dot{s} \models \ll \mathrm{shown}, \dot{y}, \dot{x}, \dot{l}, \dot{t}, 1 \gg] \\ S' &= [\dot{s}|\dot{s} \models \ll \mathrm{watch}_1, \dot{p}, \dot{y}, \dot{l}, \dot{t}, 1 \gg] \\ c &= \ll \mathrm{involves}, S, S', B, 1 \gg \end{aligned} \quad (4)$$

With these definitions, the context of the talk, say Γ, contains the encyclopedic information (3) and the rule (4). Using anchoring f:

$$\begin{aligned} f(\dot{p}) &= \mathrm{Korta} \\ f(\dot{x}) &= \mathrm{TV5} \\ f(\dot{y}) &= \mathrm{the\_football\_game} \end{aligned}$$

we achieve a contextual effect (i.e., the invocation of c). In order to achieve this contextual effect, we have first used the encyclopedic knowledge to satisfy the left hand side of the contextual rule. Thus, this time, the size of the anchoring is larger than the case in Example 1, though a size of 3 is still *small*. □

Example 3: Saturating Unsaturated Infons.
Let Γ contain the unsaturated infon

$$\ll \mathrm{goes}, \mathrm{Korta}, \dot{l}, \dot{t}, 1 \gg$$

which states that in context Γ, Korta goes to some place at some time. Over this context, the infon

$$\sigma = \ll \mathrm{goes}, \mathrm{Korta}, \mathrm{Donostia}, \dot{t}, 1 \gg$$

[4] In our thinking and examples about context, natural language has a primary role. We believe that in order to be successful a theory of context should at least be successful in the domain of natural language. As van Benthem remarked [40, p. 159]: "Natural language is the most characteristic human vehicle for conveying information [...] Linguistic structures somehow mirror informational structures, but also, linguistic processing consists of mechanisms for transmitting and transforming such structures." We think that context is one such fundamental structure. A similar view is apparent in Stalnaker [35].

is relevant to Γ, since, with anchoring

$$f(\dot{l}) = \text{Donostia}$$

we can saturate the location parameter of an existing infon in Γ, and the processing effort is minimal. □

After all, our definition of oracle may seem quite similar to that of Devlin, but has the following critical differences and clarifications. First, Devlin considers oracles to be *intensional* constructs. In our approach we are trying to define oracles extensionally in order to render this notion more useful in computational tasks. Second difference is on the conception of Γ. We consider Γ to be the context in which an oracle query takes place, whereas Devlin envisions Γ as a set of parametric infons which are to be saturated in $Oracle_\Gamma(a)$. This gives us the power of using contextual information in the measurement of relevance. Third, Devlin does not give the explicit definitions of terms such as *relevance of information*, *body of knowledge*, and *genuinely involves*; in our approach the counterparts of these terms are explained reasonably explicitly. Finally, in our approach the related information is presented together with the situation in which the information appears. This will enable one to make further analysis of the information if he is so required.

As mentioned in the previous sections, our work is not yet fully developed and the following remain as short-term research problems.

- *Accessible information:* In Figure 1, we are talking about collecting information over existing situations. In reality, it is clear that the scope of the search is limited since nobody is expected to act omnisciently. Furthermore, although some information may be reachable in principle, access to this information may be denied or use of this information may be illegal.
- *Comparison of* Γ *and the situation which contains* σ*:* In the measurement of relevance, we have always considered the relevance of σ to Γ. However, although it may prove to be difficult, comparing (and searching for relevance) Γ and the situation containing σ might result in a better understanding and measurement of relevance.
- *Size of* Γ*:* So far, we have not talked about the size of Γ. Yet, this definitely effects the measurement of relevance.

6 Conclusion

In this paper we have tried to give an outline and highlights of our research on context and the relevance of information. While doing this, we have used the idea of oracle to broaden our view of the subject. Since our formalization is not yet complete, we offer the following tentative conclusions:

- Our approach is, we believe, quite similar to the human's way of measuring relevance. In other words, we are using a cognitive theory of relevance in a technical way.
- We employ a formalization of context within Situation Theory. This interpretation of context is compatible with Sperber and Wilson's Relevance Theory. Using our formalization of context together with Relevance Theory might be beneficial. The reader may find computer applications of situation-theoretic concepts shallow or cursory, but we are encouraged by the preliminary results in Computational Situation Theory [30; 39; 38].

Acknowledgments

We would like to thank Fausto Giunchiglia for his encouragement and expert advice. Akman's research is supported in part by a NATO SFS grant (TU-LANGUAGE).

References

[1] V. Akman and E. Tın. What is in a context? In E. Arıkan, editor, *Proceedings of the 1990 Bilkent International Conference on New Trends in Communication, Control, and Signal Processing, volume II*, pages 1670–1676. Elsevier, Amsterdam, Holland, 1990.

[2] A. R. Anderson and N. D. Belnap, Jr. (with contributions by J. M. Dunn et al.). *Entailment: The Logic of Relevance and Necessity (2 volumes)*. Princeton University Press, Princeton, NJ, 1975.

[3] G. Attardi and M. Simi. Proofs in context. In *Principles of Knowledge Representation and Reasoning: Proceedings of the Fourth Conference*, 1994.

[4] T. Baker, J. Gill, and R. Solovay. Relativizations of the $P = ?NP$ question. *SIAM Journal on Computing*, 4(4):431–442, 1975.

[5] J. Barwise. Conditionals and conditional information. In E. C. Traugott, C. A. Ferguson, and J. S. Reilly, editors, *On Conditionals*, pages 21–54. Cambridge University Press, Cambridge, UK, 1986. Reprinted in [6].

[6] J. Barwise. *The Situation in Logic*. Number 17 in CSLI Lecture Notes. Center for the Study of Language and Information, Stanford, CA, 1987.

[7] J. Barwise and J. Etchemendy. *The Liar: An Essay on Truth and Circularity*. Oxford University Press, New York, NY, 1987.

[8] J. Barwise and J. Perry. *Situations and Attitudes*. MIT Press, Cambridge, MA, 1983.

[9] P. Bouquet and F. Giunchiglia. Reasoning about theory adequacy: A new solution to the qualification problem. Technical Report 9406-13, Istituto per la Ricerca Scientifica e Tecnologica, Trento, Italy, 1994. To appear in *Fundamenta Informaticae*.

[10] S. Buvač and I. A. Mason. Propositional logic of context. In *Proceedings of the Eleventh National Conference on Artificial Intelligence*, pages 412–419, Washington, DC, 1993.

[11] S. Buvač. Ambiguity via formal theory of context. Available via anonymous ftp from `sail.stanford.edu` as `/buvac/ambiguity-via--formal-theory-of-context.ps`, 1995.

[12] S. Buvač, V. Buvač, and I. A. Mason. Metamathematics of contexts. *Fundamenta Informaticae*, 23(3), 1995. To appear.

[13] S. Buvač and R. Fikes. A declarative formalization of knowledge translation. Technical Report KSL-94-59, Stanford University, Knowledge Systems Laboratory, Stanford, CA, 1994. Available via anonymous ftp from `sail.stanford.edu` as `/buvac/a-declarative-formalization-of--knowledge-translation.ps`.

[14] K. Devlin. *Logic and Information*. Cambridge University Press, New York, NY, 1991.

[15] K. J. Devlin. Oracles in situation semantics. In J. Barwise, J. M. Gawron, G. Plotkin, and S. Tutiya, editors, *Situation Theory and Its Applications, volume 2*, chapter 3, pages 41–49. Center for the Study of Language and Information, Stanford, CA, 1991.

[16] K. J. Devlin. Situations as mathematical abstractions. In J. Barwise, J. M. Gawron, G. Plotkin, and S. Tutiya, editors, *Situation Theory and Its Applications, volume 2*, chapter 2, pages 25–39. Center for the Study of Language and Information, Stanford, CA, 1991.

[17] K. J. Devlin. Infons as mathematical objects. *Minds and Machines*, 2(2):185–201, 1992.

[18] F. Giunchiglia. Contextual reasoning. *Epistemologia*, XVI:345–364, 1993. Short version in *Proceedings of the IJCAI-93 Workshop on "Using Knowledge in Its Context"*, Chambery, France, 1993.

[19] F. Giunchiglia and L. Serafini. Multilanguage hierarchical logics, or: How we can do without modal logics. *Artificial Intelligence*, 65(1):29–70, 1994.

[20] F. Giunchiglia and R. W. Weyhrauch. A multi-context monotonic axiomatization of inessential nonmonotonicity. In P. Maes and D. Nardi, editors, *Meta-Level Architectures and Reflection*, pages 271–285. Elsevier, Amsterdam, Holland, 1988.

[21] H. P. Grice. Logic and conversation. In S. Davis, editor, *Pragmatics: A Reader*, chapter 19, pages 305–315. Oxford University Press, New York, NY, 1991.

[22] R. V. Guha. *Contexts: A Formalization and Some Applications*. PhD thesis, Computer Science Department, Stanford University, Stanford, CA, 1991.

[23] R. V. Guha and D. B. Lenat. CYC: A midterm report. *AI Magazine*, 11(3):32–59, 1990.

[24] M. A. Hearst. *Context and Structure in Automated Full-Text Information Access*. PhD thesis, Computer Science Division, University of California, Berkeley, CA, 1994.

[25] V. Lifschitz, editor. *Formalizing Common Sense: Papers by John McCarthy*. Ablex, Norwood, NJ, 1990.

[26] J. McCarthy. Generality in artificial intelligence. *Communications of the ACM*, 30(12):1030–1035, 1987. Reprinted in [25].

[27] J. McCarthy. Notes on formalizing context. In *Proceedings of the Thirteenth International Joint Conference in Artificial Intelligence*, Chambery, France, 1993. Available via anonymous ftp from `sail.stanford.edu` as `/pub/jmc/context-2.tex`.

[28] J. McCarthy and S. Buvač. Formalizing context (expanded notes). Technical Note STAN-CS-TN-94-13, Stanford University, Stanford, CA, 1994. Available via anonymous ftp from `sail.stanford.edu` as `/buvac/formalizing-context.ps`.

[29] M. Minsky. *The Society of Mind*. Simon and Schuster, New York, NY, 1985.

[30] H. Nakashima, H. Suzuki, P. K. Halvorsen, and S. Peters. Towards a computational interpretation of situation theory. In *Proceedings of the International Conference on Fifth Generation Computer Systems*, pages 489–498, Institute for New Generation Computer Technology, Tokyo, Japan, 1988.

[31] T. K. Park. Toward a theory of user-based relevance: A call for a new paradigm of inquiry. *Journal of the American Society for Information Science*, 45(3):135–141, 1994.

[32] Y. Shoham. Varieties of context. In V. Lifschitz, editor, *Artificial Intelligence and Mathematical Theory of Computation: Papers in Honor of John McCarthy*, pages 393–408. Academic Press, Boston, MA, 1991.

[33] D. Sperber and D. Wilson. Mutual knowledge and relevance in theories of comprehension. In N. V. Smith, editor, *Mutual Knowledge*, pages 61–82. Academic Press, London, UK, 1982.

[34] D. Sperber and D. Wilson. *Relevance: Communication and Cognition*. Harvard University Press, Cambridge, MA, 1986.

[35] R. Stalnaker. Pragmatics. In D. Davidson and G. Harman, editors, *Semantics of Natural Language*, pages 380–397. Reidel, Dordrecht, Holland, 1972.

[36] M. Surav and V. Akman. Contexts and situations. Technical Report BU-CEIS-9433, Department of Computer Engineering and Information Science, Bilkent University, Ankara, Turkey, 1994. Available from the *Computation and Language E-Print Archive* (`http://xxx.lanl.gov/cmp-lg`).

[37] M. Surav and V. Akman. Modeling context with situations. In P. Brezillon and S. Abu-Hakima, editors, *Working Notes of the IJCAI-95 Workshop on "Modelling Context in Knowledge Representation and Reasoning"*, number LAFORIA 95/11, pages 145–156. Institut Blaise Pascal, Paris, France, 1995.

[38] E. Tın and V. Akman. BABY-SIT: A computational medium based on situations. In C. Martín-Vide, editor, *Current Issues in Mathematical Linguistics*, pages 299–308. Elsevier, Amsterdam, Holland, 1994.

[39] E. Tın and V. Akman. Computational situation theory. *SIGART Bulletin*, 5(4):4–17, 1994.

[40] J. van Benthem. Natural language: From knowledge to cognition. In E. Klein and F. Veltman, editors, *Natural Language and Speech*, pages 159–171. Springer-Verlag, Berlin, Germany, 1991.