

Towards a theory of natural conceptualization

A.J.J. van Breemen, J.J. Sarbo, and Th.P. van der Weide

Radboud University Nijmegen, The Netherlands

Abstract. The focus of this paper is on the early phases of ER-modeling consisting of the primary conceptualization of the underlying application domain. To this end we introduce a process model for the generation of meaningful concepts for a domain description. In virtue of its close relation with cognitive activity, this process model also enables the modeler as well as the user to comprehend the concepts of the resulting domain in a natural way. Beyond this goal, natural conceptualization opens the possibility for the introduction of a uniform representation enabling the efficient combination of knowledge obtained from different stake holders during a modeling process.

1 Introduction

At the start of writing a collaborative paper for an ER conference the authors are gathered around a table. All of them have a general notion of the problem at hand –conceptualization–, but each of them is experiencing the problem differently. Interaction Man experiences the problem in terms of stake holders, roles and utility. He has been working on structured methods for retrieving as much reliable information out of his informants as possible and to get it validated. Model Man experiences the problem in terms of concepts swimming around in clusters, that need to be brought to greater unity. He has been working on a theoretically informed model that enables him to get as much unity as the concepts allow. Also present is Reflection Man. He has been talking with his both co-authors separately and got convinced that both have something of value to add to the conceptualization issue, but he also learned how they value each others work. Interaction Man claims that Model Man ventures a model that is far too theoretical and will not yield any fruit of practical value. It is just a big head resting on weak legs. Model Man on the other hand claims that Interaction Man is far too practical to yield any sustainable result on the long term in our quickly changing culture. It is like a body moving around almost without brains. At this point Reflection Man is ready to settle for a sequencing of both contributions: Interaction Man gathers the clusters of concepts that are brought to unity by Model Man. That is how they experience the problem at the start of their collaborative enterprise.

During the first meetings everything goes according to the expectations of Reflection Man. Since both stake holders are mainly concerned with their own contribution and regard the work of the other as being not their responsibility,

they simply divide the writing space and discuss the content in terms of paragraph headings. But, as work progresses and the pages get filled, Model Man and Interaction Man start to get an interest in the line running through their collective argument. They have to, if only because there is no chance of acceptance of the paper without a line of argument. So they start close reading each others contribution. This greatly affects the next meetings. The stake holders start to raise pertinent questions about each others work from their own point of view. Thus they greatly enlarge their focus. This also is the moment they try to look again at their own work through the eyes of the other. As a result they sort out the ingredients they have to work with. Then they start to figure out in what possible ways both approaches complement each other. As a consequence they start questioning Reflection Man's assumption that the output of Interaction Man is the input for Model Man. The relation might be far more intimate and they find out that the relation will lie somewhere on the continuum between a strict sequence, on one extreme, and a complete merge on the other. Eventually Interaction Man, with his practical bend, notices that the writing process they went through seems to match the model proposed by Model Man and suggests that this offers a possibility for a first practical test. Model Man in response satisfies his need for harmony by adding that in that case they must not forget to pay attention to the interactional aspects and the practical constraints that forced them not to postpone decisions indefinitely.

The above sketch shows some important characteristics of conceptualization. It is a process in which we aim to make explicit and reconcile knowledge in a (semi-)formal model. The knowledge involved has been generated with the aim of realizing *goals* in processes that themselves serve a common goal. Conceptualization itself also serves the common goal. Its aim implies that from the viewpoint of the common goal, it offers an excellent occasion to execute *control*.

The knowledge involved in each process is tightly knit with the goal it serves to realize. This shapes the attitude of the stake holders of different domains. It is only natural that, as a rule, they will understand and value other processes from their own habitual point of view. Reconciliation starts with the decision to consciously subsume knowledge needed for the different processes under the common goal. But this only expresses the intention. In order to realize the intention, it is necessary to explicate the domain specific knowledge in such a way that the stake holders of related domains are able to understand and value what is going on. This is most easily achieved if each domain is explicated according to a *uniform* natural model. For, in that case the stake holders get acquainted with the form in which knowledge is conceptualized when their knowledge processes are modeled. This familiarity with the form comes in handy when stake holders of different domains partake in Information System (IS) building. It is here that our process model for natural conceptualization fits in. Its main characters are:

- 1) It reduces the continuous stream of reality to a finite amount of *types* of meaningful interpretation moments;

- 2) It conceptualizes *hierarchically*, not just by means of collections. The hierarchy is not dictated by a previously accepted ontology, but arises naturally in the course of the analysis [7];
- 3) Its *goal* orientedness provides means to ensure terminological consistency by keeping track of the processes in which common terms figure.

In actual practice the conceptualizer is not either of the type Interaction Man or Model Man. (S)he must have reconciled both attitudes. This person can be best compared with the cultural- or social anthropologist doing fieldwork: (s)he must be a Jack of all Trades. The main difference being that whether it will be a representation oriented or an intervention oriented project does not depend on the choice of the investigator, but on the assignment agreed upon with the contractor.

Below we will present our process model for natural conceptualization. Since the model is founded upon the processual semiotics of the American polymath Ch. S. Peirce (1839-1914) we will start in section 2 with a treatment of some basic notions: process (2.1) and sign and habit (2.2). This treatment is also the occasion to argue that our way of conceptualization offers organizations the opportunity to execute self control. The more technical part of the paper starts in section 3. In this section we outline the model, treat it from the perspective of cognitive activity (3.1) and from the perspective of sign recognition (3.2). Next we give it a logical interpretation (3.3). An illustrative example is presented in section 4. Finally, in section 5 we will draw conclusions.

2 How to control conceptualization?

If we confine our scope just to ER-modeling, it is possible to say that ER-models are used to create an accurate reflection of the structure of part of the real world in a database. But already if we only modestly extend our scope and ask what did trigger the modeling of a domain, we are forced to a more elaborate answer. The first thing to ask is what we do when we model, is it capturing real world events in a reflection or is it a matter of modeling our organization's future interactions with the world? Even with a narrow view on ER-modeling the answer must be that, at the very least, it is not only a matter of representation. For, although ER developed in order to remove redundancy and keep data consistent, if we implement a (normalized) model, we affect the behavior of everybody who is going to work for some reason with the realized information system.

With the rapid integration of information systems across departments, the growing model of entities and relations became hard to read. Contact with the users' demands was restored by specifying separate business processes, followed by a de-normalization of the overall ER-model from the perspective of the needs of the different roles engaged in those processes. Dimensional Modeling [6], for instance, nicely illustrates this trade off between data-demands and user-demands. Notice that this procedure presupposes the existence of a normalized ER-model.

But how to proceed if we do not have a normalized model and/or the domain to be modeled is extensive, hosting different perspectives, stated in terms that

may be vague or that lead to contradictions, in an environment where differences of interest seriously hamper progress [3]? Here *controlled* conceptualization may make a difference. In the remainder of this section we will deal with our background notions as a preparation for the presentation of the model with which we deem controlled conceptualization feasible.

2.1 Processes and conceptualization

The concept of process is central to our approach. This is not primarily so because conceptualization itself can be regarded a process or because it is part of an embracing process aiming at the realization of an IS. The primary reason for taking process as a central concept resides in the fact that all we feel, think or do, is only felt, thought or done as part of some process(es). From the point of view of meaning as actualized meaning, processes are the basic units of meaning, not terms or concepts. The latter generalize over different processes by dissociation from any particular process. This will be clear if we realize that disambiguation as a rule proceeds by providing a proper context, thus embedding the ambiguous sign in the process intended by the utterer. Stake holders will be inclined to think about their concepts from the perspective of the processes in which they are embedded. So, explicating processes is a quite natural starting point for conceptualization. Since processes figure so prominently in our approach some words about our Peirce inspired understanding of the concept of process is called for. Hulswit [4] (p. 193) succinctly summarized the Peircean concept of process thus:

A process is “a continuous sequence of events that derives its unity or internal order (distinguishing it from other processes) from a final cause, which directs the sequence to some end state which itself may evolve.”

Several notions in this definition deserve some elucidation.

A process is a continuous sequence of events:

- a) Continuous because events in a process are not separate from each other.
- b) Events because there are distinguishable parts of a process that can be articulated as facts expressed in propositions.

Unity or internal order is derived from a final cause:

- a) The goal of the process, the desired end state, provides the Reason of the process.
- b) Only what contributes to the realization of the goal ought to be part of the model that describes the process.

The final cause directs the sequence to some end state:

An event starts up a goal oriented habit that governs the process towards its terminating event.

The end state itself may evolve during the process (recursively):

The terminating event needs not to be fixed beforehand in detail, it may be general or vague.

Compare the process of registering a new client at an insurance company, with the process of designing and building a house or a process of learning from experience by repetition.

Organizations or interest groups always exist in order to achieve some end by a diversity of means; this makes it fairly possible to regard the organization as an encompassing process embracing all kind of sub-processes that are at any stage more or less subservient to the common goal. Conceptualization is one of those sub-processes. It is however a rather peculiar sub-process that ultimately aims at an increased articulation of all involved processes in order to be able to increase the quality and efficiency of the overall system. Much more can be said about the consequences of our interpretation of processes for conceptualization. Here we suffice with one last remark. By taking the final goal as criterion for distinguishing what does belong to a given process from what does not belong to it, we are better able to make conflicts of interest visible at an early stage in IS design. This applies to interest conflicts in processes that serve the common goal (e.g. between management and teachers in education going digital). But this also applies to conflicts as a result of goals that do not serve the common goal in a direct way (e.g. oil companies and environmental obligations) or even hinder it (e.g. issues of authority).

In this section we looked at processes in a very general way, now we zoom in on the different kinds of aspects that we must distinguish in all the kinds of sign processes, including the sign processes pertinent to IS.

2.2 Signs, habits and conceptualization

Peirce did hold from very early on that we have no power to think without signs (cf. CP 5.265).¹ But if every thought always partakes in a sign, then the flow of thought, as far as it is connected and influential, is representable by an argument. So, it need not wonder us that throughout his life he kept refining his thought on signs, leaving us a tremendous corpus devoted to the study of signs in many of its aspects. This most certainly is not the place to trace the development of that thought or to lay bare the ramifications of the assumption.² In the space at our disposal we will not even have the possibility to deal with the bare essentials of the core of his sign theory, but at the same time we will have to present an advancement of the basic theory which comes down to a transfer of the concept of ‘sheet’ from his logic to his semiotics. We start with the later and will afterwards introduce the relevant terminology ‘on the fly’. In section 3 we will go into more detail.

In analogy with the ‘Sheet of Assertion’ (S_A) Pierce introduced in the context of his logic (cf. CP 4.396), we introduce a ‘Semiotic Sheet’ (S_S). The difference between the two is that the S_A represents what is taken for granted to be true of the Universe of Discourse by all participants in an exchange of ideas, while the S_S represents all that is believed by a group of participants or by an individual.

¹ A reference to [8] is given by volume and paragraph, separated by a point.

² For more information on Peirce: <http://members.door.net/arisbe/arisbe.htm>.

Let us assume that the S_S has at least two modalities. The *actual* S_S holds what at a given moment actually is believed, it is the state the S_S is in at any given moment. The *potential* S_S contains all beliefs that may become actualized some time.

Now, under the assumption that all thought is in signs, lets try to describe what happens if something demanding attention gets inscribed on a S_S in the state it is actually in. See fig. 1 for a summarization of technical terms (left) and their approximate counterparts in mundane language (right). Although the 9 terms suggest that they are signs on their own, this is not the case. They are best regarded as analytical moments that surface if we are going to analyze goal oriented processes or argumentations. To give an impression of our broad interpretation of an argumentation: Taking a defensive stance is the conclusion derived from an approaching fist as a minor and a suppressed major, i.e. the habit that determines this particular reaction.

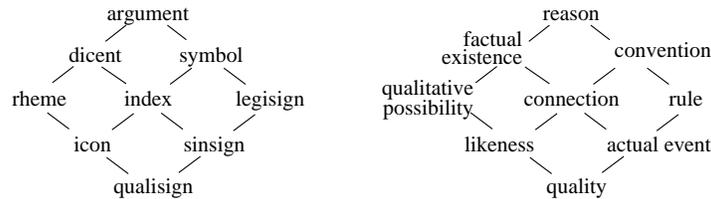


Fig. 1. Peircean sign aspects and corresponding mundane terms

The terms of fig. 1:

The *qualisign* position indistinctly involves all that is needed to reach a conclusion. Analytically it involves what is offering itself for interpretation (the effect), the actual state the S_S is in and all that may be contributed by the S_S written upon in the process of interpretation. State and effect appear as indistinct *qualities*.

The *icon* position discriminates the state and effect according to their form, but without any interpretation. It is the single occurrence of a potential *likeness*. Looked at it from a bottom up perspective the process can go on many directions, none of which is recognized at this point.

The *sinsign*³ position indicates the co-occurrence of the state and effect on a particular S_S as an *actual event*. Commonly known as Token.

The *index* position indicates the actual *connection* of all moments involved in the interpretive process, by this it expresses the permanence of a process on a S_S . At the same time it expresses the latent connection between what is actualized in this process and what is potentially contained in this actualized S_S .

³ 'Sin' as in 'singular'.

The *legisign* generalizes the occurrence (sinsign) of the instances of a form (icon). In learning this is a bottom up process, but if the form is known the actual event triggers a *rule*-like habit that classes all the instances of that form with a type. A legisign only surfaces by means of its instances, but as a habit it has a real influence.

The *rheme* position indicates the moment the form (icon) in its here and now actuality (sinsign) addresses the S_S in order to retrieve all possible ways in which to deal with it, to contextualize it, but only potentially so. Creativity heavily draws upon this moment: figuring out how to possibly deal with the input (*qualitative possibility*).

The *symbol* position indicates the moment a familiar sign (legisign, involving a sinsign-icon) gets connected (index) with a habit of interpretation that connects the sign implied in the effect with its object. Object here used as in “What is the object of your remark?” The stronger this *convention*, the harder it is to be creative. At this point however the habit is still not operational.

The *dicent* indicates the position in which a specific possibility is asserted for actual use in the argument in a particular case. Out of all possible interpretations one is actually asserted (*factual existence*).

The *argument* position indicates at least two moments. On the one hand it indicates the whole process leading to the conclusion/response as its *reason*, on the other it states the conclusion/response, ready to enter a subsequent process.

What does this approach offer to conceptualization? In order to establish that it is useful to introduce a third modality of the S_S , the modality of *law* or operational habit. What is inscribed on a S_S in this modality has real influence and as such fulfills the essential function of a sign, according to Peirce:

“It appears to me that the essential function of a sign is to render inefficient relations efficient, – not to set them into action, but to establish a habit or general rule whereby they will act on occasion” (CP 8.332).

But isn't that the kind of sign we want to develop if we are building an IS, at a level as much integrated as possible? In order to realize that, we regard our Domain Experts as Semiotic Sheets that contain the knowledge we need. The Conceptualizer can be regarded a sheet on which the habits of Interaction Man and Model Man are written. When we analyse we write the results on different sheets, for each process to be modeled one. We do this by combining the sheets of domain experts and the Conceptualizer until all that is needed for the processes is made explicit. We merge and sequence the different sheets on a sheet in which the common goals are regulative. Practical constraints, conflicts of interest and ambiguities will arise in the process. But they have a *reason* and can be discussed in an understandable way, since they can be connected to what is going on. That offers prospects for resolution. The question we pose may be summarized thus: Is it possible to develop a robust strategy for conceptualization by taking a Peircean semiotic account of cognitive activity as a starting point?

3 The model: a recap

The theory of knowledge representation assumed in this paper was first introduced in [1]. Promising applications have been presented in [9] and [10]. According to [1], knowledge arises from the observation of ‘real’ world phenomena. Such a phenomenon is an interaction between an external *effect* and the observer, appearing as a *state* (the actual S_S). State and effect, which are in principle independent hence a *duality*, are two types of entities that are qualities. Let us mention that there may be any number of qualities involved in an interaction, but according to this theory those qualities are always distinguished in two collections, which are treated as single entities. The interpretation of the simultaneous appearance of the input qualities, as knowledge, is modeled as a *process*.

The interaction between qualities is a pre-requisite for sign recognition. But a quality functions as a sign, and therefore is a sign, *if and only if* it will become interpreted as such. Hence the interpretation of a quality as well as of an interaction of qualities, as a potential sign, is the other condition for sign recognition, as a process. An interaction between qualities which develop to signs will be called a *sign interaction*. Following the received view of cognitive theory [2], perceived qualities will be called *qualia*.

3.1 The processing schema of cognitive activity

The ‘goal’ of cognitive activity, as a process, is the establishing of a meaningful relation between the input qualities. In order to recognize such a relation, we have to separate the simultaneously appearing two types of input qualia from each other, realize their meaning with information from the context ([C]), the potential state of the actualized S_S (*complementation*) and establish a relation between their contextualized meaning (*predication*). Assuming the input appears as a ‘primordial soup’ ($[q_1 \ q_2]$), the separation of the state and effect qualia additionally requires their type-wise identification (*sorting*), and the representation of their collections independently from each other (*abstraction*).

In sum, the stages of sign recognition, as a process, can be defined as follows (cf. fig. 2). We use square brackets to indicate that an entity is not yet interpreted as a sign, and no bracketing or the usual bracket symbols, if *some* interpretation is already available.

- (1) *sorting*: $[q_1], [q_2]$
the identification of the two types of qualia in the ‘primordial soup’;
- (2) *abstraction*: q_1, q_2
the separation of the collections of the two types of qualia;
- (3) *complementation*: $(q_1, C), (q_2, C)$
the linking of the qualia with information from the context ([C]);
- (4) *predication*: $(q_1, C)-(q_2, C)$
the establishment of a relation between the completed qualia.

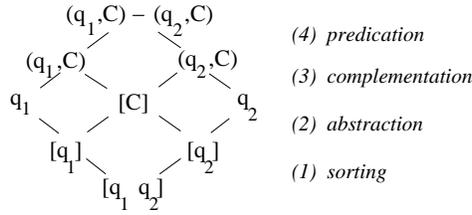


Fig. 2. The processing schema of sign recognition

3.2 The process model of sign recognition

In order to establish the meaning of a phenomenon, the perceived qualities have to be analyzed:

- a) in themselves (what do we know about them from earlier observations) and,
- b) in relation to each other (what is the *reason* of their interaction).

In [9] we pointed out that the two types of analyzes can be defined as *isomorphic* instances of the processing schema. Accordingly, we will refer to them as process-*a* and process-*b*.

In process-*a* the context ([C]) is assumed to be the memory, representing combinatory information about the perceived qualia (which qualia occurred with which other qualia in earlier observations). The functioning of the memory is modeled as follows. The input qualia trigger the memory, which in turn generates a response. This response can be either in the sense of agreement, referring to the recognition of the input as ‘known’; or, in the sense of possibility, indicating a potential meaning of the input, which can be qualified as ‘not known’ (the input quale is not ‘known’ as such an entity). This is in conformity with the assumption of cognitive theory that in an observation the brain may selectively focus on some of the input qualia, hence not focus on others. In the end, this process generates a representation of the input state and effect qualia, which is denoted by A and B , respectively, if the qualia are ‘known’ or *focused*, and, by $\neg A$ and $\neg B$, if they are ‘not known’ or *complementary*. Notice the ambiguous use of A and B , as an instance as well as a type of quale. A ‘ \neg ’ prefix indicates that a quale is not recognized as ‘known’ (it is recognized as ‘not known’).

Process-*b* is an exact copy of process-*a*, except its ‘goal’, which is the interpretation of the relation between the input qualia that are in the focus (A , B), in the light of the qualia which are complementary ($\neg A$, $\neg B$). In accordance with this ‘goal’, the context ([C]) contains relational information about the input qualia.

In process-*b*, the important representation moment is step 3. In this step, the input qualia are linked with the complementary signs, in accordance with the ‘goal’ of this process and the duality of phenomena. This explains why there can be a relation between A and $\neg B$, and $\neg A$ and B , and why there is no relation between A and $\neg A$, or B and $\neg B$. Indeed, A and $\neg A$ (but also B and $\neg B$) arise

due to the same input trigger, indicating that the two signs are *not* independent. In process-*b*, sign recognition is completed in step 4, in which, a relation between *A* and *B* is established. The three relations, which correspond to the three types of interactions between the input qualia, can be characterized by means of the meaning of their constituents, as follows (an interaction, which is a relation, computationally, is denoted by a ‘-’ symbol):

$A-\neg B$:

A is ‘known’, but *B* is ‘not known’ (completion of the input state);

$B-\neg A$:

B is ‘known’, but *A* is ‘not known’ (completion of the input effect);

$(A,\neg B)-(B,\neg A)$:

both *A* and *B* are ‘known’ (establishment of the input relation).

If both *A* and *B* are ‘not known’, interpretation terminates, meaning that sign recognition, as a process, does not actually happen.

Control can be executed at three different levels. First, in the classification of the qualia as complementary or in the focus. Second, in the mediation by the relevant context, in which the context *implicitly* determines the relation between *A* and *B* through the common meaning of $\neg A$ and $\neg B$ (both are complementary signs of the same phenomenon). Third, in actualizing that relation in a ‘proposition’, which is a hypothesis (*A* ‘is related to’ *B* or, briefly, *A* ‘is’ *B*) that can be subject to validation in a next process.

The processing schema, which is our *uniform* representation, can be used recursively, by means of interpreting the final sign degenerately as a quality. Recursive analysis can be necessary for the recognition of complex phenomena.

3.3 Logical analysis

The above interpretation of process-*b* as three types of meaningful interactions illustrates the *completeness* of the processing schema. We interpret this aspect as an indication of a *meaningful* conceptualization of the observed phenomenon. The completeness of sign recognition, as a process, becomes even more clear from the logical analysis of the processing schema. The hidden agenda of this section is a tacit introduction of logical concepts, in the process model of sign recognition. What makes this step especially important is that logical concepts have a *precise* meaning. We pursue a logical analysis of process-*b*, but the results trivially apply to process-*a* as well.

An essential element of a logical interpretation of process-*b*, is the abstraction of a common meaning of the two types of input qualia, which is the concept of a *logical variable*. Because we have two types of qualia, which are independent, the logical interpretation requires the introduction of two variables. These are denoted by *A* (state) and *B* (effect). The collections of qualia which are in the focus and which are complementary, are represented, respectively, by means of a logical variable which is stated positively and negatively. State and effect qualia which are in the focus are denoted, respectively, by *A* and *B*; those which are

complementary by $\neg A$ and $\neg B$. Notice the use of ‘ \neg ’ as logical negation, which is relative difference with respect to the collection of a type of qualia, represented as a set. For example, the complementary subsets of the A -type qualia are denoted by A and $\neg A$, ambiguously. Conform the above mapping, the logical meaning of the sign interactions can be defined as follows.

Step 1 consists in the identification of the two types of input qualia which are in the focus, A and B , that we consider to be *synonymous*. As now state and effect are indistinguishable, as variables, the difference between $[q_1]$ and $[q_2]$ is expressed by means of their interpretation as a constituent and event, respectively. As both uniformly refer to A and B , in their logical representation we make use of a logical operator, indicating the type of relation between the input qualia. $[q_1]$ is denoted by $A+B$ and $[q_2]$ by $A*B$, representing their meaning as a co-existence and co-occurrence, respectively.

Step 2 is the abstraction of the input qualia which are in the focus as constituents (q_1), irrespective of the actually co-occurring other type of qualia or, alternatively, in relation with any other qualia which are not in the focus. This is logically denoted by the expression: $A*\neg B$, $\neg A*B$. It is *this* perspective that makes the two signs synonymous, which is indicated by the “,”. The other result of abstraction is an expression of the input as an abstract event, defined by the possible co-existence or ‘compatibility’ of its abstract constituents (which are now interpreted differently). The corresponding logical expression is: $A*\neg B+\neg A*B$. The context ($[C]$) is defined by the complementary qualia represented as a co-existence ($\neg A+\neg B$) and a co-occurrence relation ($\neg A*\neg B$). The synonymous representation of these signs is an expression of the secondary meaning of the context.

Step 3 amounts to the complementation of the abstract constituents (q_1) with the context ($[C]$) or, alternatively, the actual meaning of the input qualia as constituents ((q_1,C)). For example, the meaning of $\neg A*B$ in context, is defined by the qualia complementing this abstract representation, which are: A and $\neg B$. Alternatively, the actual meaning of A , as a constituent, is defined by A itself and also by $\neg B$, the complementary qualia, linking A with B implicitly (as the relation between A and B is not yet realized, the qualia denoted by B cannot contribute to this interpretation of A). As the two expressions of A , as an actual constituent, are related to each other by the relation of co-existence, the logical meaning of (q_1,C) can be represented as $A+\neg B$. For the same reason, as in q_1 , the two representations of (q_1,C) are interpreted in the model as synonyms. The representation of the abstract compatibility relation (q_2) in context obtains the interpretation of the input as a characteristic property appearing as an event ((q_2,C)). That event, which is a representation of the interaction between A and B , is alternatively signified by the interaction between $\neg A$ and $\neg B$. The two signs signify the interaction which is in the focus, respectively, positively and negatively. This is logically denoted by the expression: $A*B+\neg A*\neg B$.

Finally, step 4 is the generation of the relation between the input qualia which are in the focus ($(q_1,C)-(q_2,C)$). This final sign is logically represented as a syllogistic proposition, A is B (a potential premise for a subsequent argument).

The logical signs of the processing schema are displayed in fig. 3. Notice the presence of *all* Boolean relations on two variables, reinforcing our conjecture about the completeness of the cognitive process ('0' and '1', which are omitted, can be defined as representations of a 'not-valid' and a 'valid' input, respectively). The results of our analysis show that logical signs, hence also the concepts of cognition, as a process, can be defined as an interaction (relation) between *neighboring* signs. In fig. 3, such signs are connected with a horizontal line.

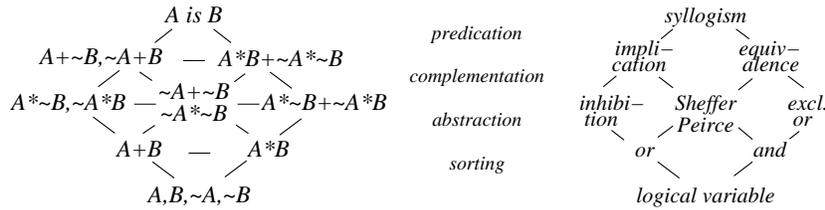


Fig. 3. The logical interpretation of the processing schema

The above classification can be interpreted as a formal logic, defined by a single operation, *relative difference* ('\'), which has three different types. The first is relative difference with respect to the type of qualia themselves (*sorting*); the second is with respect to the other type of input qualia (*abstraction*); the third is with respect to the input as a whole (*complementation*). For example, $[q_2]=: A*B$; $q_2=[q_1]\backslash[q_2]: \neg A*B+A*\neg B=(A+B)\backslash(A*B)$; $(q_2,C)=\neg q_2: A*B+\neg A*\neg B=\neg(A*\neg B+\neg A*B)$. In the last case, relative difference with respect to the context ([C]) is interpreted as relative difference with respect to the universe ('1'), which explains the use of negation (' \neg ') in the definition of (q_2,C) .

The processing schema can be offered a semiotic interpretation as well (see sect. 2.2). For example, $[q_1]$, which is a representation of the input as a constituency relation, is expressive of the *likeness* aspect of the input constituents, as 'parts', with respect to the input, as a 'whole'; $[q_2]$, representing the simultaneous occurrence of the input qualia, their appearance as an event that happens now, has the aspect of *simultaneity*. A more semiotic analysis of sign recognition, as a process, can be found in [11]. The isomorphism between the logical relations and Peirce's classification is another indication for the completeness of the processing schema, as well as of the importance of its interpretation moments.

4 An example

There are two dimensions along which our theory of conceptualization can be explored. The first is related to its potential for merging knowledge (concepts) obtained by the different stake holders. This can be illustrated with the specification of a *common* problem, by a conceptual specialist and some domain

expert(s). Full individual specifications, each expressing a certain point of view, are merged to a single meaningful relation, by merging concepts of identical meaning aspect to a single sign, through coordination. The second way of exploration capitalizes on the potential of the theory to be applied recursively, which is beneficial for the conceptualization of complex phenomena. This section give an illustration of such a problem. Let us emphasize that both types of exploration are a direct consequence of the uniform representation provided by the theory outlined in this paper.

4.1 Conceptualization of a complex problem

Assume our task is the generation of a meaningful representation of a company's strategy. From an abstract point of view such a strategy can be characterized by the company's reaction on supply and demand. According to the theory of this paper, such a reaction can be conceptualized as a relation between supply and demand where supply *can* be interpreted as the products produced by the company, which are available for some time (state), and demand as the suddenly appearing request dictated by the customers (effect).

The simultaneously appearing supply and demand (input qualia) define our company as a phenomenon. The interpretation of the two types of qualia in themselves (process-*a*) provides us with a definition of *A* and *B*,⁴ and through the interpretation of the relation between them (process-*b*) we may obtain a meaningful conceptualization of our company. That relation may also rely on complementary factors like the economic perspective taken, either profit maximization or purchasers satisfaction, which views in turn are intimately related to the more primitive concepts of price and selling events, that themselves can be specified as phenomena. The next section is an attempt to introduce a specification for selling events [5]; an analogous treatment of the concept of price is left to the reader.

4.2 Nested phenomenon: selling event

A selling event is between a dealer and a customer. The specification below is given from the dealer's perspective; a description from the purchaser's point of view is possible, but it may be less general, as the purchaser's motivation for buying or not buying a product can be more difficult to set out.

We assume that a selling event, as a phenomenon, is defined by the purchaser's selection of certain products and services. The products can be interpreted as a *state*, in which, the services appear as an *effect*. For example, dealer and customer are having a conversation while they are looking at collection of products, when suddenly the dealer recognizes (e.g. 'sees' in a saccade), the purchaser's selection of the services he or she wishes to be delivered with the products. Additionally we assume that a selling event may also be subject to

⁴ We refer to the status of a sign by means of the corresponding logical expression.

complementary factors like the different forms of the transfer of ownership and the various judicial conditions a service may have to comply with.

The goal of this section is not the analysis of a concrete selling event, however. Our focus is on the interpretation of the relation between product and service, in general. Accordingly, we will assume that the input contains abstract qualia, which are: product (A), service (B), form of ownership ($\neg A$) and judicial conditions ($\neg B$). For example, $\neg A$ may refer to purchase, lease or rental, and $\neg B$ to copyright, know-how and organizational conditions (e.g. a computer network, as a product, can be sold only as a whole, not in parts). The representation of abstract concepts, as qualia, is a non trivial question that we cannot discuss in detail, because of lack of space. A possible solution of this complex problem has been presented in [10]. The abstract concepts of selling are introduced as stages of the recognition process of an abstract selling event, as a sign (cf. fig. 2).

[q_1]= $A+B$: A product (A) can be sold with a service (B); or, a service (B), for example, a know-how, can be available through a certain product (A).

[q_2]= $A*B$: A product may provide a certain service only; or, a service can be available only in combination with a certain product (product-service dependency, as an actual event).

[C]= $\neg A+\neg B$, $\neg A*\neg B$: The relation between the complementary factors of selling events like the various forms of ownership and the corresponding judicial issues ($\neg A+\neg B$), and how those forms of ownership are regulated by law ($\neg A*\neg B$).

$q_1=A*\neg B$, $\neg A*B$: Product specification like catalog information about products ($A*\neg B$) and services ($\neg A*B$), in general.

$q_2=A*\neg B+\neg A*B$: The reconciliation of purchaser's need for products and services, from the dealer's point of view. This is the concept of *bargaining*, which is a law-like relation or *rule* between products ($A*\neg B$) and services ($\neg A*B$) that can be sold in combination. This rule includes the possibility that products and services can be sold independently from each other (in conformity with the 'or' operation in the logical expression of q_2). For example, one may buy A without B ($A*\neg B$) or, B without A ($\neg A*B$), or both.

Notice that in this case the law-like relation identified as 'bargaining', a concept that most of us are familiar with, is *automatically* revealed by this conceptualization process (except its name, of course). This indicates that in cases in which we are not familiar with the concepts they are generated by the model.

(q_1,C)= $A+\neg B$, $\neg A+B$: Product and service are the actual constituents of a selling event. This includes the full specification of product and service according to the purchaser as well as the dealer, tacitly indicating their agreement (cf. bargaining in context).

(q_2,C)= $A*B+\neg A*\neg B$: The characteristic property involved in any selling event, defined by a suitable combination of product and service ($A*B$) and nothing else ($\neg A*\neg B$) or, alternatively, the relation between the selected product and service ($A*B$), in the light of their ownership forms and judicial conditions like the transfer of ownership through leasing or contract ($\neg A*\neg B$).

$(q_1, C) - (q_2, C) = A \text{ is } B$: The meaningful description of the input phenomenon, as a selling event.

4.3 Main phenomenon continued

The meaning of the nested phenomenon can be represented in the nesting phenomenon, degenerately, as a quale. As a result, the meaningful concept of a ‘selling event’ is reduced to a complementary information $(\neg B)$ ⁵ mediating supply and demand to their interpretation as the *company’s supply* and the *customers’ demand* and, finally, to a meaningful representation of our company’s strategy. The context of price $(\neg A)$ and selling events $(\neg B)$, is an expression of their close interaction in a market mechanism $(\neg A * \neg B)$, but also of their possible influence as additional regulative factors, like environmental and political demand $(\neg A + \neg B)$.

The interpretation of the context along a continuum marked by ‘liberal’, on the one side, and ‘state controlled’, on the other, may provide a representation of our company (*A is B*’ as a conclusion), as a more commercial profit oriented enterprise or a more non-profit organization, respectively.

5 Conclusions and further research

For systematic conceptualization of a domain, a process model is needed. This has several reasons. It minimizes unwanted ambiguity of the terms used by consistently extracting them from the processes in which they function. It enhances communication with and between the end-users about their domains by the use of concepts in context. And it facilitates strategic decision making in an early stage of IS building by enabling the re-thinking of the processes that make up organizations; after all goals can be reached in different ways.

This approach is natural due to its close relation with cognitive activity interpreted as a sign process. The finite amount of types of meaningful interpretation moments (cf. fig. 1) facilitates the development of a common language between designer and user. The order relation between the interpretation moments as well as between the processes increases the possibility to control conceptualization. The goal oriented nature of our processual model is well equipped to uncover hidden tensions by systematically keeping check of the relations between ends and means.

Our task for the future can be briefly stated thus: In order to *really* estimate the value of our approach for conceptualization we will have to put our model to *actual* use. In order to make our approach of practical use in IS building we will have to study the relations between our model and formalisms of ER. Since our semiotically derived model allows a logical interpretation, prospects are good. But as the saying goes, the proof of the pudding is in the eating.

⁵ Now $\neg B$, but also A , B and $\neg A$ refer to the qualia of the *nesting* phenomenon.

References

1. J.I. Farkas and J.J. Sarbo. A Logical Ontology. In G. Stumme, editor, *Working with Conceptual Structures: Contributions to ICCS'2000*, pages 138–151, Darmstadt (Germany), 2000. Shaker Verlag.
2. S. Harnad. *Categorical perception: the groundwork of cognition*. Cambridge University Press, Cambridge, 1987.
3. S.J.B.A. Hoppenbrouwers, H.A. (Erik) Proper, and Th.P. van der Weide. A fundamental view on the process of conceptual modeling. In L. Delcambre, C. Kop, H.C. Mayr, J. Mylopoulos, and O. Pastor, editors, *Conceptual Modeling (ER-2005)*, volume 3716 of *LNCS*, pages 128–143. Springer-Verlag, 2005.
4. M. Hulswit. *From Cause to Causation : A Peircean Perspective*, volume 90 of *Philosophical Studies Series*. Kluwer, 1999.
5. H. Jonkers, L. Groenewegen, M. Bonsangue, and R. Buuren. A language for enterprise modelling. In M. Lankhorst, editor, *Enterprise architecture at work: Modelling, Communication, and Analysis*, pages 83–113, Cologne, Germany, 2005. Springer-Verlag.
6. R. Kimball. *A dimensional modeling manifesto*. <http://www.dbmsmag.com/9708d15.html>, 1997.
7. D. Mayhew and D. Siebert. Ontology: The discipline and the tool. In G. Büchel, B. Klein, and Th. Roth-Berghofer, editors, *Proceedings of the First Workshop on Philosophy and Informatics*, pages 57–64, Cologne, Germany, 2004.
8. C.S. Peirce. *Collected Papers of Charles Sanders Peirce*. Harvard University Press, Cambridge, 1931.
9. J.J. Sarbo and J.I. Farkas. A linearly complex model for knowledge representation. In U. Priss and D. Corbett, editors, *Conceptual Structures: Integration and Interfaces (ICCS'2002)*, volume 2193 of *LNAI*, pages 20–33, Borovets (Bulgaria), 2002. Springer-Verlag.
10. J.J. Sarbo and J.I. Farkas. Towards a theory of meaning extraction. In H. D. Pfeiffer, H. Delugach, and K. E. Wolff, editors, *Proceedings of ICCS'04*, pages 55–68, Huntsville (Alabama, USA), 2004. Shaker-Verlag.
11. J.J. Sarbo, J.I. Farkas, and A.J.J. van Breemen. Natural Grammar. In R. Gudwin and J. Queiroz, editors, (*in press*) *Semiotics and Intelligent System Development*, pages 1–20. Idea Group Publishing, 2006.