A Triadic Model of Information Flow

Uta Priss

Indiana University Bloomington, upriss@indiana.edu

Abstract. This paper combines aspects of formal concept analysis, Barwise & Seligman's information flow theory and a Peircean notion of triadic signs. The goal of the paper is to formalize a model of communication that includes explicit, formal representations of signs, conceptual entities that are denoted by the signs and contexts in which the signs are used. Each triad of signs, conceptual entities and contexts contains a concept lattice for the relationship between signs and conceptual entities. Any basic communicative act, such as speaking or naming of an object, usually consists of several triads whose contexts are connected via infomorphisms in the sense of information flow theory. By explicitly describing signs and their conceptual entities, a theory is developed that can describe how linguistic units are disambiguated in natural communication. An example demonstrates how this theory can be applied to merging of ontologies that could be used by software agents in e-commerce.

1 Introduction

Several theories have been established in the past to formalize the relationships between concepts, contexts and situations. Formal concept analysis (FCA, [7]) describes how concepts and concept lattices can be defined within a context or concept lattice in terms of objects and attributes (extensions and intensions); how they relate to each other and which logical clauses can be stated about the attributes. Information flow (IF, [2]) describes how information is transmitted between different concept lattices (or "classifications"). This can lead to a network of concept lattices connected via communication channels which are themselves concept lattices. Conceptual graphs (CG, [11]) consist of concepts and their relations and thus facilitate the description of conceptual structures such as complex concepts, events, processes and situations. Situation theory (ST, [1], [5]) describes relations between contexts (or "situations"). While all of these theories can be used to represent meta-information, such as information about the relationship between concepts and their representation as signs, this relationship is not explicitly modeled in any of these theories. This paper describes a Peircean sign triad that can be used to model this relationship explicitly in combination with tools and methods provided by the existing theories. The model reported in this paper is related to triadic concept analysis ([9], [8]) but the emphasis is not on constructing triadic lattices (or "trilattices") but instead on utilizing standard binary lattices to represent the information flow among triads.

This paper presents neither a strictly philosophical analysis nor a strictly formal mathematical theory nor a typical computer science description of an application. However, it has aspects of all three of these. The goal of the paper is to describe a semiformal ontological foundation that facilitates an explicit representation, use and differentiation of representations, conceptual entities and contexts in applications. From a computer science viewpoint this paper could be considered work in progress because no system has been built yet to test the assertions made here. The goal of this paper is not to establish a new system but instead to discuss the ontological foundations in section 2.

Concerning the philosophical aspects, it should be remarked that numerous people have in the past used sign triads with different meanings. A summary of some of the triads can be found in [10]. It is not attempted in this paper to argue that any interpretation of Peirce's sign theory is more appropriate than or superior to another. The interpretation of Peirce's sign triad developed in this paper is selected because it fits best with the applications of this paper. It is irrelevant for this paper whether that is an accurate interpretation of Peirce or not. In the interpretation used here, Peirce's triad does not correspond to one triad but to two: one for the sign processing in an individual's mind and one for the sign representation in a consensual sign system such as a language. Other triads are possible and not excluded in this paper.

The aim of this paper is to propose a foundation for solving some of the problems that occur in the context of communication among software agents. An important application is the development of software agents in e-commerce applications. To facilitate communication between such artificial agents, some representation of "context" is required because context is what humans use in resolving ambiguities in communication. Current approaches to represent context rely heavily on human labor (compare for example the CYC project [4]). This paper thus attempts to provide some structures that improve an understanding of what is involved in information flow among contexts and that might improve the efficiency and feasibility of contextual representations.

2 The elements of the sign triad

As mentioned in the introduction, in this paper the Peircean triad is mapped to two related triads: The first one for the sign processing in an individual's mind is called the *individual (sign) triad*. The second one for the sign representation in a consensual sign system is called the *consensual (sign) triad*. The reason for this is because interaction with an external world, such as perception and experience, is restricted to individuals. Even if humans collectively perceive the same object (event, situation, or other), it is not certain whether they internally represent it in the same manner. To the contrary, it is more likely that two humans who are looking at the same object, are listening to the same piece of music or report the same event, will perceive, store and recall it differently. On the other hand, language is a collectively built, standardized sign system that is developed around a collectively assumed consensual reality, which individuals use to communicate among each other. This leads to the construction of two closely related types of sign systems: one for individuals, one for consensual groups.

The notions of "individual" and "consensual sign" require clarification. Individuals can be human agents as well as software or other artificial agents or even animals. Different subtypes of triads can be distinguished for the different types of individuals. The term "consensual sign" is used instead of "(conventional) sign" to include innate and other communication mechanisms. Traditionally signs are defined to be conventional, i.e., intentional and learned. But the communication mechanisms of non-human agents are sufficiently similar to human communication to justify a weaker definition of a “consensual sign” that can be innate, learned or conventional. Artificial agents usually communicate via consensual signs that are not established by the agents themselves but instead by their human creators. Animals usually communicate via innate sign systems, such as tail movement of dogs or cats. The sign triads presented in this paper are in general applicable to any kind of communicating agent, including biological agents (cells, microorganisms) and other living beings.

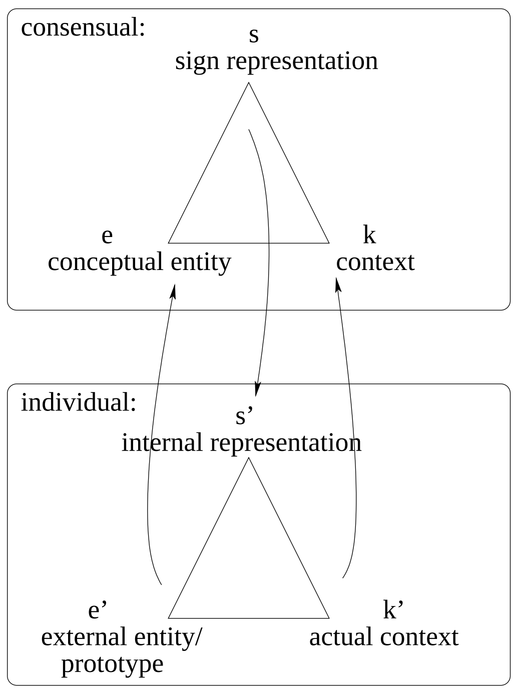

Fig. 1. The two sign triads

2.1 The individual sign triad

The Peircean object, representamen and interpretant correspond to the following notions in the individual sign triad. The object corresponds to an entity of an external world or a prototypical representative of such an entity as perceived, remembered or reflected on by an individual agent. It is called *external entity*. This entity can be any kind of unit: a physical or imagined object, an emotion or sensory perception, an experience, an observed or imagined relation, a remembered event or situation, and so on.

An external entity has attributes (features or characteristics) that are perceived by the individual. But a complete list of features is not usually generated or may be impossible to be generated in general.

The representamen corresponds to the mental or *internal representation* of the sign in the individual's mind. An internal representation can be in form of words of a language, in form of 0's and 1's in a computer or in form of other visual, auditive, tactile or other sensory-based representations. Internal representations are individually defined and for human agents their study belongs into the realm of psychology.

The interpretant corresponds to a so-called *actual context*. The notion of "context" is a difficult one. For example, there have been two international conferences on Context [3] in recent years. The papers in their proceedings contain research on modeling applications with context, philosophical and linguistic implications of context and so on but essentially no attempts of definitions or of precise descriptions of context. In this paper, "context" is used similarly to "situation" in situation theory, which does not ultimately define what contexts are but gives some indication of how to formalize and describe them. In general there is no accepted convention about what necessarily constitutes a context. It can be assumed that an actual context contains information about the individual agent, time, space and social and other rules that describe the appropriateness of using an internal representation for an external entity. Contexts can be nested. A context can have several distinct subcontexts, for example, when an individual reflects about entities that are not currently present. In that case, there would be a subcontext related to when the external entity was originally encountered and another subcontext for the current physical and mental context of the individual.

Most aspects of information representation and processing are related to consensual sign use, such as natural or artificial languages. Apart from user and usability issues, which are of course important but not relevant for this paper, the individual triad can be ignored in many information representation formalisms. The main reason for mentioning the individual triad in this paper is to explicitly state that external objects, actual contexts and internal representations are not directly contained in a consensual sign triad.

2.2 The consensual sign triad

The second and more important triad used in this paper is a consensual sign triad. The elements of a consensual triad are not directly concerned with actual experiences. The elements are easier to formalize because the formation of a consensual sign system, such as a natural language, relies on implicit standardization and rules. The relationships within and between the constituents of a consensual triad can thus be expected to follow some regularities. The consensual sign triad consists of a *conceptual entity* (in the object position), a *consensual context* (in the interpretant position) and a *sign representation* (in the representamen position), which are defined as follows.

A conceptual entity can be any kind of unit, such as a physical or abstract object, a concept, a conceptual relation, a situation, an event or a process, and so on. Conceptual entities (or "concepts") are best represented within some formal representation system, such as FCA or CGs. Within a consensual context, conceptual entities have an extension, which is a set of exemplary or prototypical instances, and an intension, which is a list of attributes or logical constraints.

Even though conceptual entities, their instances and attributes are usually represented in some natural or formal language, they can be considered language-independent. Concepts may arise within a specific cultural and linguistic context. It may not be possible to translate such concepts using a single word or phrase but it is usually possible to sufficiently describe concepts in other languages with the help of explanations and by providing background information. For example, the German notions of "bunt", "Gemütlichkeit" and "Schadenfreude" are considered lexical gaps in English but their meaning can be described to English speakers. Connotations and artistic value (such as in poetry) may be lost in language-independent representations but they are phenomena that involve more complex networks of sign triads and are not discussed in this paper.

The second component of consensual sign triads are sign representations, which are physical representations of signs. A physical representation can be visual, auditive, tactile or otherwise sensory-based. It must have some consistence over at least a short amount of time and must be communicable. Examples are linguistic signs, traffic signs, conceptual graphs and symbols. Linguistic signs can be words, phrases, sentences or longer texts. The meaning of a sign representation depends on the consensual context in which it is used. For example, a traffic sign on a rubbish dump has either no meaning (in which case it is not a sign) or a different meaning than the same sign on a street corner. Similarly, sign representations from extinct languages whose original meaning is lost have either no meaning or a different meaning in a modern context.

As mentioned in section 2.1, contexts are not easily defined. A consensual context contains information about the community that establishes the consensual use of a sign and the conditions for the appropriate use of the sign within the community. Part of the problem of describing contexts is the question as to what to include and what to omit. In natural communication as little information as possible about contexts is explicitly mentioned. For the communication among artificial agents, this poses a significant challenge. How can artificial agents "know" what constitutes a context and whether contexts used by different agents are compatible? Specifying everything that could possibly be relevant is an enormous task even if rules and reasoning mechanisms are used because an ontology such as CYC [4] would be required.

Even though the notion of "triad" suggests a symmetry, the relationships among the three components are not fully symmetrical because the relationship between conceptual entities and sign representations depends on the establishment of a consensual context. The next section investigates these relationships further.

2.3 Mappings within and between sign triads

Given the three components of a sign triad, (sign) representation, conceptual entity and context, the questions arise as to how the three are related and how they connect different sign triads into a network. In the easiest case, two assumptions are fulfilled. The first assumption is that the sign triad consists of a set of conceptual entities, a set of signs and exactly one context. In this case, the context can be represented as an FCA context and the conceptual entities as FCA concepts. The second assumption is that the sign representations are disambiguated, which means that each representation refers to exactly one conceptual entity. In this case the sign representations can be mapped onto the concept lattice. If several sign representations are mapped to the same concept, they are called synonyms.

The second assumption is only reasonable with respect to clearly delimited micro-contexts. For example, a word such as "window" can in general refer to all kinds of windows and metonymically to parts of windows. In a micro-context, which can be presumed by a sentence or part of a sentence, the word is disambiguated. For example, in "he looked through the window", "window" refers to the glass of a specific window whereas in "she painted the window", it refers to the frame of a specific window. Plural words refer to sets of instances, which can usually be considered extensions of a single conceptual entity. A normal conversation can involve numerous sequential or co-occurrent micro-contexts, sometimes even in one sentence. For example in "he looked through the window which she painted", the word "window" shifts between its glass meaning and its frame meaning. This can be considered a shift between two micro-contexts. If the contexts are represented as FCA contexts then shifting between contexts can be modeled as infomorphisms [1]. The special cases where one context is part of or an instance of another context can equally be modeled by infomorphisms, which fulfill additional conditions.

If the two assumptions hold, the relationship between individual and consensual sign triad can be modeled as such an infomorphisms between contexts. In that case, the infomorphism maps the types of the consensual context (i.e. the attributes of the conceptual entities) onto the types of the actual context (i.e. the attributes of the external entities). Meta-conditions about contexts, such as conditions of time, space, culture etc are modeled in a separate concept lattice, in which the actual context has to be an instance of a subconcept of the consensual context. The actual context thus instantiates the consensual context and the external entities instantiate the conceptual entities. The sign representations can then be mapped onto the internal representations. The arrows in figure 1 refer to these instantiations and mappings.

Without the two assumptions, cases can be modeled where several contexts jointly contribute to a shared set of conceptual entities. Triadic FCA [9] provides mechanisms that formalize these more general cases by using conceptual entities as FCA objects, sign representations as FCA attributes, and contexts as FCA conditions. As another simpler case, a global context could be assumed, such as a specific language, time and cultural group and so on. Then instead of the first assumption from above, triads with single sign representations and sets of contexts and conceptual entities could be studied. The set of contexts of such a triad would be a list of contexts in which a specific sign representation is used. In this way the "connotation" of a sign can be represented because a representation can connote a context (for example, the use of offensive words implies certain contexts) or a representation can connote other conceptual entities besides the ones it denotes based on polysemy and homography.

3 An application to e-commerce ontologies

In e-commerce applications, software agents are supposed to fulfill some of the duties of traditional sales people. Human agents can easily resolve ambiguities that might arise in sales offers, order requests or in other business negotiations even if the agents use slightly differing terminology. The ambiguity resolution is achieved by using contextual clues and/or by asking the negotiating partner for clarification. In contrast to human agents, software agents need explicit representations of contextual information, which can, for example, be provided by using AI ontologies. The following example illustrates some of the problems.

A consortium consisting of members of Edinburgh University and industrial partners developed a high-level ontology for commercial applications called *Enterprise Ontology* [6]. It provides about 90 abstract terms with definitions and relations. For example, "Customer" is defined as a "union of Actual Customer and Potential Customer" where "Actual Customer is the Role of the Legal Entity agreeing to exchange a Sale Price for a Product in a Sale". Companies can use the Enterprise Ontology as a high-level modeling framework. For more practical applications, the high-level ontology has to be combined with other ontologies that are more detailed with respect to specific business transactions. An example is the XML standard *cXML* developed by the company Ariba. Figure 2 shows a cXML representation of a fictitious transaction.

It is not straightforward to merge cXML and the Enterprise Ontology. Figure 3 shows some potential mappings between terms. The easiest problem solved is the renaming of terms, such as Supplier/Vendor. More significant is the problem that conceptual entities and/or contexts cannot directly be mapped. First, explicit information in one ontology can be implicit in the other. For example, cXML does not explicitly specify a customer. The information about the customer is implicitly contained in the ShipTo and the BillTo address fields of an Order Request. Second, conceptual entities can be subdivided differently in each ontology. For example, cXML differentiates the BillTo and ShipTo aspect of a customer, probably because these two can be different in a large organization which has a separate department responsible for billing. The Enterprise Ontology does not consider ShipTo and BillTo aspects. Third, the overall structure of the organization of the ontologies can be different. The top level structure in cXML is the distinction between Catalog, which contains all information about the supplier and the products (under Index), and Order, which is the primary activity in this ontology. The reason for this is because cXML is created for on-line transactions. The top level structure of the Enterprise Ontology, on the other hand, is more general according to activity, organization, marketing and so on. (This is not obvious in figure 3 which only contains elements of marketing.)

It is noticeable however that the underlying conceptual structures are not entirely different between the two ontologies. If it was possible to separate the general conceptual entities that are involved in sales transactions from the contextual aspects and the viewpoint that each ontology imposes, then a merging should be straightforward. This is not surprising because both ontologies model the same processes. Thus in IF terminology the conceptual entities (or "tokens") are identical, the difference lies in the sign representations (or "types") and the contexts (or the type hierarchies). It should thus be possible to construct an IF channel between the two ontologies. While the precise formulation of such a channel is left for future research, the essential step of separating conceptual entities, sign representations and context is described in an example in the rest of this section.

```xml
<OrderRequest>
   <OrderRequestHeader orderID="jdg43" orderDate="2000-11-10"
    type="new">
       <Total>
            <Money currency="USD">50.00</Money>
       </Total>
       <ShipTo>
            <Address>
                <Name xml:lang="en">John Doe</Name>
                <PostalAddress name="doe">
                    <DeliverTo>John Doe</DeliverTo>
                    <Street>315 10th St</Street>
                    <City>Indianapolis</City>
                    <State>IN</State>
                    <PostalCode>47400</PostalCode>
                    <Country>USA</Country>
                </PostalAddress>
            </Address>
       </ShipTo>
       <BillTo>
            <Address>
                <Name xml:lang="en">John Doe</Name>
                <PostalAddress name="doe">
                    <DeliverTo>John Doe</DeliverTo>
                    <Street>315 10th St</Street>
                    <City>Indianapolis</City>
                    <State>IN</State>
                    <PostalCode>47400</PostalCode>
                    <Country>USA</Country>
                </PostalAddress>
            </Address>
       </BillTo>
       <Shipping>
           <Money currency="USD">5.00</Money>
           <Description>Priority Mail</Description>
       </Shipping>
       <Tax>
           <Money currency="USD">3.00</Money>
           <Description>IN Sales Tax</Description>
       </Tax>
       <Payment>
           <PCard number="123456789" expiration="2001-01-01">
        </Payment>
    </OrderRequestHeader>
    <ItemOut>
        <ItemID>
            <SupplierPartID>1234</SupplierPartID>
        </ItemID>
    </ItemOut>
</OrderRequest>
```

Fig. 2. An Order Request in cXML

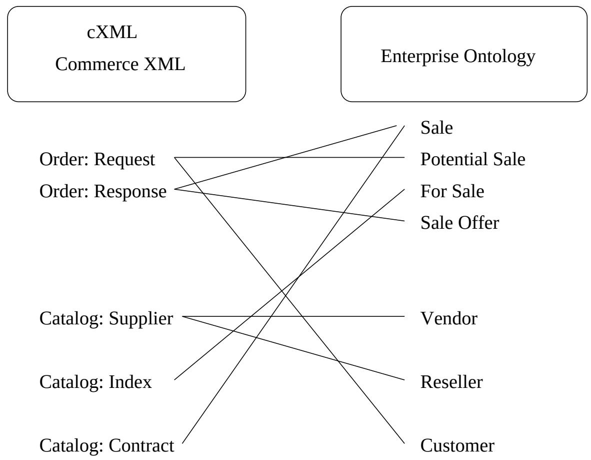

Fig. 3. Merging cXML and the Enterprise Ontology

The Enterprise Ontology considers four different types of Sales: (actual) Sale, Potential Sale, For Sale and Sale Offer. The Order Request in figure 2 is not a sale because, for example, the credit card has not yet been validated. It is neither a For Sale nor a Sale Offer because those are modeled from a vendor's viewpoint, which means that the vendor is willing to agree to the sale but the customer is either unknown or only a Potential Customer. So, the only choice left is a Potential Sale but that conceptual entity is more general because in a Potential Sale everything is potential in contrast to the Order Request in which the customer is certain. With other words, there is no direct equivalent to a cXML Order Request in the Enterprise Ontology.

The sign representations (i.e. the terms defined in the ontology), the conceptual entities (i.e. the formal definitions in terms of relations to other entities and features or constraints) and information about contexts can be separated as follows. A sale according to the Enterprise Ontology, involves a fixed set of potential conceptual entities: a customer, a vendor, a price and a product (compare figure 4). Whether each of these is potential or actual depends on the specific context. Here "context" does not refer to the global context of "English language, global time, culture etc", but instead to very specific contexts, such as whether the sale is potential in the future, currently negotiated or successfully performed.

This information is separately stored in a lattice of contexts and represented in figure 5. Each of the six contexts could be further instantiated by actual contexts as described in the previous sections. The subconcept-superconcept relation in the lattice is based on the conditions that all constraints are inherited by more specific contexts and that all conceptual entities of specific contexts either instantiate their higher level counter parts or are not present in the higher level context. The lattice shows the viewpoint differences between the two ontologies. The Enterprise Ontology is modeled from a vendor's viewpoint whereas cXML is modeled from a customer's viewpoint. The lattice demonstrates how the cXML Order Request is related to the Sale contexts described in the Enterprise Ontology.

The example also demonstrates that sign representations need to be kept separately because they change depending on context and they can even be misleading. For example, in the contexts Order Search, Order Request and Sale, the Customer is called "Actual Customer" according to the Enterprise Ontology. In the other ones, Customer is either missing or called "Potential Customer". According to the Enterprise Ontology Customer is the union of Actual Customer and Potential Customer but conceptually every actual customer is also a potential customer in future sales or potential sales.

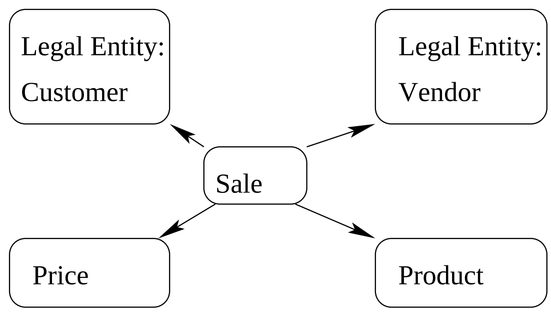

Fig. 4. The conceptual entities involved in a Sale

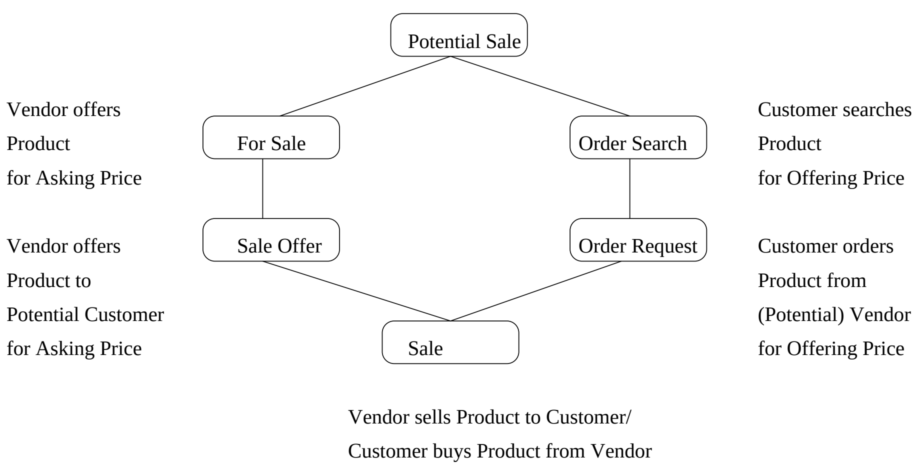

Fig. 5. A hierarchy of contexts of a Sale

4 Conclusion

This paper describes information representation and communication in terms of two sign triads, an individual and a consensual sign triad. While individual triads are important for psychological aspects of information processing, consensual sign triads describe natural and artificial language systems. The triads separate conceptual entities, sign representations and contexts. The relationships between these three constituents and between different triads can be formalized using techniques from FCA, CG's and IF. This paper establishes a foundation for using these insights in e-commerce applications. Further analyses, including a more precise specification of relations and information channels are left for future research.

References

1. Barwise, Jon; Perry, John: Situations and Attitudes. Cambridge, MA: Bradford Books/MIT Press, 1983.
2. Barwise, Jon; Seligman, Jerry: Information Flow. The Logic of Distributed Systems, 1997. Cambridge University Press.
3. Context'99: International and Interdisciplinary Conference on Modeling and Using Context. Conference webpage on-line available at http://www.cs.unitn.it/CONTEXT-99/
4. Lenat, Douglas: The Dimensions of Context Space. On-line available at http://www.cyc.com/publications.html.
5. Devlin, Keith: Logic and Information. Cambridge University Press, 1991.
6. Enterprise Ontology (2000). On-line available at http://www.aiai.ed.ac.uk/project/enterprise/
7. Ganter, Bernhard; Wille, Rudolf: Formal Concept Analysis. Mathematical Foundations. Berlin-Heidelberg-New York: Springer, 1999.
8. Groh, Bernd; Wille, Rudolf: Lattices of Triadic Concept Graphs. In: Ganter; Mineau (Eds.): Conceptual Structures: Logical, Linguistic, and Computational Issues. Proceedings of ICCS2000. Berlin-Heidelberg-New York: Springer, 2000.
9. Lehmann, Fritz; Wille, Rudolf: A Triadic Approach to Formal Concept Analysis. In: Proceedings of the Third International Conference on Conceptual Structures, 1995.
10. Priss, Uta: Relational Concept Analysis: Semantic Structures in Dictionaries and Lexical Databases. (Dissertation) Verlag Shaker, Aachen 1998.
11. Sowa, John: Conceptual Structures: Information Processing in Mind and Machine. Addison-Wesley, Reading, MA, 1984.