A Fundamental View on the Process of Conceptual Modeling

S.J.B.A. Hoppenbrouwers, H.A. (Erik) Proper, and Th.P. van der Weide

Institute for Computing and Information Sciences, Radboud University Nijmegen,
Toernooiveld 1, 6525 ED Nijmegen, The Netherlands, EU
{S.Hoppenbrouwers, Th.P.vanderWeide, E.Proper}@cs.ru.nl

Abstract. In an ongoing effort to better understand the process of creating conceptual models (in particular formal ones), we present a fundamental view of the process of *modeling*. We base this view on the idea that participants in such a process are involved in a deliberate and goal-driven effort to share and reconcile *representations* of their personal conceptions of (parts of) the world. This effort takes the shape of a *modeling dialogue*, involving the use of *controlled language*. We thus take a fundamental approach to *subjective* aspects of modeling, as opposed to traditional approaches which essentially consider models as *objective* entities. We position and present our initial theory of modeling, and briefly discuss how we intend to validate and further develop it.

1 Introduction

The view on conceptual modeling as presented in this paper is rooted in a number of different modeling practices and theories. We take the point of view that the act of conceptual modeling should be understood (1) at a fundamental level, (2) in context of what models are for, and (3) taking into account the capacities and goals of the individuals who create or use them [1]. In taking this view we have, first and foremost, been inspired by the conceptual modeling approach called ORM (Object Role Modeling) [2]. We have been involved in studies of the application of ORM in domain modeling and requirements engineering [3, 4]. In addition, we have drawn from theory and practical experience acquired through the ArchiMate project, which focused on Enterprise Architecture Modeling [5].

1.1 Focus and Fundamental Assumptions

The central question we address is: "Why do we model?". We attempt to answer this question in a generic fashion that nevertheless clears a path for further and more specific research. Further elaborations and directions for solutions are based on our answer to the question, which is: "We model because modeling answers questions". While this is too generic an answer to solve much, it does directly clarify our approach to conceptual modeling. By asking: "*Who asks the questions that need to be answered?*" and "*Why these people ask those questions?*", we immediately arrive at a view on conceptual modeling that is deeply rooted in *communication*, involving *language as a means to achieve communication*

L. Delcambre et al. (Eds.): ER 2005, LNCS 3716, pp. 128–143, 2005.

[6, Chapter 3]. This entails that we are especially interested in *cooperative* aspects of modeling: we focus on what a model, and the process of creating it, achieves in terms of communication between people. Essentially, modeling concerns creative learning. It is akin to communication techniques encountered in education and knowledge management. Modeling is a learning process in which cooperating participants together construct a view on (and a model of) reality [7]. Ultimately, therefore, we see modeling as a tool for developing and sharing knowledge.

Our communication-based approach is largely inspired by the desire to understand (and eventually improve) the modelling *process*. We believe that indeed, for this purpose, communication aspects of modeling are crucial. However, we also believe that many if not all other approaches can be combined with or even integrated into such a view.

Modeling as we see it may or may not involve the use of a (semi-)formal modeling language, i.e. a modeling language the syntax and (in case of fully formal languages) semantics of which can be coherently formulated in a mathematical language. While we focus primarily on formal modeling, we include informal and semi-formal modeling in our view, and are strongly interested in the differences and commonalities that hold between the various flavours.

1.2 Questions and Answers Underlying Modeling

The vast majority of literature on modeling concerns restrictions on the form, structure, and meaning that a model (expressed in a certain language) should respect. Such restrictions may range from an iconic vocabulary for conveying a coherent set of informal notions, to a fully formal set of restrictions on syntax and semantics. We do not argue for or against any form of modeling or modeling language, but emphasize the importance of asking why a certain restriction is imposed, and what its relation is to the questions asked and answered in context of the modeling process and the use of the finished model. In fact, we are less interested here in the modeling languages per se than in the questions asked as part of the modeling process.

We observe that many of the questions asked during actual modeling are not answered if a complete, finished model is "read". Instead, many questions are asked and answered *during the process of modeling*. The finished model corresponds to the minutes of a meeting that has taken place [8]. Reading the minutes certainly answers some questions, but provides no further opportunities for asking new questions, nor to add to the answers or to verify whether what has been said is well understood (i.e. truly learned) by all parties involved. In addition, we observe that in many cases, people tend to adapt the modeling language used (the "Way of Modeling" [9]) to the needs that occur during the modeling process [10]. We can only hope to understand all these aspects of modeling by looking at the details of the process. We therefore propose a view on modeling that respects its *product* (and the intended usage thereof), while also clarifying the nature of the modeling *process* and what it might involve and achieve apart from the product as such. Our view thus is process-oriented, yet aspires to be complementary to product-oriented views.

The questioning-and-answering that takes place during modeling can fruitfully be seen as a *dialog* or *conversation*. Given the assumptions presented above, understanding the goals of modeling, and the means to match them, boils down to understanding the questions people ask during modeling, and the means they deploy to get them answered. Once this becomes clear, we can begin to work towards the formulation of basic *modeling strategies*. These are ways of proceeding in a modeling dialogue that are optimally fit to fulfill two main goals:

1. Answering all the questions the participants in the modeling process might have (explicitly or implicitly).
2. Answering all the questions asked by those who use the product (i.e. the completed model).

In this paper, we will not discuss modeling strategies as such, but merely pave the way for study of actual modeling dialogue and the strategies it involves. We focus here on the essentials of our view on modeling.

1.3 Positioning Verification and Validation

Restrictions on models are generally related to one of two sets of *demands on quality*: those related to *verifiability* (a.k.a. "internal quality") of a model, and those related to *validity* (a.k.a. "external quality") of a model. In formal modeling literature, emphasis lies mostly on verifiability, but clearly a good formal model must also be *valid*. However, in many cases validity is not a matter that can be resolved by any means of objective validation in the mathematical sense. It usually depends on subjective judgments passed and viewpoints held by humans. Because of this, though validation is considered an important and problematic issue, it is often discarded because it cannot be handled very well within the realm of mainstream computer science.

Based on our extensive personal experience in modeling, as well as our theoretical work in that area [1, 3, 11, 12, 6, and more] we expect that the validation of models (in both informal and formal modeling) can be much improved by means of better modeling processes and strategies, within a communicative approach. Along similar lines, it should also be possible to formulate dialogue-based strategies that lead to verifiable correctness in completed models.

It is not just the quality of models we are concerned about. We also hope that by finding detailed modeling strategies, we can eventually help deal with an increasingly problematic bottleneck that occurs in AI and system development: a growing demand for constant creation of formal models in specific and dynamic operational contexts, combined with a lack of people who are capable and willing to perform the modeling required.

Our main focus is on formal conceptual modeling because in terms of combined validation and verification, it poses the biggest challenge and is most urgent. Also, the modeling bottleneck mostly concerns formal models. We strive for an integrated approach to achieving validation and verification: a good process, resulting in a valid model which is also verifiably correct in the end. The key then is to achieve a careful and systematic exchange of questions and answers, guided and restricted by the particular demands on both validity and verifiability as posed by the context in which and for which a model is created.

1.4 Approach

Though science has since long embraced and studied the product of formal modeling (the models and modeling languages), the details of the underlying production process (modeling) still lie mostly in the realm of art. We aspire to be more scientific about the modeling process as such. This requires a study of modeling in terms of participant behavior. More in particular, we intend to find and develop well-formulated *strategies* as a means to describe modeling processes, in order to better understand what courses of action lead to good (valid, verifiable) formal models in line with specific demands posed by their contexts.

In finding answers to the questions raised above, we are currently in the stage of developing a theoretical framework. The *plausibility* and *soundness* of this framework is argued initially in terms of its linguistic [6, 13], epistemic [14] and semiotic [15, 16] foundations. Once a plausible and sound theoretical framework is created, we will conduct a series of modeling experiments to confirm the *validity* of the framework.

2 Conceptual Modeling

The aim of this section is to more closely investigate the process of conceptual modeling. In defining precisely what we mean by modeling a domain, we first need to introduce a framework describing the essential process that takes place when a person (for example, a stakeholder) observes a domain (for example, a work situation to be supported by an information system).

Let us first consider what happens if some *viewer* observes 'the universe'. Our central underlying assumption is that viewers perceive a universe and then produce a *conception* of that part they deem relevant. The conceptions harbored by a viewer cannot be communicated and discussed with other viewers unless they are articulated somehow (the need for this ability in the context of system development is evident). In other words, a conception needs to be represented. Following Peirce, we embrace the idea that both perception and conception of a viewer are strongly influenced by her interest in the observed universe. The *viewer* is an actor perceiving and conceiving the universe (the 'world' around the viewer), using her senses. A *conception* is that which results, in the mind of a viewer, when she observes the universe –using her senses– and interprets what she perceives. Finally, a *representation* is the result of a viewer denoting a conception, using some language and medium to express herself.

2.1 Viewers and Their Frame of Reference

From a modeling point of view, a viewer could metaphorically be seen as an observation tool (a telescope) used to obtain information from the observed universe. The modeler may observe the universe directly, but still depends on the viewer *and the representations she brings forth* to get (more) accurate information. An observation tool should provide a trustworthy image of the universe in such a way that structure that can be derived from the image corresponds to the structure of the observed universe. Different observation tools (or even different observations) may yield different images (representations), all reflecting the same universe.

In our context, a viewer is assumed to be *competent* (i.e. knowledgeable) [12] and *trustworthy* (i.e. not tell lies). The viewer is also capable of providing a *verbalized image of the observed world*, consisting of *statements* in some language (either verbal or graphical; ORM, for example, covers both). We also assume that the structure of the statements uttered has at least some correspondence with the structure of the world observed. Without referring to particular universals, we assume there to be some underlying commonality in how people perceive and conceptualize the world. Both the bio-cognitive make-up of people and their experiences of living as and among humans create at least some common ground, reflected in their language [6, 13].

As mentioned above, in conceiving a part of the universe, viewers will be influenced by their particular interest in the observed universe. In the context of system development (more in particular, enterprise architecture), this corresponds to what tends to be referred to as a *concern* [17]. For example, a viewer may be concerned with safety issues within a domain. Though we acknowledge that a concern may influence the choice of modeling language, we abstract for the moment form such peculiarities, and see a viewer purely as a *language source with a personal syntax*. Sentences delimited by this syntax convey the meaning of the associated (personal) world. The underlying semantical function is an unknown and possibly informal function. We call a language (intended for communication) informal if it has no well-defined syntax, or no semantic interpretation in terms of some underlying formal (i.e. mathematically expressed) model.

Concerns are not the only factors that influence a viewer's conception of a domain. Another important factor concerns the pre-conceptions a viewer may harbor as they are brought forward by their social, cultural, educational and professional background. More specifically, in the context of formal modeling, viewers will approach a domain with the aim of describing it in terms of some predefined set of meta-concepts, such as classes, activities, constraints, etc. The set of meta-concepts a viewer is used to using (or trained to use) when modeling a domain will strongly influence the conception of the viewer (the well known Sapir-Whorf hypothesis in its weak and commonly accepted form [18]). We therefore presume that when viewers model a domain, they do so from a certain perspective; their *Weltanschauung* (German for "view of the world") [19]. The Weltanschauung can essentially be equated to the notion of a *viewpoint* [17].

Viewers may decide to zoom in on a particular part of the universe they observe, or to state it more precisely, they may zoom in on a particular part of *their* conception of the universe. This allows us to define the notion of a *domain* as: any subset of a conception (being a set of elements) of the universe, that is conceived of as being some 'part' or 'aspect' of the universe. In the context of (information) system development, we have a particular interest in unambiguous abstractions from domains. This is what we refer to as a *model*: a purposely abstracted and unambiguous conception of a domain. Note that both the domain and its model are *conceptions* harbored by the same viewer. We are now in a position to define more precisely what we mean by *modeling*: the act of purposely abstracting a model from (what is conceived to be) a part of the universe.

For practical reasons, we will understand the act of modeling to also include the activities involved in the *representation* of the model by means of some language and medium. In line with [16], we consider modeling to boil down to "making statements in some language".

2.2 Participants in the Modeling Process

In this and the following sections, we will use the generic term *participant* for all actors taking part in the modeling process. Importantly, all such participants are viewers as defined above.

For the sake of the argument, let us consider a basic (and admittedly over-simplified) situation in which two participants in the modeling process play the following roles. One is the *domain expert*, who is competent and trustworthy; she knows all there is to know about the target domain, or can find out more if need be. In other words, she can *generate* and *validate* statements about the domain,

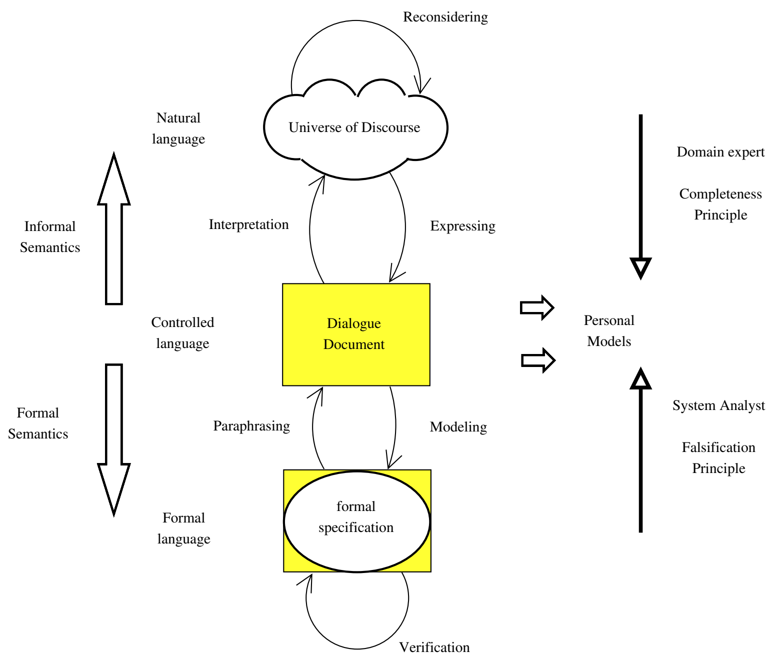

Fig. 1. The classic view on the conceptual modeling process

but she has no formalization skills. The other participant is the *system analyst*, who has no knowledge of the target domain but does know how to create a verifiably correct formal model. The interactive relation between the example roles of domain expert and system analyst is depicted in figure 1, which represents the classic view on modeling. The upper half of the figure shows the "informal" world of the domain expert, statements about this world which are typically expressed in natural language. The lower half of the figure shows the "formal" world of the system analyst, statements about which are typically expressed in some formal language. The link between the two worlds is achieved through a dialogue (and a *dialogue document* that records it). We presume the dialogue to be conducted using some form of controlled language (see discussion below).

In the activity of cooperatively creating a formal model on the basis of informal information, there is a parallel and a symmetry between the Completeness principle and the Falsification principle (positioned on the upper and lower right in figure 1). In the formal world, a model may be deemed falsifiable because it is semantically or syntactically incorrect. While such formal falsifiability is impossible in the informal world, this world allows for judgments of (in)completeness: has everything that needs to be said been said (and has no more than what is relevant been said)? Though the parallel may not be an ultimate, philosophical one, it does hold for the practice of formal modeling: the best we can do as we provide informal input is be complete (within the boundaries of relevance); the best we can do in formal articulation is be formally correct. The marriage between the two makes for good formal models.

In our view, then, the domain expert typically harbors an informal semantic function (natural language), while the system analyst's language may be expected to be governed by a formal semantic function. However, both are "language sources"; it is just that their syntax and semantics differ in structure and nature. Thus, beyond this example, it seems justified to indeed use the more neutral concept *participant* as a generic term for domain expert and system analyst. Participants all have their personal syntax and a formal or informal semantic function, depending on the roles they play in modeling.

In the context of communication resulting in formal models (in particular, the traces of communication recorded in the dialogue document; see figure 1), we strictly focus on written expressions. Though the document is linear, and therefore the order in which the text has been uttered is captured, further aspects of communication and medium (time, location, gestures and facial expressions, technologies used to communicate, etc.) are discarded and abstracted from. There is one exception to this: it is recorded which participants uttered a particular sentence. Also, we consider the possibility to accept, at a more advanced stage or our research, dialogue logs involving the use of graphical utterances (drawings) that are (in syntactic terms) translatable 1:1 to verbalizations.

2.3 Controlled Language

Formal and informal language may be hard to fully reconcile, but a classic meeting point between natural and formal language lies in similarities between the basics of their grammar and meaning, in particular in predicates and predication. After all, formal languages have historically been derived from their natural counterparts. It has since long been recognized that when we use simple, elementary sentences in natural language, we can relatively easily bridge the gap between formal and informal [20, 10], even if the bridge can only bear very light traffic. Such simple, elementary language can be described by a relatively simple grammar and can yet be realistically used in a modeling conversation (see, for example, [21]). We referred to it as *controlled language* [22]. Our notion of controlled language is related to that of *simplified English*; see [23]. The syntax (not necessarily the semantics!) of a controlled language is limited and fixed, and therefore it can expected to be both known to and agreed on by its various users [6, p26-31].

The competency of a participant [12] may then be defined as:

1. transform model into controlled language,
2. validate a description in controlled language.

It is assumed that a participant can express statements in this controlled language, but is also capable to express statements *about* that language. In system development, it is crucial to reach clarity and agreement about terminology, concepts, and sometimes syntax used in communication between members of the development team [4]. Controlled language can be used to reach clarity and agreement about any other type of language that might be used (for example, full-fledged natural language, schematic language, or (semi-)formal language). Thus it becomes possible to discuss any model through controlled language, but also to discuss the modeling languages –both formal and informal– through controlled language.

2.4 The Goal of Modeling

The goal of the modeling process can be described as: trying to reach a state where all participants agree that they have some degree of common understanding (for a similar stance in context of requirements engineering, see [24]). The participants try to derive from their personal semantics a group semantics. Participants will be convinced this goal has been achieved if they have validated their assumptions to contentment of everyone involved. For example, a system analyst will be convinced that the derived model is complete if the model has been validated against the real situation. In our view, this means that the domain expert, harboring the semantics of her conception of the universe, has positively responded to the controlled language description of the model provided by the system analyst, which may be rooted in a formal language. Various semantic functions come into play, but the shared, controlled language (which may be cooperatively constructed) performs an intermediary function.

The goal of the interaction can thus be seen as the construction of (1) a grammar for representations that are acceptable to all participants, and (2) semantic interpretation(s) in terms of some model(s). The grammar produced in the interaction is a generative device. It can also be used as a parsing device.

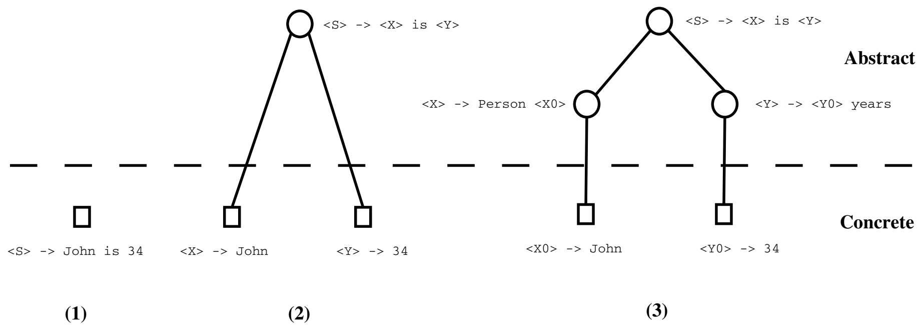

Fig. 2. Parsing levels

The grammar is correct when all sample sentences can be generated. From the point of view of the system analyst, the target model is restricted by the (formal) semantics of the modeling technique used. From the point of view of the domain expert, validation of the model may be seen as assigning meaning (interpretation) to the representations generated by the system analyst. A more symmetric way of putting this is that for each party (a), the other party (b) agrees with the controlled language statements provided by party (a).

Example 1. A simple sentence like *John is 34* is the initial statement verbalizing a fact occurring in the domain, provided by a domain expert (see case 1 in figure 2). The parsing structure of this sentence leads to case 2 of this figure. A simple modeling strategy is to defoliate parse trees, i.e. remove labels and constants like John and 34. These defoliated parse trees provide an example of the grammatical structure of the expert language. The leaves are concrete instances. During modeling, we are interested in acquiring the expert grammar, and therefore we are (only) interested in the defoliated parse trees.

The 3rd case provides a fully qualified version of the sentence in NIAM normal form. Fully qualified sentences are well suited as a basis for modeling. However, domain experts are more inclined to produce statements as in case (2). A main goal of a modeling dialog aiming for a NIAM model is to detect and resolve unqualified constants, hence asking and answering questions related to qualification.

Initially, the dialogue may use a modeling technique that accepts statements that do not convey all the information needed for a formalization. However, the system analyst will eventually require statements that match a more demanding and restrictive modeling technique like NIAM. During modeling, the model will have to migrate from the first (informal) to the second (formal) modeling technique.

2.5 Modeling as Interaction Between Viewers

The modeling process is seen as a goal-driven dialogue between a number of participants. Each participant is a viewer. The only way the participants can achieve their modeling goals is to communicate with each other, and remember and build on what has been discussed. An explicit way to do this would be to keep "modeling minutes" that are agreed on by the participants (figure 1). Taking a dialogue perspective on a modeling process is in line with a constructionist world view [25, 16].

As discussed, formal modeling can be succesfully captured by recording restricted aspects of communication. In this vein, in the NIAM method, the type of communication that takes place between modeling participants depicted as the *telephone methaphor* (two participants who communicate via a telephone line). Following this image, the modeling minutes consist of a recording or logbook of the telephone conversation. In order to capture the more rich and complex dialogue patterns in larger groups, we propose the so-called *chatbox metaphor*, assuming the participants communicate as in a chatbox. This is a real-time, teletype like communication channel that has become immensely popular among internet users; famous public applications of this type are, for example, Jabber, ICQ and MSN. In advanced use, chatbox conversations may branch off of (and rejoin) other chatbox conversations.

In view of the chatbox metaphor, the communication between the participants is assumed to be conducted entirely through a chatbox. If we view only the sentences of a particular participant, then it makes sense to interpret this view as a description of the model put forward by this participant. As the participant's model may evolve during the chat, obtaining the model from a participant involves the dynamics of dialogue, and is certainly not trivial.

For practical reasons, we will make some assumptions about the language that is used by the participants during the chat. The underlying controlled language model should be such that putting a sentence in some chosen normal form (for example, NIAM normal form) is an activity that does not require other skills from the participant than elementary knowledge of language and sufficient knowledge of the domain observed by this participant.

2.6 Some Types of Modeling Dialogue

As an illustration of various kinds of dialogue that may occur, consider the three modeling situations described below:

tabula rasa – This kind of modeling process is roughly comparable with the way in which a baby learns the basics of how the world is, from its parents and environment –with its developing language as a key item. *This is roughly the type of process compatible with the simplified domain expert - system analyst scenario discussed above. However, in that scenario the demands posed by a formal language of course imply a much stricter set of questions-to-answer than a child would harbor*

open mind – This kind of modeling process takes place when two people with their own, well-developed views on the world, are eager to learn about their mutual views –as reflected in their languages. *This type of modeling starts off with two unconnected representations/models, after which the commonalities between them may soon be discovered. This may or may not lead to a full reconciliation of the initial models in terms of a translation between them.*

colliding views – This kind of modeling process will occur when the participants have different views on the world, and the participants' priorities force them to maintain (part of) that view, while at the same time a mutually acceptable model is needed. This will usually lead to conflicts –modeling conflicts, possibly rooted in language conflicts, reflecting world view conflicts. *Modeling of this type will have to involve negotiation or argumentation about a common model. In some cases, one participant will impose his model upon the other participant; in others, one participant will be able to convince another by rational argumentation; in yet other cases, pure negotiation will take place: seeking a compromise both parties can agree with.*

Importantly, either the "open mind" or "colliding views" situation occurs whenever pre-existing conceptualizations or models (possibly, rooted in existing systems) are relevant to the modeling process, which in fact is very often the case.

The modeling strategies followed in the three types of conversation mentioned above are quite different. In addition, how the basic strategies are executed will strongly depend on the sort of model that is aimed for, and the modeling languages involved. It seems most realistic to start with investigating the "tabula rasa" type of dialogue and work up from there. The "colliding views" strategies are the most complex but seem representative of many real-life modeling situations in; arguably, they are ultimately the most interesting.

3 Modeling as a Dialogue

In the previous section (as illustrated by figure 1), we discussed our view on modeling as an exchange of statements between participants, in a modeling *dialogue*. In this section we introduce our core model for such dialogues-for-modeling.

3.1 Basics

As a starting point, we assume two participants in the modeling process, referred to as a and b respectively. Each participant harbors some volume of knowledge. For example, in a modeling process a domain expert has knowledge of the universe of discourse (i.e. some domain); as discussed in the previous section, we assume a domain expert to be fully knowledgeable: we do not question the validity of the expert's knowledge as such. We further assume the falsifiable basis (the test, as it were) for having acquired knowledge is *the capability to demonstrate it.* As a consequence, assuming participant p to have knowledge S_p corresponds to assuming p to be capable to somehow demonstrate the knowledge elements from S_p. Whether this demonstration is considered convincing depends on the judgment of the other participant, the initial contributor of the knowledge. This judgment will, generally, also depend on the goals and demands driving the modeling dialogue (as viewed by the participants).

Epistemically [14] we choose to view knowledge indeed as *knowledge* and not as *belief.* Thus, we abstract from such philosophical questions as whether or not the domain expert has knowledge that does not match reality. We collapse the notions of knowledge and belief to keep our current argument transparent.

As for the demonstration of knowledge, consider the following illustrative example (loosely based on [7]). If a teacher attempts to teach a student about something (for example, a historical episode), she tries to create in the student's mind a conception of the item taught that is equal or very similar to her own. How can she be convinced that the student has conceived (learned) the item? By asking for a demonstration of that knowledge. Such a demonstration can come in various forms. Discarding non-verbal demonstrations, we distinguish *exemplification* and *paraphrasing* as the most common techniques for convincing a teacher. These techniques can also be applied within modeling dialogues.

Traditional approaches to communication in modeling simply assume a participant who is willing to transfer domain knowledge to another participant who is eager to learn about this domain, i.e. to create a conception of it. The dialogue that brings about the transfer is assumed to be objectively meaningful: to be de-contextualized, making the message independent of time, location, and participant. This motivates a view of the modeling process restricted to the actual *symbols* used in communication, abstracting from the participants as such.

In the traditional view, subjectivity is only a relevant notion if there are disagreements between domain experts, which should be resolved. Contrary to the traditional view, we are neutral about the subjective conceptions of the participants involved in modeling.

We assume that participants are willing to communicate about their knowledge, and are willing to listen to the others. Thus, both a and b must be willing to take turns in playing the leading role of contributor and or the more passive role of receiver.

3.2 Characterizing the Dialogue

Let us first consider a the simple, "tabula rasa" view on knowledge transfer. We assume, without loss of generality, participants a and b to be in different roles (a is the contributor, b the receiver). The basic assumptions that must underly the dialogue are the following:

1. participant a is willing to transfer its knowledge to participant b,
2. participant b is willing to assimilate a's knowledge.

The above assumptions directly relate to the dialogue between a and b. The pragmatic assumptions with respect to the statements made in context of the modeling dialogue can be phrased as follows:

1. a has the intention to talk; a makes a statement s under the assumption that b seeks to know s.
2. b has the intention to listen; if a states s, then b assumes this is done with the ultimate intention to enable b know s.

Note that during modeling the participants will take turns in playing the leading role of contributor and or the more passive role of receiver. In the dialogue document the transferred statements will be registered, including the participants involved and the roles they play at that moment. This model allows dialogues to have sub-dialogues, with very specific goals that are sub-goals of the main dialogue. This is in line with the chatbox model for communication.

This analysis of the "tabula rasa" type of modeling dialogue can be extended to cover the "open mind" and "colliding views" types as introduced in the previous section:

3. a has the additional intention enable b to translate his representation to s; a makes a statement s under the assumption that b seeks map his representation to s where possible.
4. b has the additional intention to translate his representation to s; if a states s, then b assumes this is done to enable him to map his representation to s where possible.
5. a has the additional intention to negotiate, argue in favor of, or impose s on b; a makes a statement s under the assumption that b wants to negotiate, needs to be convinced of, or will have to be forced to accept s.
6. b has the additional intention to negotiate about s in view of his own, preferred representation, defend his own representation through argumentation, or resist accepting s instead of his own representation; if a states s, then b assumes this is done in order to negotiate, argue in favor of, or impose s upon him.

We consider understanding of the above intentions, and the strategies that follow from them, to be fundamental to the understanding of the modeling process. Matters may be complicated by unawareness of one participant of some goal or strategy of another participant; also, various goals and strategies may become entangled.

4 Conclusion and Future Work

We have presented a fundamental view on (formal) modeling rooted in knowledge creation and exchange, in which communication plays a central role. We have argued that to achieve a high quality combination of validity and verifiability in models, we need to look not only at the product, but *specifically* also at the process of modeling. In line with our communicative approach, the modeling process is viewed as a dialogue between participants. We have described a first, general analysis of the essential properties of modeling dialogues. In particular we discussed the central role controlled language can play in modeling dialogues, and the basic underlying intentions of such dialogues, rooted in the sharing, translation, negotiation, argumentation, and imposing of (participant-based) knowledge representations. This should provide a good starting point for more detailed, domain-specific or application-specific exploration of modeling dialogues, with as a central goal the discovery of modeling strategies and optimal selection of such strategies depending on the goals for particular situations.

As possible domains of application of the controlled use of modeling strategies, the following flavors of modeling seem particularly interesting: domain modeling, information modeling, architecture modeling, ontological modeling, and interactive querying. We plan on focused research activities in all of these areas.

Possibly, the range of application areas can be extended to include more complex forms of modeling, such as software modeling, formal business rules specification, and numerous AI applications.

Validation and improvement of our theoretical framework will be a crucial aspect of our further research. After further investigating the plausibility and soundness of our framework, we will test its validity by starting a substantial experimentation programme to validate our initial theory, the chatbox metaphor, and the use of controlled language. We thus intend to lay an empirical foundation under our exploration of the basic dynamics of modeling and the use of controlled language therein. We intend to start our experiments by investigating "tabula rasa" type modeling, and quickly move into "open mind" modeling. Understanding and improving "opposite views" modeling is more challenging, and may be successful only in the long run, but is also the Main Prize. Core focus of our theory development will be on eliciting, describing, and testing *strategies* for formal modeling (possibly also other forms of modeling).

One of our long term objectives is to investigate ways of developing a new brand of CASE tools that involves the interactive monitoring and guidance of some dedicated (i.e. situationally fitted) modeling process, integrated with the IS development process at large. Two factors underly this idea: solving the modeling bottleneck and improving model quality and grounding. The new brand of case tools can be expected to be a blend of classic, product-oriented CASE tools on the one hand, and cooperative, interactive dialogue systems on the other (probably involving issues as studied in the field of Computer Supported Cooperative Work or CSCW).

We eventually hope to extend the range of our experiments by providing an increasingly attractive digital environment for people to use during modeling, providing added value for the participants as well as data for the researchers, and enabling insightful interaction between the two groups.

References

1. Proper, H., Verrijn-Stuart, A., Hoppenbrouwers, S.: Towards utility-based selection of architecture-modelling concepts. In Hartmann, S., Stumptner, M., eds.: Proceedings of the Second Asia-Pacific Conference on Conceptual Modelling (APCCM2005), Newcastle, New South Wales, Australia. Volume 42 of Conferences in Research and Practice in Information Technology Series., Sydney, New South Wales, Australia, Australian Computer Society (2005) 25–36.
2. Halpin, T.: Information Modeling and Relational Databases, From Conceptual Analysis to Logical Design. Morgan Kaufman, San Mateo, California, USA (2001).
3. Proper, H., Bleeker, A., Hoppenbrouwers, S.: Object-role modelling as a domain modelling approach. In Grundspenkis, J., Kirikova, M., eds.: Proceedings of the Workshop on Evaluating Modeling Methods for Systems Analysis and Design (EMMSAD'04), held in conjunctiun with the 16th Conference on Advanced Information Systems 2004 (CAiSE 2004). Volume 3., Riga, Latvia, EU, Faculty of Computer Science and Information Technology, Riga Technical University, Riga, Latvia, EU (2004) 317–328.

4. Bleeker, A., Proper, H., Hoppenbrouwers, S.: The role of concept management in system development – a practical and a theoretical perspective. In Grabis, J., Persson, A., Stirna, J., eds.: Forum proceedings of the 16th Conference on Advanced Information Systems 2004 (CAiSE 2004), Riga, Latvia, EU, Faculty of Computer Science and Information Technology, Riga Technical University, Riga, Latvia, EU (2004) 73–82.
5. Lankhorst, M., others: Enterprise Architecture at Work : Modelling, Communication and Analysis. Springer, Berlin, Germany, EU (2005).
6. Hoppenbrouwers, S.: Freezing Language; Conceptualisation processes in ICT supported organisations. PhD thesis, University of Nijmegen, Nijmegen, The Netherlands, EU (2003).
7. Pask, G.: Conversation, cognition, and learning: a cybernetic theory and methodology. Elsevier (1975).
8. Veldhuijzen van Zanten, G., Hoppenbrouwers, S., Proper, H.: System Development as a Rational Communicative Process. In Callaos, N., Farsi, D., Eshagian-Wilner, M., Hanratty, T., Rish, N., eds.: Proceedings of the 7th World Multiconference on Systemics, Cybernetics and Informatics. Volume XVI., Orlando, Florida, USA (2003) 126–130.
9. Wijers, G., Heijes, H.: Automated Support of the Modelling Process: A view based on experiments with expert information engineers. In Steinholz, B., Sølvberg, A., Bergman, L., eds.: Proceedings of the Second Nordic Conference CAiSE'90 on Advanced Information Systems Engineering. Volume 436 of Lecture Notes in Computer Science., Stockholm, Sweden, EU, Springer-Verlag, Berlin, Germany, EU (1990) 88–108.
10. Hoppenbrouwers, J., Vos, B.v.d., Hoppenbrouwers, S.: Nl structures and conceptual modelling: Grammalizing for KISS. Data & Knowledge Engineering **23** (1997) 79–92.
11. Proper, H., Hoppenbrouwers, S.: Concept evolution in information system evolution. In Gravis, J., Persson, A., Stirna, J., eds.: Forum proceedings of the 16th Conference on Advanced Information Systems 2004 (CAiSE 2004), Riga, Latvia, EU, Faculty of Computer Science and Information Technology, Riga Technical University, Riga, Latvia, EU (2004) 63–72.
12. Frederiks, P., Weide, T.v.d.: Information modeling: the process and the required competencies of its participants. In Meziane, F., Métais, E., eds.: 9th International Conference on Applications of Natural Language to Information Systems (NLDB 2004). Volume 3136 of Lecture Notes in Computer Science., Manchester, United Kingdom, EU, Springer-Verlag, Berlin, Germany, EU (2004) 123–134.
13. Pinker, S.: The Language Instinct. Allen Lane/Penguin Press, London, United Kingdom, EU (1994).
14. Meyer, J.J., Hoek, W.v.d.: Epistemic Logic for AI and Computer Science. Cambridge University Press, Cambridge, United Kingdom, EU (1995).
15. Peirce, C.: Volumes I and II – Principles of Philosophy and Elements of Logic. Collected Papers of C.S. Peirce. Harvard University Press, Boston, Massachusetts, USA (1969).
16. Krogstie, J.: A semiotic approach to quality in requirements specifications. In Kecheng, L., Clarke, R., Andersen, P., Stamper, R., Abou-Zeid, E.S., eds.: IFIP TC8/WG8.1 Working Conference on Organizational Semiotics – Evolving a Science of Information Systems, Montréal, Québec, Canada, Kluwer Academic Publishers, Dordrecht, The Netherlands, EU (2002) 231–250.

17. The Architecture Working Group of the Software Engineering Committee, Standards Department, IEEE: Recommended Practice for Architectural Description of Software Intensive Systems. Technical Report IEEE P1471-2000, The Architecture Working Group of the Software Engineering Committee, Standards Department, IEEE, Piscataway, New Jersey, USA (2000).
18. Chandler, D.: The Act of Writing: a Media Theory Approach. University of Wales, Aberystwyth, United Kingdom, EU (1995).
19. Wood-Harper, A., Antill, L., Avison, D.: Information Systems Definition: The Multiview Approach. Blackwell Scientific Publications, Oxford, United Kingdom, EU (1985).
20. Frederiks, P.: Object-Oriented Modeling based on Information Grammars. PhD thesis, University of Nijmegen, Nijmegen, The Netherlands, EU (1997).
21. Rolland, C., Souveyet, C., Ben Achour, C.: Guiding goal modeling using scenarios. IEEE Transactions on Software Engineering **24** (1998) 1055–71.
22. Schwitter, R.: Controlled Natural Languages. Centre for Language Technology, Macquary University, Sydney, Australia. (2004).
23. Farrington, G.: An Overview of the International Aerospace Language. (1996).
24. Pohl, K.: The three dimensions of requirements engineering: a framework and its applications. Information Systems **19** (1994) 243–258.
25. Berger, P., Luckmann, T.: The Social Construction of Reality. Doubleday, New York, New York, USA (1966).