Technical Report AI-TR-96-1, I.S.I.R., Osaka Univ

Towards Ontology Engineering

Riichiro MIZOGUCHI and Mitsuru IKEDA
The Institute of Scientific and Industrial Research, Osaka University, 567 Japan

Abstract. The main objectives of this paper include to propose a new research field called "Ontology Engineering" and to show it can be a basic research of content-oriented research and provides such technologies that are badly needed. We begin the paper by discussing what an ontology is. We next analyze the depth of the ontology use in eight levels followed by the discussion on what concrete advantages ontology can give in the real-world problem solving. The next topic is the classification of ontologies. On the basis of the discussion, we present the scope of ontology engineering. Finally, we exemplify ontology engineering by summarizing our work.

1. Introduction

In AI research history, we can identify two types of research. One is "Form-oriented research" and the other is "Content-oriented research". The former deals with logic and knowledge representation and the latter content of knowledge. Apparently, the former has dominated AI research to date [Mizoguchi, 95]. Recently, however, "Content-oriented research" has become to gather much attention because a lot of real-world problems to solve such as knowledge reuse, facilitation of agent communication, media integration through understanding, large-scale knowledge bases, etc. require not only advanced theories or reasoning methods but also sophisticated treatment of the content of knowledge.

Formal theory such as predicate logic provides us with a powerful tool to guarantee sound reasoning and thinking. It even enables us to discuss the limit of our reasoning in a principled way. However, it cannot answer to any of the questions such as what knowledge we should have for solving problems given, what is knowledge at all, what properties a specific knowledge has, and so on.

The knowledge principle proposed by Feigenbaum is to the point in that he stresses the importance of accumulation of knowledge rather than formal reasoning and logic. This has been proved by the success of the expert system development. Of course, his idea of knowledge accumulation should be further deepened. Representation of expertise in production rules is very preliminary. It should be in-depth analyzed to make it sharable and reusable among computers and human agents. The source of problem solving knowledge is of various. Advanced knowledge processing technology should cope with these various knowledge sources and elicit, transform, organize, and translate knowledge to enable the agents to utilize it. Thus, the knowledge base technology should contribute to the next knowledge medium[Stefic, 86]. Ontology engineering provides us with a basis of the knowledge medium research.

Importance of "Content-oriented research" has been recognized a bit these days. Unfortunately, however, we do not have sophisticated methodologies for content-oriented research now. In spite of much effort devoted to such research, major results were only development of KBs. It does not have been considered as "academic research". We could enumerate the reasons for this as follows:

1) Content-oriented research tends to be ad-hoc, and
2) It does not have a methodology which enables the research results to accumulate.

It is necessary to overcome these difficulties in order to establish the content-oriented research. We would like to propose Ontology engineering for that purpose. It is a research methodology which gives us design rationale of a knowledge base, kernel conceptualization of the world of interest, strict definition of basic meanings of basic concepts together with sophisticated theories and technologies enabling accumulation of knowledge which is dispensable for modeling the real world.

We begin the paper by discussing what an ontology is. Although ontology is becoming popular within a community, it is not well understood in AI community in general. We carefully explain what an ontology is. To our knowledge, how to use an ontology is one of the crucial issues in ontology research. Therefore, we analyze the depth of the ontology use in eight levels followed by the discussion on what concrete advantages ontology can give in the real-world problem solving. The next topic is the classification of ontologies. On the basis of the discussion made thus far, we present the scope of ontology engineering. Finally, we exemplify ontology engineering by summarizing our work done to date.

2. What Is an Ontology?

2.1 Simple Definitions

Three simple definitions are given below.

(1) Ontology is a term in philosophy and its meaning is "Theory of existence".
(2) A definition of an ontology in AI community is "An explicit representation of conceptualization"[Gruber, 92].
(3) A definition of an ontology in KB community is "a theory of vocabulary/concepts used as building artificial systems"[Mizoguchi, 93].

Although these are compact, it is not sufficient for in-depth understanding of what an ontology is. A more comprehensive definition is given in the next subsection.

2.2 Comprehensive definitions

(1) Ontology: Following Guarino[Guarino, 95], we use the convention in which capital letter "O" is used to distinguish the "Ontology" in philosophy from others. "Ontology" is a theory which can answer questions such as "what is existence", "what properties can explain the existence", "How these properties explain the differences of existence", etc.

(2) ontology: The design methodology is like one for Ontology, but the target is different from it. Not the "existence" but smaller and concrete thing such as enterprise, thermo-dynamics, problem solving, etc. are discussed. We define an ontology as an explicit and less ambiguous description of concepts and relations among them appearing in the target thing. Such ontologies exist as many as the possible target things. We do not have to use logic to describe it.

(3) Formal ontology: Axiomatic description of an ontology. It can answer questions about the capability of ontology.

(4) Axiom: Declaratively and rigorously represented knowledge which has to be accepted without proof. In predicate logic case, a formal inference engine is implicitly assumed to exist. But, one seldom mentions it.

Axioms have two roles as follows in ontology description:

1) To represent the (partial) meaning of concepts rigorously.
2) Within the scope of the knowledge represented declaratively, to answer the questions on the capability of the ontology and things built using the concepts in the ontology.

Questions about the capability of ontology plays an important role in its evaluation and they are divided into the following two:

1) Questions about the formal properties of the ontology and things designed using ontology.
2) Questions about the behavior of the things designed using the ontology.

The former is called "competence" questions and the latter "performance" questions. Axioms written in predicate calculus are sufficient for answering the former. To answer the latter questions, however, we often need procedural engines to interpret the meaning of concepts in the ontology because declarative knowledge with a formal prover cannot answer all the questions. To cope with such situations, we introduce axiom equivalents defined as follows:

(5) Axiom equivalent[Forbus, 95]

An axiom equivalent is not a rigorous or declarative axiom based on formal inference engine, but it is partially declarative knowledge based also on interpretation by a procedural engine to answer performance questions. Axiom equivalents do not have to be formalized completely.

The difference between axioms and axiom equivalents is essential. "Axioms" can be also interpreted as "small number of rules which are represented in a declarative form and can derive all the facts from them". It is true they contribute to making the characteristics of technology clear and explicit. This also applies to ontology. In fact, many researchers have been trying to represent ontology formally. However, we could say such an attempt neglects the reality. It is obvious that declarative and formal methodology cannot cope with the performance of the knowledge required by knowledge engineering. For example, if we adopted the first order predicate calculus, we have to abandon dealing with the knowledge such as "mathematical induction is sound for all the predicates". What we should do for knowledge engineering is to adopt not only formal approaches but also informal ones such as natural language representation and axiom equivalents based on procedural interpretation. This will enable ontology research to contribute to the future knowledge engineering community.

3. Roles of Ontology Engineering

First of all, we would like to declare the ultimate purpose of ontology engineering is:

"To provide a basis of building models of all things in which computer science is interested".

And, ontologies have to be intelligible both to computers and humans.

3.1 Ontology As a Design Rationale

In the mechanical design setting, previous designs are often used as a reference for design of new products. One of the critical issues in such cases is how to understand the intentions and justifications of various decisions made in them by different designers. They are collectively called "Design Rationale"(hereafter referred to as DR). DR information is often implicit and the implicitness often causes difficulties in reusing the designs. Thus, DR is as important as design drawings.

An ontology plays a role similar to DR in reusing knowledge bases. In order to reuse knowledge in a knowledge base, we have to know underlying conceptualization which reflects the assumptions and requirements made in the problem solving using the knowledge base. Although many KBs have been built to date, no such information has been described. Ontologies as DR information of knowledge bases will contribute to reuse of knowledge bases and play the roles of backbones of knowledge bases. The future knowledge bases should be built with explicit representation of ontologies.

3.2 How To Use an Ontology

Although there have been many discussions on ontology, how to use it has not been fully discussed. This section discusses the levels of usage of ontology. The following is a list of how to use ontology(The shallowest first).

Level 1: Used as a common vocabulary for communication among distributed agents.

Level 2: Used as a conceptual schema of a relational data base. Structural information of concepts and relations among them is used. Conceptualization in a data base is nothing other than conceptual schema. Data retrieval from a data base is easily done when there is an agreement on its conceptual schema.

Level 3: Used as a backbone information for a user of a certain knowledge base. Levels higher than this plays roles of the ontology which has something to do with "content".

Level 4: Used for answering competence questions.

Level 5: Standardization

5.1 Standardization of terminology(at the same level of Level 1)

5.2 Standardization of meaning of concepts

5.3 Standardization of components of target objects(domain ontology).

5.4 Standardization of components of tasks(task ontology)

Level 6: Used for transformation of data bases considering the differences of the meaning of conceptual schema. This requires not only structural transformation but also semantic transformation.

Level 7: Used for reusing knowledge of a knowledge base using DR information.

Level 8: Used for reorganizing a knowledge base based on DR information.

Thus, variety of ontology use is deep and wide. Those higher than level 3 is innovative and suggest future style of knowledge manipulation by computers, which demonstrates the utility of ontology engineering.

3.3 Standardization: Bolts&Nuts in Knowledge Bases

Needless to say, industries have attained high productivity due to standardization of components, say, bolts and nuts. It is a pity that we have no such standardized components in knowledge base technology. In order to model target objects, such components would help a lot and facilitate model-based problem solving. For example, standardization of a pipe and pumps in qualitative modeling of a plant, that of enterprise ontology, and that of task ontology. Standardization of components does not necessarily imply that of knowledge in general. We are not claiming that all the knowledge should be standardized. Using standardized basic components, one can easily design their own knowledge by configuring them, which is proved by the current engineering production.

4. Typology of Ontology

To consider how knowledge is used helps us understand what an ontology is. Little discussion on how to use ontology has been made to date. Many researchers say they "use" knowledge/ontology without defining what they mean by "use", that is, who uses it in what ways. We here discuss this issue considering the usage levels discussed in 3.2. An ontology is further divided into subcategories from the knowledge reuse point of view as follows:

Ontology:

Workplace ontology[Vanwelkenhuysen, 94, 95b]

This is an ontology for workplace which affects task characteristics by specifying several boundary conditions which characterize and justify problem solving behavior in the workplace. Workplace and task ontologies collectively specify the context in which domain knowledge is interpreted and used during the problem solving.

Examples(Circuit troubleshooting): Fidelity/Efficiency/Precision/High reliability/etc.

Task ontology[Mizoguchi, 92, 95a][Hori, 94][Wielinga, 93]

Task ontology is a system of vocabulary for describing problem solving structure of all the existing tasks domain-independently. It does not cover the control structure but do components or primitives of unit inferences taking place during performing tasks. Task knowledge in turn specifies domain knowledge by giving roles to each objects and relations between them.

Examples(Scheduling tasks): Schedule recipient/schedule resource/goal/constraint/ availability/load/select/assign/classify/remove/relax/add/etc.

Domain ontology

Task-dependent ontology[Mizoguchi,95b]

A task structure requires not all the domain knowledge but some specific domain knowledge in a certain specific organization. We call this special type of domain knowledge T-domain ontology because it depends on the task.

T-Domain ontology

Examples(Job-shop scheduling): Job/order/line/due date/machine availability/ tardiness/load/cost/etc.

Task-independent ontology

Because object and activity ontologies are related to activities, we call them activity-related ontology and field ontology activity-independent ontology.

Activity-related ontology

This ontology is related to activities taking place in the domain and is designed having simulation of the domain activity in mind such as enterprise ontology. There are two major activities exist in a domain. One is behavior of an object and the other is organizational or human activities. Verbs play an important role in this ontology, however, they are different from those in task ontology. The subjects of the former verbs are objects, components, or agents involved in the activities of interest, while those of the latter are domain experts.

Object ontology[Vanwelkenhuysen,95a]

This ontology covers the structure, behavior and function of the object.

Examples(Circuit boards):component/connection/line/chip/pin/ gate/bus/state/role/etc.

Activity ontology(Enterprise ontology)[Gruninger,94]

Examples: use/consume/produce/release/state/resource/commit/enable/ complete/disable/etc.

Activity-independent ontology

Field ontology

This ontology is related to theories and principles which govern the domain. It contains primitive concepts appearing in the theories and relations, formulas, and units constituting the theories and principles.

Units[Mars,94]

Examples: mole/kilogram/meter/ampere/radian/etc.

Engineering Math[Gruber,94]

Examples: Physical quantity/physical dimension/unit of measure/ Scalar quantity/linear algebra/physical component/etc.

General/Common ontology

Examples: *Things/Events/Time/Space/Causality* [Lenat,90]
Behavior/Function [Sasajima,95], etc.

This shows there are many categories of ontologies. But this does not mean ontology research diverges but mean the richness of the real world. The identification of the variety of knowledge, and hence that of ontology itself can be one of the research topics which deepens our understanding of knowledge.

The above classification of ontology distinguishes task-dependent ontology from task-independent one. The latter has been often discussed in the literature. The former of which importance the authors have stressed is the authors' original concept.

5. Scope of Ontology Engineering

We here demonstrate the subjects which should be covered by ontology engineering. It includes basic issues in philosophy, knowledge representation, ontology design, standardization, EDI, reuse and sharing of knowledge, media integration, etc. which are the essential topics in the future knowledge engineering. Of course, they should be constantly refined through further development of ontology engineering.

Basic division

- Philosophy(Ontology, Meta-mathematics)
  Ontology which philosophers have discussed since Aristotle is discussed as well as logic and meta-mathematics.
- Scientific philosophy
  Investigation on Ontology from the physics point of views, e.g., time, space, process, causality, etc. is made.
- Knowledge representation
  Basic issues on knowledge representation, especially on representation of ontological stuff, are discussed.

Division of ontology design

-General(Common) ontology

General ontologies such as time, space, process, causality, part/whole relation, etc. are designed. Both in-depth investigation on the meaning of every concept and relation and on formal representation of ontologies are discussed.

- Domain ontologies

Various ontologies in, say, Plant, Electricity, Enterprise, etc. are designed.

Division of common sense knowledge

- Parallel to general ontology design, common sense knowledge is investigated and collected and knowledge bases of common sense are built.

Division of standardization

- EDI(Electronic Data Interchange) and data element specification

Standardization of primitive data elements which should be shared among people for enabling full automatic EDI.

- Basic semantic repository

Standardization of primitive semantic elements which should be shared among people for enabling knowledge sharing.

- Conceptual schema modeling facility(CSMF)
- Components for qualitative modeling

Standardization of functional components such as pipe, valve, pump, boiler, register, battery, etc. for qualitative model building.

Division of Data/knowledge interchange

- Translation of ontology

Translation methodologies of one ontology into another are developed.

- Database transformation

Transformation of data in a data base into another of different conceptual schema.

- Knowledge base transformation

Transformation of a knowledge base into another built based on a different ontology.

Division of knowledge reuse

- Task ontology

Design of ontology for describing and modeling human ways of problem solving.

- T-domain ontology

Task-dependent domain ontology is designed under some specific task context.

- Methodology for knowledge reuse

Development of methodologies for knowledge reuse using the above two ontologies.

Division of knowledge sharing

- Communication protocol

Development of communication protocols between agents which can behave cooperatively
under a goal specified.

- Cooperative task ontology

Task ontology design for cooperative communication

Division of media integration

- Media ontology

Ontologies of the structural aspects of documents, images, movies, etc. are designed.

- Common ontologies of content of the media

Ontologies common to all media such as those of human behavior, story, etc. are designed.

- Media integration

Development of meaning representation language for media and media integration through understanding media representation are done.

Division of ontology design methodology

- Methodology
- Support environment

Division of ontology evaluation

Evaluation of ontologies designed is made using the real world problems by forming a consortium.

6. Examples of Ontology Research

Some of the research the authors' group have done to date can be summarized as examples of ontology engineering research. The following introduces a brief overview of them in order to exemplify what ontology engineering is.

6.1 Conceptual Level Programming And Task Ontology

The major objectives of MULTIS[Mizoguchi, 92][Tijerino, 93] project is to bridge the gap between domain experts and computers to enable computers to elicit domain experts' ways of problem solving using task ontology. This can be interpreted in another way: MULTIS can help end user describe how they perform a task at the conceptual level without considering how computer works. This interpretation gives us a new idea that "Conceptual level programming supported by task ontology" which is an advancement of utility of task ontology research[Seta, 96].

Our research in this direction has been made extensively and we reformalized task ontology as follows:

1) We explicitly represent the vocabulary intelligible both to end users and computers. To do this, we formalized task ontology in two levels such as knowledge level for humans and symbol level for computers.
2) Function as a syntax of the sentences composed using vocabulary in the task ontology.
3) Semantics at the conceptual level execution is defined in the knowledge level ontology
4) Semantics at the symbol level execution is defined in the symbol level ontology.
5) A language based object-oriented and logic paradigms is designed together with a sophisticated environment.

The last ontology is defined as axiom equivalents, while the second and third ones are defined as axioms.

6.2 Ontologies of Function And Behavior And Explanation Generation[Sasajima, 95]

We have been involved in the research on function and behavior representation. This research is motivated by a strong desire to know what function is and what behavior is. In spite of the long history of the research about this topic, no satisfactory model of them is not obtained yet. Needless to say, well-established understanding of them is indispensable to qualitative modeling, and hence model-based problem solving. Our research has been conducted under the goals as follows:

1) Deep understanding of function and behavior

2) To develop a powerful representation language for them aiming at standardization of basic components for qualitative modeling

3) Designing necessary and sufficient vocabulary for explanation generation at both of the behavioral and functional levels.

6.3 Ontology of Time And Causality of Fluid for Qualitative Simulation[Kitamura, 96]

Qualitative simulation based on sophisticated ontology of causality is indispensable for model-based problem solving. This research is deeply concerned with ontology of time and causality. Qualitative simulation has a different ontology from the real time. But it is not so clear how they are different. We identified 7 different time resolutions which our simulator can identified according to our naive understanding of causality:

1) locally simultaneous, 2) globally simultaneous, 3) fast transition not represented in the system explicitly but recognizable by humans, 4) slow transition through components, 5) normal time transition represented in the system using differential equations, 6) time until a partial equilibrium, 7) complete equilibrium. We demonstrated these time units are simulated by using our qualitative simulator implemented.

7. Conclusion

We described ontology engineering which should be attacked for the future of knowledge based technology. We hope ontology engineering contributes to promotion of content-directed research, and hence to coping with real world problem solving.

References

[Forbus, 95] E-mail communication on August 7th, 1995.

[Gruber, 92] Gruber,T.: A translation approach to portable ontology specifications, Proc. of JKAW"'92, pp. 89-108, 1992.

[Gruber, 94] Gruber, T.: An ontology for engineering mathematics, Proc. of Comparison of implemented ontology, ECAI'94 Workshop, W13, pp.93-104, 1994.

[Gruninger, 94] Gruninger, R. and M. Fox: The design and evaluation of ontologies for enterprise engineering, Proc. of Comparison of implemented ontology, ECAI'94 Workshop, W13, pp.105-128, 1994.

[Guarino, 95] Guarino, N. and P. Giaretta: Ontologies and knowledge bases towards a terminological clarification, Proc. of KB&KS'95, pp.25-32, 1995.

[Hawes, 94] ISO/IEC standard 11179: Specification and standardization of data element, Working Draft 5, ISO/IEC SC14, 1994.

[Hori, 94] Hori, M. and Y. Nakamura: Reformulation of problem-solving knowledge via a task-general level, Proc. of the Third Japanese Knowledge Acquisition for Knowledge-Based Systems Workshop: JKAW'94, pp.1-15, 1994.

[Kitamura,96] Kitamura, Y., M. Sasajima, M. Ikeda, and R. Mizoguchi: A method of qualitative reasoning for mode-based problem solving and its application to a nuclear plant, International J. of Expert Systems with Applications, Vol. 10, No. 3/4, pp.441-448, 1996.

[Lenat, 90] Lenat, D. and R. Guha: Building Large Knowledge-Based Systems, Addison-Wesley Publishing Company, Inc., 1990.

[Mars, 94] Mars, N.: An ontology of measurement, Proc. of Comparison of implemented ontology, ECAI'94 Workshop, W13, pp.153-162, 1994.

[Mizoguchi, 92] Mizoguchi, R. et al.: Task ontology and its use in a task analysis interview systems -- Two-level mediating representation in MULTIS --, Proc. of the JKAW'92, pp.185-198, 1992.

[Mizoguchi, 93] Mizoguchi, R.: Knowledge acquisition and ontology, Proc. of the KB&KS'93, Tokyo, pp. 121-128, 1993.

[Mizoguchi, 94] Mizoguchi, R. and J. Vanwelkenhuysen: Task ontologies for explaining expertise in problem solving context, Proc. of ECAI Workshop on Comparison of Implemented Ontologies, Amsterdam, pp.163-174, 1994.

[Mizoguchi, 95a] Mizoguchi, R. and Johan Vanwelkenhuysen: Task ontology for reuse of problem solving knowledge, Proc. of KB&KS'95, pp.46-59, 1995.

[Mizoguchi,95b] Mizoguchi, R, M. Ikeda, K. Seta, et al.: Ontology for modeling the world from problem solving perspectives, Proc. of IJCAI Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal, 1995.
[Nishida, 95] Nishida, T., H. Takeda, K. Iino, and M. Nishiki: A knowledge media approach to ontology development, Proc. of KB&KS'95, Enschede, pp.84-94, 1995.
[Sasajima, 95] Sasajima, M., Y. Kitamura, M. Ikeda, and R. Mizoguchi: FBRL: A Function and behavior representation language, Proc. of IJCAI-95, pp. 1830-1836, Montreal, 1995.
[Seta, 96] Seta, K., M. Ikeda, O. Kakusho, and R. Mizoguchi: Design of a conceptual level programming environment based on task ontology, Proceedings of the First International Conference on Success and Failures of Knowledge-Based Systems in Real-World Applications, Bangkok, pp.11-20, 1996.
[Schank, 75] Schank, R: Conceptual information processing, North Holland Pub. Co., Amsterdam, 1975.
[Skuce, 95] E-mail communication on July 24th, 1995.
[Sowa, 95] Sowa, J.: Distinctions, combinations, and constraints: Proc. of IJCAI Workshop on Basic Ontological Issues in Knowledge Sharing, Montreal, 1995.
[Swartout, 93] Swartout, W. et al.: Knowledge sharing: Prospects and challenges, Proc. of KB&KS'93, Tokyo, pp.95-102, 1993.
[Vanwelkenhuysen, 94a] Vanwelkenhuysen, J. and R. Mizoguchi: Maintaining the workplace context in a knowledge level analysis, Proc. of JKAW'94, Hatoyama, Japan, pp.33-47, 1994.
[Vanwelkenhuysen, 95a] Vanwelkenhuysen, J. and R. Mizoguchi: Ontologies and guidelines for modeling digital systems, Proc. of 9th KAW, Bannf, Canada, 1995.
[Vanwelkenhuysen, 95b] Vanwelkenhuysen, J. and R. Mizoguchi: Workplace-Adapted Behaviors: Lessons Learned for Knowledge Reuse, Proc. of KB&KS '95, pp.270-280, 1995.
[Wielinga, 93] Wielinga, B. et al.: Reusable and sharable knowledge bases: A European perspective, Proc. of KB&KS'93, Tokyo, pp.103-115, 1993.