Granularity of Knowledge from Different Sources

Maria A. Mach and Mieczyslaw L. Owoc

University of Economics

Komandorska 118/120

53-345 Wroclaw, Poland

maria.mach@ue.wroc.pl

mieczyslaw.owoc@ue.wroc.pl

Abstract: The paper deals with the problem of knowledge granularity in case of building intelligent systems. The origin of the problem is discussed, some knowledge classifications are presented, next the links between types of knowledge and knowledge granularity are shown. In the last part of the paper the question of knowledge granularity types and their usage in intelligent systems is presented and discussed.

Keywords: knowledge based systems, knowledge heterogeneity, knowledge granularity.

1. Introduction

Knowledge existing in modern information systems usually comes from many sources and is mapped in many ways. There is a real need for representing "knowledge pieces" as rather universal objects that should fit to multi-purpose acting systems. According to great number of information system's tasks, knowledge representation is more or less detailed (e.g. some level of its granularity is assumed). The main goal of this paper is to present chosen aspects of expressing granularity of knowledge implemented in intelligent systems. One of the main reasons of granularity phenomena is diversification of knowledge sources, therefore the next section is devoted to this issue.

2. Heterogeneous Knowledge as a Source for Intelligent Systems

Knowledge, the main element of so-called intelligent applications and systems, is very often heterogeneous. This heterogeneity concerns the origin of knowledge, its sources as well as its final forms of presentation. In this section the selected criteria of knowledge differentiation will be presented, in the context of potential sources of knowledge acquisition. In Fig. 1 an environment of intelligent systems is shown, divided into different knowledge sources for the system.

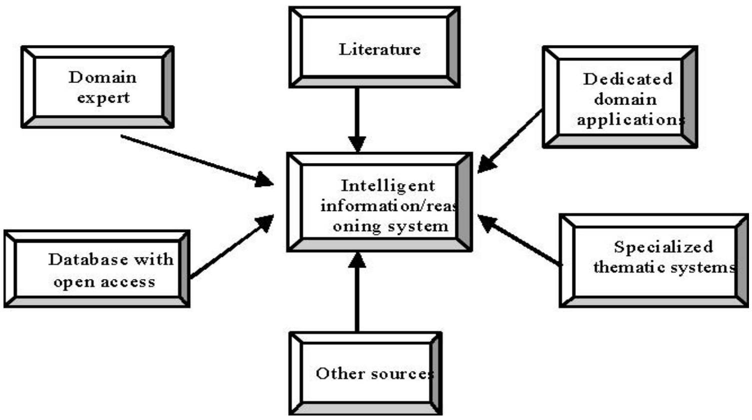

Fig. 1. Potential knowledge sources for intelligent information/reasoning system. Source: own elaboration based on (Mach, 2007) p. 24.

Classical knowledge sources are as follows: literature concerning the problem to be solved; domain experts possessing knowledge on the way of preparing decisions; databases, from which – using appropriate techniques – it is possible to acquire knowledge; specialized knowledge base systems, providing useful intermediate expertise (see Nycz&Smok, 2000). Knowledge gained from each of the above mentioned sources may have a specific form (e.g. report on discussion with an expert, description of a problem solving method or a set of rules generated on the basis of former system's correct reasoning procedures).

The sources pointed out above constitute the main factor, enabling to differentiate the acquired knowledge. Nevertheless, assuming more precise criteria it is possible to obtain different forms of knowledge, taking into account e.g. the domain and type of knowledge application, or the way in which knowledge is transformed during the process of building an intelligent system.

With the first criterion – domain of knowledge application – we may point out the following types of knowledge: the one supporting management (manufacturing and different kinds of business) and other forms of human activity, e.g. medicine or military activity. This criterion is strictly connected to the next one – category of tasks being supported. In this context, we may speak of: knowledge used for classification, for diagnostics, for monitoring etc. Computer applications making use of particular knowledge types, usually concern a concrete domain (or several domains) and generate solutions according to previously defined tasks.

According to the third criterion – the type of domain knowledge – we may mark off several specific forms of knowledge, encompassing among others:

– declarative knowledge – that is knowledge on what is already known about the problem,
– procedural knowledge – stating how the problem may be solved, and finally,
– heuristic knowledge – describing expert's experience, gained during previous problem solving procedures (see e.g. Durkin, 1994).

The type of knowledge is in this case identified according to the source knowledge comes from.

If we use the criterion of knowledge representation form, we have two main classes of solutions: the symbolic and non-symbolic knowledge (see Durkin, 1994). In case of symbolic representation, we deal with so called explicit representation (it will be discussed below), while non-symbolic representation concerns such representations, as neural networks, genetic algorithms, or algorithms of inductive training.

The next criterion concerns the way knowledge is represented: knowledge representation techniques. It leads to a more detailed diversification, because it is strictly connected with the previous classification (symbolic and non-symbolic representations). According to the classical depictions of the topic, one may speak about knowledge expressed in the form of:

– propositional logic,
– semantic networks,
– rules,
– mixed techniques (hybrid knowledge, e.g. frames).

The last criterion – method of knowledge creation – allows for denoting knowledge represented according to the selected methodology, and knowledge generated with the aid of selected tools. The first type of knowledge (bases) is built in three main steps: knowledge acquisition from an expert, knowledge representation using one of the above mentioned techniques, and knowledge implementation. It may be therefore said that this type corresponds with symbolic representations. On the other hand, generated knowledge bases correspond generally with non-symbolic representations and are built up automatically.

Summing up, we have to deal with so-called multi-sourced knowledge, which in intelligent systems may be represented using different methods, may have different forms and different structures. In particular, these structures may have a different level of details, which in turn is connected with the phenomenon of knowledge granularity.

3. Granularity Concepts in Information Technologies

One of the crucial aspects of information architecture is how to define data structures for multi-purpose usage and to process it effectively. Some of these data structures can include repeatable elements while some cover different levels of data details. At least two main trends in information technologies touch such issues (see: BitsOfPower, 2007) :

1. increasing recognition of the importance of standards and
2. growing acceptance of a need for cooperation in monitoring and controlling network.

In both cases a concept of creation universal data "storages" seems to be a potential solution.

In practice such approach is typical for two advanced technologies: object-oriented (in databases, programming languages, knowledge representation techniques) and distance learning (e.g. in formulation of learning units, sharing common "knowledge pieces"). Modeling information and knowledge architecture in such a context a category of an object or a learning object are defined respectively in mentioned technologies.

Nevertheless of data/information/knowledge unit's content the problem of its range and level of details is fundamental for encapsulation of data collections. In other words, granularity phenomena in informational infrastructures can be observed and investigated.

Granularity is a term used in photography to describe accuracy or measure of pictorial presentation on film (the higher level means more details). There are several interpretations of this category in physics, computing and risk management for example. Granularity of information resources refers to size, decomposability and the extent to which a resource is intended to be used as part of a larger resource (see: Wagner, 2002).

It should be stressed certain aspects of “sizing” or “dimensioning” that are present in particular definitions of granularity.

First, granularity refers to **temporal** context of a defined object e.g. provides service of an acting objects across the time. Sometime we need to gather and to process information structures daily, weekly, monthly and the like. Smaller time units allow to represent more details of the investigated phenomena.

Second, granularity relates to **spatial** dimension of a determined entity e.g. cover its functioning in a particular space. Again, the smaller space unit the more details can be achieved.

Third, granularity considers ways of **assuring consistency** of objects belonging to information architecture. Two approaches represent this type of granularity: abstraction and aggregation. In a case of abstraction generated objects are a result of generalization procedures while aggregation deals with defining conditioning of joint objects.

More detailed taxonomy of granularity including formal interpretation of interrelated objects is presented by C.M. Keet – see (Keet, 2006).

General concepts of granularity implemented in IT sector very often are oriented at intelligent systems. The process starts with representation of knowledge pieces in targeted applications. Of course solutions based on particular granularity concept allow (or not) to obtain defined goals of such systems.

4. Granularity and Ontology

In Section 3 some concepts of granularity have been briefly outlined. The question is, where these granularities come from. In our opinion, granularity is strictly connected with the notion of ontology. Therefore, let us now investigate different granularities in context of different ontologies, as suggested in previous section.

First, the temporal ontology typically concerns ontology of time, that is, what is time composed of: points, intervals or both, as basic temporal entities. But in more common, everyday context, ontology of time concerns calendar units, and here one deals with such units as e.g. years, quarters, months, days, referred to as time granularities (Goralwalla et al., 2001). Granularities are unanchored durations, which can be used as units of time (ibidem). The most complete formal framework for manipulating temporal granularities was described by (Bettini et al., 2000).

Spatial ontology strictly depends on the definition of space adopted. For example, if one assumes a metric space, that is a generalization of n-dimensional Euclidean space, the only ontological elements of this space are points and there is no granularity problem. However, this is of course not the only model of space possible.

In the field of artificial intelligence, where time and space are often considered together, the region-based theory of space is often assumed and employed (see e.g. Bennett, 2001). It is a theory of spatial regions based on parthood relation P(r1, r2), and the sphere predicate S(r). If we adopt the RGB theory of space, every region may be treated as a granule in space and the problem here is whether these granules may be compared or not. If so, what conditions have to be fulfilled by the space granules to compared them?

Information architecture ontology deals with objects considered as primitive units of information architecture. They depend on the level of architecture that are considered as a basic one, e.g. logical and physical one.

Granularity phenomena can be defined including many purposes. One of the most demanding approaches comes from the e-learning area. Learning courses can be divided into "knowledge pieces" according to audience familiarization with presented topics or aims of the course. Therefore from logical point of view we may separate knowledge presenting definition of some phenomena, put some procedures how to classify some objects or give examples of this sort procedural knowledge. Of course ontologies mentioned before (spatial or temporal) are actual in may respects e.g. a definition of a database can be detailed for particular users taking into account their properties and ways of applications. On the other hand physical "pieces of knowledge" can be identified as different files or cluster of some media aggregations.

Nowadays knowledge granularity is strictly connected to knowledge grid ideas. Any concept of knowledge grid acquires resolving problems with knowledge aggregation and its distribution. In every case we are obliged to divide the whole domain knowledge into units sometimes at many levels.

Granularity concepts presented earlier can be applied in different areas. Let's try to investigate their usability in the knowledge acquisition process.

One of the mentioned granularity types stressed importance of time category. This type of knowledge granularity is connected with the question of representing knowledge about dynamic heterogeneous environment in the intelligent system's knowledge base. In this case, it may happen, that particular elements of the environment present different pace of changes. This leads to the need of representing temporal knowledge that is heterogeneous in the temporal context: each part of knowledge (concerning a particular environment element) may have a different time granularity. In this situation a solution may consist of using different temporal formalisms for knowledge representation, which in turn leads to the problem of representation integration. This question is beyond the scope of this paper, more details may be found in (Mach, 2005).

Spatial aspect of knowledge granularity refers to gathering domain knowledge from many sources. They can represent different subjects allocated in separated sites. The main problem is how to implement standards of knowledge representation that will lead to universal form of knowledge mapping in case of diversification of sourced data. This is a typical challenge for hybrid intelligent systems.

Procedures of knowledge refinement use the third type of the mentioned phenomena, namely abstraction and aggregation. Looking for more efficient knowledge bases we try to discover new knowledge pieces to generalize initially introduced information or to elaborate useful extended "knowledge grains".

5. Conclusions

Aspects of knowledge granularity presented in this paper refer to rather specific way of information processing called granular computing. In a more philosophical sense, granular computing can describe an approach that relies on the human and computer ability to perceive the real world under different levels of granularity (i.e., abstraction). In order to abstract and consider useful from the defined goals knowledge pieces should be represent and switch among different granularities. Focusing on different levels of granularity, one can obtain various levels of knowledge, as well as a greater understanding of the inherent knowledge structure. Granular computing can be treated as an essential way of human problem solving and hence should have a very significant impact on the design and implementation of intelligent systems.

In order to establish a proper granularity, one has to investigate the context of the problem and the ontology of the domain. In our opinion ontology plays the decisive role in establishing knowledge granularity. In case of ontology and/or knowledge sources heterogeneity, there is the need of unification before establishing the final granularity. These problems are discussed e.g. in (Mach, 2003).

References

Bennett B., Space, Time, Matter and Things. Proc. FOIS'01. ACM, USA, 2001.

Bettini, C., Jajodia, S. & Wang, S. X. (2000). Time Granularities in Databases, Data Mining, and Temporal Reasoning. Berlin Heidelberg: Springer-Verlag.

www.nap.edu/readingroom/books/BitsOfPower/2.html; 2007

Durkin J.: Expert Systems: Design and Development. Prentice Hall, Enlewood Cliffs 1994.

Goralwalla, I. A., Leontiev, Y., Ozsu, T. M. & Szafron, D. (2001). Temporal Granularity: Completing the Puzzle. Journal of Intelligent Information Systems, 16, 41-46.

Keet, C.M. A taxonomy of types of granularity. IEEE Conference in Granular Computing (GrC06). 10-12 May 2006, Atlanta, USA.

Mach M., Integrating Knowledge from Heterogeneous Sources. "Argumenta Oeconomica", No. 1-2(14) 2003, University of Economics, Poland Publishing House, PL ISSN 1233-5835, pp. 189-210.

Mach M., Analysing Economic Environment with Temporal Intelligent Systems: the Union of Economic and Technical Perspectives, [in:] Tadeusiewicz R., Ligeza A., Szymkat M. (eds.), Computer Methods and Systems, Volume I, Kraków 2005, ISBN 83-916420-3-8, p. 355-366.

Mach M. : Temporalna analiza otoczenia przedsiebiorstwa. Techniki i narzedzia inteligentne. [Temporal analysis of enterprise's environment. Intelligent tools and techniques]. Wydawnictwo Akademii Ekonomicznej Wroclaw, 2007 (in Polish).

Nycz M., Smok B.: Knowledge Sources for Expert Systems. Transactions in International Information Systems. Systems Analysis and Development Theory and Practice. Nowicki A. and Unold J. (eds.) Wroclaw university if Economics

Wagner, E.: Steps to Creating a Content Strategy for Your Organization. eLearning Developers' Journal. eLearning Guild. October 29, 2002. Retrieved from http://www.elearningguild.com/pdf/2/102902MGT -H.pdf on 12/07/2007