Hanfei Bao: Conceptual, Experimental and Compass-beacon Ontologies —— **The Principles, concepts and methodology of BMKI(Ⅰ),Medical Information, 18(8), 2005**

Conceptual, Experimental and Compass-beacon Ontologies

——The Principles, concepts and methodology of BMKI(Ⅰ)

Hanfei Bao

The Lab. Of Informationization and Standardization of TCM
Shanghai University of Chinese Traditional Medicine
201203,bhf2002@online.sh.cn

Abstract This article discussed the basic three components of an information unit(IU), ie the Carrier(C), Feature(F) and Value(V), and presented and several different types of feature, including the static, relation and process features(SF ,RF and PF respectively). It also discussed about the knowledge for statistical group(KP) and for individual(KI) and the transformation mechanisms between them. Then both the philosophy and application attributes of ontology have been explored and two new concepts, ie ontology of existence(EO) and ontology of application(OA) were created. The paper gave an integrated sketch of the whole medical knowledge engineering cause, which is thought to be composed of three areas: conceptual, experimental and compass-beakon ontoloies (CO, EO, CBO, respectively). And based on those considerations BMKI achieved its fundamental formula

$$BMKI = \Theta[(CO \oplus EO) \otimes KI] = \Theta[(CBO) \otimes KI].$$

Key words Artificial intelligence Medical informatics Theory of BioMedical Knowledge Integration Medical ontology

1. Starting from Information Unit(IU)

Radically speaking, any information unit(IU) needs to answer these questions: (1)What object we are going to know? (2) Which aspects of the object we have chosen to be observed? (3) What is the result of the observation? We call these three attributes of an IU the Carrier(C), Feature(F) and Value(V) , respectively. Namely, the three parts construct an IU and IU = （C×F×V）(see Fig. 1). In SNOMED, feature is called “observable”, explained as “Entities that are the subject of an observation”.

Carrier(C)	Feature(F)	Value(V)

Information carrier can be divided further into two types, ie of individual and statistical group. Accordingly the knowledge has also two types, the knowledge for individual(KI) and for group(KP). T

Fig. 1　The basic three parts of an information unit(IU): Carrier, Feature, Value.

(1) ***Defini***

Definition 1 knowledge for individual(KI) the knowledge about an individual or the knowledge the carrier of which is an individual.

Definition 2 knowledge for group(KP) the knowledge about a statistical group or the knowledge the carrier of which is a statistical group.

(2) ***The transformation between KI and KP:***

In our cognitive processes, the transformations between KI and KP are very high frequent intelligent actions.

a. Finding and assigning of a knowledge pattern: The processes to find the knowledge patterns are the transformation from KI into KP, to determine which parts and their organizational patterns of a data set of many individuals(KI) are the relatively common or stable content(KP); Contrarily, the processes to assign the data of an individual(KI) to a knowledge pattern(KP) are the ones under the guidance of KP to determine which pattern(KP) a KI belongs to or is diagnosed as, based on the restricts or particular values provided by KI(see the Fig. 2) .

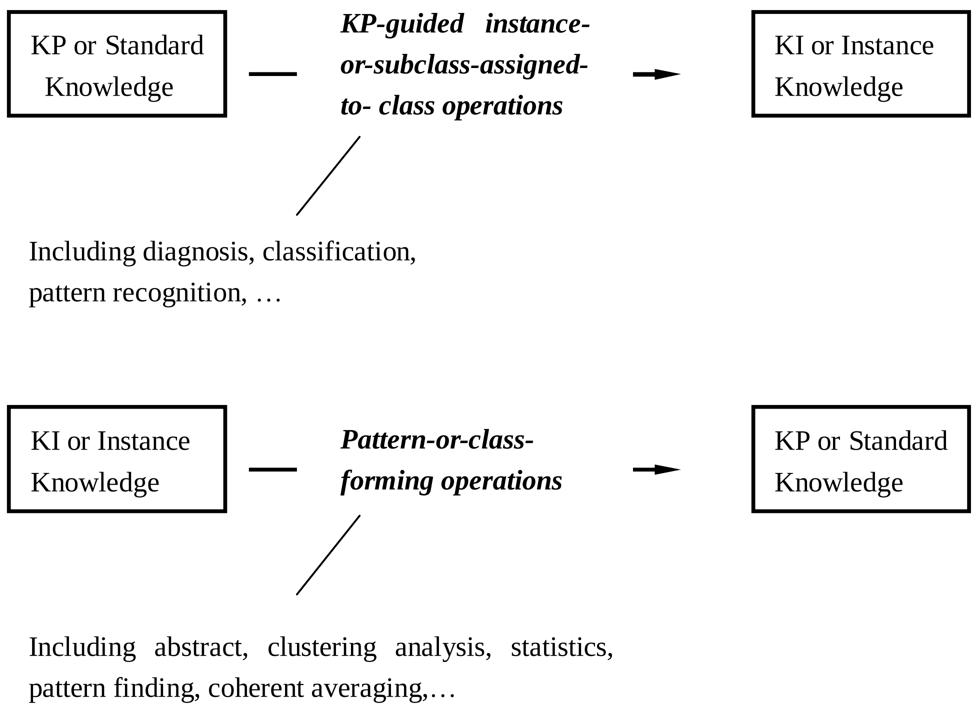

Fig. 2　reciprocal transformation between KI and KP.

b. As we known, when we make an abstract we will abnegate some attributes which are different between the classes and reserve what are common for classes, thus we get a new and lager class which is, mathematically, an union of the old classes. Let $\xrightarrow{a}$ and $\xrightarrow{s}$ **represent abstraction (or generalization) and reification (or specification), respectively, and if A class and B class have the relationships** A$\xrightarrow{a}$B or B$\xrightarrow{s}$A, **then A is the subclass or instance of B, B is the parent class of A. The knowledge on A is more specified or individual-quality compared with B and, reversely, the knowledge on B is more generalized or group-quality compared with B.**

Likewise, there also exists the same transformation between the relatively generalized and relatively specified propositions(see Fig 3).

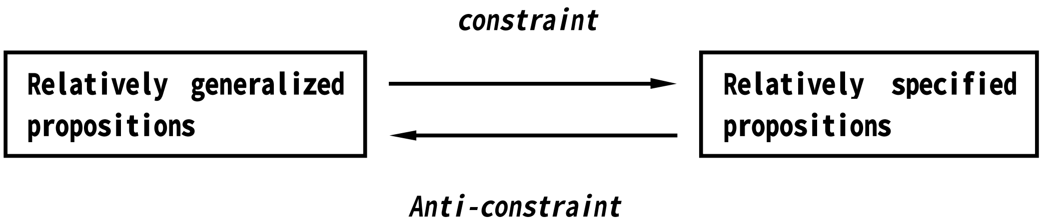

Fig. 3 The transformation between relatively generalized and specified propositions.

(3) *Other operators or transformations between the parts of IU:*

We can apply many other types of operators such as ***-is a-*** relation, ***-a part of-*** relation, ***function*** relations, ***mapping*** relations, ***physical semantics*** relations of different levels**,** etc to process or calculate data, information and knowledge, making the reciprocal transformations. Additionally, we can by choosing and adding some reasonable and necessary constraints realize the transformation from the relatively general, rough-granularity and uncertainty knowledge to the relatively particular, fine-granularity and certainty knowledge, or do reversely.

2. The Types of Features

As stated above, so called the features of IUs are the those aspects observed. As a matter of fact, the aspects of sth to be observed are a colourful and endless world. To choose which feature to be observed or measured, of course, is determined by the problem space to be resolved, and sometimes it reflects the wisdom of the observer. But, nevertheless, they may roughly classified into tree main types, according to their change-ability, combination-degree, etc.

(1) ***Definition 3 Static feature(SF)*** Static features usually cover those characteristics which are relatively unchangeable while you observe or measure them and therefore having a value of constant, which may be euther numerical or descriptive. For instances, intraocular pressure(Schiotz orphthalmotonometer): right eye 6.6kPa（49.76mmHg）, left eye 2.7kPa

（ 22.55mmHg ） or those with descriptive or textual value: bulbar conjunctiva(C)-congestion(F)-(exists(V)) , cornea(C)-opacity (F)-(exists(V)), etc. The form of SF usually is ***&f***, ***&*** is the evaluating or assigning operator.

At bottom, any types of feature are relationships (binary, multi-elemental, dynamic, etc.) in nature. But under many situations, we don't need to point out all participants of the relations. For examples, we don't need to call "height" the relation as "one between the upright length and ruler", and "color" as "one between the incident light and reflecting light".

(2) ***Definition 4 Relation feature(RF)*** Most of them are binary relations, such as "the sizes of the right and left pupils are equal" (***equality*** relation), "eye has a foreign body" (***having*** or ***existence*** relation)，the reaction of pupils to light(***reaction*** relation), "resolving power of eyesight for red colour"(***relation between red colour and eyesight***), etc. The form of this type of feature is ***&(a₁*a₂)***, a_1，a_2 are two participants of the relation , ***** is the relation and ***&*** is the evaluating or assigning operator.

Similarly, any thing has numberless observable or measurable relations, obvious or potential，to be chosen, and which are determined by the questions under discussion or based on the question space, as well. For instance when you want to valuate "the reaction ability" of a driver, you can observe the performance when an traffic suddenness happens, eg. a dog abruptly appears in front of the car. Otherwise, this ability is something potential.

(3) ***Definition 5 Process feature(PF) or dynamic relation feature*** This kind of feature can be viewed as a consequence relation chain or continuum, which has, already, been started up and carried out. The form of them may be $\& (r_1 * r_2 * \ldots * r_n)$, where $r_i = (a_1, a_2 \ldots a_n)$ be the relations and a_1，a_2，…a_n the participants of those relations,***** is the second-order-relations between relations, ***&*** is the evaluating or assigning operator.

The generalized value of a PF usually is not a constant and may be a set of digit pair. Fig 4 shows the value generalized of x = 2(x)(a dynamic relation), which is a function indicating a series of consequent states.

1	**2**	**3**	**4**	**5**	**6**	**7…**
2	**4**	**6**	**8**	**10**	**12**	**14…**

Fig 4 The generalized value of a PF, showing a dynamic mapping among a series of consequent states.

3. Ontology of Existence and Ontology of Application

(1) what is ontology:

T. R. Gruber[15] pointed out:

The word "ontology" seems to generate a lot of controversy in discussions about

AI. It has a long history in philosophy, in which it refers to the subject of existence. It is also often confused with epistemology, which is about knowledge and knowing.

In the context of knowledge sharing, I use the term ontology to mean a specification of a conceptualization. That is, an ontology is a description (like a formal specification of a program) of the concepts and relationships that can exist for an agent or a community of agents. This definition is consistent with the usage of ontology as set-of-concept-definitions, but more general. And it is certainly a different sense of the word than its use in philosophy.

Thus we have two almost opposite interpretations of the word ontology. When it refers to "the subject of existence", it answers "What is the world". Hence we may call it "ontology of existence"(OE). When it means "a specification of a conceptualization" in the context of knowledge sharing, however, it answers "How the world is used", thus we may call it "ontology of application"(OA). The OEs are developed, being driven by our curiosity, whereas the OAs by our will for benefit.

Perhaps the distributions of OE and OA can be symbolically shown as Fig. 5, where OE = Environment ontology plus Organism ontology, OAs = Trophology ontology + Neurology ontology + Pneumology ontology + Cardiology ontology + …

The goodbye each other of OE and OA is apparently because of the infinity quality of OE and the high efficiency of localized and repeated work(OA).

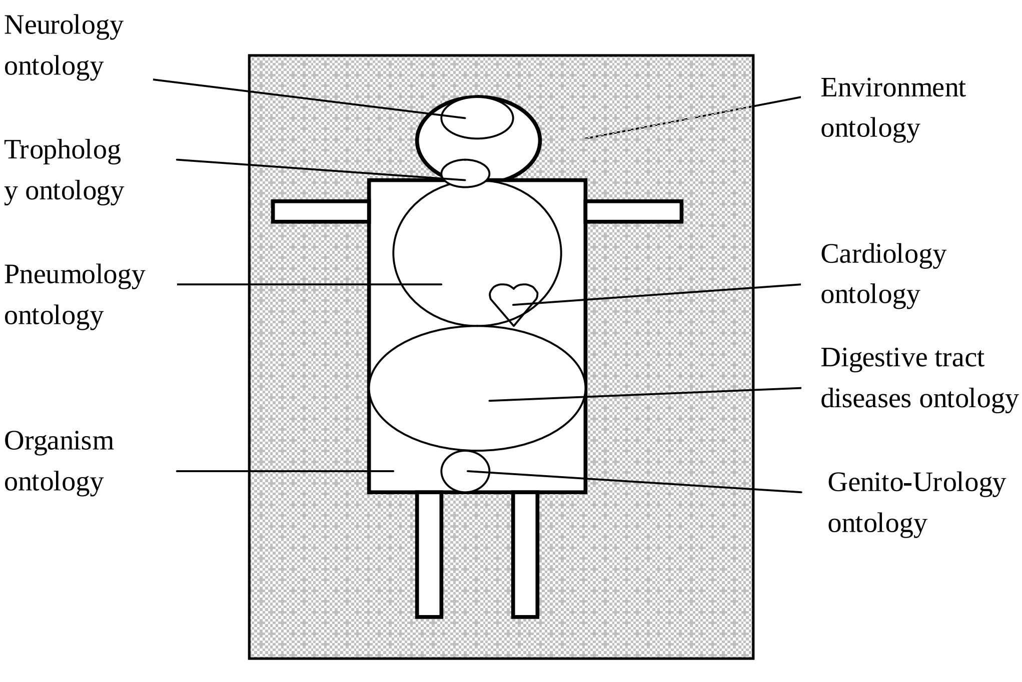

Fig. 5 Two kinds of ontology with the different essences in BioMedical Informatics: OE and OA. The former is consisted of environment and organism, whereas the latter of serious application domains.

4. Conceptual and Experimental Ontologies

But the author is most interested in some other qualities of an ontology. Let's see again what **T. R. Gruber**[15] **talked about:**

What is important is what an ontology is for. My colleagues and I have been designing ontologies for the purpose of enabling knowledge sharing and reuse. In that context, an ontology is a specification used for making ontological commitments.

Because the important characteristics of an ontology are its sharability and reusability, as understood by the author, an ontology should be, after all, the standard (or typical) and conceptual knowledge on the universality of a statistical group rather than of an individual.

And based on this interpretation, the author extended the idea of ontology to the *statistically standard (or typical)* and experimental (rather than conceptual) knowledge on the universality of a group. Thus the current mainstream ontology is called the conceptual ontology（CO） and the new "ontology" is the experimental ontology(EO).

Definition 6 Conceptual Ontotogy(CO) The specifications of a conceptualization, which have the conscious nature.

Definition 7 Experimental Ontotogy(EO) The descriptions of a standard mapping, which are resulted from the physical and statistical experiments.

CO is a hot development area of AI, cognitive sciences, and knowledge engineering of medical informatics, nowadays. By dint of some new ontology language such as **DAML+OIL,OWL(Web Ontology Language), COs have been developed as distributed, frame-structured, sharable and reusable knowledge sources. And because of the Internet-locating ability, it is quite possible that they may be developed in world-wide-joint way**[16-17] **.**

Contrastly，EOs are a wide desolate virgin land and people don't know how to use them more effectively. Theoretically, without the participation of OEs, COs are only a "**skeleton beauty dancing**" or a "**reinforced-concrete-framed building**"**. And when we can see something more perfect and enjoy a** "**nice girl dancing**"**?**

In the aspect of how to use Eos, the Theory of BioMedical Knowledge Integration(BMKI)[1-14,18] **has following ideas:**

Many biologic mechanisms or relations in BioMedicine were found by means of scientific experiment, and they are, at may cases, non-interpretable or non-explainable but conditionlessly accepted. It is unimaginable that we talk about biomedical knowledge integration without their participation. In BMKI we call these mechanisms or relations the biological non-explainable operators. They are the very complicated fields, including: (1) various physiological features and their ranges of normal values. (2) digital functions between substances, times, conditions, etc. Some functions have the shape of cycle: Let y = $f(t)$ **is a function on time and** $f(t) = f(t + \triangle t)$, $\triangle t$ **is a fixed period of time, then** $f(t)$ **is a time-cycle**

function; If we have an order relation chain(eg causality chain), *f={(x_i,x_{i+1})|(x_i,x_{i+1})∈f and (x_i,x_{i+1})= (x_{i+θ},x_{i+1+θ})}*, θ is a positive integer, then *f* is a process-cycle function. (3) Where or in which spaces those digital functions become effective.

According to time-axis, we may have a time-synchronous integration(TSI).It is a primary or syntax-level knowledge integration or a mapping-type digital "human-body" :

If we have a set of function $g_{ci} = f_i(t- t_0)$, i =1,2,3,…, g_{ci} are lung-volume-to-time cycle, left-ventricle-volume-to-time cycle, blood-sugar-to-time cycle……, and they can be expressed in the form of discreteness. t_0 is a common start-point of time determined arbitrarily. Let G_c be human-body-state-to-time function, then $G_c = I_n(g_{ci})$, i =1,2,3,…, is a kind of time-synchronous virtual human body, I_n is a time-synchronous and syntactic integration operator. I may be a series of different integration operators and $I_n \in I$. Fig.6 is a sketch map indicating time-synchronous virtual human body.

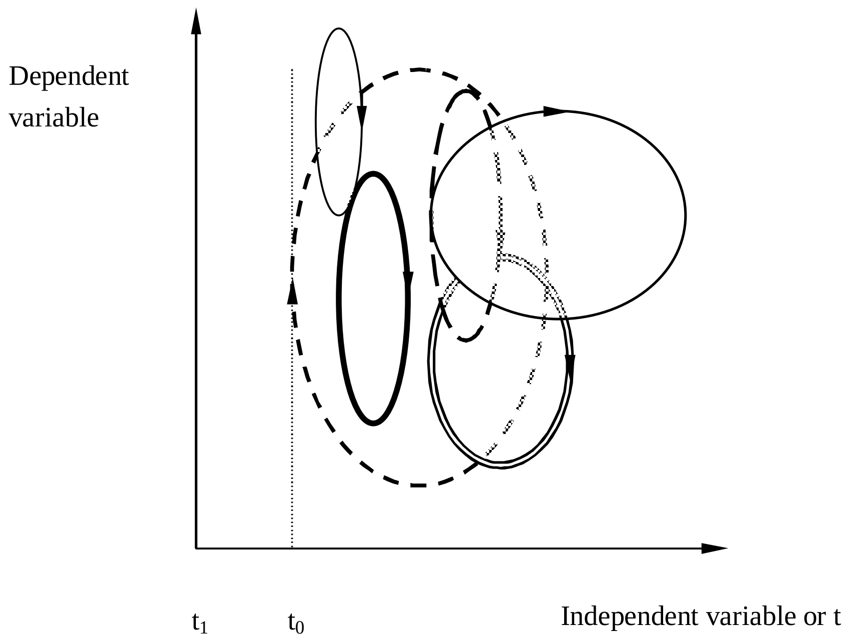

Fig. 6 The sketch map of a kind of time-synchronous virtual human body. t_0 is a common start-point of time determined arbitrarily,t is general time axis.

5. Compass-beacon Ontologies

As stated above, If we wish our Medical Knowledge Engineering Building is not only a "**reinforced-concrete-framed building**", **but also with windows, doors, floors,**

ceilings, etc, we should try to combine COs and EOs together and we may be called the combination Compass-beacon Ontology(CBO), which is the overall stratagem of BMKI. Otherwise we can never get rid of a pity of “a unperfect sky”.

Definition 5 *Compass-beacon Ontology(CBO)* The combination of CO and EO.

6. Compass-beacon Ontologies

Apart from the standard or typical knowledge(CO and EO), which is from statistical group, we also have the huge resource of knowledge, which is from individual and stored in the various dadabase or electronic medical record(EMR). Thus the whole of medical knowledge mainly includes the three basic parts: two types of conceptual onlology CO and EO, and data or knowledge on the individuals(KI))(see Fig. 7).

As described in my previous papers[1-14,18], the role of BMKI is to explore the relationships, reciprocal transforms and links between the hetero-background or hetero-spatial biomedical knowledge. From discussion above, we can get the basic formula of BMKI,

$$BMKI = \Theta[(CO \oplus EO) \otimes KI] = \Theta[(CBO) \otimes KI],$$

where *CBO is* Compass-beacon Ontology, symbols $\Theta, \otimes, \oplus$ mean the relationships, reciprocal transforms and links between the hetero-background or hetero-spatial knowledge.

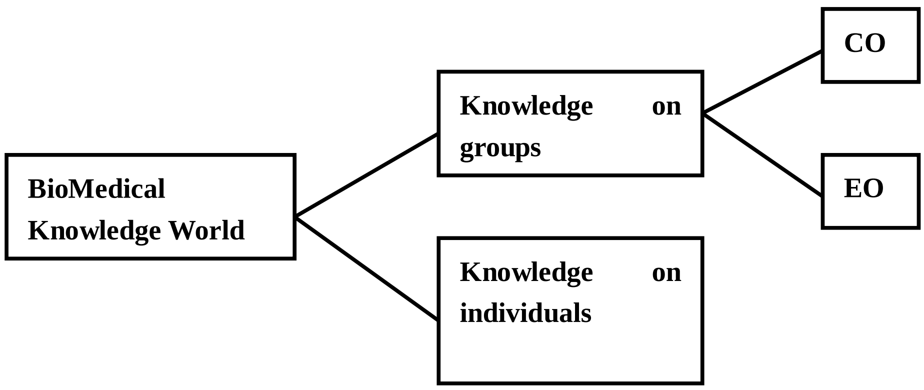

Fig. 7 The main tree parts of BioMedical Knowledge.

References

1.包含飞：生物医学知识整合论(Ⅰ),《医学信息》杂志，16（6）：174-279，2003
2.包含飞：生物医学知识整合论(Ⅱ),《医学信息》杂志，16（8）：410-415，2003
3.包含飞：生物医学知识整合论(III),《医学信息》杂志，16（11）:602-605,2003
4.包含飞：生物医学知识整合论(Ⅳ),《医学信息》杂志，16（12）：666-670,2003

5.包含飞：生物医学知识整合论(Ⅴ)，《医学信息》杂志，17（5）：244-250,2004
6.包含飞：生物医学知识整合论(Ⅵ)，《医学信息》杂志，17（8）：452-457,2004
7.包含飞：生物医学知识整合论(Ⅶ)，《医学信息》杂志，17（11）：685-692,2004
8.包含飞：生物医学知识整合论(Ⅷ)，《医学信息》杂志，18（1）：11-15，2005
9.包含飞：生物医学知识整合论(Ⅸ)，《医学信息》杂志，18（3）：172-176，2005
10.包含飞：生物医学知识整合论(Ⅹ)，《医学信息》杂志，(待发表)
11. Bao HF: The Theory of Biomedical Knowledge Integration(Ⅰ), Chinese Journal of Medical Treatment, 2(13): 1-6, 2003
12. Bao HF:The Theory of Biomedical Knowledge Integration(Ⅱ) ——What is underlying the domain concepts, Chinese Journal of Medical Treatment, 2(21): 1-6, 2003
13. Bao HF: The Theory of Biomedical Knowledge Integration(Ⅲ), Chinese Journal of Medical Treatment, Vol(3),No(7),2004
14. Bao HF: The Theory of Biomedical Knowledge Integration(Ⅳ), Medical Information，(accepted to be published)
15. T. R. Gruber. What is an Ontology?
http://www-ksl.stanford.edu/kst/what-is-an-ontology.html
16.梅婧 刘升平 林作**: Chpt-63-**语义**Web** 语言的逻辑分析**,**
http://www.is.pku.edu.cn/~lz/publications/semanticweblang.pdf
17.M. Taboada1, at al, 包含飞, 王晓幸翻译:DAML+OIL 应用实例：眼科学领域的一个本体,**http://www.miforum.net/bhf/ykbt.doc**
18.王晓幸等：眼科信息学：眼科知识的整合研究（Ⅰ），(已接受，待发表)