First CJKMI(Oct., 1999, Suzhou, China)
ZR Li et al. (Eds)
Beijing: Xue Yuan Publishing house
© 1999. All right reservedProceedings of The
First China-Japan-Korea Joint Symposium on Medical Informatics (CJKMI'99),1999,177-181

HCSL: a Human-Computer Commonly Understandable and

Communicatable Medical Language

Han-Fei Bao

computer network center Shanghai University of Traditional Chinese Medicine, China

Abstract

This article presented an exploratory program to format a kind of half-structured and half-natural medical language which would be understandable by both human and computer, and communicatable between them. The purpose of it is to develop a new type of medical knowledge base by which human and computer could "talk" more thoroughly. Some philosophic discussion, the classification and exemplification of the elements and relations of the language were given as well.

Keywords

Medical Knowledge Integration; Artificial Medical Intelligence; Internet

1. Some Philosophic Thoughts

Medical knowledge system is an infinite system which covers all the relations in and between two spaces, the space inside the body and the space outside the body.

Any individuals should, for the purpose of their living and health, know some basic medical knowledge, whereas a doctor should, for the purpose of helping people with their living and health, know all the knowledge in his or her field.

In this aspect, people are challenged by the features of medical knowledge and their relevant actual processes in the body:

- The hiddenness. The events or processes actually happen in the body are usually unobservable directly.
- The indefiniteness. The multiple possibilities of inner processes of the body exist for the same set of symptoms.

- The boundlessness. Namely, there are no definite boundaries between the knowledge scopes of the specialty fields. They are connected each other.
- The inseparateness. The discovery of medical knowledge is needed to do by way of "social division of labour", but the nature and the application of medical knowledge are of inseparateness, it is an organic whole.

Perhaps because doctors are the noble kings in the medical kingdoms, little researches of medical informatics in the past were aimed at the limitedness of the doctor's memory and shortness of their thought. The fact is that a good doctor spends most of his or her spare time in reading medical books and journals. The knowledge scope in doctor's memory is never large enough to do their works as well as desired (a similar view point can see [3]). Doctors need always to "talk" with books and journals to extend their knowledge area. Nevertheless, the doctors are isolated in the narrow knowledge world by the invisible ignorance-wall and the area outside is still something like unknown "stars of outer space" for them.

IREMB (Integratable Relationized Electronic Medical Book) has tried to develop a research program of "linking engineering" between various medical knowledge elements, units (or relations) and fields. It is considered that in the computer era the medical knowledge system should be no longer traditionally pile-typed, but somewhat nervous-system-typed or communication-system-typed. For this goal, a new kind of half-structured and half-natural language which can be understood and processed by both human and computer is basically necessary. We might call it human-computer shared language (HCSL).

Due to the different aims of problem observation, analysis and process in the various knowledge fields, the degree of simplification, generalization and abstraction of the simplification, generalization and abstraction of the knowledge are different. From lower degree to higher degree, the knowledge might be divided into three levels: physical level, systems-logic-mathematics-computer sciences level (SLMC level) and philosophy level. The knowledge of physical level covers all the practical problems, being generalized according to its physical methodology and expressed mainly by the natural language. That of SLMC level is generalized depending on the formalized methods and expressed usually by the structured language [1-2], such as various personal, medium, large data bases. It is difficult to use these absolutely restricted and formatted languages to express the majority or whole of medical knowledge. As for the knowledge of philosophy level, it is always generalized and abstracted on basis of the different thought methods.

In accordance with the three-level-division, our HCSL should be inserted into between SLMC level and physical level, resulting in a four-level-knowledge-division:



The IREMB medical knowledge base structured by HCSL language consists of three parts over all [4-16]. ①The element files which consist of the sets of encoded medical knowledge elements at physical level. ②The relation files each of which is composed of a set of medical knowledge units being formated by one of so called "knowledge driving verbs (*KDVs*)". Now at the initial stage ("one-person huge program" stage) of IREMB, only ten *KDVs* have been specified. They are ten generalized verbs: (pan) creating (-*CREAT*-), (pan) increasing (-*INCRS*-), (pan) decreasing (-*DECRS*-), (pan) containing (-*CONTN*-), (pan) passageway (-*PASTO*-), (pan) accompanying (-*COMPN*-), (pan) transforming (-*TRANS*-), (pan) equating (-*EQUAL*-), (pan) ordering (-*PORDR*-), (pan) irrelating (-*PNULL*-). In future, along side the development of IREMB, there might be tens or hundreds of *KDVs* needed to be specified. ③A series of softwares of IREMB, including editing system, maintaining systems, searching systems, automated logic-guided reading functions and other artificial intelligent functions.

2. The Sorts of Medical Elements

Any medical objects, e. g., substances, concepts, movements, functions, processes etc. are splittable, endlessly and infinitely. Therefore we mention "medical elements" always in relative sense. In IREMB, a simple element contains only one element, whereas a complex element contains two or more elements, linked with complex sign "-".

The complex medical elements might be exampled by "glucokinase-synthesis", "monilia albicans-septicemia", "expiratory center-excitability", "fatty liver-cause", "aspirintoxic or side reaction", "aspirin-careful use", "therapy", etc.

The medical elements might in HCSL be farther subdivided into the subelements so called "meaning elements", which are the concepts of high occurrence rate in medical knowledge, but not the integral medical objects. The examples of "meaning" elements are "excitability", "activity", "predilection site", "susceptible population", "predilection age", "after meals", "before meals", "excretion", "therapy", etc.

Based on their physical nature, the medical elements might be classified into following aspects: substance, concept, structure, movement, time, place, state (behaviour, form, feature, attribute, quantity, degree, parameter...), ability, domain, process (complex of state and time),

relation 'substance-substance relation, process-process relation), etc. They are exemplified in the following section.

Concept: "first cause of death", "contraindication", "inversion", … State: "hypothermia", "weakness", "refraction change", … Domain: "individual patient", "patient with disease history of more than 20 years", "patient of attack age more than 20 years", … Process: "gluconeogenesis", "fat decomposition", "ketone bodies generation", "protein synthesis", "absorption of Vit B12", "intravenous injections, …Parameter: "secretion index of insulin", "lipid clearance rate", "diameter of blood vessel", "respiratory tract resistance",… Quantity or Degree: "large dose", "excessive dose", "number of B cell", "slight abnormality", … Place: "HLA-control gene-site", "in vivo", … Time: absolute time (A. D. time), relative time (from beginning of the events) [17], e.g. "after meals", "before meals", … Compounded sort: self-compounded, reciprocal compounded, compounded of compounded element sorts.

On the basis of their operation attributes, the elements might be devided into following types. (1) The elements with *KDV* meaning component: e.g. "analgesia", "antithermic function", "anti-inflammatory function", "antirheumatic function", … (2) The elements might be transformed into *KDV*, such as "indication", "therapy", "contraindication", "noneatable foods", "less eatable foods", … (3) Editing elements: those strings with high occurrence rate in medical knowledge. (4) General elements.

3. The Sorts of Medical Relations

Medical relations, or knowledge units, are constructed in IREMB by three sets, i.e. subject element set (SES), object element set (OES) and condition element set (CES), which are linked by relation words *KDV*s and/or condition word (*CWD*): "UNDER PANCONDITION". Thus they have the same form "SES-*KDV*-OES -(-*CWD*- CES)". Again based on their physical meaning, the medical relations could be roughly categorized and exemplified as followings.

3.1 Pan-excluding relations

The subject element set and object element set repel each other. e. g., "headache due to hypertension"-"abstain from" -*CONTN-* "ergotamine". "asthma"-"noneatable food"- *CONTN-* "hairtail", "yellow croaker", "razor clam", "shrimp", "crab", " leaf mustard " . "hypertension"-"less eatable food" - *CONTN-* "yolk", "cream", "pork liver", "pig brain". "rifampin"-"abstain from"- *CONTN-* "paminosalicylic acid (PAS)", "oral antihyperglycemic", "fregnant woman", "alcoholism", "hepatic insufficiency". The elements "abstain from", "noneatable food" have the same or approximate operational attribute.

3.2 Pan-accompanying relations

They are not definite causalities, but coexistence relations. e. g., "rheumatoid nodules" -*COMPN*- "rheumatoid factor titer increased". "V type hyperlipemia"-*COMPN*- "obesity", "hyperuricemia", "diabetes", "acute pancreatitis" "hepatomegaly", "splenomegaly", "chronic active hepatitis (CAH)"-"complication" -*CONTN*- "nephritis", "arthritis", "angitis", "dermatomyositis", "myocarditis", "sicca syndrome", "endocrine dysfunction", "ulcerative colitis".

3.3 Pan-construction relations

They are whole-part relations. e. g., "hemorrhoid"-"class" -*CONTN*- "internal hemorrhoid", "external hemorrhoid", "mixed hemorrhoik". "acute pancreatitis"- "stage" -*CONTN*- "edematous pancreatitis", "necrotic pancreatitis". "yolk", "cream", "brain", "liver" -*CONTN*- "cholesterol" –*CWD*- "content:"-"rich".

3.4 Pan-creation relations

They are causalities, the "causes" preexist and keep unchanged, whereas the "results" are new born. e. g., "chronic alcoholism" –*CREAT*- "hepatocirrhosis". "anger", "mental stress" –*CREAT*- "spasm of gastric muscle", "spasm of gastric microangium". "uterine prostaglandin" –*CREAT*- "dysmenorrhea". "liver" –*CREAT*- "coagulation factor II", "coagulation factor VII", "coagulation factor IX", "coagulation factor X". "Vit B12"-"deficiency" –*CREAT*- "spinal cord degeneration", "intelligence going down". "indolol"-"toxic and side effect" -*CONTN*- "headache", "tinnitus", "sleepless", "nausea", "vomiting", "diarrhea". These relations can be formatted in opposite direction as well. e. g., "Raynaud's disease"-"inducing factor" -*CONTN*- "emotion excitement", "coldness". "Vit K 2"-"source" -*CONTN*- "synthesis by intestinal bacterium". "fatty liver"-"cause" -*CONTN*- "alcoholism", "obesity", "diabetes".

3.5 Pan-equivalence relations

In these relations, the subject elements get or partly get the values, meanings or their equivalences. e. g., "secondary"-"dysmenorrhea"-"definition" -EQUAL- "dysmonorrhea due to other gynecopathy". "latent jaundice"-"icterus index" -EQUAL- "7-15(". "serum"-"total bilirubin"-"normal concentration" -EQUAL- "0.1-1.0mg/dl".

"I type hyperlipoproteinemia" –*EQUAL*- "hyperchylomicronemia". "Vit B12"-"physiologic meaning" -*CONTN*- "take part in formation of pyrimidine and purine", "coenzyme of DNA synthesis". "total cholesterol"-"decrease"-"clinical meaning" -*CONTN*- "acute hepatonecrosis", "hepatocirrhosis", "hypothyroidism", "severe malnutrition", "severe anemia". "Vit B12"-"absorption site" –*EQUAL*- "end of ileum". "Raynand's disease"-"predilection sex" –*EQUAL*- "female".

3.6 Pan-recognition relation

The relations represent the symptoms, features, appearances. etc. of the subject elements. e. g., "urinary tract infection"-"symptom" *-CONTN-* "fever", "chilly", "aching pain in lumbar region", "frequent micturtion", "urgency of micturition", "pain of micturtion", "pyuria", "bacteruria".

"bronchial asthma"-"forewarning" -CONTN- "nasal itch", "sneezing", "laryngeal itch", "chest

distress", "cough". "acute pancreatitis"-"pathologic change" -*CONTN*- "pancreatic acinus destruction", "pancreatic duct rupture". "chronic active hepatitis"-"laboratory examination" -*CONTN*- "V factor decreased", "VII factor decreased", "white bood cell decreased", "thrombocyte decreased", "IgG positive", "lupus cell positive",

3.7 Pan-transform relations

They are another kind of causality, the process is "subject elements preexist? Object elements new born and subject elements died". e. g., "fatty liver" **-TRANS-** "hepatocirrhosis" **-CWD-** "long term alcoholism". "folic acid" **-TRANS-** "tetrahydrofolic acid" **-CWD-** "explanation:"-"in vivo"-"folic acid reductase"-"dihydrofolate reductase". "tetrahydrofolic acid"-"methyl" **-TRANS-** "N5-methyl tetrahydrofolic acid".

3.8 Pan-inhibition relations

For example, "adrenosem salicylate" –*DECRS*- "capillary permeability". "metronidazole" –*DECRS*- "trichomonas vaginalis", "anaerobic bacteria", "bacillus coli", "bacillus proteus", "pneumobacillus". "Vit B6" –*DECRS*- "cholesterol". "cimetidine"-"indication" -*CONTN*-"duodenal ulcer", "acute hemorrhagic lesion of gastric mucosa", "hirsutism", "gastric ulcer". The followings are pan-inhibition relations in opposite direction.

"chronic bronchitis"-"suitable food" -*CONTN*- "bean products", "green vegetable", "carrot", "functional maldigestion"-"therapy" -*CONTN*- "metoclopramide", "domperidone (motilium)", "cisapride (prepulside)". "fatty liver"-"prevention" -*CONTN*- "high-protein diet", "limit diet total heat".

3.9 Pan-promotion relations

For example, "cell decomposition" –*INCRS*- "potassium in urine". "cold", "eat raw or cold food" –*INCRS*- "dysmenorrhea". "insulin" –*INCRS*- "glucose"-"getting into cell".

3.10 Pan-displaement relations

These relations represent the passageway by which the substances (or process) move from one site to another site. e. g., "liver"-"bile acid" –*PASTO-* "intestinal tract"-"bile acid". "glucose-6-phosphate" –*PASTO-* "phosphopentose pathway", "glucuronic acid metabolism pathway", "glycogen synthesis".

3.11 Pan-order relations

These relations mean time-order, quantitative order, degree order, comparison, unequivalence, etc. –*PORDR*- represents "the subject elements are "stronger than" (or "more than", "ealier than", …) object elements". e. g., "after meals"-"LPL activity" –*PORDR*- "fasting"-"LPL-activity". "left atrial"-"depolarization"-"time-order" –*EQUAL*- "anterior wall of right atrium", "interatrial fiber

linking anterior walls of left and right atriums", "anterior wall of left atrium", "left aurcle", "posterior wall of left atrium".

3.12 Pan-mixed relations

The physical meanings of these relations are dependent on the operational components of object elements. e. g., "salbutamol"-"function" - *CONTN*- "dilate bronchus". "ibuprofen (brufen)"-"function" - *CONTN*- "anti-inflammation", "analgisia", "antitherm", "antirheumatism".

3.13 Pan-undefinite-or-negation relations

These relations are represented by undefinite elements such as "cause unknown" or negative elements, e. g. "no symptom". For example: "prostatic calculus"-"cause" - *CONTN*- "cause unknown". "prostatic calculus"-"symptom"- *CONTN*- "no symptom".

3.14 Other relations

The relations beyond the above relations.

Application and Future of HCSL

This is the broad outline of the primary HCSL of IREMB. It is believed that HCSL such constructed could cover the most substantial part of the medical knowledge system and might be operational by computer and, simultaneously, understandable by human. Based on HCSL, powerful medical intelligent functions could be explored and developed. Any relations (or knowledge units) in any medical fields could be reachable, connectable, 'some of them) deducible and, to some extent, integratable. Some of operation functions have been developed in personal computer [12-16]. The seaching functions have been running on the web sites of Internet: 202.121.32.10/bhf/tcmbke.htm (English) and 202.121.32.10/bhf/tcmbkc.htm (Chinese) (mainly for Traditional Chinese Medicine, TCM), and www.sh.col.com.cn/bhf/bhf/ebao.htm (English) and www.sh.col.com.cn/bhf/bhf/cbao.htm (Chinese) [16]

4. Acknowledgement

The author would like to express his heartful thanks to Miss Ying Jie Bao for her secretarial work.

5. References

- Wu X.M.: The Pansystems View of the World. Beijing: Publishing House of China People University, 1990, pp32-35
- [2] Wu X.M.: Pansystems Philosophy and Its Encyclopedia-Connecting Application.Advances in Systems Science and Applications (Inanguration Issue), 1995, pp198-203
- [3] B. de Farie Leao: Synopsis: Knowledge Processing, Year Book of Medical Informatics, 1986, pp453-455

- [4] Bao H.F.: The Structure Characteristics of the New Research QMSOC and Its Relevant Operators, J Tongji Med Univ, 1989, 9 (4), pp235-238
- [5] Bao H.F.: The New Functions of Quantitatively Medicine Simulating and Operating by Computer —New research QMSOC (III), 1990, J Tongji Med. Univ, 10 (1), pp52-56
- [6] Bao H.F., Geng J.H. and Su Z.F.: Pansystems Methodology (PM) and a New Research on Large-scale Integration of Biomedicine-An Introduction of QMSOC and Its Recent Progresses, Acta of Jiansu Industrial College, Journal of Jiangsu Institute of Technology, 1991, 4 (2), pp69-75
- [7] Bao H.F.: Information Medicine and Its Primary Function of Exploitation of Medical Resources, J Tongji Med Univ, 1991, 11 (1), pp59-64
- [8] Liu S.C., Geng J.H.and Bao H.F: The Differentiation Model of Human-Embryo of Computerized Multi-Field Integrated Instructing System-QMSOC and Computer-aided Medicine Education (I), Special Issue of Application Research of Computer (in Chinese), 1992, pp137-138
- [9] Bao H.F.: Integratable Relationized Medical Electronic Book (IRMEB)-QMSOC and Computer-aided Medicine Education (II), China High-Education Research (in Chinese), 1993, vol.4, pp651-654
- [10] Bao H.F., Wang Y., and Ao Z.J.: Auto-reasoning Function of Simple Transmissible Knowledge Chain of Integratable Relationized Medical Electronic Book (IRMEB)-QMSOC and Computer-aided Medicine Education(III), Medicine Education, 1998, No.8, pp17-23
- [11] Bao H.F., Wang Y., and Liu S.C.: The Background, Structure, Program-Development and Functions of Integratable Relationized Medical Electronic Book (IRMEB)-QMSOC and Computer-aided Medicine Education (IV), Special Issue of Application Research of Computer (in Chinese), 1995, No.2, pp34-40
- [12] Bao H.F., Yang L., and Dai Y.: The Present Progresses of the Integration Engineering of Medical Knowledge Developed under the Inflneuce of Pansystems Theory-QMSOC and Computer-aided Medicine Education (V), China High-Education Research (in Chinese), 1996, vol.9, No.3, pp56-69
- [13] Bao H.F., Wang Z.F.: The Exploration of the Method to Counter Madical Knowledge Explosion-QMSOC and Computer-aided Medicine Education (VI), China High-Education Research (in Chinese), 1997, vol.11, No.2, pp13-19
- [14] Bao H.F., Ni X.W., Lou S.: Integratable Relationized Medical Electronic Book (IRMEB)— An Exploration of an New Type of Intellegent Knowledge Medium under the Influence of Pansystems Theory. Advances in Systems Science and Applications (Inauguration Issue) 1995, pp304-309
- [15] Bao H.F: The Integrating Engineering of Medical Knowledge by Computer under the Influence of Pansystems Theory, General Systems Studies and Applications, 1997, pp358-363
- [16] Bao H.F., Wang Z. F., Ji H. L.: The Intelligent Medical Book IREMB on Internet Developed under the Influence of Pansystems Theory. The Third Workshop of the International Institute for General Systems Stidies (IIGSS): Systems Sciences and Its Application. pp478-481, Beidaihe, China, 1998
- [17] J. H. van Bemmel, M. A. Musen: Handbook of Medical Informatics, Bohn Stafleu Van Loghum, pp453-454, 1997