Subsumption Principles Underlying Medical Concept Systems and their Formal Reconstruction

Jochen Bernauer, Institute for Medical Informatics, University of Hildesheim, Germany

Conventional medical concept systems represent generic concept relations by hierarchical coding principles. Often, these coding principles constrain the concept system and reduce the potential for automatical derivation of subsumption. Formal reconstruction of medical concept systems is an approach that bases on the conceptual representation of meanings and that allows for the application of formal criteria for subsumption. Those criteria must reflect intuitive principles of subordination which are underlying conventional medical concept systems. Particularly these are: The subordinate concept results (1) from adding a specializing criterion to the superordinate concept, (2) from refining the primary category, or a criterion of the superordinate concept, by a concept that is less general, (3) from adding a partitive criterion to a criterion of the superordinate, (4) from refining a criterion by a concept that is less comprehensive, and finally (5) from coordinating the superordinate concept, or one of its criteria. This paper introduces a formalism called BERNWARD that aims at the formal reconstruction of medical concept systems according to these intuitive principles. The automatical derivation of hierarchical relations is primarily supported by explicit generic and explicit partitive hierarchies of concepts, secondly, by two formal criteria that base on the structure of concept descriptions and explicit hierarchical relations between their elements, namely: formal subsumption and part-sensitive subsumption. Formal subsumption takes only generic relations into account, part-sensitive subsumption additionally regards partive relations between criteria. This approach seems to be flexible enough to cope with unforeseeable effects of partitive criteria on subsumption.

INTRODUCTION

Conventional medical concept systems have the following characteristics: (1) They combine the ordering of concepts with a coding schema which represents concept relations by meaningful codes. (2) Their use presupposes a certain amount of intuitive knowledge [1]. These characteristics limit the services of computer-based concept representations for the following reasons: Often, the structure of the coding scheme is constraining the structure of the concept system, and terminological principles are sacrificed in favour of coding principles. Common phenomena in this respect are: The potential number of hierarchical levels is constrained by hierarchical coding principles, i.e. hierarchical, group sequential or combinatorial codes, with fixed length. Typical compensations are the parallel representation of independent criteria on the same level of subordination, or the introduction of sibling concepts although child concepts were adequate. The potential number of siblings on a particular hierarchical level is restricted by the length of the coding alphabet. This sometimes leads to "artificially" balanced concept formation and ordering. Finally, common hierarchical coding principles support only strict hierarchies. This may be circumvented by assigning several codes to a concept as realized by MeSH, but gives rise to the problem of redundancy and consistency of views [2].

Conventional medical concept systems comprise the following implicit knowledge types. Commonly, they make no difference between generic and partitive relations, sometimes they mix both in a transitive manner. In general, the criteria for subordination are hidden, often they are partitive. This situation is acceptable if the coding schemes are for intuitive use, but is an obstacle for computer-based services. A global effect is the limited potential for automatical subsumption.

Formal subsumption has been investigated in the context of term subsumption languages in the tradition of KL-ONE [3,4]. These languages focus on the formal definition and automatical classification of concepts. In general, they distinguish between definitorial and assertional knowledge and base subsumption strictly on definitorial knowledge.

In contrast to term subsumption languages, the concept representation formalism developed in the GALEN project emphasizes the formation of sensible concepts, instead of classification of already defined ones. The classificator makes no difference between definitorial and assertional knowledge and takes also partitive criteria into account [5,6].

This paper introduces a constrained representation formalism for medical concepts based on conceptual graphs [7] which focuses on the principles of subordination underlying conventional medical concept systems and which aims at their formal reconstruction. Formal criteria for subsumption are given that reflect the intuitive principles of subordination.

PRINCIPLES OF SUBSUMPTION

Subordination in conventional medical concept systems is based on several intuitive principles. They are illustrated by the generic concept ladder of Figure 1, which is in the style of a conventional concept system. In principle, a concept description is assumed to be composed by a base concept and a set of criteria. In the Aristotelean sense of a definition the base concept refers to the primary category which the concept belongs to, and the set of criteria refer to the differentiae, which distinguish the concept from the primary category. (cf. [8])

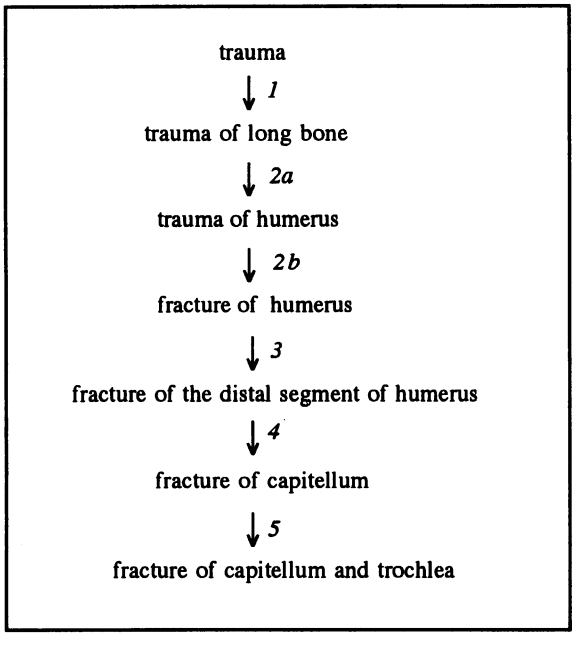

Figure 1: A generic concept ladder presenting different intuitive principles of subordination

Informally, a composite concept description can be superordinate to another concept for any of the following reasons:

Introduction of a specializing criterion
The subordinate description includes a specializing criterion of the base concept (1), or of a criterion (numbering refers to Figure 1), which is not present in the superordinate description.

Generic refinement of a concept element
A criterion (2a), or the base concept (2b) in the superordinate description is more general than one in the subordinate description.

Introduction of a partitive criterion
The subordinate description includes a partitive criterion of a criterion, which is not present in the superordinate description (3).

Partitive refinement of a criterion
A criterion in the superordinate description is more comprehensive than one in the subordinate description (4).

Conjunctive coordination
The primary category or a criterion of the subordinate description is a conjunctive coordination of one in the superordinate description (5). The inverse situation holds for disjunctive coordination (not depicted in Figure 1).

Of these principles the introduction of a specializing criterion and the generic refinement of a concept element are typically realized by the classifier of the languages of the KL-ONE-family [4].

FORMAL RECONSTRUCTION

A model for the formal reconstruction of medical concept systems must be capable for representing the intuitive principles of subordination and must support the automatical classification of concept descriptions by formal subsumption criteria.

Structure of conceptual descriptions

Concept descriptions in BERNWARD (Building Essential concept Representations in Well-Arranged Restricted Domains) are of the following types: primitive, composite, coordinated, negated and defined. A primitive concept description consists of a name of a primitive concept type, and has no formal substructure. A composite concept description is constructed from other concept descriptions by means of concept-forming operators. The structure reflects the definition of the concept and serves for the formal classification of conceptual descriptions. Formally, a composite concept consists of a base concept and a set of criteria, where a criterion consists of a role symbol and a concept description. The base concept reflects the primary category of the concept, the set of criteria reflects its essential characteristics. A composite concept may have one particular criterion, called "partitive criterion", which consists of the symbol designating the part-of-role and a concept description, which is called the "whole" of the composite concept. A coordinated concept description is a set of concept description combined by a logical conjunction (AND, OR), a negated concept description is combined by a negator (NOT). A defined concept description is of composite, coordinated or negated type, that is given a name. It can be used like a primitive type. Concept descriptions can be represented in the conceptual graphs notation [7].

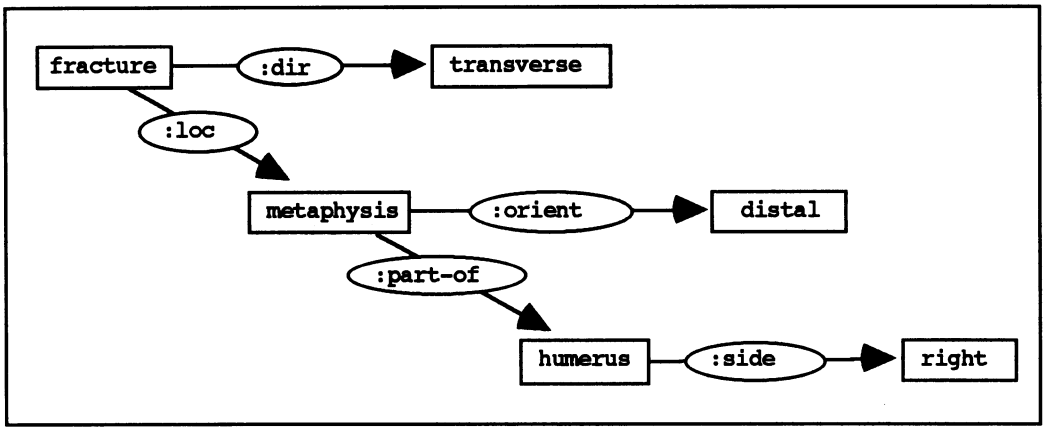

Figure 2: Representation of the concept "transverse fracture of the distal metaphysis of the right humerus"

Figure 2 shows the graphical representation of a fracture type as a composite concept description. Its base concept is "fracture", its criteria are "has direction transverse" and "has location distal metaphysis of right humerus". The concept which is in the latter criterion has the partitive criterion "is part of right humerus". Its whole is "right humerus".

Explicit hierarchical relations

Concept descriptions can be explicitely organized in a generic and a partitive hierarchy. The generic hierarchy (taxonomy) reflects the a priori subordination of concepts because of reasons that are not further relevant. Its ordering relation is designated by <. For example `fracture < trauma` means: a fracture is a trauma. In parallel, the partitive hierarchy (partonomy) reflects the a priori part-whole relation between concepts, which is designated by $<_p$. For example `capitellum` $<_p$ `dist_segment_humerus` means that the capitellum is a part of the distal humerus segment. The relation $<_p$ is supposed to be transitive. This approach is partly in contrast to the results of [9], where six different types of part-whole-relations are identified with potential intransitivity in case of mixing particular types. It is justified by the observation, that firstly the predominant medically related types of part-whole relations are "component/integral whole" and "place/region" ("segment/organ"). Secondly, the transitive mixing of these types is almost intuitively acceptable [10].

Constrained composition of concepts

The formal reconstruction of medical concept systems must realize the intuitive principles of subordination described above. Therefore, BERNWARD provides different types of restrictions that control the formation of subordinate concept descriptions. The function of these restriction types shall be illustrated on reconstructing sections of the AO/ASIF classification of fractures of long bones [11].

Local role restrictions restrict the introduction of sensible and relevant criteria. These are either specifying or partitive. For instance, the following local role restriction

```
lrestr(:dir;fracture_extraarticular_distal_
metaphysis_humerus_simple;fracture):=
        {oblique-inwards,oblique-outwards,
        transverse}
```

restricts the direction of a extra-articular simple fracture of the distal humerus to the concepts "oblique-inwards", "oblique-outwards", and "transverse".

The local role restriction

```
lrestr(:fragm;fracture_extraarticular_
distal_metaphysis_humerus_wedge;wedge):=
        {fragmented,non-fragmented}
```

states that the concept "wedge" in the context of the concept "extra-articular wedge fracture of the distal humerus metaphysis" can be modified by "fragmented" or "non-fragmented".

The local role restriction:

```
lrestr(:part;fracture_long_bone;long_bone):=
        {proximal_segment,diaphysis,
        distal_segment}
```

restricts the introduction of a partitive criterion for "long bone" in the context of "long bone fracture" to the concepts "proximal segment", "diaphysis" and "distal segment".

The generic refinement of concept elements can be restricted by subconcept restrictions. A subconcept restriction defines the set of concept candidates which might used for generic refinement of a concept description. For instance, the subconcept restriction:

```
subrestr(long_bone_fracture;long_bone):=
{humerus,radius,ulna,femur,tibia,fibula}
```

restricts the refinement of the concept "long bone" in the context of "long bone fracture" to "humerus", "radius", "ulna", "femur", "tibia", and "fibula".

In analogy to subconcept restriction, part restrictions define the set of concept candidates, which might be used for partitive refinement. The following part restriction:

```
partrestr(fracture_partial-articular-
distal_humerus_medial-sagittal;
distal_humerus):=
        {medial-trochlea,
        trochlea-grove}
```

means, that for the concept "partial articular fracture of the distal humerus with medial-sagittal direction" the concept "distal humerus" can be refined by the partitive concepts "medial-trochlea", or "trochlea-grove".

Coordination restrictions define which concepts are allowed for coordination with an element of a concept description. For instance, the coordination restriction:

```
coordrestr(fracture_partial-articular_
frontal_capitellum;capitellum):=
        {trochlea}
```

allows the conjunctive coordination of "capitellum" and "trochlea" in the composite description "partial-articular frontal fracture of the capitellum".

Formal subsumption

For conventional medical concept systems automatical classification is restricted to those relations that are represented by hierarchical coding principles. Common obstacles are the inconsistent use of different coding principles, i.e. the mixing of hierarchical, group sequential or combinatorial codes within particular hierarchical ladders, the mixing of generic and partitive relations, or the disregard of hierarchical relationships. The main goal of formal reconstruction of medical concept systems is to allow for the automatical classification of composite concept descriptions. In contrast to conventional coding schemes, concepts are not represented by meaningful codes, but by formal definitions.

The principle of formal subsumption is to derive logical relationships between concept descriptions from their structure and from explicit hierarchical relationships between their elements. Criteria for formal subsumption have been developed for term subsumption languages [4], and for conceptual graphs (canonical formation rules) [7].

A peculiar problem especially for medical concepts is the effect of partitive relations between criteria on subsumption. There are different approaches for coping with this problem. Doyle and Patil [12] suggest to include axioms of the kind "a disease of a part of an organ is a disease of the organ" in the axiomatic component of a terminological representation system, and thus, to deal with the problem outside of formal subsumption. In GRAIL [6] subsumption over partitive criteria is integrated into formal subsumption. There are syntactical means for specifying for a particular role (e.g. "has location") to be refineable along a transitive role (e.g. "is part of"). However, this approach seems to be too general in certain situations, like for instance the following: A scoliosis of the thoracic spine is not a scoliosis of the spine, a revision of the colon is not a revision of the gastro-intestinal system, an avulsion of the humerus apophysis is not an avulsion of the humerus, etc. In these cases pathological conditions regarding body structures do not subsume pathological conditions regarding parts.

The unforeseeable implications of partitive relations between criteria justify the introduction of two separate criteria for formal subsumption and formal part-whole-relation. This allows for the distinction between a formal criterion for subsumption, which disregards partitive relationships between criteria, called "formal subsumption" and a different, which takes also partitive criteria into account, called "part-sensitive subsumption".

A conceptual description c_1 is formally subsumed by a conceptual description c_2, iff the base concept of c_1 is explicitely subsumed by the base concept of c_2 and the criteria set of c_2 is a subset of the criteria set of c_1, and all the concepts of the criteria set of c_2 formally subsume the concepts of the criteria set of c_1. (A formal definition is given in [10].)

A conceptual description c1 represents a formal partitive concept of one represented by the conceptual description c2, iff either c2 is the whole of c1, or the base of c1 is an explicit partitive concept of the base of c2 and the criteria sets of both are equal. (A formal definition is given in [10].)

Part-sensitive subsumption is equal to formal subsumption, but additionally, takes formal part-whole-relations between criteria into account [10].

Figure 3 illustrates the difference between formal subsumption and part-sensitive subsumption.

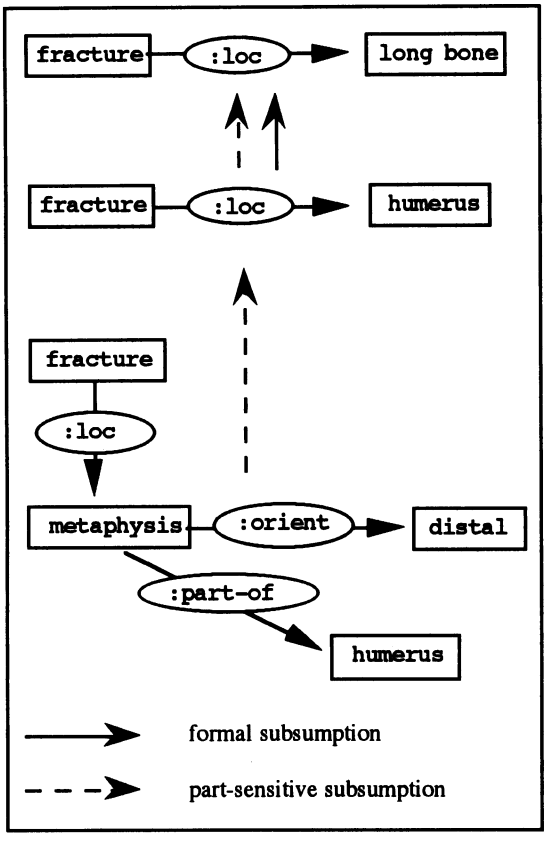

Figure 3: Formal subsumption and part-sensitive subsumption between conceptual descriptions.

SUMMARY

There are several intuitive principles of subordination that are underlying conventional medical concept systems. Some of them involve part-whole relations. A model for the formal reconstruction of medical concept systems has been outlined which considers these principles by mechanisms for constraining subconcept formation. For automatical classification two formal criteria for subsumption are introduced: formal subsumption and part-sensitive subsumption. The latter takes partitive relations between concept forming criteria into account.

REFERENCES

[1] Rossi Mori A., et al. (1993). Model for Representation of Terminologies and Coding Systems in Medicine, in: de Moor G.J.E., McDonald C.J., Noothoven van Goor J., (eds.), Progress in Standardization in Health Care Informatics, IOS Press, 1993; 92-104

[2] Cimino J.J., Hripcsak G., Johnson S.B., Clayton P.D. (1989). Designing an introspective, multipurpose, controlled medical vocabulary. In: Kingsland L.C. (ed.), Proc. 13th Symposium on Computer Applications in Medical Care (SCAMC), McGraw-Hill pp 513-18

[3] Woods W.A., Schmolze J.G. (1992). The KL-ONE family. Computers and Mathematics with Applications, 23; pp 133-177

[4] Woods W.A. (1991). Understanding Subsumption and Taxonomy: A Framework for Progress. In Sowa J.F., (ed.), Principles of Semantic Networks. Morgan Kaufmann Publishers, San Mateo CA, 1991, pp 45-94

[5] Rector A.L., Nowlan W.A., Glowinski A. (1993). Goals for concept representation in the GALEN project. In: Safran C. (ed.), Proc. 17th Symposium on Computer Applications in Medical Care (SCAMC), McGraw-Hill pp 414-418

[6] GALEN: The GRAIL Kernel, Version 1. 1993 GALEN Deliverable 6. The GALEN Consortium.

[7] Sowa J.F. (1984). Conceptual Structures: Information Processing in Mind and Maschine, Addison-Wesley

[8] Campbell K.E., Das A.K., Musen M.A. (1994). A Logical Foundation for Representation of Clinical Data, J Am Med Informatics Assoc. 1:218-232

[9] Winston M.E., Chaffin R., Herrmann D. (1987). A Taxonomy of Part-Whole Relations. Cognitive Science 11: 417-444

[10] Bernauer J., Modelling Formal Subsumption and Part-Whole Relation for Medical Concept Descriptions, European Conference on Artificial Intelligence 94 (ECAI'94), Workshop Parts and Wholes: Conceptual Part-Whole relations and Formal Mereology

[11] Müller M.E., Nazarian S., Koch P., Schatzker J. (1990). The Comprehensive Classification of Fractures of Long Bones. Springer-Verlag

[12] Doyle J., Patil R.S. (1991). Two theses of knowledge representation: language restrictions, taxonomic classification, and the utility of representation services. Art. Intelligence 48; 261-297