Background knowledge in diagnosis

E.T. Keravnou^a, F. Dams^b, J. Washbrook^b, R.M. Dawood^c, C.M. Hall^c and D. Shaw^c

^a Department of Computer Science, University of Cyprus, 75 Kallipoleos str., Nicosia, T.T. 134, Cyprus

^b Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK

^b Department of Radiology, The Hospital for Sick Children, Great Ormond Street, London WCN 3JH, UK

Received October 1991 Revised March 1992

Abstract

Keravnou, E.T., F. Dams, J. Washbrook, R.M. Dawood, C.M. Hall and D. Shaw, Background knowledge in diagnosis, Artificial Intelligence in Medicine 4 (1992) 263-279.

A diagnostic framework in which there is a clear separation between the expertise per se and relevant background knowledge is discussed. We argue for the need to have an explicit representation of background knowledge. Background knowledge is domain foundational knowledge or common-sense knowledge; it is brought into play in the diagnostic context activated by the foreground knowledge, the diagnostic expertise. In the diagnostic framework discussed, background knowledge is of two types: foundational knowledge related to diagnostic findings and common-sense knowledge about time. The explicit representation and integration of expert and background knowledge is essential for achieving competent behaviour, both from the perspectives of the conversational context and the diagnostic performance of the system. The framework presented is being applied successfully to the domain of skeletal dysplasias.

Keywords. Background knowledge; diagnostic expert system; data-model; common-sense reasoning; SDD.

1. Introduction

The overriding motivation of any creator of a diagnostic expert system is to build a system that achieves accurate and timely diagnosis. For a medical diagnostic system to gain acceptance this must certainly be so.

To achieve this goal, one is faced with the challenge of formulating/eliciting the relevant model and of constructing a system whose architecture supports it adequately. The diagnostic model for the domain of skeletal dysplasias and syndromes amalgamates foreground domain knowledge (expertise per se) with relevant background knowledge. The architecture of the Skeletal Dysplasias Diagnostician (SDD) system supports the separation and co-operation between these two distinct bodies of knowledge [24]. The current clinical assessment of the system shows that this diagnostic framework is promising, with the SDD

Correspondence to: E.T. Keravnou, Department of Computer Science, University of Cyprus, 75 Kallipoleos str., Nicosia, T.T. 134, Cyprus.

system consistently performing significantly better than radiologists who are not expert in the domain [10]. Good performance is of course a major step towards achieving the ultimate objective of timely and accurate diagnosis. To maintain this achievement in the long term, the system knowledge will undoubtedly need to be continuously upgraded and extended. This is when the viability of the system architecture will be truly tested. We are confident that the SDD architecture will pass the test.

The need for a diagnostic framework to support the explicit integration of background domain knowledge is certainly not specific to the domain of skeletal dysplasias. It is relevant to any diagnostic domain where temporal information is pertinent or to any diagnostic domain where findings (manifestations) have internal structure of their own. In reality, most domains fall in this category. The following discussion revolves around the domain of skeletal dysplasias but care is taken to bring out the generalities.

2. Background knowledge

Many studies have been undertaken by Cognitive Scientists, and continue to be so, which aim to reveal the differences between the memory organisations and reasoning strategies of experts and novices in some area, as a means of formulating and testing theories regarding the role of experience in acquiring expertise (e.g. [12, 13, 15, 25]). Such theories, if valid, would be of paramount importance since they could point to changes in current training practices which would speed up the conversion process from novice to expert.

Such studies have revealed that novices possess a large body of factual (or text-book) knowledge but that they do not necessarily know how to apply this knowledge in an efficient and effective way. Experts on the other hand have a richly interconnected body of knowledge and they are capable of applying this knowledge effectively during problemsolving – they can immediately retrieve the pieces of knowledge relevant to a particular context. For example medical graduates have a detailed knowledge of human anatomy and physiology and detailed knowledge of the manifestations of rare and common ailments (the former usually being more interesting to remember) but they are not usually competent diagnosticians immediately on completion of their basic training. It is through experience that rich interconnections are incorporated in their originally 'flat' body of knowledge and problem-solving skills are developed. For example, it is through experience that disease classifications and differentiations are formulated and detailed causal chains are collapsed into associations in context.¹ A novice may model a disease in terms of a detailed list of manifestations of equal importance. Through experience this 'flat' model will be refined and modified, i.e. some manifestations will be demoted to coincidental abnormalities. Novices have detailed knowledge but this knowledge is not properly interconnected or abstracted for the purposes of applying a particular task such as diagnosis effectively, whilst experts have both the detailed knowledge and the higher level structures which facilitate the usage of this knowledge for specific tasks. Using Chandrasekaran's terminology, experts have compiled portions of their detailed knowledge for effective use in the context of a particular task [4]. The compiled knowledge, the reasoning skills, and the underlying detailed knowledge (from which the compiled knowledge was generated) constitute the expertise per se, i.e. the competence of the expert. The actual interaction between the compiled and underlying 'deep' knowledge, during problem-solving, could vary from problem-case to problem-case. Since

¹Usually, eliciting an association, e.g. between cause and effect, from an expert is relatively easy – but maybe all that can be obtained! What may prove difficult is the elicitation of the context in which the association is valid; even the expert who has learned to think in a 'knowledge-based' way may have great trouble in formulating and articulating the context.

the compilation is performed incrementally, through the accumulation of problem-solving experiences, in the early stages the deep knowledge will play an active role but gradually its role will become one of providing secondary support, i.e. of providing detailed explanations/justifications of solutions without necessarily having participated in any way in the construction of these solutions.

The knowledge organisation (structure of compiled knowledge) and reasoning strategies acquired through experience are rarely documented, they simply reside in the heads of the practising experts. However, not all of the original detailed knowledge will be compiled, in the sense of incorporating a higher level structure to it through new entities and relations - only the knowledge which is directly applicable in the context of the relevant diagnostic practices is compiled, for example only the knowledge of the given set of diseases is compiled. Domain foundational knowledge, for example knowledge of the human anatomy, normality ranges for different parameters, medical concepts, etc., essentially remain invariant through experience. Experience will certainly assist in consolidating concepts which may seem vague or abstract, or even extend this knowledge, for example experience will teach someone how to abstract useful, qualitative, diagnostic information from clusters of quantitative data, but essentially this knowledge will be the same detailed text-book knowledge taught as part of one's basic training. This detailed foundational knowledge, which has not been compiled because it is not a direct component of the expertise but which provides indispensable support in applying the acquired expertise, is a core type of background knowledge but still specialist knowledge. Another type of background knowledge is common-sense knowledge. Everyone possesses world knowledge. However, the objective is not to capture all world knowledge but rather to capture those aspects of world knowledge which are directly relevant to the task at hand, i.e. diagnostic reasoning. In this respect the relevant aspects are temporal, spatial, taxonomic and meronomic common-sense knowledge. Humans are capable of reasoning with incomplete or uncertain temporal information and of reasoning with different temporal granularities. Time is intrinsically related to diagnostic reasoning. Similarly anyone understands and can reason with spatial, taxonomic and meronomic relations. The aim is therefore to capture common-sense about time, space and taxonomic relations which would support a deep understanding of diagnostic knowledge and findings. This objective is considerably restrained in relation to the objective of the CYC project team [26]. Their objective is to capture all world-knowledge necessary to enable a computer program to understand, say, an entry in the Encyclopaedia Britannica; this objective essentially boils down to fully fledged natural language understanding. In addition it is possible that none of the knowledge incorporated in the CYC knowledge-base will be deemed as specialist knowledge in the sense of detailed knowledge possessed by a graduate of that discipline but rather it will present a knowledgeable layman's understanding of the incorporated concepts. Thus the background knowledge advocated in this paper consists of specialist foundational knowledge and common-sense knowledge directly relevant to the application of diagnostic tasks. Although the required common-sense is limited in comparison with the CYC objectives, the specialist type of the background diagnostic knowledge is at a level of sophistication beyond the objectives of the CYC project; the scope of the background diagnostic knowledge is narrower but much deeper than the scope of the CYC knowledge-base which is very broad (and much more ambitious) but relatively shallower. The two approaches share the belief that common-sense (of some sort) is necessary for a knowledge-based system to achieve intelligent behaviour.

The foreground knowledge therefore consists of the diagnostic expertise and the background knowledge consists of the domain foundational knowledge and relevant common-sense which collectively support deep understanding and hence intelligent handling of diagnostic information. The foreground and background knowledge are distinctly different bodies of knowledge serving different roles during a diagnostic consultation – they address different aspects of the problem. Foreground and background knowledge co-exist in a collaborative fashion; this division of labour is in accordance with the generalised model of diagnostic behaviour discussed in [19].

A diagnostic system with an explicit division between the foreground and background knowledge does not necessarily incorporate different reasoning (i.e. diagnostic) modalities. It is now recognised that to achieve problem-solving flexibility the integration of different models, e.g. a heuristic model and a causal model, is the way forward [20] and many integrated diagnostic systems or frameworks have been proposed (e.g. [14, 16, 31, 34, 36]). For example, if the detailed diagnostic knowledge is applied in conjunction with the compiled knowledge during problem-solving, or if more than one compilation of the same detailed knowledge is possible and each compiled model may be used during problem-solving, then the system exhibits multiple modalities. Although these different modalities can be applied in a collaborative fashion, they are essentially opposing modalities since they present different means for dealing with the same aspect of the problem. It is important to appreciate that the reasoning modalities associated with the foreground knowledge and the background knowledge are not opposing modalities; they are strictly collaborative modalities dealing with different aspects of the problem.

Background knowledge is therefore *directly relevant* to the specific diagnostic task. The elicitation of background knowledge is not onerous or an additional overhead. Once the model(s) for the background domain knowledge have been formulated, the acquisition of the relevant background knowledge falls naturally within the acquisition of the foreground knowledge. After the initial stages whereby the foreground knowledge drives the acquisition of the background knowledge, the acquisition of the two bodies of knowledge is done in parallel. Both bodies of knowledge have their corresponding models which may be modified independently.

Skeletal dysplasias and malformation syndromes are generalised abnormalities which may affect the growth of bone and cartilage. These groups are of especial interest to geneticists and radiologists. The former are interested in the hereditary potentials and the latter in the bone features. Diagnostic skills in this field are scarce, and although individually these syndromes occur infrequently, collectively they constitute a frequent problem, e.g. about 1% of the population in the UK is affected. Accurate and timely diagnosis is vital to allow proper treatment and genetic counselling. An expert system that achieves this would have significant social and economic benefits. The SDD system is constructed as a collaboration between computer scientists and radiologists. Two of the radiologists are experts in the field of skeletal dysplasias. Since the system captures diagnostic skills of radiologists (rather than geneticists) the relevant background knowledge is essentially foundational knowledge about the human skeleton and its temporal development. This is a constrained, but a useful view of the background knowledge which would be conceivably relevant to any diagnostic domain dealing with bony abnormalities. Unlike foreground knowledge, background knowledge has wider applicability. At this point one may argue that if the unconstrained view of background knowledge had been adopted then there would be a corpus of knowledge with even wider range of application. This might be so assuming that the elicitation task is manageable. In our diagnostic context it would mean eliciting all foundational radiological (and possibly genetic) knowledge. Even if this were achievable a high proportion of the knowledge would be utterly irrelevant to any of the specific relevant tasks. Acquiring knowledge for the sake of it and without any specific bias in mind is not the best way to spend resources. The acquisition of background knowledge should be focused, and this focus must be provided by a model which constitutes an integral part of the overall competence model [18] for the particular task. Constrained bodies of background knowledge can, if necessary, be amalgamated to obtain background knowledge with wider scopes.

In our specific domain and task, the overall competence model was elicited from the two consultant radiologists who are expert in diagnosing skeletal dysplasias. This model includes a model for the background knowledge. It is interesting to note that the initial body of background knowledge acquired from the domain experts is now extended and refined by a general radiologist, always in accordance with the elicited model. Thus even more parallelism, between the acquisition of the foreground and background knowledge, can be achieved.

3. Illustrating background knowledge

In this section concrete examples of background knowledge from the domain of skeletal dysplasias are given.

Suppose that in a diagnostic consultation it was given that the carpal-centres of the patient showed delayed-ossification, and suppose that the question 'Are the femora abnormal?' is raised. Any physician can immediately answer this with 'yes'. For a computer system to be able to derive that 'femora abnormal' holds if 'carpal-centres delayed-ossification' holds, the following relationships and common-sense reasoning strategies must be made explicit in the system:

- 1. carpal-centres delayed-ossification \Rightarrow epiphyses small
- 2. femoral-epiphyses are a kind of epiphyses
- 3. femoral-epiphyses are part of femora
- 4. If a component exhibits a property which denotes divergence from normality then the component is abnormal
- 5. If a class of concepts exhibits a property then all types of that class exhibit the same property
- 6. If a part of a component is abnormal then the whole component is abnormal.

Derivation

carpal-centres delayed-ossification \Rightarrow epiphyses small carpal-centres delayed-ossification holds

- epiphyses small holds
- (using strategy 4) epiphyses abnormal holds

femoral-epiphyses are a kind of epiphyses

(using strategy 5) femoral-epiphyses abnormal holds

femoral-epiphyses are part of femora

(using strategy 6) femora abnormal holds.

Suppose that in the same consultation the system knows that the patient has generalised platyspondyly, and that it wants to find out whether the patient's dorsal-vertebral-bodies are flat and whether the patient's trunk is normal. A radiologist would immediately answer

these questions with 'yes' and 'no' respectively. Once again the system requires the explicit representation of the following foundational knowledge to enable it to reach the right conclusions.

Platyspondyly is a condition of the vertebrae whereby they are flattened. The vertebrae are components of the spine, 29 in all. The regions of the spine are thoracic (or dorsal) spine, cervical spine and lumbar spine. Generalised platy-spondyly means that all vertebrae are affected from this condition. When there is generalized platyspondyly the trunk is short.

The taxonomy of *Fig.* 1 depicts some of the relationships mentioned above. There are a number of routes for deriving that 'flat dorsal-vertebral-bodies' holds given that 'generalised platyspondyly' holds. Two are shown below:

- 1. Generalised platyspondyly means that all vertebrae are flat. Since dorsal-vertebralbodies (or dorsal-vertebrae) are a kind of vertebrae then they are flat.
- 2. Generalised platyspondyly means that the entire spine exhibits the condition platyspondyly, i.e. the *locality* of condition platyspondyly is the whole spine. Thoracic-spine is a part of the spine and hence 'platyspondyly thoracic-spine' holds. The components of thoracic-spine are dorsal-vertebral-bodies and hence the conclusion 'flat dorsal-vertebralbodies' may be reached.

Often there exists a number of derivation paths for some query, 'holds(F)?', where F is a finding, and shorter paths should be attempted first.

To refute the query 'holds(trunk normal)?' the system uses the implication

generalised platyspondyly \Rightarrow trunk short (or short trunk dwarfism).

Suppose further that the system is told that the patient's joints are prominent and that the question 'are the metaphyses flared?' is subsequently raised by the system. Again a radiologist would immediately reply 'yes'. The two expressions, 'flared metaphyses' and 'prominent joints', may be treated as synonymous in the sense that they describe the same abnormality at different levels. 'Prominent joints' is a clinical observation which is explained by the radiological feature 'flared metaphyses'. There is a number of such associations between clinical findings and their radiological counterparts or simply between radiological features, or between clinical features, e.g. sloping acetabula and hypoplastic iliac-bases or hitch hikers thumbs and thumbs which are proximally set and hypermobile.

These examples use knowledge which is general radiological/clinical knowledge rather than knowledge specific to the domain of skeletal dysplasias. Strategies for manipulating this knowledge are generic common-sense strategies such as 'if a component is absent then all its parts are absent'.

Temporal reasoning is relevant to skeletal dysplasias, as it is to most diagnostic tasks. Skeletal dysplasias involve bone (and cartilage) malformations. Such malformations may be concerned with the ossification process of parts of the skeleton, and more precisely with divergencies from the normal ossification processes, e.g. whether a particular ossification process is delayed or advanced. Again this requires general knowledge about the normal progression of an ossification process (when does it normally start and when does it normally finish?). This knowledge enables one to determine the age at which a particular bone is normally expected to have ossified and hence to be visible in an X-ray image. Apart from normal ossification processes, skeletal abnormalities may have a generic temporal progression, e.g. a given abnormality may only be visible from a given age and similarly specific levels of severity may only be reached after a certain age. In addition to general (i.e. dysplasia-independent) temporal knowledge, the descriptions of features for some dysplasia involve time; it is not uncommon for features to appear and then disappear. For example, in the case of Dyggve Melchior Clausen (DMC) dysplasia a significant feature of the dysplasia is that the iliac-crests are described as lace-like from about the age of 3 years until about the age of 15 years. Thus if the system explores the possibility of DMC for a 2 year old it should not expect to establish that the iliac-crests are lace-like. Similarly the system exploring the same possibility for an 8 year old should expect the iliac-crests to be lace-like and should in fact exclude the possibility if this is not so. For a 20 year old the iliac-crests should be expected to be currently not-lace-like but to have been so before.

The temporal screening of a dysplasia to tailor it to the specific diagnostic context is absolutely essential for the performance of the diagnostic system. This is based both on background domain knowledge and temporal information specific to the particular dysplasia [22].

4. Background knowledge models

In this section the models for the background knowledge in the SDD system are discussed. These are:

- Feature subject model
- Description model
- Temporal model.

The above models are collectively referred to as the *Data-Model* of the system, where data are diagnostic findings. The discussion on the models which follows is at a conceptual level. A formal axiomatic discussion is given in [23].

4.1. Feature subject model

Feature subjects are of two types: *component* subjects, i.e. anatomical components (spine, metaphyses, hands, face) or body functions (renal function, cardiac function, pulmonary function), and *condition* subjects describing abnormalities of component subjects (platy-spondyly, polydactyly, micrognathia).

A finding asserts that various descriptions hold for a particular subject at a particular locality of the subject and for a particular temporal period. Example findings would be 'platyspondyly mild generalised from-birth', 'polydactyly absent hands', 'femoral-metaphyses irregular medially', 'iliac-crests lace-like from about 3 to 15 years', 'epiphyses delayed-ossification'. The descriptions in the above findings are the words mild, absent, irregular, lace-like and delayed-ossification. The localities are the words generalised, hands and medially. The localities of a condition subject like platyspondyly and polydactyly are the component subjects which constitute their scopes, e.g. spine (or vertebral-bodies) for platy-spondyly, and hands and feet for polydactyly. The localities of a component subject are its parts, e.g. lumbar-spine, thoracic-spine and cervical-spine for component subject spine, and its subtypes, e.g. dorsal-vertebrae, cervical-vertebrae and lumbar-vertebrae for subject vertebrae, or ulnar-metaphyses and radial-metaphyses for subjects in their own right, it is more intuitive to use the specific component subjects as the subject of the finding, rather than as its locality, e.g. it is better to say 'ulnar-metaphyses irregular' rather than 'metaphyses

irregular ulnar'. Strategies operating on the taxonomic relationships can relate more specific findings to their generalisations and vice versa.

In addition, there are generic localities such as generalised, throughout, widespread, localised, medially, laterally, posteriorly, anteriorly, left, right, etc. These are the localities more commonly used. In the case of component subjects, a generic locality specifies the generic region of the given component. In the case of condition subjects the generic localities often used are generalised and localised, respectively indicating that the specific abnormality holds for all or some (exactly which not being specified) of the component subjects in the abnormality's scope.

Findings may have explicit temporal aspects, e.g. 'from birth' and 'from about 3 to about 15 years', in the above findings.

The Feature subject model consists of the relationships given in Table 1.

Table 1. Feature subject model.

• Subject type	• Linguistic relations
– negative-value – status-value	– synonyms – synonymous-subjects
Taxonomic relations	• User information directives
 isa (and its inverse types) 	 clarifying questions
- part-of (and its inverse consists-of)	 prompting questions images
• Temporal screening	 Inferential relations
- when-to-ask	- implications
	- locality

4.1.1 Subject type

The subject type relations reveal whether a subject is a component or a condition subject. If the negative-value of a subject is *normal* then it is a component subject and if it is *absent* then it is a condition subject. Thus 'hands normal' and 'polydactyly absent' both describe normal situations or negative findings. If a negative finding holds then all the positive findings for the given subject do not hold. Status-value is only relevant for component subjects. The status-value indicates the absence of the given subject, e.g. 'hands absent' or 'skull-vault absent-ossification'. Such status findings describe abnormalities and hence are positive findings. They are, however, special positive findings since they exclude all other findings for that subject.

4.1.2 Taxonomic relations

Taxonomic relations *isa*, *part-of* and their respective inverses enable the application of various common-sense strategies, e.g. if a part of a whole is abnormal the whole is also abnormal, as illustrated above. Both isa and part-of are tangled taxonomies enabling a subject to be of more than one type and to be a part of more than one component (*Fig. 1*).

4.1.3 Temporal screening

The temporal screening relation when-to-ask is very important. This relates findings of the given subject with the earliest time that is intelligent to ask about them. If all the findings of a subject must be screened out prior to a given time, then the subject on its own is associated with the particular time as well. This is so for subject IQ. The relation when-to-ask makes explicit the temporal progressions of relevant processes since by registering the earliest times that is intelligible to ask whether a particular process is delayed or advanced the system can deduce the normal initiation and termination times of ossification processes. Similarly, significant temporal progressions of abnormalities may be recorded, e.g. progressive kyphoscoliosis is only detectable after the age of 1 year.

270



Fig. 1. Taxonomic relationships in SDD.

4.1.4 Linguistic relations

The linguistic relations associate groups of findings or subjects which can be used interchangeably because at any relevant context they may be treated as synonymous even if they are not synonymous in the strict sense of the word. For example, dumbbell-shaped longbones and flared metaphyses are not strictly synonymous but for diagnostic purposes they can be used interchangeably. Other examples are short stature and dwarfism. Examples of synonymous subjects are dorsal-spine and thoracic-spine, bone-age and skeletal-maturation. The relations *synonyms* and *synonymous-subjects* are included because the aim is to support a user interface which is as flexible as possible. These relations can be used by a front end system which translates user volunteered input to a standard internal vocabulary or they can be used by a reasoner that dynamically correlates expressions which mean the same thing; the latter approach is computationally more expensive than the former.

4.1.5 User information directives

Other important relations for the user interface are prompting and clarifying questions, and in the future X-ray images. Prompting questions are used in order to elicit more complete information in response to a user volunteered finding. For example, if the user volunteers that there is platyspondyly, the system immediately prompts the question 'Is the platyspondyly generalised?'. Similarly, in response to the finding 'polydactyly' the system will prompt the questions 'Polydactyly feet?' and 'Polydactyly hands?'. Prompting questions are useful for findings whose subjects are condition subjects since they enable the system to get complete information about the locality of the given abnormality. If the locality is not given the system cannot assume that it refers to all relevant localities, e.g. if the user volunteers platyspondyly the system should not assume that the platyspondyly affects the entire spine. This is not the only context in which prompting questions are raised. Prompting questions are used when independent findings tend to occur together or in order to get more specific information about the reported abnormality, e.g. if the user says that the iliac-crests are irregular the system will prompt the question 'Are they lace-like?', that being a more specific (and diagnostically very significant) type of irregularity.

Clarifying questions are raised in response to a finding which is highly likely to be an inaccurate description of the actual situation. Such findings can easily be volunteered by users

who are not expert in the field (as is usually the case with expert system users). In the case of skeletal dysplasias a non expert may misinterpret an X-ray image² resulting in inaccurate or even erroneous information (see below). An example of a clarifying question is that if the user volunteers that the face is flat the system will ask 'Do you mean that the mid-face is hypoplastic?'. The difference between prompting and clarifying questions is that prompting questions aim to elicit information in addition to the volunteered information albeit more specific information whilst clarifying questions query the validity of volunteered information and aim to trap as quickly as possible common misconceptions of misinterpretations of an actual situation. Thus information elicited through a clarifying question will replace the information originally volunteered by the user. In the domain of skeletal dysplasias erroneous data are possible and even likely because the non-expert will not 'see' the subtle abnormalities which may constitute significant features. Clarifying questions are one way of detecting such errors; they are based on the experience that particular subtle abnormalities are commonly misinterpreted by novices in the field as such and such abnormalities. On-line display of X-ray images is a more effective way of combating this problem and at the same time provides a form of justification for the final diagnosis. X-ray images can be displayed in response to user information as a means of validating the user input and of course system questions can be illustrated through X-ray images. The database of X-ray images for the SDD system is currently under design.

4.1.6 Inferential relations

The system has a number of generic 'taxonomic' axioms which are instantiated on demand from taxonomic relations, to generate relevant implications (see [23]), e.g.

dorsal-spine abnormal or lumbar-spine abnormal or cervical-spine abnormal \Rightarrow spine abnormal

metaphyses irregular \Rightarrow femoral-metaphyses irregular

The axioms encode context-free (generic) implications. In addition to the generic implications there are specific implications representing dependencies between specific findings. A specific implication is associated with the subject of its consequent finding under the *implications* relationship. For example, subject kyphoscoliosis has the following implications:

scoliosis and kyphosis \Rightarrow kyphoscoliosis scoliosis absent or kyphosis absent \Rightarrow kyphoscoliosis absent

Kyphoscoliosis is an abnormality whereby both scoliosis and kyphosis co-occur. Presently implications are used in a backwards chaining fashion.

The *locality* relation makes explicit the relevant localities for condition subjects such as platyspondyly and polydactyly. The locality gives the set of highest-level relevant component subjects, e.g. spine for platyspondyly and {hands, feet} for polydactyly. This information can be used to generate dynamically even more implications, e.g.

hands normal and feet normal \Rightarrow polydactyly absent spine normal \Rightarrow platyspondyly absent

²The input to SDD is textual information, not images.

4.2. Description model

Descriptions are attribute-values for feature subjects and they include generic localities. Examples of descriptions are small, flared, hypoplastic, generalised, medially, etc. The Description model includes the following relations:

- Synonymous descriptions: descriptions which could be used interchangeably under specific contexts (i.e. feature subject), e.g. small and hypoplastic, or flared, broad and wide.
- Opposing descriptions: mutually exclusive sets of descriptions, e.g. small, medium and large, or present and absent.
- Generic localities: all the descriptions which constitute generic localities.
- Normal descriptions: descriptions which describe a normal situation under any context, e.g. regular, normal, normal-ossification, normal-size, well-modelled, etc. This relation is important in a diagnostic system because it makes explicit findings which describe normalities rather than abnormalities. Suppose that a finding for a given patient is that the stature is normal. This finding does not require any explanation and hence any diagnostic possibility which does not expect the stature to be abnormal accounts for it. However, the finding provides evidence against a diagnostic possibility which, say, expects the stature to be short.

4.3. Temporal model

As discussed above temporal reasoning is an integral aspect of skeletal dysplasias diagnosis. The temporal reasoning is common-sense reasoning. For example, the temporal screening of a dysplasia is common-sense reasoning – do not expect to observe currently a future manifestation. Modelling common-sense reasoning is not necessarily straightforward. The temporal expressions in SDD include uncertainty, e.g. something occurs *from about* a certain time or something terminates *at about* a given time. This uncertainty is modelled in terms of open time-intervals, a time-interval being the ontological primitive. A time-interval has an open base, or an open limit, or both. The openness means that the interval can get smaller in the relevant direction. Recently Console et al. [9] have reported on a causal temporal framework allowing 'variable' time-intervals.

The temporal model includes absolute temporal relations, relative temporal relations, temporal trends, and context-free time (which embraces the relations discussed above under when-to-ask). The reader is referred to [22] for an extensive discussion of the temporal model and its implementation.

The background knowledge models discussed above have gone through a number of refinements and future extensions are probable. Such changes are possible because the models are explicit in the system.

5. Explicit background knowledge: Advantages

Background knowledge is foundational domain knowlege and common-sense knowledge directly related to the domain and task. Its role in an expert diagnostic system (or any expert system) is to support the execution of the particular expert task. Its absence seriously undermines the competence of the expert system [20]. In this section we do not argue for the *presence* of background knowledge in a diagnostic system, as we believe that this does not need any justification, but for the *explicit* presence of background knowledge in a diagnostic system. Most diagnostic systems, even early ones like Mycin include background domain knowledge in their knowledge bases. However, this knowledge is implicit in the rules or other structures used to represent the system's expert knowledge. If the background knowledge is implicit then it is not possible to have an explicit reasoner incorporating all the generic (common-sense and other) strategies operating on background knowledge.

Explicit representation means that the background knowledge and its associated reasoning is an identifiable entity (component) in the expert system. This is the case with the SDD system. At the top level there are two co-operating tasks, a *hypotheses reasoner* and a *findings reasoner*. The co-operation is of an investigator-assistant mode where the hypotheses reasoner is the investigator and the findings reasoner is the assistant. The assistance provided by the findings reasoner is invaluable since without it the hypotheses reasoner would be grossly incompetent. The functions of these two reasoners are overviewed below.

The hypotheses reasoner performs differential diagnosis on the domain of skeletal dysplasias. It operates on the dysplasia model and incorporates strategies for triggering hypotheses, for focussing on the current differential, for evaluating the current differential, for acquiring additional information and for deciding when to terminate a consultation. The hypotheses reasoner and its associated dysplasia knowledge base constitute the expert skills and knowledge in the diagnostic system.

The findings reasoner, as its name implies, deals with patient findings. It performs the temporal screening of hypotheses profiles. It supports a flexible user interface, it acquires information (both user volunteered findings or findings requested by the hypotheses reasoner). It detects, and eliminates, redundancy in the patient findings so that each finding represents 'independent' evidence. It can detect conflicts in the user information. It ensures that none of the questions raised by the system is redundant, either because its answer can be deduced from already given information, or because the user has already said that this is unknown, or because it could not be possibly known in the particular temporal context.³ It structures a set of questions requested by the hypotheses reasoner into a comprehensive and meaningful sequence by grouping questions for the same subject together and preceding each group with relevant general questions. Each answer from the user is immediately processed and hence follow-up questions may be screened out. The findings reasoner therefore contributes significantly to the naturalness of the system's dialogue. The overall conversation sequence is the result of the activations of the hypotheses reasoner. During a consultation questions are asked for many different reasons, depending on which tasks are being executed; the findings reasoner ensures that individual question subsequences are comprehensible and that the user is appropriately prompted, either for additional, more specific, information or for validations, when information is being volunteered. Hence the findings reasoner's contributions to the dialogue are local but nonetheless critical. Collectively the hypotheses and findings reasoners achieve a natural dialogue structure exhibiting a focused and methodical approach to information acquisition. Adequate dialogue structure is important for interactive systems. Lastly the findings reasoner provides assistance to the hypotheses reasoner in the evaluation of hypotheses.

6. Background knowledge representation: Classification

Some diagnostic systems and frameworks, both early and recent, do not include background knowledge in any form. For example, Internist-I has no knowledge of the depen-

³Such questions should never arise if the temporal screening of the hypotheses profiles is done properly.

dencies between its manifestations with adverse effects on its scoring function, nor does it model temporal knowledge in any way, again with adverse effects on its scoring function as well as resulting in non-intelligent questions [28]. The parsimonious covering theory of diagnosis [30] does not accommodate any structural relationships between manifestations. This applies to other domain-independent diagnostic frameworks such as de Kleer and William's [11] and Reiter's [33].

Other systems have an implicit representation of background knowledge. For example, Mycin does not represent 'world-facts' explicitly but implicitly in terms of screening clauses [6, 21]. Some of the properties in Mycin's clinical-parameter tables though, e.g. the condition property, represent background knowledge. Amongst the first generation of diagnostic systems, Prospector, a contemporary of Internist-I and Mycin, is possibly the only system which has background domain knowledge in a relatively explicit form. The expertise in the system is in terms of a number of ore-deposit models, each model consisting of a number of inference rules. These models are supported by a classification hierarchy of rock types, minerals, physical forms and geological ages, which represents general domain knowledge. The creators of the system say that this is a critical structure for the performance of the system [32]. It gives flexibility in expressing data, enables the system to co-relate expressions and to detect inconsistencies in the user evidence. The Casnet diagnostic model [37] represents relationships between observations that are used to establish local control over the sequence of questions in a fashion consistent with medical practice; such relationships specify truth values of observations that can be directly deduced from already conducted observations. Eliminating redundancy in evidence, though, is not necessary because for each pathophysiological state, Casnet only uses the single piece of evidence with the highest associational value. Neomycin [8] and Centaur [1] which are respectively reconstructions of Mycin and Puff (where Puff has the same architecture as Mycin since it was constructed through Emycin), have a more explicit representation of background knowledge than their predecessor systems. Yet their representations do not present the background knowledge and its associated reasoning as a separate entity within a co-operative framework. In Centaur there is a distinct knowledge structure, the component frame, which is used to represent knowledge about pulmonary tests. The component frame gives a range of possible values for the test, an importance measure and a set of inference rules for deducing values from other test results. In Neomycin 'world facts' are represented explicitly in terms of screening rules which are applied by tasks such as Findout.

Explicit representation of background knowledge in the sense discussed in Section 5, i.e. as a separate, distinct, identifiable component in the system which co-exists with an expert reasoner within a co-operative framework, is exhibited in few cases. A primary example is the co-operative system Mdx/Patrec [5] where Mdx represents the expert reasoner and Patrec [29] is the auxiliary reasoner that knows all about the relevant medical concepts and has a rudimentary temporal model, which collectively enable it to manage case findings intelligently. The generalisations of the functions of Mdx and Patrec figure as separate generic tasks, CSRL [2] and IDABLE [17], in Chandrasekaran's Generic-Task Architecture [3]. This work has been a major source of inspiration for our work. Another influential framework is Clancey's heuristic classification method [7] embodied in the skeletal system Heracles, and more precisely the co-operative structure between the data taxonomy and the solution taxonomy, the former capturing essentially general knowledge, enabling the abstraction of quantitative data into useful qualitative information as well as other forms of inference between data, and the latter together with the heuristic links representing the expert knowledge. Lastly ARBY [27] is another diagnostic framework which exhibits a co-operating structure, the co-operative modules being a hypotheses generator and an interaction frames manager. The former generates and tests hypotheses, and the latter

manages the action requests in the course of testing hypotheses, the objective being to ask for the most informative actions; in this respect the interaction frames manager may possess expert knowledge too.

Figure 2 gives a classification of diagnostic expert systems from the perspective of background knowledge representation.



Fig. 2. Classification of diagnostic expert systems from the perspective of background knowledge representation.

7. Conclusion

The primary objective behind a diagnostic task, whether it is performed by a human or a computer system is to achieve timely and accurate diagnosis, enabling the application of effective treatment regimes. The primary objective of a knowledge engineer for a diagnostic expert system is to elicit the diagnostic skills of the particular experts and to formulate an accurate model of their diagnostic competence. In so doing a knowledge engineer should not overlook the fact that a human expert's diagnostic skills are supported on a solid foundation of general domain knowledge and common-sense reasoning. Such background knowledge is second nature to experts, more often than not being subconsciously applied. The coupling between the foreground and background knowledge is strong and unless background knowledge is made explicit in the diagnostic system, the diagnostic competence of the system will be vulnerable. Lack of relevant background knowledge in a diagnostic system quickly manifests itself in terms of redundant or erroneous questions or wrong decisions. Such manifestations may be 'fixed' somehow, e.g. pieces of background knowledge may be implicitly incorporated into the expert knowledge to prevent the specific manifestation from reoccurring. This is not a proper solution since it does not attack the cause but simply its effects.

In summary the explicit representation of background knowledge and its associated reasoning has the following significant advantages:

- The overall functionality of the system is more perspicuous which is absolutely necessary for future extensions. There is a clear demarcation between the expertise of the system and its foundational support. The system's knowledge-base and associated reasoning strategies are not hindered with obscure control constraints which are the side effects of not having explicit background support. The hypotheses and findings reasoners may be independently extended. In particular their underlying models may be independently extended and/or new strategies may be added to either reasoner.
- The problem of knowledge redundancy is controlled. Each piece of knowledge is stored once and may be used in different contexts.
- Reusability is achieved since the strategies incorporated in the findings reasoner are domain independent and also the same findings base can be used in more than one application.
- The system achieves a competent dialogue structure in a natural and extensible way. An implicit representation of the background knowledge would forever require the incorporation of unnatural 'fixes' in order to alleviate specific dialogue or user interface problems.
- The overall diagnostic performance of the system is enhanced, since the findings reasoner goes towards ensuring that all evidence is independent, and that hypotheses profiles are properly tailored for the specific temporal context.

The identification of relevant background domain knowledge and its incorporation as an integral component of the particular competence model must be included in the primary design objectives for a diagnostic system. Past experiences have demonstrated that unless an expert system is designed from the start with the purpose of providing adequate explanations, simply incorporating an explanation facility subsequently on an existing architecture as an add-on *does not* work. Similarly the need to make the background knowledge adequately explicit must influence the diagnostic architecture from the start; it cannot be expected that adding background knowledge subsequently will be sufficiently effective.

Acknowledgements

The support of the Leverhulme Trust for the SDD project is gratefully acknowledged.

References

[1] J. Aikins, Prototypical knowledge for expert system, Artificial Intelligence 20 (1983) 163-210.

^[2] T. Bylander and S. Mittal, CSRL: a language for classification problem solving and uncertainty handling, AI Mag. (Aug. 1986) 66-76.

- [3] B. Chandrasekaran, Generic tasks as building blocks for knowledge-based systems: the diagnosis and routine design examples, *Knowledge Engrg. Rev.* 3 (1988) 183-210.
- [4] B. Chandrasekaran, Models versus rules, deep versus compiled, content versus form: some distinctins in knowledge systems research, IEEE Expert (Apr. 1991) 75-79.
- [5] B. Chandrasekaran and S. Mittal, Conceptual representation of medical knowledge for diagnosis by computer: Mdx and related systems, Advances in Comput. 22 (1983) 217-293.
- [6] W.J. Clancey, The epistemology of a rule-based expert system: a framework for explanation, Artificial Intelligence 20 (1983) 215-251.
- [7] W.J. Clancey, Heuristic classification, Artificial Intelligence 27 (1985) 289-350.
- [8] W.J. Clancey and R. Letsinger, Neomycin: reconfiguring a rule-based expert system for application to teaching, Proc. IJCAI-81 (1981) 829-836.
- [9] L. Console, A. Janin Rivolin and P. Torasso, Fuzzy temporal reasoning on causal models, Internat. J. Intelligent Syst. 6 (1991).
- [10] R.M. Dawood, C.M. Hall, D. Shaw, E.T. Keravnou, J. Washbrook and F. Dams, Model-based expert system for diagnosis in paediatric skeletal radiology, *Radiology* 177 P (1990) 208.
- [11] J. de Kleer and B. Williams, Diagnosing multiple faults, Artificial Intelligence 32 (1987) 97-130.
- [12] A.S. Elstein, L.A. Shulman and S.A. Sprafka, Medical Problem Solving: An Analysis of Clinical Reasoning (Harvard University Press, MA, 1978).
- [13] P.J. Feltovich, P.E. Johnson, J.H. Moller and D.B. Swanson, LCS: The role and development of medical knowledge in diagnostic expertise, in: W.J. Clancey and E.H. Shortliffe, eds., *Readings in Medical Artificial Intelligence: The First Decade* (Addison-Wesley, Reading, MA, 1984) 275-319.
- [14] P.K. Fink, J.C. Lusth and J.W. Duran, A general expert system design for diagnostic problem solving, IEEE Trans. Pattern Analysis and Machine Intelligence PAMI-7 (1985) 553-559.
- [15] G.J. Groen, Modelling clinical explanation, Proc. Advanced Models of Cognition for Medical Training and Practice (NATO Advanced Scientific Research Workshop, Italy, June 1991).
- [16] J. Hunt, A task specific integration architecture for multiple problem solver, model-based, diagnostic expert systems, Ph.D. Thesis, Department of Computer Science, The University College of Wales, Aberystwyth, 1991.
- [17] K. Johnson, J. Sticklen and J.W. Smith, IDABLE application of an intelligent data base to medical systems, Proc. AAAI Spring Artificial Intelligence in Medicine Symp. (Stanford University, March 1988) 43-44.
- [18] E.T. Keravnou and L. Johnson, Competent Expert Systems: A Case Study in Fault Diagnosis (Chapman and Hall, London, 1986).
- [19] E.T. Keravnou and L. Johnson, Towards a generalized model of diagnostic behaviour, Knowledge-Based Syst. 2 (1989) 165-177.
- [20] E.T. Keravnou and J. Washbrook, What is a deep expert system?: an analysis of the architectural requirements of second-generation expert systems, *Knowledge Engrg. Rev.* 4 (1989) 205-233.
- [21] E.T. Keravnou and J. Washbrook, Deep and shallow models for medical expert systems, Artificial Intelligence in Med. 1 (1989) 11-28.
- [22] E.T. Keravnou and J. Washbrook, A temporal reasoning framework used in the diagnosis of skeletal dysplasias, Artificial Intelligence in Med. 2 (1990) 239-265.
- [23] E.T. Keravnou, J. Washbrook and F. Dams, Explicit data-modelling for second-generation diagnostic expert systems, Proc. 11th Internat. Conf. on Expert Systems and their Applications 2 (General Conference on Second-Generation Expert Systems, Avignon 1991) 185-198.
- [24] E.T. Keravnou, J. Washbrook, R.M. Dawood, C.M. Hall and D. Shaw, A model-based diagnostic expert system for skeletal dysplasias, Proc. AIME 89 (Springer, Berlin, 1989) 47-56.
- [25] J.L. Kolodner and R.M. Kolodner, Using experience in clinical problem solving: introduction and framework, *IEEE Trans. Systems Man Cybernet.* SMC-17 (1987) 420-431.
- [26] D.B. Lenat, R.V. Guha, K. Pittman, D. Pratt and M. Shepherd, CYC: Towards programs with common sense, Commun. ACM 33 (1990) 30-49.
- [27] D. McDermott and R. Brooks, ARBY: diagnosis with shallow causal models, Proc. AAAI-82 (1982) 370-372.
- [28] R.A. Miller, H.E. Pople and J.D. Myers, Internist-I: an experimental computer-based diagnostic consultant in general internal medicine, New England J. Med. 307 (1982) 468-476.
- [29] S. Mittal, B. Chandrasekaran and J. Sticklen, Patrec: a knowledge-directed database for a diagnostic expert system, *IEEE Comput.* (Sep. 1984) 51-58.
- [30] Y. Peng and J.A. Reggia, Abductive Inference Models for Diagnostic Problem-Solving (Springer, Berlin, 1990).

- [31] H.E. Pople, Heuristic methods for imposing structure on ill-structured problems: the structuring of medical diagnosis, in: P. Szolovits, ed., Artificial Intelligence in Medicine (AAAS Selected Symposium Series, West View Press, Boulder, CO, 1982) 119-185.
- [32] R. Reboh, Knowledge engineering techniques and tools for expert systems, Linköping Studies in Science and Technology, Dissertations No. 71, Software Systems Research Centre, Linköping University, S-581 83 Linköping, Sweden 1983.
- [33] R. Reiter, A theory of diagnosis from first principles, Artificial Intelligence 32 (1987) 57-95.
- [34] J. Sticklen, MDX2: an integrated diagnostic system, Ph.D. Dissertation, Department of Computer & Information Science, The Ohio State University, 1987.
- [35] P. Torasso and L. Console, Diagnostic Problem Solving: Combining Heuristic, Approximate and Causal Reasoning (North Oxford Academic, 1989).
- [36] W. van de Velde, Learning heuristics in second-generation expert systems, Proc. Sixth Internat. Workshop on Expert Systems and their Applications (Avignon, 1986).
- [37] S.M. Weiss, C.A. Kulikowski, S. Amarel and A. Safir, A model-based method of computer-aided medical decision making, Artificial Intelligence 11 (1987) 145-172.