Inference and Advisory System for Medical Diagnosis
Second Report STW-NGN55.3614

WAJJ Wiegerinck MJ Nijman WJPP ter Burg EWMT ter Braak
YL O JP Neijt HJ Kappen

April 15, 1999

This report describes the progress of the project Inference and Advisory System for Medical diagnosis (STW NGN55.3614) over the period October 1997 until March 1999.

Contents

1 Summary

1.1 Overview 1996-98

Initially, the goal of this project was to use neural networks to learn a medical domain from data. The idea was that since explicit modelling is very tedious, and since patient data is available, a learning approach could be successful. Several medical sub-domains were considered. In the domain of iron metabolism disorders good results were obtained on the basis of data only (a specialists' database of about 500 cases was available - 30 (combinations) of diseases were to be distinguished): the system outperformed all general internists that collaborated in the test. On a subpart of the domain anaemia, results were less good. The available patient data (only 400 out of a total of 10000 cases were considered suitable for training) were not sufficient to learn the relation between all 100 different medical findings and 80 diagnostic categories. Only the most frequently occurring diagnoses could be modelled adequately. This result is disappointing, because physicians are most interested in a system that helps them with the rare diagnoses. We therefore concluded that the approach to use data for model building crucially depends on the availability of high quality data. We estimated that for medical domains in general the availability of high quality data is rather exception than rule. A successful medical decision support system requires explicit modelling effort by human experts in order to reach sufficient performance. Subsequent improvement of the performance of the system can be made by learning from (selected) patient cases.

1.2 Motivation for 1998-99

We therefore decided to abandon our original plan to learn a model of anaemia using data only and to model a subpart of anaemia (namely megaloblastic anaemia) using a Bayesian network. Bayesian networks are explained in section 3. The reason for using belief networks are that according to current AI wisdom, they are the proper way of representing uncertain knowledge, and they are ideal for learning from data. The probabilistic framework can in fact be seen as a very general methodology that includes rule based systems and neural networks as special cases. Due to this integrating aspect, belief networks are becoming the common language of both the rule based AI and neural network communities.

The idea to use a Bayesian network to build a Decision Support System for medical diagnosis is not new. However, a drawback is that complex networks are intractable for exact computation. This has obstructed the development of a useful system for internal medicine, which should cover a broad medical domain at a detailed level. Several studies have shown that existing systems for internal medicine fail in part due to lack in the required detailed level at which the domain is modelled.

Advances in approximation techniques, in particular using variational methods, have opened new possibilities for computation in a large class of complex belief networks which are intractable for exact methods. Our project group addressed the question how these methods could be exploited and/or extended for a system for medical diagnosis.

Although the formalism of Bayesian networks is very powerful, the construction of actual networks is not straightforward. A question is whether pathophysiological knowledge, as found in the medical literature, offers a sufficient basis to guide the development of probabilistic models. Although the medical literature on anaemia indeed includes elaborate descriptions of the mechanisms underlying diseases causing anaemia, no complete pathophysiological models are available. Such models have to be designed from scratch, where one has to stay within the formalism of Bayesian networks. Apart from this generic modelling issue, there are also practical problems of knowledge management in large, complex models. Our aim was to show that these issues can by dealt with by demonstrating the feasibility to make a complex model for medical diagnosis. The demonstration consists of a detailed Bayesian network for megaloblastic anaemia.

1.3 Results 1998-99

- We have developed novel generic methods for approximate inference. These methods open the possibility to use Bayesian networks for detailed modelling of large medical domains. [1, 2, 3], More details are described in section 6.
- We have developed a software environment, called BayesBuilder, for development and (standard) inference of Bayesian networks. The software contains graphical tools for network construction and evaluation, and several tools for network maintenance. BayesBuilder is written in Java and C++. BayesBuilder is described in section 4.
- Using BayesBuilder, we have developed a demonstration version of a diagnostic decision support system, called Promedas. Promedas consists of a graphical user interface (GUI) to enter patient data and a GUI for diagnostic consultancy. Promedas consultancy is based on inference in a Bayesian network. We have realized a network covering megaloblastic anaemia. The model consists of 80 variables. Promedas is written in Java. Promedas is described in section 5, see also [4]. A demonstration version of Promedas will be available on CD-ROM in April 1999. The CD-ROM will be distributed to target users at the "Internisten Dagen" and other events. A web-page with more information about Promedas is launched. (`www.mbfys.kun.nl/snn/Research/promedas`).

1.4 Conclusions 1998-99

Our results sofar indicate that it is indeed possible to obtain very detailed and very accurate diagnoses with the approach using Bayesian networks. A great advantage of this approach over the neural network approach is the transparency of the knowledge. It allows for the automatic generation of explanation facilities which will be of great value for user acceptance.

1.5 Utilisation/Future

The Bayesian network tool BayesBuilder is in a stage where it could be commercially exploited (outside the medical domain). Bolesian BV, (full daughter of Cap Volmac) showed interest in the tool.

The DSS software Promedas is too premature for commercial exploitation. We hope that we can extend Promedas in a follow-up project to a broader medical domain. Several groups, namely Department of Internal Medicine and the Department of Endocrinology, both at Utrecht University Hospital, and companies, namely HISCOM BV, Bolesian BV, and Siemens Germany showed serious interests in a follow-up. We currently have contacts with 'de Nederlandsche Internisten Vereninging' to raise more interest. We will have presentations of Promedas at the 'Internisten Dagen', AIMDM'99 (Aalborg, DK), and Biomed (Basel, CH).

2 Introduction: Decision Support Systems for Medical Diagnosis

2.1 Why use decision support in medicine?

Decisions made by physicians are arbitrary and highly variable (within one physician and between physicians) and often lacking explanation or "rationalisation" [5, 6]. Problems in modern medicine are often very complex, but evidence for the best choice to be made is often lacking. Clinical examples of this phenomenon in diagnosis making are abundant and easy to understand.

The body of potentially useful knowledge that is relevant to even a relatively narrow diagnostic area may be too large to make the optimal (diagnostic) decision on the spot. Ironically, modern information technology (especially through the Internet) increases the amount of available knowledge even more, probably further complicating this situation. Moreover, individual patients need "individualised" decisions, because their characteristics differ from the "average" and because of their individual wishes[7]. Apparently, individualising the general results of research may be cumbersome and time consuming, while on the other hand modern medical practice demands efficiency, cost-effectiveness and high technical quality.

The derivation of diagnostic protocols is a main problem in health care. But in some environments diagnostic support is simply not likely to influence physician's decisions, e.g. on a neurological intensive care unit, since the diagnosis is often obvious [8]. In contrast, general internal medicine covers an enormous range of sometimes relatively rare diagnostic categories. Hence the tendency of medicine to be differentiated in super-specialisations, (e.g. gastro-entrologists, rheumatologists, cardiologists. etc.). A diagnostic DSS covering general internal medicine may be appreciated by both generalists and super-specialists alike: by the generalist because this field of work typically covers a very broad range of diagnoses, by the super-specialist because he/she may not feel completely at ease outside his/her specific field of expertise.

It is readily understandable that the above comprises an enormous task and challenge for modern medicine in general and individual doctors in particular, illustrating the need for decision support techniques. Obviously, automated DSSs may be very promising from a theoretical point of view.

2.2 What are the problems in current decision support systems?

The currently available systems (e.g. Meditel [9], QMR [10], Dxplain [11] and Iliad [12]) have not yet been very successful. Certainly their use is still not widespread and not established in daily routine. A variety of reasons may be responsible for this:

Lack of accuracy

Current systems that intend to cover a broad diagnostic domain of medicine lack diagnostic accuracy [13, 14]. This is not due to the method that is used, but rather due to the levels of detail (e.g. diagnostic categories at the level of ICD-9) and completeness in the knowledge base. Systems that are based on detailed modelling of knowledge may have a good performance. Up to know, however, such systems are restricted to a relatively narrow field [15, 16]. The crucial problem with a consistent detailed model covering a broad domain is that it would be computational intractable. The next section will discuss how can be dealt with this problem.

Lack of transparency

In the era of evidence based medicine the advice of "a black box" is unacceptable. An advice must be motivated and preferably accounted for on the basis of research published in the peer reviewed literature.

Users attitude

A subset of (potential) users may have a misunderstanding about what computers can and cannot do for them. Generally, DSSs need educated and responsible users, who are able to interpret the advise given and estimate its merit [17]. This, however, is not exclusively a matter of users attitude. Producers of decision support tools should take this issue into account as well, especially when designing the user interface and deciding which facilities are needed.

Lack of integration of information

Patient specific decision support needs input data from several sources. A DSS will generate new information (e.g. a diagnostic advise) through inference. For this it uses specific information about a patient, given "patient-independent" knowledge (e.g. about diagnosis making) stored in the knowledge base of the DSS. Integration of information, multiple usability of patient data, integration of databases and knowledge bases are common problems when using a heterogeneous Hospital Information System (HIS). Unfortunately, the completeness of patient information and the accuracy and level of detail of diagnoses stored in a HIS is in general very poor [18].

Lack of a controlled terminology

This is a problem that might not even be solved completely in the near future. Most standard classification systems are at a general level [19, 20], thus lacking the required detail, or specialised [21] and therefore too limited to meet the needs for a DDS covering a broad domain.

Careful introduction

Introduction of a DSS should be done as careful and thorough as is use for drugs that are new on the market. After implementation, the use of a DSS will need meticulous surveillance resulting in further improvement. Maintenance is needed to keep up with the last results of research.

2.3 Why develop a new diagnostic decision support system?

In conclusion, modern medicine is in need of computerised decision aids both to meet its own high standards and to keep pace with the stage of development in other domains such as manufacturing or the services industry. Taking into account the need for decision support and *diagnostic* decision support in particular, we strongly believe that a diagnostic DSS for a broad medical domain is viable and, eventually, marketable.

To avoid a "gold rush style" in the search for these tools, the foremost thing to do is the development of safe and sound methods. The expertise of our multidisciplinary group primarily focuses on three main parts of the methodology typically needed in the development of decision support tools:

- *Inference algorithms that are able to deal with large complex systems.*
- *Knowledge modelling in the medical domain.*
- *User aspects.*

3 Belief networks

Reasoning in the medical domain is a typical example of reasoning with uncertainty. This uncertainty has different sources: uncertain, incomplete or even missing patient information, uncertainty in medical tests, and the inherent uncertainty in physiological processes. Clearly, the model on which a DSS is based should be able to deal with these uncertainties.

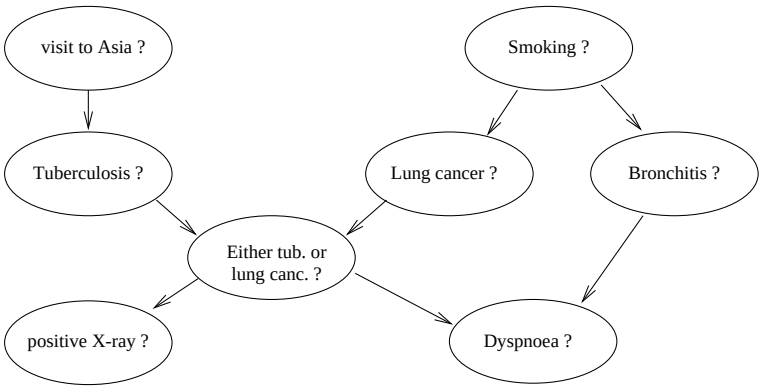


Figure 1: Belief network. Variables are represented by the nodes. Conditional independencies are represented by the arrow structure. The graph of this network implies the following structure of the joint probability model: $p(A,T,X,E,D,L,B,S) = p(A)p(T|A)p(X|E)p(E|T,L)p(D|E,B)p(L|S)p(B|S)p(S)$

The different DSSs that have been developed sofar use a variety of modelling approaches which can be roughly divided into two categories: rule-based approaches and probabilistic methods. The rule based approach can be viewed as an attempt to simplify the probabilistic approach in order to reduce computational complexity. The probabilistic approach has the advantage of mathematical consistency and correctness. Belief networks [22] in particular provide a powerful and conceptual transparent formalism for probabilistic modelling.

A belief network – also called Bayesian networks or causal probabilistic networks – is a graph of nodes and arrows. See fig. 1. The nodes represent random variables whereas arrows between nodes represent direct influences. For each node a conditional probability table quantifying the effect of the parent nodes, i.e. the nodes that directly point to it. The total graphical structure specifies a probability distribution over the state space of all variables. Belief updating is done using the rules of probability (Bayes theorem).

Belief networks have several distinguishing features:

Transparency

A belief network has an appealing, transparent and intuitively clear structure which can be graphically visualised. Expert knowledge can be made explicit, while users can have insight into how the system operates. Belief networks have a more modular representation of uncertain knowledge than rule-based systems. This makes them easier to maintain.

Accuracy

Because all relations between variables are described by the rules of probability, there are no assumptions made by the methodology. The definition of the variables and the structure of the network contain all assumptions in the network. If the accuracy of the network is too much hindered by a particular assumption this can easily be removed by restructuring the network.

"Hidden" variables

Large numbers of observable variables can be related in the model via "hidden", e.g. pathophysiological, variables. Besides the modelling advantages and the improvement of transparency, the use of this type of variables makes the network in general less complex and therefore less sensitive to over-fitting.

Learning

Using standard learning algorithms for belief networks, the system can be fine-tuned using historical patient data and learn further "from experience" on the basis of prospectively gathered patient data.

The progress that has been made during the last decade in exact computation in belief networks makes the argument in favour of rule based approaches less and less persuasive - at least for relatively small and simple models. Indeed, most modern approaches for medical diagnosis are based on the probabilistic approach. A drawback is that complex probabilistic models are intractable for exact computation. New techniques are very promising to overcome this drawback.

4 BayesBuilder

BayesBuilder is a tool for development and (standard) inference of Bayesian networks.

We have chosen to develop our own Bayesian network tool rather to use a commercial package. Main advantage of developing our own software is flexibility: the possibility to extend the tool for our own purposes, e.g. to include new algorithms, or to cope with specific modelling issues in our project. An additional point is that commercially available tools have a rather awkward user interface. We paid in particular attention to optimise the user interface of BayesBuilder in close collaboration with the users in our project group.

4.1 Description

All functionalities in Bayesbuilder are menu or mouse driven. BayesBuilders functionalities include

File management Creating, opening, and saving of networks. If security is needed, networks can be encrypted.

Node management

- Standard manipulations for Bayesian networks: Graphical insertion and deletion of nodes, moving of nodes by drag and drop, graphical linking nodes with other nodes, setting of conditional probability tables. On-line computation of marginal probabilities of all nodes in the networks (see figure 2).
- Additional node management: Storage of additional node information for later reference(see figure 3). Several ways to get convenient overviews of the modelled nodes are implemented, (see figure 3). A standardised reference managers can be used include references to the literature in the node information.

View With this feature, the user can select a smaller amount of nodes of the network to be drawn on the screen. A selected view on the network is useful if the user wants to concentrate on a smaller part of the network for modelling or evaluation. Creation of several views is possible.

Forms This feature enables the user to organise variables in a familiar folder-like way (see Promedas software). Creation and definition of folders is completely done by the user.

In order to get an optimal, platform independent user interface, BayesBuilder's interface is written in Java. For speed and maintenance reasons, BayesBuilders engine is written in C++. Exact and approximate (Gibbs sampling, mean field) inference algorithms have been implemented. Although BayesBuilder development is not yet complete (it probably never will), the tool is in a stage where it can be used in practice (as we did for the Promedas program).

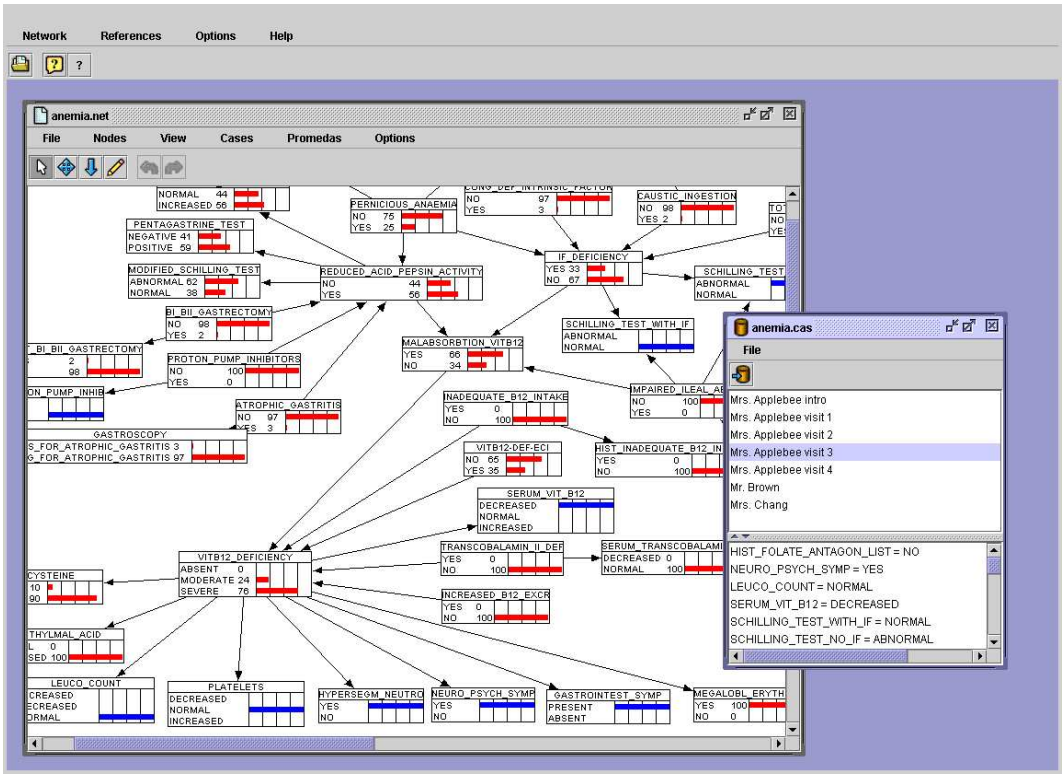
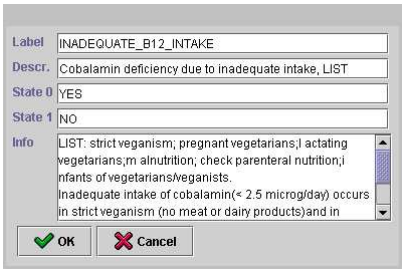
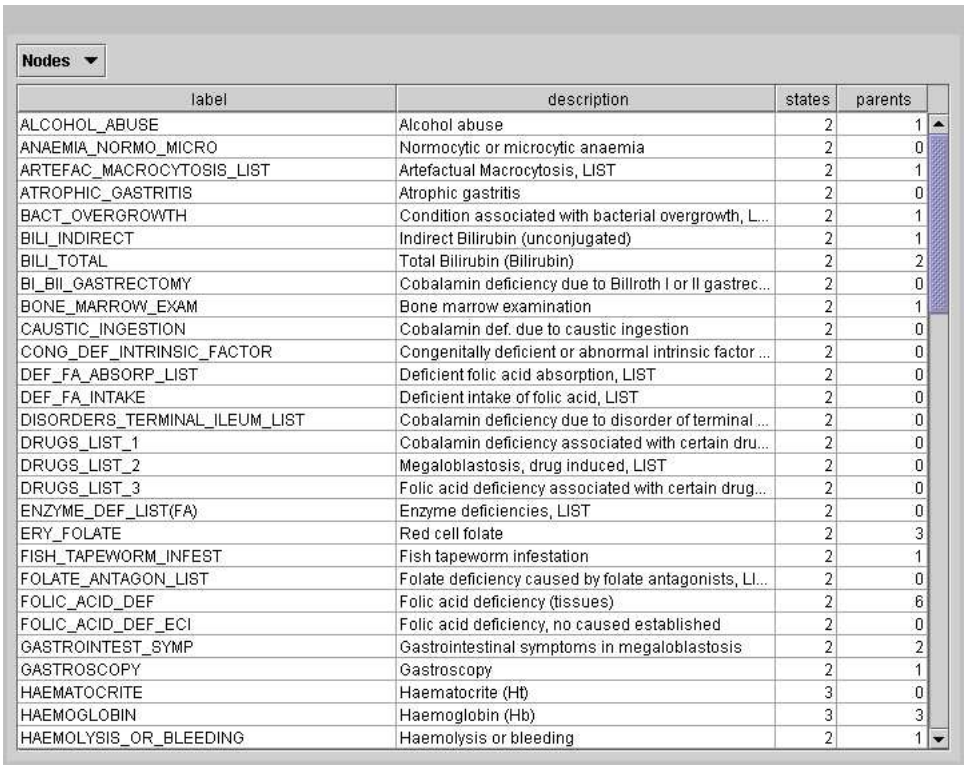


Figure 2: BayesBuilder's interface. A view on the Bayesian network. Nodes represent variables and arrows represent conditional dependencies. The red histograms show the (marginalised) probabilities of the variables. The blue histograms display the clamped nodes. These are externally provided values, in this example loaded from Mrs. Applebee visit 3. .





label	description	states	parents
ALCOHOL_ABUSE	Alcohol abuse	2	1
ANAEMIA_NORMO_MICRO	Normocytic or microcytic anaemia	2	0
ARTEFAC_MACROCYTOSIS_LIST	Artefactual Macrocytosis, LIST	2	1
ATROPHIC_GASTRITIS	Atrophic gastritis	2	0
BACT_OVERGROWTH	Condition associated with bacterial overgrowth, L...	2	1
BILI_INDIRECT	Indirect Bilirubin (unconjugated)	2	1
BILI_TOTAL	Total Bilirubin (Bilirubin)	2	2
BI_BII_GASTRECTOMY	Cobalamin deficiency due to Billroth I or II gastrec...	2	0
BONE_MARROW_EXAM	Bone marrow examination	2	1
CAUSTIC_INGESTION	Cobalamin def. due to caustic ingestion	2	0
CONG_DEF_INTRINSIC_FACTOR	Congenitally deficient or abnormal intrinsic factor ...	2	0
DEF_FA_ABSORP_LIST	Deficient folic acid absorption, LIST	2	0
DEF_FA_INTAKE	Deficient intake of folic acid, LIST	2	0
DISORDERS_TERMINAL_ILEUM_LIST	Cobalamin deficiency due to disorder of terminal ...	2	0
DRUGS_LIST_1	Cobalamin deficiency associated with certain dru...	2	0
DRUGS_LIST_2	Megaloblastosis, drug induced, LIST	2	0
DRUGS_LIST_3	Folic acid deficiency associated with certain drug...	2	0
ENZYME_DEF_LIST(FA)	Enzyme deficiencies, LIST	2	0
ERY_FOLATE	Red cell folate	2	3
FISH_TAPEWORM_INFEST	Fish tapeworm infestation	2	1
FOLATE_ANTAGON_LIST	Folate deficiency caused by folate antagonists, LI...	2	0
FOLIC_ACID_DEF	Folic acid deficiency (tissues)	2	6
FOLIC_ACID_DEF_ECI	Folic acid deficiency, no caused established	2	0
GASTROINTEST_SYMP	Gastrointestinal symptoms in megaloblastosis	2	2
GASTROSCOPY	Gastroscopy	2	1
HAEMATOCRITE	Haematocrite (Ht)	3	0
HAEMOGLOBIN	Haemoglobin (Hb)	3	3
HAEMOLYSIS_OR_BLEEDING	Haemolysis or bleeding	2	1

Figure 3: BayesBuilder's interface. Screendump of the node properties fame (left), and the node manager (right).

5 Promedas

Promedas (PRObabilistic MEdical Diagnostic Advisory System) is a patient-specific diagnostic Decision Support System (DSS). This DSS is a computer program that produces a differential diagnosis using a set of patient findings. Moreover, it suggests additional tests that may be performed to make the differential diagnosis more precise.

The distinguishing features of Promedas will be its accuracy and transparency. Promedas is based on medical expert knowledge, acquired from the literature by our medical specialists. The acquired knowledge will be categorised to fit with international standards, such as ICD-10 and Unified Medical Language (not yet implemented). Using BayesBuilder, this knowledge is translated into a Bayesian network, which serves as inference engine of the system. The system's transparency is enhanced by several explanatory and clarifying facilities, including the availability of the appropriate references to the literature.

The goal of Promedas is to demonstrate that an accurate diagnostic DSS covering a large diagnostic repertoire in medicine is possible. The current demonstration version on the CD is restricted to the megaloblastic anaemias and is intended to give an impression about the future capabilities of the system.

5.1 Description

Promedas' graphical user interface (GUI) consists of five folder-like tab panel, which are called 'Introduction', 'Form', 'Consultant','Network', and 'Cases'.

The user can switch between tab panels via a mouse click on the desired folder tab.

Introduction This tab panel contains general information about Promedas.

Form (see figure 4) This tab panel can be used to view and/or enter patient data. The Form tab panel organises clinical data of individual patients in categories and subgroups in a similar way they might be organised in a conventional Patient Record file. Starting point for the design of this panel was the current conventional hospital patient record, which use is wide-spread and accepted, if not obligatory. In the Form tab panel, clinical data of individual patients is organised in categories and subgroups in a similar way. If the system would be connected to an EPR, the screen display of Forms might be adapted accordingly. Specific (medical) background information can be accessed via 'i' buttons.

Consultant (see figure 5) This panel gives access to the diagnostic advice of the system.

The Consultant tab panel contains four frames:

- *Diagnostic Categories (D):* This frame displays the probabilities of potentially relevant diagnoses as percentages (ranked in descending order), given the values of the variables previously entered.
- *Mechanisms (D):* This frame displays the probabilities of potentially involved underlying mechanisms (e.g. pathophysiologically) as percentages (ranked in descending order), given the values of the variables previously entered.

  The user can select a diagnostic category or mechanism for which he/she wants test proposals by a mouse click.
- *Test Proposals (T):* This frame displays a relative measure (no units) for the information that is to be expected by performing a proposed test *in relation to the selected diagnosis or mechanism,* (given the values of the variables previously entered and the context). This information is defined as

$$I(D,T) = \sum_{D,T} P(D,T) \ln \frac{P(D,T)}{P(D)P(T)} \tag{1}$$

with $P(D,T)$ the joint probability of diagnosis and test result, and $P(D), P(T)$ the marginal probabilites of diagnoses and tests, respectively. The information is normalised, and displayed in descending order.

The user can select a a test proposal T for which he/she wants more information by a mouse click.

- *Test Information:* This frame displays a table with additional information concerning the selected test proposal, *in relation to the selected diagnosis or mechanism,* (given the values of the variables previously entered and the context). This information can be accessed via three different probability tables, namely, the joint probabilities of possible test results and diagnosis states $P(T,D)$; the conditional probabilities of possible test results given the possible diagnosis states $P(T|D)$; the conditional probabilities of possible diagnosis states given the possible test results $P(D|T)$.

In the Consultant tab panel, specific (medical) background information can be accessed via 'i' buttons.

Network This tab panel gives access to view the corresponding state of the underlying Bayesian network. This view is actually a view to the interface of BayesBuilder (see figure 2).

Cases This tab panel gives the names of demonstration cases that can be selected. These demonstration cases can be found in section A. See figure 7

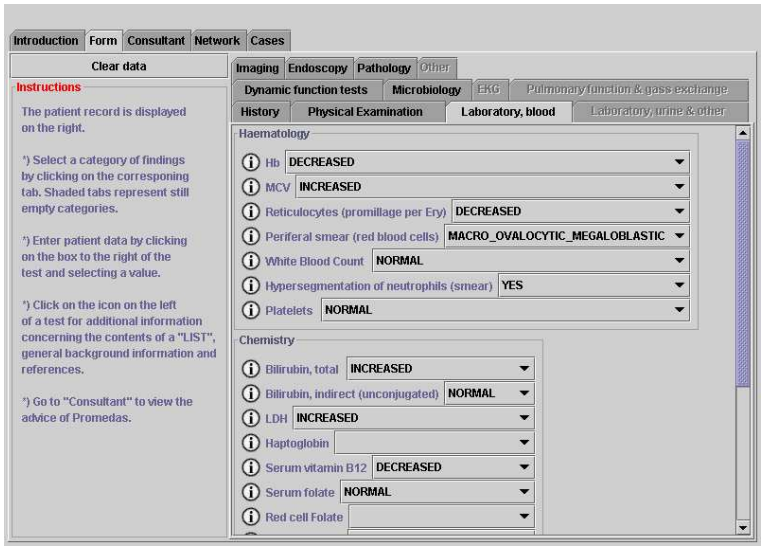


Figure 4: Promedas' interface: the Forms tab panel.

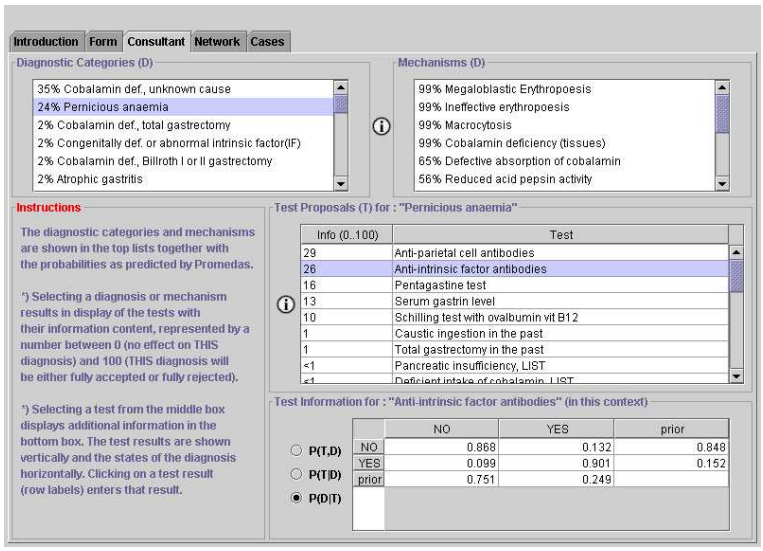


Figure 5: Promedas' interface: the Consultant tab panel. In this example, the most likely diagnosis (given the entered patient-data), Cobalamin def. unknown cause. The selected diagnosis is pernicious anaemia (ranked secondly in the differential diagnosis diagnosis). The test proposals are in relation with this selected diagnosis. The most informative test proposal is anti-parietal cell antibodies. The second most informantive test proposal, anti-intrinsic factor antibodies, is selected for test information. Test information is provided in terms of probablities. In this example, the conditional probability of the diagnosis given the test-result (and all entered patient-data) is provided. For example, the probability of diagnosis pernicious anemia (i.e. D = yes) given the a negative test result of anti-intrinsic factor antibodies (i.e. T = no) is 0.132 (is 13.2 %). The prior indicate $P(D)$ and $P(T)$, e.g. the probability of pernicious anaemia without knowing the test-result (but in the context of all entered patient data) is 0.249 (which is - rounded to below - equal to its probability provided in the upper left panel).

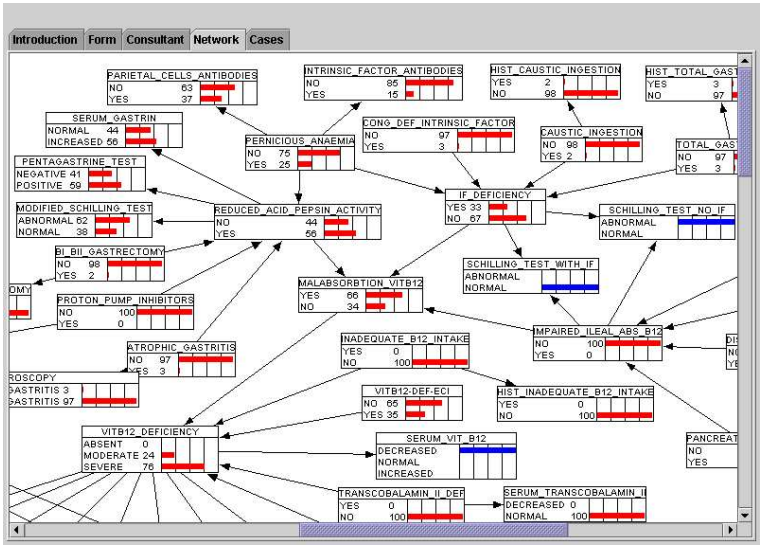


Figure 6: Promedas' interface: the Network tab panel. Cf. figure 2

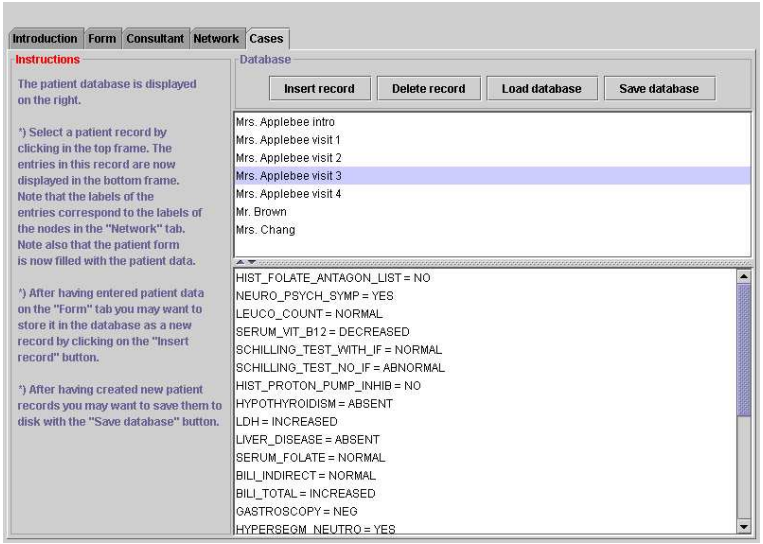
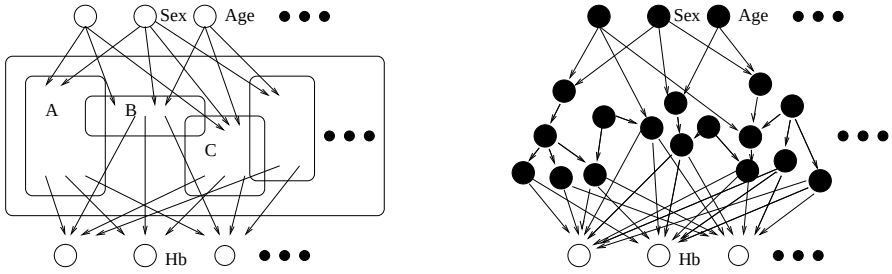


Figure 7: Promedas' interface: the Cases tab panel.

6 Modular structures and Variational methods

In this technical section we outline how the structure of a broad and detailed belief network constructed by human experts will typically look like, based on an extrapolation of the current modelling experiences of the physicians in our group, and explain in more detail why variational methods are likely to be suited to deal with such a network.

Medical experts tend to subdivide knowledge concerning a medical domain, e.g. anaemia, into sub-domains with a relatively small overlap. As a result, the network that models the full domain will typically have a modular structure (cf. fig. 8). Each module represents knowledge about a sub-domain and is modelled by a reasonably small belief network in which the nodes have only a small number of parents. The direct interconnectivity between the sub-domains is also small (the indirect interconnectivity is, however, large as we will discuss later). There are two types of variables outside the sub-domains which link the many sub-domains together. One such type are variables like, for instance, 'age' or 'sex', which determine prior probabilities of diseases. These variables are common ancestors of a large number of sub-domains. The other type of variables are influenced by causes in many sub-domains. An example is the variable 'hemoglobin level' (Hb) in the domain of anaemia. There are many sub-domains within anaemia, each impacting on Hb. That is, variables like Hb are common children of a large number of sub-domains. Since these nodes then have parents in many sub-domains, modelling using explicit probability tables is not feasible (the sizes of the tables grow exponentially in the number of parents). Fortunately, it is not necessary to define these large tables explicitly, since medical experts are likely to have in mind a more compact functional relation between these variables and their direct parents. Such a compact relationship is typically a noisy-OR [22] or a similar parametrised relationship.



(a) Modular structure

(b) Graphical structure

Figure 8: Modular and graphical network structure. Left: modular structure of the network. A, B, C ... represent (overlapping) sub-domains. Each sub-domain is modelled by a number of nodes (cf. right figure) representing variables that are relevant in that domain. The upper nodes, e.g. 'sex' and 'age' represent common ancestors of nodes in several sub-domains. The lower nodes, e.g. 'Hb' represent common children of nodes in several sub-domains (e.g. related to anaemia). Right: underlying graphical structure of same network. Filled circles: nodes in sub-domains and their common ancestors. Open circles: common children

The inference problem is to compute probabilities in the model, given evidence. If a network includes only a few medical sub-domains, exact inference using standard algorithms is feasible. However, in a detailed *and* broad network, such as described in the previous section, exact inference is infeasible due to the connectivity between the modules via the shared variables such as Hb. For instance, exact inference involves a summation over all the parent-states of the variable Hb and this exponentially large summation cannot be performed efficiently. In the following, technical sections, we give a more formal description of the network and we explain how variational techniques can be applied to the inference problem.

6.1 Modular Networks

The considerations of the previous section motivates us to consider a modular network $\mathcal{N}$. A modular network consists of two parts (see fig. 8(b)). One part, which we call the parent network $\mathcal{N}_{\text{parent}}$, is a conventional belief network (black nodes), arranged in a modular structure. We assume that, as a consequence of the (weakly interconnected) modular structure, the parent network is tractable. Nodes of the parent network are connected via noisy-OR gates [22] (other similar gates are equally well possible) to the common children in a second network, which we call the child network $\mathcal{N}_{\text{child}}$ (white nodes in fig. 8(b)). Conditioned on the state of the parent network, the nodes in $\mathcal{N}_{\text{child}}$ are independent. One may view the modular network as a generalisation of the bipartite network of the QMR-DT database (see [23, 24, 25]). The difference is that the upper network in QMR-DT consists merely of disconnected nodes, while in our case the upper layer is a belief network with a non-trivial graphical structure.

In the modular network the probability of a state $S = (S_1, \ldots, S_n)$ therefore assumes the following factorised form,

$$P(S) = \prod_k P(S_k|\pi(S_k)) \prod_i nOR(S_i|\pi(S_i)) \tag{2}$$

with $i \in \mathcal{N}_{\text{child}}$ and $k \in \mathcal{N}_{\text{parent}}$. We denote $\pi(S_k)$ for the state of the parent set $\pi(k)$ of the node k. All the parents are in $\mathcal{N}_{\text{parent}}$. For convenience only, we focus in this treatment on binary variables $S_i \in \{0, 1\}$, representing the presence or absence of a disease or a finding. Extensions to n-ary variables are straightforward. We will use shorthand notation S_i^- for $S_i = 0$, S_i^+ for $S_i = 1$. The noisy-OR gates $nOR(S_i|\pi(S_i))$ are defined such that the probability for S_i^- is

$$nOR(S_i^-|\pi(S_i)) \;=\; (1 - q_{i0}) \prod_{j \in \pi(i)} (1 - q_{ij})^{S_j} \tag{3}$$

and $nOR(S_i^+|\pi(S_i)) = 1 - nOR(S_i^-|\pi(S_i))$. The parameter q_{i0} is the so-called leak probability. This is the probability of a positive finding S_i^+ if all its parents are 0. The parameter q_{ij} can be interpreted as the probability on S_i^+ if only the parent $S_j = 1$, while the others are 0 (if there was no leak probability).

Since $\mathcal{N}_{\text{parent}}$ is assumed to be tractable, exact computation is efficient if the nodes in $\mathcal{N}_{\text{child}}$ are not observed. Also negative findings S_i^- can be dealt with in linear time, since $nOR(S_i^-|\pi(S_i)$ factorises over the parents (see (2)). The problem in this network is that exact computation is inefficient for positive nodes S_i^+ in the child network[24, 25], since the computational costs of inference involving these nodes scales exponentially in the number of parent states. In the next section we will propose variational approximations to deal with this problem.

6.2 Variational Methods for Modular Networks

The starting point of variational methods is to transform the inference problem into an equivalent optimisation problem. The optimisation problem has a simpler structure than the original inference problem, but there are unknown parameters involved. If one is able to find the optimal set of these parameters, one has solved the inference problem. Of course, to find the optimal set of parameter is just as hard, or even harder than the inference problem itself. The trick in the variational methods is to restrict the parameter space so that the optimisation problem becomes feasible. Although the solution of the restricted optimisation problem does not lead to the exact solution of the inference problem, it is guaranteed to bound the exact solution.

In the following subsections, we show that upper and lower bounds of marginal likelihoods can be computed. The way that these bounds are derived are similar to the upper bounds for the bipartite QMR-DT network derived in [24, 25] and the lower bounds for sigmoid belief networks derived in [26]. The difference with these previous papers, however, is that we are able to exploit more fully the graphical structure of the parent network, as in [1, 2], leading to a better approximation algorithm.

Once upper and lower bounds of marginals are computed, bounds on conditional probabilities can be obtained by taking fractions of the marginal bounds, see [24, 25] for more details. Techniques to combine approximate and exact combinations can also be found in these references.

6.3 Upper Bound

Before we proceed, it is convenient to re-express the noisy-OR gates (3) using an exponential notation with parameters $\theta_{ij} = -\log(1 - q_{ij})$,

$$\begin{aligned} nOR(S_i^-|\pi(S_i)) &= \exp(-\theta_{i0} - \sum_{j\in\pi(i)} \theta_{ij} S_j) \\ &= \exp(-z_i) \end{aligned} \tag{4}$$

in which with $z_i = \sum_j \theta_{ij} S_j + \theta_{i0}$. The upper bound of the marginal likelihood is based on the following inequality, valid for any given $z \in I\!R$

$$\ln(1 - e^{-z}) \leq \xi z - F^*(\xi) \tag{5}$$

in which $F^*(\xi) = -\xi \ln \xi + (1 + \xi) \ln(1 + \xi)$. For given z, this inequality is valid for each value of the variational parameter ξ. If the right-hand-side of (5) is minimised with respect to ξ, the inequality becomes an equality (for given z).

Applying inequality (5) to the modular network (2) we obtain an upper bound on the marginal likelihood $P(S_V)$ of the 'visible' variables $V = \{V_{\text{parent}}, V_{\text{child}}\} \subset \{\mathcal{N}_{\text{parent}}, \mathcal{N}_{\text{child}}\}$,

$$P(S_V) \leq \sum_{\{S_{H\text{parent}}\}} \prod_k P(S_k|\pi_(S_k)) \exp(\sum_{i^+} \xi_{i^+} z_{i^+} - F^*(\xi_{i^+}) - \sum_{i^-} z_{i^-}) \tag{6}$$

Nodes with indices k are in $\mathcal{N}_{\text{parent}}$. Nodes with indices $i^{+/-} \in V_{\text{child}}$ have positive/negative findings. $S_{H\text{parent}}$ are undetermined (hidden) states in the parent network. Note that this bound is tractable, since the product over child nodes is log-linear in the parent states S_j (which are contained in the $z_{i^{+/-}}$'s). As a result, the graphical structure of the parent states is therefore not affected, and remains tractable. To get the bound as tight as possible, we optimise the right-hand-side of (6) with respect to the ξ_{i^+}'s, using some numerical procedure.

6.4 Lower Bound

Recently it has been proposed to use variational techniques to obtain a lower bound of the marginal likelihood in sigmoid belief networks [26, 1, 2]. Similar methods can be applied to the modular networks considered here. To derive the lower bound for a network with noisy-OR gates, we use the following expansion of the exponential function [25],

$$1 - \exp(-z) = \prod_{\kappa=0}^{\infty} (1 + \exp(-2^\kappa z))^{-1}$$

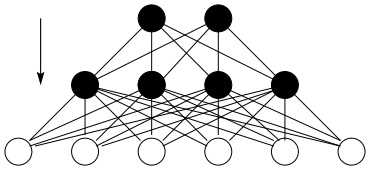
from which we deduce, using Jensen's inequality,

$$\langle \ln(1 - \exp(-z)) \rangle_Q \geq -\sum_{\kappa=0}^{\infty} \ln(\langle 1 + \exp(-2^\kappa z) \rangle_Q).$$

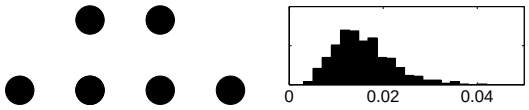
Using this bound, combined with the general theory of variational lower bounds of the likelihood we obtain,

$$\begin{aligned} \ln P(S_V) \geq \mathcal{L}[Q] = &\sum_k \Big\langle \ln P(S_k|\pi_(S_k)) \Big\rangle_Q - \sum_{i^-} \langle z_{i^-} \rangle_Q \\ &+ \sum_{i^+} \sum_{\kappa=0}^{\infty} \ln(1 + \langle \exp(-2^\kappa z_{i^+}) \rangle_Q) \\ &- \sum_{\{S_H\}} Q(S_H) \ln Q(S_H) \end{aligned} \tag{7}$$

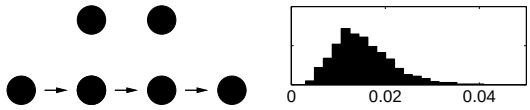
which is valid for any approximating distribution $Q(S_H)$. $\langle\cdot\rangle_Q$ is the average with respect to the so-called mean field distribution $Q(S_H)$. Since this inequality holds for any Q, one can make the bound as tight as possible by optimising $\mathcal{L}[Q]$ with respect to Q. Recently, it has been shown that if tractable belief networks are used for the mean field distribution $Q(S_H)$, then the optimisation of $\mathcal{L}[Q]$ with respect to Q can be performed efficiently using mean field equations [1, 2]. The same studies noted the increase in precision of the bound if the structure of the mean field distribution had more overlap with the structure of the network that was to be approximated (see fig. 9). In the modular networks considered here, a natural structure for the mean field distribution $Q(S_H)$ is a belief network with the same structure as $\mathcal{N}_{\text{parent}}$.



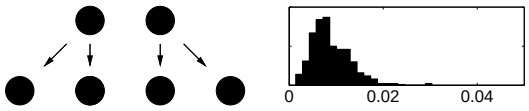
(a) Graphical structure of the 2-4-6 nodes sigmoid belief network that is approximated. Open circles: visible units S_V. Filled circles: hidden units S_H. Maximum clique size: 6.



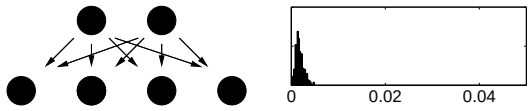
(b) disconnected ('standard mean field') - 16 parameters, mean: 0.01571(5). Max. clique size: 1



(c) chain - 19 parameters, mean: 0.01529(5). Max. clique size: 2



(d) trees - 20 parameters, mean: 0.0089(1). Max. clique size: 2



(e) network - 28 parameters, mean: 0.00183(1). Max. clique size: 3

Figure 9: Variational methods using belief networks applied to a toy benchmark problem to compare its performance with standard variational methods. The problem is to do inference in a three layer (2-4-6 nodes) sigmoid belief network in which the last 6 nodes are visible (panel(a)). 500 networks with parameters drawn randomly from the uniform distribution on [-1,1] were generated. The lower bounds $\mathcal{L}$ for several approximating structures (including 'standard mean field', panel(b)) are compared with the true log likelihood, using the relative error $\mathcal{E} = \mathcal{L}/\ln P(S_V) - 1$, plotted in the histograms panels (b-e). Note that the horizontal scale has been fixed to [0,0.05] in all plots. The results show that considerable improvements can be obtained when belief networks are used. In particular, they indicate that exploiting knowledge of the graphical structure of the model is useful. For instance, the chain (panel (c)) with no graphical overlap with the original graph shows hardly any improvement over the standard mean field approximation. On the other hand, the tree model (panel (d)), which has about the same number of parameters, but a larger overlap with the original graph, does improve considerably over the mean field approximation By increasing the overlap, as in panel (e), the improvement gained is even greater. The maximum clique size refers to the complexity of computation for each approximation, which is exponential in this quantity.

A Demonstration Cases

We present three cases in the domain of the megaloblastic anaemias to demonstrate reasoning in the Promedas program. In the following text you find the descriptions of the cases in a conventional way (and some remarks about the advice of Promedas). These cases are stored in Promedas, and can be accessed in the Cases tab panel.

CASE 1: Mrs. Applebee, a 64 years old lady with type 2 diabetes, in whom routine laboratory examination revealed slight anaemia.

CASE 2: Mr. Brown, who was operated 18 years ago for recurrent peptic ulcer disease.

CASE 3: Mrs. Chang, a vegetarian who is in the 20th week of her first pregnancy.

Note: case 1 is designed to give insight in the changing probabilities in the underlying model, each time additional information is added. Moreover, it is possible to view which test Promedas would have suggested next in that particular stage of the diagnostic process. The other cases straightforwardly offer an illustration of some other diagnostic categories within this domain.

CASE 1, MRS. APPLEBEE

Introduction

Mrs. Applebee is a 64 years old lady who is seen on a regular basis in the out-patient department by a diabetologist. Because of her type 2 diabetes she is currently being treated with insulin since 3 years.

Routine laboratory examination during a scheduled appointment detected slight anaemia.

The diagnostic process that had been prompted by the detection of the anaemia is retrospectively reviewed by Promedas. The story is subdivided in 4 visits and finishes with some"comments".

Mrs. Applebee, visit 1

This is the scheduled appointment when the anaemia was detected.

PREVIOUS HISTORY:

1967 Car accident

1986 Cholecystectomy

Type 2 diabetes since 1986, complicated by microalbuminuria, hypertension, combined dyslipidaemia, neuropathy stage 1

CURRENT MEDICATION:

Insulin, Simvastatin, Gemfibrozil, Enalapril

SMOKING HABITS:

stopped smoking 10 years ago

ALCOHOL:

1 unit per day

ROUTINE LABORATORY EXAMINATION:

Hb 6.9 mmol/L, creatinine 102 umol/L, HbA1c 8.1%, normal results for ALAT, AF and CK

Mrs. Applebee, visit 2

This was the next scheduled appointment. The problem was discussed with the patient by her physician and a diagnostic procedure for the anaemia was instituted.

HISTORY:

Slight loss of appetite since a few months. Bloating after dinner. Occasional loose stools. Prickling sensations of the feet (unchanged since several years).

PHYSICAL EXAMINATION:

Body weight 83 kg (unchanged). Somewhat pale looking elderly lady. Nil else of note.

ADDITIONAL AND REPEATED TESTS:

Hb 6.9 mmol/L (decreased), MCV 102 fL (increased), leukocyte (WBC) and platelet counts both near the lower boundaries of the normal values, reticulocyte count: 2 promille (decreased), total bilirubin 23 umol/L (slightly increased), bilirubin direct ¡ 20% (normal), LDH 724 (increased).

Cholesterol decreased from 6.6 mmol/L to 3.7 mmol/L during the previous year, triglycerides 1.8 mmol/L, ASAT, ALAT, AF, gammaGT: all within normal limits.

TSH: 2.3 mU/L (normal), ferritin and folic acid both normal, vitamin B12 105 pmol/l (decreased).

Gastroscopy: normal.

Mrs. Applebee, visit 3

This is the next visit, 2 weeks later. The patient was informed that her anaemia was probably caused by vitamin B12 deficiency. The physician explained that it was necessary to establish the cause of this deficiency.

SUBSEQUENTLY ORDERED ADDITIONAL TESTS:

Periferal smear: hypersegmentation of neutrophils and macro-ovalocytic erythrocytes.
Schilling test without intrinsic factor: abnormal.
Schilling test with intrinsic factor: normal.

Mrs. Applebee, visit 4

This was another 3 weeks later.

ADDITIONAL TESTS:

Antiparietal cell antibodies: positive Anti-intrinsic factor antibodies: positive Fasting serum gastrin: increased

CONCLUSION

DIAGNOSIS OF THE PHYSICIAN IN CHARGE:

Pernicious anaemia.

DIAGNOSTIC ADVISE SUGGESTED BY PROMEDAS:

Anaemia due to vitamin B12 deficiency, probably pernicious anaemia.

COULD PROMEDAS HAVE BEEN OF HELP?

Promedas would not have needed Schilling tests nor a gastroscopy in order to suggest a correct diagnosis of Pernicious Anaemia.

CASE 2: MR. BROWN

Mr. Brown is 72 years old. He is going to have orthopaedic surgery. Pre-operative screening revealed anaemia. He was referred to the out-patient department of internal medicine and a diagnostic procedure was performed.

PREVIOUS HISTORY:

1960 appendectomy
1972 pneumonia
1979 recurrent peptic ulcer disease
1980 Billroth II operation
1984 cholecystectomy

CURRENT MEDICATION:

ibuprofen

SMOKING HABITS:

he does not smoke

ALCOHOL:

none

CURRENT HISTORY:

no other complaints beside painful hip joints.

PHYSICAL EXAMINATION:

Overweight male, difficulty walking. Blood pressure 170/95 mm.

LABORATORY EXAMINATION:

Hb 6.7 mmol/L (decreased), MCV 104 (increased), vitamin B12 decreased, folic acid normal, ferritin normal

CONCLUSION:

It was concluded that vitamin B12 deficiency due to the Billroth II operation in the past is the most likely cause of the anaemia. No additional diagnostic tests were ordered.

Promedas agrees with the physician in charge.

CASE 3: MRS. CHANG

Mrs. Chang is 20 weeks pregnant of her first child. She is 29 years old and strictly vegetarian since she was 20. The obstetrician performed a routine laboratory examination. Advice was asked from an internist concerning the results.

LABORATORY EXAMINATION:

Hb 6.0 mmol/L (decreased), MCV 115 (increased), creatinin: 58 (normal).

CURRENT MEDICATION:

Folic acid 0.5 mg daily since 8 months ago, when she stopped her oral contraceptive.

PREVIOUS HISTORY:

She has always been healthy

CURRENT HISTORY:

Some numbness of the toes and fingers, fatigue. No diarrhoea.

PHYSICAL EXAMINATION:

Absent ankle tendon reflexes. Otherwise healthy pregnant woman.

CONCLUSION:

It was concluded that vitamin B12 deficiency elicited by pregnancy in long standing vegetarianism was a quite likely cause of the anaemia and that the neurological findings were compatible with this.

Promedas agrees with the physician in charge.

B Publications in this project

[1] W.A.J.J Wiegerinck, W.J.P.P. ter Burg, P.S. van Dam, Y.L. O, J.P. Neijt, and H.J. Kappen. An advisory system for anaemia based on Boltzmann machines. In H-J Zimmerman, editor, *EUFIT '97, 5th Eurpean Congres on Intelligent Techniques and Soft Computing*, volume 1, pages 364–8, Aachen, 1997. Verlag Mainz.

[2] W.A.J.J Wiegerinck, W.J.P.P. ter Burg, P.S. van Dam, Y.L. O, J.P. Neijt, and H.J. Kappen. Lab-test selection in diagnosis of anaemia. In H.J. Kappen and C.C.A.M Gielen, editors, *Proceedings SNN*. World Scientific, 1997.

[3] W. Wiegerinck *et al.* A learning diagnostic system for anemia based on boltzmann machines. Conference on Medical Statistics, Oberwolfach, 1997.

[4] W. Wiegerinck, W. ter Burg, E. ter Braak, P. van Dam, Y. O, Neijt J., and H Kappen. Inference and advisory system for medical diagnosis. Technical report, SNN, 1997.

[5] W.A.J.J. Wiegerinck and H.J. Kappen. Mean field theory for graphical models as a small coupling expansion. In *Proceedings Snowbird*, 1998.

[6] W. Wiegerinck and D. Barber. Mean field theory based on belief networks for approximate inference. In L. Niklasson, M. Bodén, and T. Ziemke, editors, *ICANN'98: Proceedings of the 8th International Conference on Artificial Neural Networks, Skövde, Sweden, 2-4 September 1998*, volume 2, pages 499–504, London, 1998. Springer.

[7] W. Wiegerinck and D. Barber. Variational belief networks for approximate inference. In H. La Poutre and J. van den Herik, editors, *Proceedings of the Tenth Netherlands/Belgium Conference on Artificial Intelligence (NAIC'98)*, pages 177–83, Amsterdam, 1998. CWI.

[8] W. Wiegerinck *et. al.* Variational approximations in a broad and detailed probabilistic model for medical diagnosis. In H. La Poutre and J. van den Herik, editors, *Proceedings of the Tenth Netherlands/Belgium Conference on Artificial Intelligence (NAIC'98)*, pages 185–91, Amsterdam, 1998. CWI.

[9] H.J. Kappen. Variational approximations in a broad and detailed probabilistic model for medical diagnosis. In *Workshop on Highly Structured Stochastic systems, European Science Foundation*, Tirano, Italy, 1998.

[10] W.A.J.J. Wiegerinck and E.W.M.T ter Braak*et. al.* Consultant, a prototype decision support system for medical diagnosis. Technical report, SNN / UMCU, Nijmegen, The Netherlands, 1998.

[11] H.J. Kappen, W. Wiegerinck, E.W.M.T ter Braak, W.J.P.P ter Burg, M.J. Nijman, Y.L. O, and J.P. Neijt. Approximate inference in medical diagnosis. *Pattern Recognition Letters*, 1999.

[12] H.J. Kappen. Approximate inference in medical diagnosis. In *Biomed.* Basel, 1998.

[13] W.J. ter Burg *et al.* A diagnostic advice system based on pathophysiological models of diseases. In *Medical Informatics Europe*, Ljubljana, Slovenia, 1999.

[14] Y.L. O, W.J.P.P. ter Burg, E.W.M.T. ter Braak, J.P. Neijt, W.A.J.J. Wiegerinck, M.J. Nijman, and H.J. Kappen. A development protocol for a diagnostic dss. In *Medical Informatics Europe*, Ljubljana, Slovenia, 1999. submitted.

[15] Promedas. CD-ROM, April 1999.

C Project team

Dr. M.M.H.J. Theeuwen (Neural networks)	1-1-96 until 1-9-96 full time
Dr. W.A.J.J. Wiegerinck (Neural networks)	1-9-96 until present full time
Ir. M.J. Nijman (Scientific programmer)	1-3-98 until present 0.8 fte
Drs. W.J.P.P. ter Burg (Medical informatics)	1-9-96 until present full time
Dr. P.S. van Dam (Internal Medicine)	1-1-96 until 1-8-97 full time
	1-8-97 until present 0.2 fte
Dr. E.W.M.T. ter Braak (Internal Medicine)	1-8-97 until present 0.8 fte
Dr. M.G. Nieuwenhuis (Internal Medicine)	1-8-97 until present 0.0 fte
Dr. H.J. Kappen (Project coordinator)	1-1-96 until present 0.2 fte
Dr.ir. Y-L. O (Supervisor Medical informatics)	1-1-96 until present 0.1 fte
Dr. J.P. Neijt (Supervisor Internal Medicine)	1-1-96 until present 0.1

D Acknowledgement

The assistance of dr. P. Lucas (Dept. Computer Science, Utrecht University) in the process of implementing medical findings in a Bayesian network is acknowledged.

References

[1] W. Wiegerinck and D. Barber. Mean field theory based on belief networks for approximate inference. In L. Niklasson, M. Bodén, and T. Ziemke, editors, *ICANN'98: Proceedings of the 8th International Conference on Artificial Neural Networks, Skövde, Sweden, 2-4 September 1998*, volume 2, pages 499–504, London, 1998. Springer.

[2] W. Wiegerinck and D. Barber. Variational belief networks for approximate inference. In H. La Poutre and J. van den Herik, editors, *Proceedings of the Tenth Netherlands/Belgium Conference on Artificial Intelligence (NAIC'98)*, pages 177–83, Amsterdam, 1998. CWI.

[3] W. Wiegerinck *et. al.* Variational approximations in a broad and detailed probabilistic model for medical diagnosis. In H. La Poutre and J. van den Herik, editors, *Proceedings of the Tenth Netherlands/Belgium Conference on Artificial Intelligence (NAIC'98)*, pages 185–91, Amsterdam, 1998. CWI.

[4] W.J. ter Burg *et al.* A diagnostic advice system based on pathophysiological models of diseases. In *Medical Informatics Europe*, Ljubljana, Slovenia, 1999.

[5] D. M. Eddy. The challenge. *Journal of American Medical Association*, 263:287–290, 1990.

[6] M. Berg. *Rationalizing Medical Work, Decision-support Techniques and Medical Practices.* MIT press, 1997.

[7] R.J. Lilford, S.G. Pauker, D.A. Braunholz, and J. Chard. Decision analysis and the implementation of research findings. *BMJ*, 317:405–409, 1998.

[8] B. Brigl, P. Ringleb, T. Steiner, P. Knaup, W. Hacke, and R. Haux. An integrated approach for a knowledge-based clinical workstation: Architecture and experience. *Methods of Information in Medicine*, 37:16–25, 1998.

[9] Meditel, Devon, Pa. *Meditel: Computer assisted diagnosis*, 1991.

[10] CAMDAT, Pittsburgh. *QMR (Quick Medical Reference)*, 1992.

[11] Massachussetts General Hospital, Boston. *DXPLAIN*, 1992.

[12] Applied Informatics, Salt Lake City. *ILIAD*, 1992.

[13] AS Elstein, CP Friedman, FM Wolf, G Murphy, J Miller, P Fine, P Heckerling, T Miller, J Sisson, S Barlas, K Biolsi, M Ng, X Mei, and T Franz. Performance of four computer-based diagnostic systems. *N-Engl-J-Med.*, 330(25):1792–6, 1994.

[14] ES Berner, JR Jackson, and J Algina. Relationships among performance scores of four diagnostic decision support systems. *J-Am-Med-Inform-Assoc.*, 3(3):208–15, 1996.

[15] D.E. Heckerman, E.J. Horvitz, and B.N. Nathwani. Towards normative expert systems: part I, the Pathfinder project. *Methods of Information in medicine*, 31:90–105, 1992.

[16] D.E. Heckerman and B.N. Nathwani. Towards normative expert systems: part II, probability-based representations for efficient knowledge acquisition and inference. *Methods of Information in medicine*, 31:106–116, 1992.

[17] J.J. Dijkstra. *On the use of computerised decision aids.* PhD thesis, State University of Groningen, 1998.

[18] W. Wiegerinck, W. ter Burg, E. ter Braak, P. van Dam, Y. O, Neijt J., and H Kappen. Inference and advisory system for medical diagnosis. Technical report, SNN, 1997.

[19] WHO. *International Classification of Diseases, 10th revision, Clinical Modification ICD-10*, 1992.

[20] J. Boersma, R.S. Gebel, and H. Lamberts. *International Classification of Primary Care, Short titles (vertaald in het Nederlands) en Nederlandse subtitels.* Nederlands Huisartsen Genootschap, 1993.

[21] WHO. *International Classification of Diseases for Oncology.*

[22] J. Pearl. *Probabilistic Reasoning in Intelligent systems: Networks of Plausible Inference.* Morgan Kaufmann Publishers, Inc., 1988.

[23] M.A Shwe, B. Middleton, D.E. Heckerman, M. Henrion, Horvitz E.J., H.P. Lehman, and G.F. Cooper. Probabilistic Diagnosis Using a Reformulation of the Internist-1/ QMR Knowledge Base. *Methods of Information in Medicine*, 30:241–55, 1991.

[24] T.S. Jaakkola and M.I. Jordan. Variational methods and the QMR-DT database. MIT Computational Cognitive Science Technical Report 9701, Massachusetts Institute of Technology, 1997.

[25] T. Jaakkola. *Variational Methods for Inference and Estimation in Graphical Models.* PhD thesis, Massachusetts Institute of Technology, 1997.

[26] L.K. Saul, T. Jaakkola, and M.I. Jordan. Mean field theory for sigmoid belief networks. *Journal of Artificial Intelligence Research*, 4:61–76, 1996.