A connectionist approach to the diagnosis of dementia

Benoit H. Mulsant, Emile Servan-Schreiber

Carnegie Mellon University, Pittsburgh, PA

Abstract

This paper describes an implemented connectionist network that performs clinical diagnosis in the domain of dementia. During the past decade, connectionism --also called parallel distributed processing or neural processing-- has been established as a new cognitive and computational paradigm, with strong claims that it provides powerful mechanisms to bring solutions to problems previously intractable. To study the suitability of connectionist networks to perform a sequential diagnostic classification task under uncertainty, we have implemented a network that learns to diagnose cases of dementia. We describe in detail the implementation, training, and behavior of this network. We also discuss directions for future research suggested by the limitations of this network.

Background

In the early 70s, analyses of the limitations of Bayesian diagnostic programs led to the adoption of Artificial Intelligence (AI) techniques to develop clinical decision support systems [1, 2, 3]. As a result, the first generation of medical expert systems was developed [4, 5, 6, 7, 8, 9]. These systems perform diagnostic classification tasks by matching findings with stored profiles of diseases using statistic or heuristic associations between findings and diagnoses. Despite high levels of performance, these systems suffer from serious limitations: "[They] are virtually unable to cope with variations in the clinical picture. In particular they have difficulties in recognizing variations in the way that a disease can present, in terms of both the spectrum of findings and severity" [10].

To overcome these limitations several researchers have proposed the adoption of causal reasoning based on representations of pathophysiologic mechanisms [7, 11]. During this decade, causal reasoning has become one of the major focuses of research in Artificial Intelligence in Medicine (AIM) [10]. However, some authors [7, 12] pointed out early the severe limitations of causal reasoning to initiate and guide the diagnostic process: "[First] principles are good for summarizing arguments and good to fall back on when you've lost grasp on the problem but they don't drive the process of medical reasoning... It may be difficult or impossible to expect a set of diagnostic rules both to serve as concise, "clincher" methods for efficiently getting to the right data and still represent a model of disease" [12]. In addition, in many (if not most) important clinical domains, the relevant underlying causal mechanisms are not known.

The appeal of connectionism resides in its potential to overcome some of the shortcomings of traditional methods that directly match clinical findings with stored profiles of diseases and of causal reasoning. A connectionist network can be viewed as a device constituted by many units, organized in distinct layers. Its input is constituted by the activation of the units of the lowest layer; this activation vector is propagated bottom-up to units of higher layers through activating and inhibiting connections; in the process, the activation vector is internally transformed; the activation pattern of the units of the highest layer constitutes the output.

To date, connectionist networks have been successfully applied to several low level cognitive tasks involving perception, language processing, or motor control [13]. Connectionist techniques seem to be suited to the modeling of higher level tasks such as classification problem-solving under uncertainty of which medical diagnosis is a typical example [14]. Some characteristics of such connectionist networks are particularly relevant: in a connectionist network information is combined simultaneously and interactively. This allows the implementation of what has been characterized as flexible frames [15]. These flexible frames can cope with variations in the clinical presentation of a disorder (due for instance to various degree of severity) without having to specifically represent every variation. Similarly, connectionist networks can model how various clusters of findings generate and support a diagnostic hypothesis, without having to implement each possible cluster. A connectionist network supports generalizations; this of course is essential in clinical domains where we need to reason on incomplete and noisy data. Connectionist networks also provide a framework to manipulate predictive and diagnostic beliefs and to manage uncertainty. Finally, connectionism provides several learning procedures [16, 17, 18, 19]; learning is a very attractive feature that may have the potential to help overcome the knowledge-acquisition bottle-neck [20].

Implementation

To study the validity of these theoretical claims and the feasibility of developing a connectionist network to perform sequential medical diagnosis, we have implemented a network that deals with the clinical diagnosis of dementia. In this section, we describe in detail the implementation and the operation of this network. This description illustrates concretely what the application of connectionist techniques entails and what results can be expected.

Task analysis: The diagnosis of dementia

Elstein, Shulman, & Sprafka [21] have shown that physicians' styles of problem-solving are often highly dependent on the specific problem at hand. The diagnosis of dementia is a problem fairly well circumscribed and of moderate complexity (but of great practical importance).

Dementia is a clinical syndrome defined as an acquired

0195-4210/88/0000/0245$01.00

persistent impairment of intellectual function with compromise of language, memory, visuospatial skills, emotion or personality, and cognition (abstraction, calculation, judgement, etc.) [22]. When dementia is suspected, the physician's goal is to establish its presence and to identify either a dementing disorder or a non-dementing disorder that mimics the syndrome. Particular attention has to be given to treatable conditions. As with many other diagnostic tasks, the diagnostic process is guided and biased by available therapeutic and management options [23, 24]. For instance, psychiatrists tend to diagnose treatable disorders (e.g., depression) more readily than irreversible disorders (e.g., Alzheimer's disease).

For purpose of simplification, we have ignored several rare dementing diseases (e.g., syphillitic meningoencephalitis, vitamin B12 deficiency, thyroid and parathyroid disorders, tuberculous meningitis, etc.) that are rarely ruled in clinically but routinely ruled out by laboratory tests [25]. We have retained seven classes of dementing disorders that can be clearly associated with distinct management decisions (including further evaluation): dementia of Alzheimer's type (DAT); vascular dementia (including multiple-infarct dementia -- MID); hydrocephalic dementia (NPH); Parkinson's dementia; dementia with space-occupying lesion (SOL; i.e, primary and secondary neoplasms; sub-dural hematoma); delirium; dementia syndrome of depression (pseudodementia). Technically neither delirium nor pseudodementia are dementias but they are (with normal aging) major elements of the differential diagnosis [26]. These seven classes constitute the diagnoses known by our model. It is important to notice that many dementing disorders evolve over several years and may present very differently at different stages. For example, a patient with an early DAT could present with a mild memory deficit, some difficulties in performing complex construction tasks (e.g., copying a Necker cube), anxiety, and sadness because of his increasing difficulties at work. Five years later the same "Alzheimer's patient" may become mute, bedridden, incontinent and totally oblivious of his condition. The same diagnostic label can correspond to quite different sets of manifestations.

Design: Structure of the network.

A study using high-fidelity clinical simulations [21] and protocol analysis [27] confirmed that in this domain psychiatrists use a hypothetico-deductive approach (unpublished data, Mulsant & Servan-Schreiber). This study also demonstrated that, since knowledge on the pathophysiology of dementing disorders is very limited, causal reasoning based on "first principles" plays only a minimal role. Therefore, we postulated a fairly simple knowledge structure associating observed data (e.g., stopped working; takes Valium), data interpreted contextually into findings (e.g., difficulty working; insomnia), explicit or "hidden" diagnostic facets corresponding to clusters of findings (e.g., dementia; cognitive impairment), and high level diagnostic concepts (e.g., Alzheimer's disease; dementia syndrome of depression).

Our network has the simplest possible architecture necessary to represent this simple knowledge structure and to use it to perform a sequential diagnostic classification task using a *hypothesize and test method*. The network contains four layers of units: an input layer; two intermediate layers; and an output layer.

Input layer. The input layer contains 80 input units that represent abstracted clinical features; we haven't attempted to account for the contextual interpretation of observed data into findings. The 58 *finding units* are grouped into 22 clinical attributes selected for their relevance to the diagnostic task. For instance, "age" or "memory" are attributes; "old -- over 70" or "severe memory impairment" are findings.

In addition, 22 other input units code for the confidence in the value attributed to each attribute; we will refer to them as the *confidence units*.

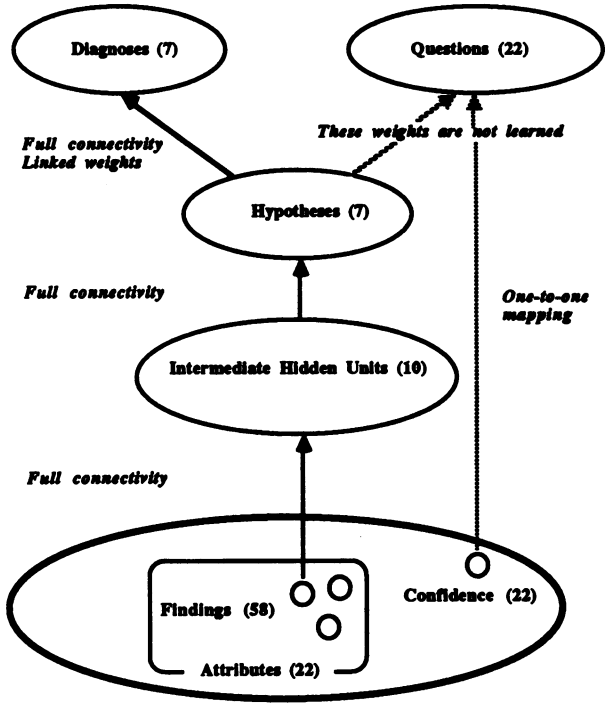

Figure 1: The structure and connection pattern of the network

Output layer. The output layer represents diagnostic classes and the request for further information. It contains two types of units: 7 *diagnosis units* represent the 7 "target" classes of disorders such as "dementia of Alzheimer's type" or "dementia syndrome of depression"; 22 *question units* represent requests for information about a specific patient's attribute as for instance the presence or absence of memory impairment and its degree of severity.

Intermediate layers. There are two layers of intermediate units. The lowest intermediate layer contains 10 units referred to as *hidden units*. On this layer, the findings of a case are recoded into a distributed pattern of activation, corresponding to the internal representation of the case; this internal representation is derived by the network during the training phase (see below). The highest intermediate layer contains 7 units referred to as *hypothesis units*. Each unit is constrained to represent a specific diagnostic hypothesis corresponding to a specific diagnostic class.

Activation at the input level. The user can set the activation of finding units to any value between 0 and 1; 0 means "no", 1 means "yes" and 0.5 means "unknown". Confidence units code whether an attribute is already known or not; their activation is either 0 (not yet known) or 1 (known).

Activation at the output level. The activation of diagnosis and question units can have any value between 0 and 1. For a diagnosis unit, an activation of 0 means "absent", 1 means "present", and 0.5 means "uncertain". At baseline, all finding and diagnosis units have an activation of 0.5 (this scheme purposely ignores baseline prior probabilities; when an individual patient is met, all diagnoses are equally likely to be considered, even though their probabilities of being confirmed

are different). Confirming a single diagnosis and ruling-out its competitors is represented by having a diagnosis unit with an activation close to 1 and the six others with an activation close to 0. Conversely, if several diagnoses are present concurrently, several units should have intermediate activation values. Similarly, the activations of question units are used to rank the questions, so that the question with the highest activation is asked (i.e., the value of the corresponding attribute is requested).

Connection pattern. The units of the network are linked ("connected") bottom-up and a weight is attached to each connection. The connection pattern is as follows: every finding is connected to all the hidden units; every hidden unit is connected to all the hypotheses; every hypothesis is connected to all the diagnoses. Finally, each confidence unit is connected to the corresponding question and each hypothesis unit is connected to all the question units. Multiple paths exist from one unit to higher level units. For instance, there are up to 10 * 7 * 7 = 490 different paths from a finding unit to a diagnosis unit.

In addition, some weights are *linked*: they are forced to take the same value. The 7 weights connecting each of the seven hypotheses to their corresponding diagnoses are linked; the remaining 42 weights connecting each of the 7 hypotheses to the 6 other diagnoses are also linked.

Learning rule and training procedure

We did not train the network to select questions. Consequently, the weights from the hypotheses to the questions were determined heuristically so that information confirming the most active hypothesis is requested. A weight from a confidence unit to its corresponding question unit is set up so that once a question has been asked and answered satisfactorily (activation of confidence unit is 1), the question cannot be asked again.

In contrast, all the other weights were learned using a training corpus of 75 cases. The purpose of this learning phase, is to discover a set of weights so that, upon presentation of all the findings of any training case, the network activates the correct diagnosis unit(s). To do so, we used the *generalized delta rule*, also called *backpropagation learning method* [28].

Training corpus: Prototypes and cases. We constructed 75 non-ambiguous *training cases* constituted by a set of findings and a single corresponding diagnosis. These 75 training cases are unevenly distributed among the possible disorders. For instance, there are 14 cases of DAT, but only six cases of NPH. It is noteworthy that even if we restrict the input activations to binary values (0 and 1), there is a total of 2^{58} (about 10^{17}) possible input patterns; with the further semantic restriction that only one finding per attribute can be true/active, we can still define $4^3 * 3^8 * 2^{11}$ (about 10^{10}) input patterns. However, only 75 of these possible input patterns are used during training.

We further reduced our initial training corpus by defining one to four *prototypes* per diagnosis. A prototype captures the regularities of a specific class of disorders across all its cases (or across sub-classes of its cases corresponding to specific presentations or stages). The activation of a finding in a prototype is grossly equivalent to the conditional probability *p(Finding/Disorder)* calculated over the training corpus. For instance, each of our 14 DAT (Alzheimer's) cases presents some degree of memory impairment. Conversely, two-thirds of the DAT cases are continent of urine and only one third at an advanced stage are incontinent. Therefore, in DAT prototypes, the finding "memory impairment absent" has a value of 0 and the findings "memory impairment mild" and "memory impairment severe" have a value of 0.3 and 0.7 respectively. Similarly the findings "incontinence absent" and "incontinence present" have a value of 0.6 and 0.4.

Finally, four "indeterminate" cases were used to create a baseline state in which all findings are unknown and all diagnoses are uncertain.

Weight initialization. Weights are assigned small random positive or negative values. We also used some initial constraints to help the network discover the desired mapping between hypotheses and diagnoses and create a strong competition between the diagnoses, while competing hypotheses can be activated concurrently.

Training procedure. The network is first trained with the 16 prototypes and then with the 75 training cases. The training is continued until all the cases have been learned according to a preset learning criterion. Roughly, this criterion imposes that if all the findings of a training case are presented to the network, the corresponding correct diagnosis has an activation above 0.90 and all the other diagnoses have an activation below 0.10. This training procedure requires only a few hours (usually, less than 6) on an IBM RT workstation with a math. coprocessor.

We have described how our network is structured and trained to select one diagnosis among seven when presented with a set of findings corresponding to a complete and unambiguous clinical case. Paraphrasing Hinton [29], we observe with satisfaction that the network is able to learn how to solve these training cases. However, this is not surprising considering previous successes in using the generalized delta rule to train other networks. The interesting questions are: Has this network discovered an internal representation that capture the regularities implicit in the training cases? How does the network behave? Can the network generalize sensibly to cases of related disorders?

Diagnostic behavior and performance

Behavior. Our network has been trained to associate diagnoses with complete descriptions of cases, yet it has to function in a normal diagnostic mode, guiding the incremental and sequential accumulation of data until a diagnosis is concluded.

Initially, when no information is known, all the findings are "unknown". The activation of all the findings are set accordingly at 0.5 and the network settles at its baseline: it produces an activation of 0.5 (uncertain) for all the diagnoses. The user volunteers a few findings by setting their activation to 1 (yes) or 0 (no). The network settles in a new state: some hypotheses become mildly activated, the activations of corresponding diagnoses increase while the activations of their competing diagnoses decrease. Hypotheses selectively activate some questions. The question with the highest activation is selected and asked. The status of the corresponding findings are provided by the user. The propagation of the new activations change the activations of the hypotheses, which in turn change the activations of the diagnoses and of the questions. The sequential nature of the task is modeled by these successive states of activation, determined by the incremental knowledge about the case.

When only a few findings are known, if a new finding "fits" the entertained hypotheses (i.e., support them), these hypotheses are strengthened; otherwise these hypotheses are weakened or even deactivated and alternative hypotheses are generated. A hypothesis can be generated in two ways: either directly by a cluster of findings supporting it, or indirectly by a cluster of findings supporting another related or competing hypothesis. This is because competing hypotheses share some intermediate units; in other words, the internal representations

of similar but competing diagnoses are similar.

A specific question is selected either because it is strongly associated with the most active hypothesis or because it is associated with several moderately active hypotheses. In turn, a change in the activations of the hypotheses can shift the focus of questioning. Usually, hypotheses are evaluated and reordered in a smooth fashion; the network requires several (three to four) discrepant findings to switch its focus from the domineering hypothesis to an alternative one. Conversely it will tend to ignore (wrongly or rightly) a single discrepant finding.

As more findings become known, the activations of the hypotheses become less and less prone to drastic revisions. A finding affects a hypothesis differently depending on the prior value of this hypothesis' activation: a finding supporting or inhibiting a hypothesis will affect its activation the most when the hypothesis is uncertain and the least when the hypothesis has a very low or a very high activation. In other words, a specific finding supports a diagnosis differently depending on the context created by other findings. Once a certain level of activation is reached, the accumulation of redundant findings tends to weaken a diagnosis instead of strengthening it. Finally, when one hypothesis dominates, its corresponding diagnosis dominates even more: the competition between diagnoses amplifies small differences of activation existing at the hypothesis level and tends to establish a "winner takes all" system.

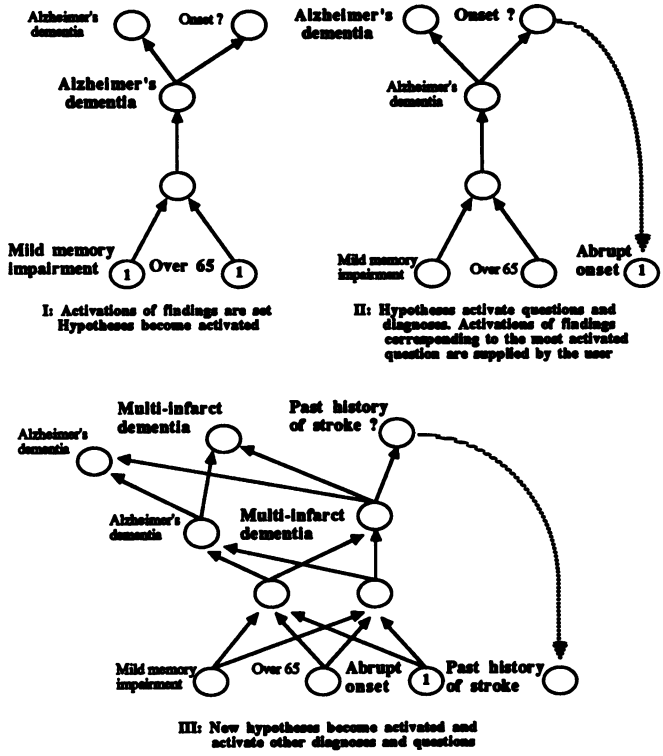

Figure 2: Sequential diagnosis

Performance. The goal of this study was not to develop a medical expert system, but to demonstrate the validity and usability of connectionist techniques to solve diagnostic problems. However, the level of performance may be an indication of the potential of similar networks to be used as inference engines in decision support systems.

By virtue of the training procedure, any complete case from the training corpus is diagnosed correctly. When the network is presented with only the critical features of such a case, this diagnostic accuracy is maintained: the network is able to generalize appropriately when a case is fairly typical. When the network is tested with new cases, its diagnostic accuracy depends on the similarity of the test case to the training cases. The network was tested with 18 cases seen during a 2 month period on a psychiatric ward; contrary to our training cases, these patients were often atypical and several presented with multiple concurrent disorders. Nevertheless, the network was able to generate an appropriate pattern of activation on the diagnosis units for 11 of these 18 cases; this corresponds to a correct classification rate of 61%.

Discussion

Network architecture and training procedure

To increase the validity of generalization and the accuracy of diagnoses made by such a network, we need to experiment with variations of the network's architecture and of the training procedure.

The validity of generalizations could be improved by implementing a more distributed representation of findings so that similar findings share more units. It would help the network to discover how various findings pertaining to the same clinical attribute are related. For instance, "absence of memory impairment" would be: (1, 0, 0, 0); "mild memory impairment": (0, 1, 1, 0) and "severe memory impairment": (0, 1, 0, 1), instead of the localized representation we used: (1, 0, 0); (0, 1, 0); (0, 0, 1). With this new distributed representation, the representation of "mild memory impairment" and "severe memory impairment" are sharing a unit together; if it is learned with some cases that "mild memory impairment" supports a diagnosis, "severe memory impairment" would do so to a certain extent; note however that the two findings have distinct representations and that, when it is appropriate, the variation in the degree of severity can be used to discriminate between two diagnoses.

The diagnostic accuracy can also be increased by the use of a larger training corpus representing a broader spectrum of the possible clinical presentations for a single disorder. Alternatively, incomplete cases conveying clinical heuristic rules could be used. These rules could prevent the network from either overgeneralizing or disregarding inappropriately discrepant findings.

During its training, this network was presented first with a few prototypical cases capturing the global structure of the domain, and then with a large number of distinct cases capturing the complexity of the domain. We need to study further how a network can first acquire the structure of a domain by being trained with a few "pre-digested descriptions" (this is similar to acquiring declarative knowledge by reading a textbook or by being told) and then, fine-tune its knowledge by being exposed to complex cases (this is similar to transforming declarative knowledge into procedural knowledge by doing).

Finally, we need to train the network so that it learns to select appropriate diagnostic questions. Once this is done, we need to reassess the adequacy of using complete cases to train the network while it has to perform sequential diagnosis and reason on incomplete cases.

Categorical and statistical reasoning

The reasoning of the network can be characterized as a blend of categorical and statistical reasoning [2]. Conceptually, some (unknown) pathophysiologic rules govern the way a given disorder manifests itself; the findings presented by patients suffering from this disorder are determined by these rules. By inspecting a set of cases we can induce some diagnostic rules that correspond to "the invert" of the deeper pathophysiologic rules. For instance, if a finding is present in

all cases of a disorder, its presence will be deemed necessary to make this diagnosis. Similarly some rules of sufficiency can be derived. No such clear-cut rules emerges from the network; nevertheless, it discovers structures implicit in the training cases. When presented with enough findings, it is able to select the right diagnosis. It does so not only on the basis of simple associations between findings and diagnoses, but also of conditional dependencies existing between findings and diagnoses. For instance, the network could learn that either "insomnia and weight loss" or "hypersomnia and weight gain" strongly support the diagnosis of depression, but that other combinations of sleep disorder and weight change don't.

This crucial ability to discover and use conditional dependencies requires intermediate units. In contrast, a two layer network associating findings and diagnoses would be equivalent to a simple Bayesian system in which all findings are conditionally independent. In such a network, biases correspond to log of prior odds; weights are equal to the log of likelihood ratios; and the graded activation is equivalent to posterior probabilities. Conversely, Pearl [30, 31, 32] has emphasized the correspondence between multi-layered Bayesian and connectionist networks: they both use parallel propagation of evidence, local updating mechanisms, and intermediate hidden variables. Pearl has proposed to use these networks as a computational architecture to model evidential reasoning; he suggested that tracing of their links could be responsible for the basic steps in querying and updating predictive and diagnostic beliefs. Obviously, far more research is needed to formally define the similarities and differences between multi-layered connectionist and Bayesian networks.

Using a connectionist network for decision support

In addition to these fundamental questions concerning the adequacy of representations, training procedures, and uncertainty management, several practical problems need to be addressed if we want to use such networks as the core of a decision support system.

The medical problem we have tackled is only moderately complex. It could probably be handled by a 200 rule expert system. We need to study larger networks handling more complex problems (e.g., concurrent multiple diagnoses). Large networks are more difficult to train but they may actually perform better than similar smaller networks: they can generalize better and what is inappropriate for a small network may be appropriate for a larger network. For instance, when only 7 diagnoses are possible, some hypotheses may seem to be erroneously generated given a cluster of findings; however, the generation of the same set of hypotheses may be adequate when the task is to discriminate among dozens of possible diagnoses.

Obviously, to be used in a decision support system, such a network would need to be embedded in an hybrid system: it would accomplish only part of the task (e.g., hypothesis generation or evaluation) while other necessary functions (e.g., data abstraction, explanation) would be provided using more conventional AI techniques. For instance, given findings and a concluded diagnostic formulation, explanations could be provided by a mechanism constructing rational justifications independently from the way at which the diagnosis was actually arrived.

Conclusion

We have implemented, trained, and tested a connectionist network that performs clinical diagnosis in the domain of dementia. Despite obvious limitations, the results of this experiment support the claim that connectionism provides powerful mechanisms that can solve problems previously intractable, including the ability to learn from experience and to generalize appropriately; this is congruent with the results of recently reported related studies [33, 34]. Far more work is necessary to demonstrate the practical usefulness of these techniques to develop clinical decision support systems. The limitations of this network suggest some directions for future research, in particular the study of various network architectures and training procedures, the comparison of connectionist and Bayesian networks, and the integration of connectionist networks in hybrid systems.

Acknowledgments

The first author (BHM) was partially supported by a fellowship from Western Psychiatric Institute and Clinic, Department of Psychiatry, University of Pittsburgh.

The connectionist network described in this paper was implemented using a software package developed by J. McClelland and D. Rumelhart and now commercially available [35].

We thank D. A. Evans and J. McClelland for their advice and comments; T. Barsalou and S. Luger for their thoughtful reviews.

Bibliography

1. Gorry GA (1973). Computer-assisted clinical decision making. Methods of Information in Medicine, 12:45-51.

2. Szolovits P & Pauker SG (1978). Categorical and probabilistic reasoning in medical diagnosis. Artificial Intelligence, 11:115-144.

3. Shortliffe EH, Buchanan BG, & Feigenbaum EA (1979). Knowledge engineering for medical decision making: A review of computer-based clinical decision aids. Proceedings of the IEEE, 67:1207-1224.

4. Buchanan BG & Shortliffe EH (1984). Rule-based expert systems. The MYCIN experiments of the Stanford Heuristic Programming Project. Reading, MA: Addison-Wesley.

5. Miller RA, Pople HE, & Myers JD (1982). INTERNIST-I, an experimental computer based diagnostic consultant for general internal medicine. New England Journal of Medicine, 307: 468-476.

6. Pauker SG, Gorry GA, Kassirer JP, & Schwartz WB (1976). Towards the simulation of clinical cognition: Taking a present illness by computer. American Journal of Medicine, 60: 981-996.

7. Pople HE (1982). Heuristic methods for imposing structure on ill-structured problems: The structuring of medical diagnosis. In: Szolovits P (ed). Artificial Intelligence in Medicine. Boulder, CO: Westview Press.

8. Shortliffe EH (1976). Computer-based medical consulations: MYCIN. New York, NY: Elsevier.

9. Weiss SM, Kulikowski CA, Amarel S, & Safir A (1978). A model-based method for computer-aided medical decision-making. Artificial Intelligence, 11:145-172

10. Schwartz WB, Patil RS, & Szolovits P (1987). Artificial intelligence in medicine: Where do we stand? New England Journal of Medicine, 316:685-688.

11. Patil RS, Szolovits P, & Schwartz WB (1981). Causal understanding of patient illness in medical diagnosis.

Proceedings of the 7th IJCAI. Los Altos, CA: William Kaufmann.

12. Clancey WJ (1983). The epistemology of a rule-based expert system: A framework for explanation. Artificial Intelligence, 20: 215-251.

13. McClelland JL & Rumelhart DE (1986). Parallel distributed processing. Explorations in the microstructure of cognition. Volume 2: Psychological and biological models. Cambridge, MA: MIT Press.

14. Clancey WJ (1984). Classification problem solving. Proceedings of AAAI-84. Los Altos, CA: William Kaufmann.

15. Rumelhart DE, Smolensky P, McClelland JL, & Hinton GE (1986). Schemata and sequential thought processes in PDP models. In: McClelland JL & Rumelhart DE (eds). Parallel distributed processing. Explorations in the microstructure of cognition. Cambridge, MA: MIT Press.

16. Anderson JA, Silverstein JW, Ritz SA, & Jones RS (1977). Distinctive features, categorical perception, and probability learning: Some applications of a neural model. Psychology Review, 84: 413-451.

17. Feldman JA & Ballard DH (1982). Connectionist models and their properties. Cognitive Science, 6:205-254.

18. Hopefield JJ & Tank DW (1986). Computing with neural circuits: a model. Science, 233: 625-633.

19. Rumelhart DE & McClelland JL (1986). Parallel distributed processing. Explorations in the microstructure of cognition. Volume 1: Foundations. Cambridge, MA: MIT Press.

20. Henrion M & Cooley DR (1987). An experimental comparison of knowledge engineering for expert systems and for decision analysis. Proceedings of AAAI-87. Los Altos, CA: William Kaufmann.

21. Elstein AS, Shulman LS, & Sprafka SA (1978). Medical problem solving. An analysis of clinical reasoning. Cambridge, MA: Harvard University Press.

22. Cummings JL & Benson DF (1983). Dementia: A clinical approach. Boston, MA: Butterworth.

23. Mulsant BH & Servan-Schreiber D (1984). Knowledge engineering: A daily activity on a hospital ward. Computers and Biomedical Research, 17:71-91.

24. Werner G (1987). Methuselah -- An expert system for diagnosis in geriatric psychiatry. Computers and Biomedical Research, 20:447-488.

25. Mulsant BH & Thornton JE (in press). The outpatient management of Alzheimer disease and other dementias. In Thase M & Edlestein B, & Hersen M (eds). Handbook of outpatient treatment of adults. New York, NY: Plenum Press.

26. APA: American Psychiatric Association (1987). DSM-III-R: Diagnostic and statistical manual of mental disorders, 3rd edition, revised. Washington, DC: APA.

27. Ericsson KA & Simon HA (1984). Protocol analysis. Verbal reports as data. Cambridge, MA: MIT Press.

28. Rumelhardt DE, Hinton GE, & Williams RJ. (1986). Learning internal representations by error propagation. In: Rumelhart DE & McClelland JL (eds). Parallel distributed processing. Explorations in the microstructure of cognition. Cambridge, MA: MIT Press.

29. Hinton GE (1986). Learning distributed representations of concepts. Proceedings of the 8th Annual Conference of the Cognitive Science Society.

30. Pearl J (1985). Bayesian network: a model of self-activated memory for evidential reasoning. Proceedings of the 7th Annual Conference of the Cognitive Science Society.

31. Pearl J (1986). Fusion, propagation, and structuring in belief networks. Artificial Intelligence, 29:241-288.

32. Pearl J (1987). The logic of representing dependencies by directed graphs. Proceedings of AAAI-87. Los Altos, CA: William Kaufmann.

33. Gallant SI (1988). Connectionist expert systems. Communications of the ACM, 31:152-169.

34. Peng Y & Reggia JA (1987). A connectionist model for diagnostic problem solving. Technical Report TR-143. Computer Science Department, University of Maryland.

35. McClelland JL & Rumelhart DE (1988). Handbook of parallel distributed processing. Cambridge, MA: MIT Press.