

EVINCE : a neuropsychiatric desktop expert system for the diagnosis of dementia

Citation for published version (APA):

Plugge, L. A. (1992). *EVINCE : a neuropsychiatric desktop expert system for the diagnosis of dementia*. Rijksuniversiteit Limburg.

Document status and date:

Published: 01/01/1992

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

SUMMARY

This thesis describes the development of the neuropsychiatric expert system Evince for the differential diagnosis of dementia, and the subsequent experiments in which the expert system was compared with clinicians from several medical disciplines who are involved in dementia diagnostics. An expert system is a computer program that is equally proficient in solving problems, in a specific domain, as a human expert in the same field.

Although many medical expert systems have been developed only a few have been subjected to an extensive formal evaluation. Even less expert systems have been developed for the psychiatric domain, and none for the neuropsychiatric field of dementia diagnostics, because this domain is generally considered too difficult to implement in an expert system, due to the use of descriptive criteria, and a multitude of data with very different origins. Furthermore, many of the expert systems that were developed were meant to be used on a mainframe, mini computer, or workstation, which makes such systems less attractive for many small or medium large institution.

In this thesis it is shown that it is possible to develop an expert system for an IBM-PC compatible desk-top microcomputer on the domain of the neuropsychiatric dementia diagnostics, that produces a performance equal to a multidisciplinary team of expert clinicians and better than the average clinician.

Chapter I gives a short introduction in the history of artificial intelligence and expert system research. Furthermore, an introduction is given into the previous efforts to develop expert system within the domain of psychiatry, and the lessons that can be learned from those earlier attempts. Chapter II identifies the lack of thorough expert system evaluations studies as one of the reasons why expert systems have been so slow in penetrating daily practice. Additionally some of the most recent ideas about expert system evaluation are discussed. An adapted three stage model from Wyatt and Spiegelhalter^{*} for expert system evaluation is proposed, and some of the problems and possibilities of this model are discussed. In chapter III the expert system development tool Acquaint that was used for the development of Evince is reviewed. Chapter IV continues with a description of the development method used, i.e., the incremental top-down development of Evince. Furthermore, in this chapter the sources of knowledge, the knowledge acquisition, and the architecture of Evince are discussed. The following chapters contain a description of the experiments performed to assess the performance of Evince.

Chapter V is a report about the first experiment in which Evince was tested in diagnosing 19 patients with varying stages of dementia and 10 patients showing other disorders except dementia. It is shown that EVINCE-I and the human expert are in perfect agreement on the diagnosis dementia and correlate highly on the diagnoses dementia of the Alzheimer type and multiple infarct dementia. The results showed that the approach followed for the development of Evince was successful. Therefore,

^{*} Wyatt J, Spiegelhalter D. Evaluating medical expert systems: what to test and how? *Medical Informatics*. 1990, 15; 3: 205-217.

Evince was developed further on the basis of these results. This revised and expanded version of Evince was to be tested by comparing it with clinicians from several disciplines from outside the institution where Evince was developed.

In order to gain more insight in the performance of the average clinicians involved in dementia diagnostics, an experiment described in chapter VI was set up in which 90 clinicians from 6 disciplines diagnosed 10 case descriptions of patients, judged by a multidisciplinary expert committee to suffer from dementia. Five cases were diagnosed before and 5 after a consensus meeting on the diagnosis of dementia. A significant change in the level of agreement between the disciplines could not be established. The analysis did show a significant difference between the disciplines in the use of etiological diagnoses. The results indicated that, in order to avoid possible bias caused by medical specialization, a multidisciplinary approach for this type of patients is recommended.

Finally, in chapter VII an experiment is described in which Evince is compared with 85 clinicians in diagnosing 10 patients suspected of suffering from dementia. A multi-disciplinary expert committee provided a standard diagnosis as reference for comparison. The results show that the syndrome and etiological diagnoses made by Evince were in very close agreement with those of the expert committee and that the diagnostic performance of Evince was better than that of the average clinician.

In the final chapter (VIII) a general discussion of the findings is presented, and the limitations and possible implications of the use of Evince for dementia diagnostics in practice are discussed. Furthermore some topics for future research are addressed.

SAMENVATTING

In dit proefschrift wordt de ontwikkeling beschreven van een neuropsychiatrisch expertsysteem (Evince) voor de differentiële diagnostiek van dementie, en de experimenten waarin het expertsysteem werd vergeleken met klinici uit verschillende medische disciplines die betrokken zijn bij dementiediagnostiek. Een expertsysteem is een computerprogramma dat, bij het oplossen van problemen in bepaald domein, prestaties levert die vergelijkbaar zijn met die van een menselijke expert op dat terrein. Hoewel in de loop der jaren veel medische expertsystemen zijn ontwikkeld, werden slechts enkele daarvan onderworpen aan een grondige formele evaluatie. Nog minder expertsystemen werden ontwikkeld voor de psychiatrie en geen op het gebied van de neuropsychiatrische dementiediagnostiek, omdat dit terrein te gecompliceerd geacht werd voor implementatie in een computer programma, onder andere door het gebruik van descriptieve criteria en de grote verscheidenheid van de vereiste gegevens. Verder waren veel van de ontwikkelde expertsystemen bedoeld voor gebruik op zogenaamde 'mainframes', mini-computers of werkstations, waardoor het gebruik van dergelijke programma's minder interessant is voor kleine(re) instellingen.

In dit proefschrift wordt aangetoond dat het mogelijk is een expertsysteem voor een IBM-compatibele microcomputer te ontwikkelen op het gebied van de neuropsychiatrische dementiediagnostiek, dat resultaten levert die vergelijkbaar zijn met die van een multidisciplinair team van experts, en beter dan de prestaties van de gemiddelde clinicus.

In hoofdstuk I wordt een korte introductie gegeven in de geschiedenis van het wetenschappelijk onderzoek van kunstmatige intelligentie en expertsystemen. Verder krijgt de lezer een kort overzicht van eerdere pogingen om expertsystemen op het gebied van de neuropsychiatrie te ontwikkelen en de lessen die daaruit geleerd kunnen worden.

In Hoofdstuk II wordt ingegaan op het gebrek aan grondige expertsysteemevaluaties en wijst het gebrek aan dergelijke evaluaties aan als één van de oorzaken waarom expertsystemen zo langzaam in de dagelijkse praktijk doordringen. Bovendien worden enkele van de meest recente vooraanstaande ideeën over expertsysteemevaluatie besproken. Er wordt een aangepast model van Wyatt en Spiegelhalter* met drie stadia geïntroduceerd, waarna de mogelijkheden en problemen van dat model worden besproken.

Hoofdstuk III beschrijft het expertsysteem-ontwikkelingsgereedschap Acquaint dat gebruikt is voor het bouwen van Evince. Hoofdstuk IV gaat verder met een beschrijving van de gebruikte ontwikkelingsmethode, i.e. de incrementele en van het algemene in de bijzonderheden afdalende methode. Bovendien worden in dit hoofdstuk de kennisacquisitie, de kennisbronnen en de architectuur van Evince besproken. De daaropvolgende hoofdstukken bevatten een beschrijving van de experimenten die werden uitgevoerd om de prestatie van Evince vast te stellen.

* Wyatt J, Spiegelhalter D. Evaluating medical expert systems: what to test and how? *Medical Informatics*. 1990, 15; 3: 205-217.

Hoofdstuk V bevat de rapportage over het eerste experiment waarin Evince vergeleken werd met een menselijke expert in het diagnostiseren van 19 patiënten in verschillende dementiestadia en 10 patiënten met andere stoornissen, maar geen dementie. Aangevoerd wordt dat Evince en de menselijke expert volledige overeenstemming vertonen over de diagnose dementie en een hoge mate van overeenkomst over de diagnose dementie van het type Alzheimer en multiple infarct dementie. Deze resultaten laten zien dat de gevolgde ontwikkelingsmethode succesvol was. Op basis van deze resultaten werd Evince verder ontwikkeld.

Deze aangepaste en uitgebreide versie van Evince werd getoetst in een vergelijking met klinici uit verschillende disciplines die afkomstig waren van buiten het instituut waar Evince werd ontwikkeld. Om meer inzicht te krijgen in de prestaties van de gemiddelde clinicus die te maken heeft met dementiediagnostiek, werd een experiment opgezet dat beschreven wordt in hoofdstuk VI, waarin 90 klinici uit 6 disciplines ieder 10 patiënten diagnostiseerden aan de hand van een casusbeschrijving. Vijf van de casusbeschrijvingen werden vóór en vijf na een consensusbijeenkomst over dementiediagnostiek beoordeeld. Op basis van de beschikbare gegevens kon geen belangrijke verandering in het niveau van overeenstemming tussen de disciplines worden vastgesteld als gevolg van de consensusbijeenkomst. De analyses lieten echter wel een significant verschil zien tussen de disciplines in het gebruik van etiologische diagnoses. De resultaten gaven aan dat een multidisciplinaire aanpak bij dementiediagnostiek is aan te bevelen, om te voorkomen dat de uiteindelijke diagnose gekleurd wordt door een medische specialisatie.

Tenslotte wordt in hoofdstuk VII een experiment beschreven waarin Evince vergeleken wordt met 85 klinici in het diagnostiseren van 10 patiënten waarvan vermoed werd dat ze leden aan een vorm van dementie. De diagnoses van een multidisciplinair forum bestaande uit drie experts op het gebied van dementiediagnostiek fungeerden als referentie voor de vergelijking. De resultaten van die vergelijking lieten zien dat de syndroom en etiologische diagnoses van Evince nauw overeen kwamen met die van het forum en beter waren dan die van de gemiddelde clinicus.

Het laatste hoofdstuk (VIII) bevat een algemene discussie over de resultaten, beperkingen van Evince en de mogelijke implicaties van het gebruik van Evince bij dementiediagnostiek, en enkele onderwerpen voor toekomstig onderzoek.