

# De ontwikkeling van een 'intelligent' computergestuurd onderwijssysteem

---

D. J. BIERMAN en P. A. KAMSTEEG  
*Universiteit van Amsterdam, Vakgroep  
Psychonomie, Amsterdam*

---

## Samenvatting

*Het onderzoeksproject 'Een computercoach voor de thermodynamica' had tot doel het ontwikkelen van een geautomatiseerd onderwijssysteem, gebaseerd op een bestaand computermodel van probleemoplossen in de thermodynamica. Het resulterende prototype bestaat uit vier componenten: een domeinkennis module, een diagnostische module, een onderwijskennis module en een module voor (gedeeltelijk grafische) interactie met de leerling*

*Tijdens het project opgedoken problemen, gedeeltelijk te wijten aan moeilijkheden bij het achterhalen van specifieke didactische kennis, leidden tot de ontwikkeling van een kennis-elicitering methode (MUSPA) voor een interactief computersysteem, gebruikmakend van het systeem-in-ontwikkeling zelf. Een leerling werkt hierin met het incomplete systeem, terwijl een menselijke leraar aan een andere computerterminal die taken overneemt die het computersysteem (nog) niet kan uitvoeren. Zowel het hardop denken van leerling en leraar als hun (ingetypte) interactie wordt geregistreerd. De methode bestaat uit drie fasen: ontwikkeling van het interactie gedeelte, ontwikkeling van een minimaal prototype, en uitbouw en verfijning van het prototype.*

## 1 Inleiding

In het onderzoeksproject *Kennisvererving in Formele Domeinen* wordt onderzocht op welke wijze problemen worden opgelost, bijvoorbeeld in de natuurkunde. De onderzoeksmethode is gebaseerd op analyse van de uitspraken van proefpersonen als hun gevraagd wordt om 'hardop te denken' bij het

oplossen van een probleem. Deze uitspraken, vastgelegd in een zogenaamd hardopdenk-protocol, zouden de onderliggende mentale processen redelijk ongestoord reflecteren (Breuker, Elshout, Van Someren & Wielinga, 1986). De analyse van dergelijke protocollen moet leiden tot een model betreffende het probleemoplosgedrag in het betreffende domein (in ons project, het domein van de thermodynamica), liefst een zodanig flexibel model dat het aanpasbaar is voor verschillende typen probleemoplosgedrag zoals dat van een expert en dat van een beginner.

In de praktijk komt dit neer op een cyclus van analyse en modelbouw, waarbij de protocollen worden geanalyseerd tegen de achtergrond van het op dat moment beste model en waarbij de verschillen aanleiding zijn voor verdere verbetering van het model. De modellen worden geïmplementeerd als een computerprogramma m.b.v. methoden ontleend aan het onderzoeksgebied van de Artificiële Intelligentie (AI). Uiteraard moet een computerprogramma dat een goed model is van een geavanceerde probleemoplosser, zelf de problemen ook kunnen oplossen. En daarbij moet het 'denkspoor' van het computerprogramma overeenkomen met het denkspoor van de menselijke probleemoplosser voor zover dat uit het hardopdenk-protocol valt op te maken. Evenzo moet een goed model van een beginnende probleemoplosser dezelfde fouten vertonen als de menselijke beginner.

Hoewel het niet expliciet de bedoeling is van deze methodologie, opent ze de mogelijkheid om op zinvolle wijze onderzoek te doen naar één enkele proefpersoon. Aangezien het tot op heden onmogelijk is om een kwantitatieve evaluatie van een eenmaal geproduceerd model te geven, lijkt er maar één soort kritische toets uitvoerbaar te zijn. Dat is om inderdaad het model op grond van een groot aantal protocollen van één proefpersoon te construeren, en dan zowel aan de proefpersoon als aan het model een nieuw probleem voor te leggen. Dit probleem zouden ze dan beide op precies dezelfde wijze moeten oplossen. In de praktijk

worden de modellen meestal ontwikkeld aan de hand van gegevens ontleend aan een beperkt aantal proefpersonen. Het model wint daardoor aan generaliseerbaarheid maar verliest aan gedetailleerdheid en toetsbaarheid.

In het begin van de jaren '80 beschikten wij in het project *Kennisverwerving in Formele Domeinen* over een werkend model voor een geavanceerde probleemoplosser<sup>1</sup> op het gebied van de thermodynamica<sup>2</sup>. Een voorbeeld van de problemen die het model kan oplossen is de volgende typische thermodynamica-opgave:

*Een ideaal gas bevindt zich in een vat dat is afgesloten met een zuiger. Het volume van het gas is 2 liter en de druk 120 kPa. Door de zuiger langzaam naar buiten te bewegen wordt het volume vergroot tot 3 liter, terwijl de temperatuur van het gas constant gehouden wordt. Wat is de druk van het gas na de beweging van de zuiger?*

Naast onderzoek op het gebied van het probleemoplossen zelf, is er binnen het project aandacht voor onderzoek naar het *leren* van het oplossen van problemen. Recentelijk werd bijv. een model geconstrueerd voor een beginnende probleemoplosser (Jansweijer, Elshout & Wielinga, 1986). De ontwikkeling van dit model naar het model van een geavanceerde probleemoplosser zou op zich weer een model voor het leren van deze vaardigheid zijn. Ook hier is validatie echter een moeilijk punt.

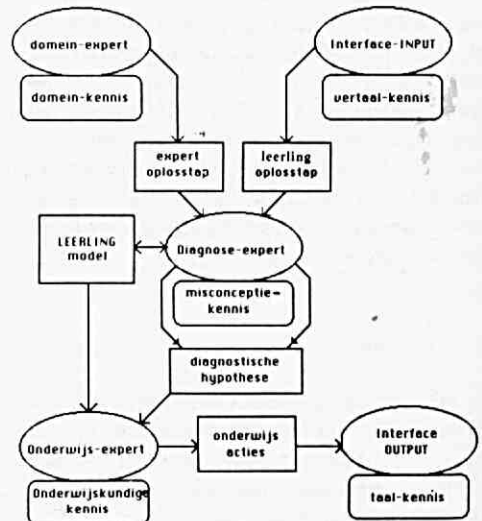
Inzicht in het aanleren van probleemoplosvaardigheden kan men ook verkrijgen door verschillende onderwijsvormen te hanteren. Men kan, op basis van een model van probleemoplossen en het leren daarvan, onderwijsacties ontwerpen die op dit model aansluiten. Als deze onderwijsacties succes hebben bij werkelijke leerlingen kan voorzichtig geconcludeerd worden dat het model correct zou kunnen zijn. Anderson (1987) meent dat het gebruik van kennisgebaseerde computergestuurde onderwijssystemen juist in dit opzicht (dus als toetsinstrument voor cognitieve leermodellen) veelbelovend is. Het recentelijk afgesloten SVO-project *Een computercoach voor de thermodynamica* had tot doel een dergelijk systeem, waarbij verschillende onderwijsvormen kunnen worden uitgetoetst, te ontwikkelen. Onderwijsacties in dit systeem zouden gebaseerd moeten zijn op een

expliciet, dynamisch model van de leerling (een beginnende dan wel een gevorderde).

Het hele systeem kan men zich voorstellen als een model van een ideale<sup>3</sup> individuele docent. Aangezien een docent zelf een geavanceerde probleemoplosser is, zal een deel van de computercoach deze functie moeten kunnen uitvoeren. Omdat we ten tijde van de start van het project beschikten over een computerprogramma dat althans deze functie uitvoerde (Jansweijer, Konst, Elshout & Wielinga, 1982) leek de hoop gerechtvaardigd dat de implementatie van overige delen van het systeem niet tot onoverkomelijke moeilijkheden zou leiden. Deze aanname bleek niet terecht. Teneinde de tijdens het project opgedoken moeilijkheden op te lossen werd een nieuwe methode van kenniseliciteratie ontwikkeld welke uiteindelijk een belangrijk product van het hele project geacht mag worden te zijn.

## 2 De computercoach voor de thermodynamica

De functionele structuur (de architectuur) van de computercoach, voor zover het de begeleiding van het maken van één enkele opgave betreft<sup>4</sup>, is geschetst in Figuur 1. Het systeem bestaat uit vier componenten welke ieder te onderscheiden functies uitvoeren.



Figuur 1 Architectuur van de computercoach

## 2.1 De domein-expert

Dit is de oorspronkelijke 'geavanceerde probleemoplosser'. Als invoer accepteert deze component een probleem, lost dit vervolgens op en produceert daarbij een 'denkspoor'. Dit spoor wordt ook wel 'norm'-spoor genoemd, vanwege de achterliggende gedachte dat de leerling uiteindelijk aan deze denk-norm<sup>5</sup> dient te voldoen. Het is van meet af aan de bedoeling geweest om alle kennis die betrekking heeft op het domein van de thermodynamica binnen deze component te localiseren. De overweging hierbij was dat eenzelfde type computercoach voor een ander domein dan te construeren zou zijn door enkel en alleen de domein-expert te vervangen.

## 2.2 De diagnose-expert

De functie van de diagnose-expert is het genereren van hypothesen omtrent de reden van afwijkend denk-gedrag van de leerling. Daarnaast is een belangrijke taak het bijhouden van een model van de leerling. Dit kan gebeuren door van elk kenniselement (feitelijke kennis, procedurele kennis dan wel kennis omtrent oplosstrategieën) dat de domein-expert tot zijn beschikking staat, te schatten in hoeverre het door de leerling beheerst wordt. De kennis van de leerling wordt in dit opzicht gezien als een deelverzameling van de uiteindelijke kennis die hij moet gaan bezitten. Het model van de leerling is dan ook een onvolledige maar verder getrouwe afspiegeling van het expert-model. Zo'n soort leerling-model wordt 'overlay model' genoemd, omdat het (op de lacunes na) het expert-model als het ware precies bedekt.

In domeinen, zoals de natuurkunde, waarvan mag worden aangenomen dat de leerling al een bepaalde (mogelijk incorrecte) voorkennis bezit is het daarnaast nodig bij te houden welke verkeerde kennis (misvattingen, verkeerde mentale modellen) er bij de leerling aanwezig is. Aangezien het hier om kennis gaat die een expert juist *niet* heeft, past deze niet in een 'overlay model', maar moet apart gerepresenteerd worden, bijvoorbeeld in een misvattingenlijst ('bug catalog') of als verwijzende uitbreidingen van het het 'overlay model' ('extended overlay model').

De eerder genoemde functie van het genereren van hypothesen m.b.t. afwijkend leerling-gedrag geschiedt geheel in de context van een dergelijk leerling-model. Dat wil zeggen dat de

waarschijnlijkheid of onwaarschijnlijkheid van bepaalde hypothesen bepaald wordt middels informatie in het leerling-model.

## 2.3 De onderwijs-expert

De taak van de onderwijs-expert is om op grond van een diagnose tot een optimale onderwijs-interventie te komen. Ook hierbij speelt het leerling-model een belangrijke rol. Naast de cognitieve aspecten van het leerling-model kunnen ook persoonlijkheidskenmerken van de leerling betrokken worden in de beslissing ten aanzien van optimale onderwijs-acties. Zo kan het persoonlijkheidskenmerk 'negatieve faalangst' een rol spelen in de beslissing veel of weinig procedurele hints te geven.

## 2.4 De interface

Een (user) interface is dat deel van een computerprogramma dat zorg draagt voor de interactie tussen de gebruiker en het programma. In ons geval betekent dat, dat de handelingen van de leerling aan de terminal in een vorm vertaald moeten worden welke bruikbaar is voor de diagnose-expert. Dat wil zeggen, motorische acties dienen geïnterpreteerd te worden als reflecties van cognitieve acties. Omgekeerd moeten onderwijs-interventies van de onderwijs-expert in de vorm van één of meer begrijpelijke zinnen aan de leerling aangeboden worden.

## 3 De resultaten van het computercoach project

In de eerste plaats heeft het project een prototype van een zogenaamd 'intelligent' computergestuurd onderwijssysteem opgeleverd. De intelligentie van het systeem is geconcretiseerd in een tweetal aspecten. Het systeem is in staat het gedrag van de leerling te begrijpen. Dat wil zeggen het kan het leerling-gedrag duiden in termen van eigen handelingen<sup>6</sup>. Daarnaast inferereert het systeem bij voortdurende nieuwe kennis, met name met betrekking tot een cognitief inhoudelijk leerling-model, op basis van overeenkomsten en verschillen tussen de eigen handelingen en die van de leerling.

In de tweede plaats, en dit is o.i. belangrijker, zijn door uitvoering van het project een groot aantal problemen aan het licht gekomen. Voor sommige problemen, met name die

van technologische aard, is een oplossing gevonden. Voor de problemen van inhoudelijke aard is een methodologie ontworpen waarvan verwacht mag worden dat deze de problemen uiteindelijk oplost.

### 3.1 *Problemen van technologische aard*

Het betreft hier problemen die o.a. te maken hebben met de vertaling van het motorisch gedrag van de leerling aan de terminal naar een cognitief gedrag. Dit probleem is opgelost door enerzijds de leerling te verbieden enig papier bij de terminal te gebruiken, en anderzijds de interface uit te breiden met een interpreterbaar 'electronisch kladblok'. Dit is een gedeelte van het beeldscherm waarop de leerling tekst (gegevens etc.) kan intypen, formules uit een bestand kan kopiëren, en tekeningen uit voorgevormde onderdelen kan samenstellen (Bierman & Anjewierden, 1986; Bierman & Kamsteeg, 1986).

Daarnaast bleek dat de structuur van het bij aanvang van het project beschikbare domein-expert programma zich op bepaalde punten slecht leende voor een soepele inpassing in het totale systeem. Met name is niet altijd inspecteerbaar welke kennis precies in een stap gebruikt wordt (Kamsteeg & Bierman, 1987).

### 3.2 *Problemen van inhoudelijke aard*

Bij de constructie van kennisgebaseerde systemen kan men drie soorten kennis onderscheiden. Algemene wereldkennis ('common sense knowledge') blijkt bij de oplossing van problemen op bijna elk terrein een rol te spelen. Met name waar het de vertaling van het probleem zoals dat zich aan ons voordoet naar een voor het domein specifieke representatie betreft. Een voorbeeld is de kennis van de kenmerken van dingen als een fietspomp of een ballon, waar in een opgavetext sprake van kan zijn.

Globale domeinkennis is kennis die gemeengoed en veelal algoritmisch van aard is, zoals definities van en relaties tussen concepten, vaste handelingen etc. Door middel van interviews met experts of direct uit tekstboeken kan men deze kennis eliciteren. Specialistische (ervarings- of intuïtieve) kennis is meer van heuristische aard en wordt zelden in geschreven vorm aangetroffen; het blijkt meestal ook niet mogelijk om deze door post-hoc interviews te verkrijgen. Een veelgebruikte techniek om deze laatstgenoemde kennis te achterhalen is om de experts problemen hard-

op denkend te laten uitvoeren. In het kader van de ontwikkeling van expert systemen is voor de elicitering van kennis zelfs een speciale methodologie ontwikkeld (Wielinga & Breuker, 1984).

Zo wordt ook de expertise m.b.t. individueel onderwijs geven (tutoring) nergens volledig beschreven. Zelfs globale kennis hierover is moeilijk te vinden (Kamsteeg, 1984). Onderwijstheorieën worden meestal in algemene termen geformuleerd en nauwelijks in termen van leraar-gedrag geoperationaliseerd (Knoers, 1973). Bovendien gaan de meeste theorieën impliciet uit van een klassikale onderwijs situatie; vanuit een individueel gezichtspunt overbelichten ze daardoor organisatorische en sturende taken, en onderbelichten ze persoonlijke diagnose en interactie. Kenmerkend is de uitspraak van Ausubel, Novak & Hanesian (1978, blz. 50): '(...) the teacher's most important and distinctive role in the classroom is still that of director of learning activities'. Voor probleemoplossen bestaat eigenlijk zelfs, op een enkele uitzondering na (Mettes & Pilot, 1980; Mettes & Roossink, 1982), geen onderwijstheorie hoewel daar wel behoefte aan is (Reif, 1980; Shuell, 1980).

Een extra moeilijkheid op dit gebied is, dat de specialistische kennis van een leraar niet direct volgens de methode van analyse van hardopdenk-protocollen is op te sporen, aangezien het hardop denken van de docent in hoge mate interfereert met het individuele onderwijs dat hij tegelijkertijd verzorgt.

## 4 *De MUSPA methode voor kenniselicitering*

De MUSPA (Multiple Source Protocol Analysis) methode is een combinatie van twee methoden. Bekend in de informatie-industrie is de methode waarbij al in een vroeg stadium van de implementatie van een informatie-systeem (meestal reeds in de specificatie fase) de toekomstige gebruiker met een 'pseudo'-systeem mag werken. Bij dit pseudo-systeem is de software vervangen door iemand aan een andere terminal die het 'nog te ontwikkelen' systeem nabootst. Deze techniek staat bekend onder de naam 'Wizard of Oz'-techniek: net als in het gelijknamige sprookje wordt de gebruiker in de waan gebracht dat hij met een

niet-menselijk systeem (in het sprookje de tovenaar, in ons geval het computersysteem) interacteert, terwijl hij dat in feite toch met een mens doet. Bij de MUSPA-methode wordt deze techniek gecombineerd met het hardopdenken van degene die het systeem nabootst (in ons geval een expert-docent) én de toekomstige gebruiker (leerling). Hierdoor ontstaan drie protocollen, namelijk de twee hardopdenk-protocollen én het interactie-protocol, die tot één protocol worden gecombineerd (vandaar de aanduiding 'multiple source'). Bestudering van deze protocollen verschaft inzicht in de door de docent gebruikte diagnostische en onderwijskundige kennis en daarnaast, als een extra bonus, inzicht over de filterwerking van de interface<sup>7</sup> welke aanleiding kan zijn tot misverstanden bij zowel docent als leerling.

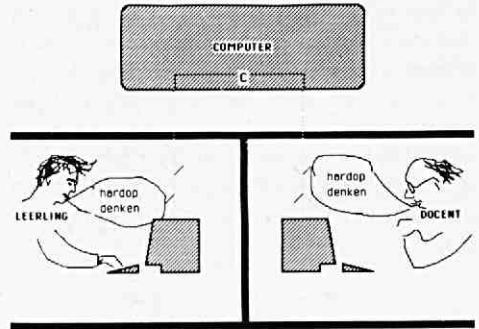
Een wezenlijk aspect van de MUSPA-methode is dat deze in een aantal fasen wordt uitgevoerd. In elke volgende fase wordt de geëliciteerde kennis in het systeem geïmplementeerd waardoor de onderwijstaak langzaam maar zeker door de computercoach wordt overgenomen. De computercoach wordt aldus gebruikt om zichzelf te verbeteren. Het is alsof het zichzelf aan zijn eigen haren uit het moeras omhoog trekt. Daarbij verschuift de taak van de menselijke docent in dit hybride systeem langzaam van het produceren van pure hardopdenk-protocollen naar het produceren van on-line (dat is: tijdens de uitvoering) kritiek op het functioneren van de reeds geïmplementeerde gedeelten van de computercoach.

#### 4.1 MUSPA – fase 1: De computercoach bestaat nog geheel niet

In dit stadium hebben de docent en de leerling met elkaar contact via alfanumerieke terminals (dat zijn terminals die alleen tekst kunnen weergeven). Zie figuur 2. De leerling heeft een aantal schriftelijke aanwijzingen en een opgave-boek naast de terminal liggen.

Een typisch stuk interactieprotocol is in Figuur 3 weergegeven. Hierin is het getal aan het begin van een regel het tijdstip (in minuten) van de uiting. 'T': staan voor een uiting van resp. de leraar (Teacher) en de leerling (Pupil).

Merk op dat de docent enige tijd nodig heeft om tot de keuze van een probleem te komen (de eerste mededeling is pas na 3½ minuut). Uit



Figuur 2 MUSPA – fase 1, experimentele opstelling

het hardopdenk-protocol van de docent blijkt dat dit wordt veroorzaakt door het feit dat hij niets omtrent de voorgeschiedenis van de betreffende leerling weet. Zijn overwegingen om tot opgave 5 te beslissen reflecteren bepaalde onderwijskundige kennis die in volgende fasen in de ontwikkeling van de computercoach geïmplementeerd kan worden.

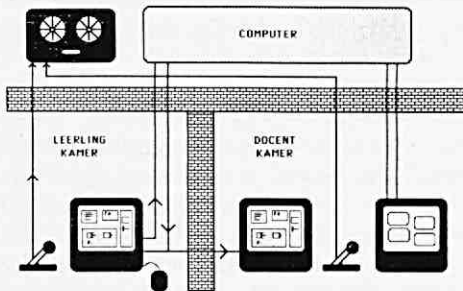
TIJD	BRON	INTERACTIE
03.30	T:	OK GA MAAR BEGINNEN MET OPGAVE 5
05.00	T:	KUN JE ME AF EN TOE LATEN ZIEN WAT JE AAN HET DOEN BENT OF WANNEER JE HULP NODIG HEBT?
06.00	P:	KUN JE ME DE FORMULE VAN DE MOL GEVEN?
07.30	T:	JE HEBT OP PAPIER ENKELE HINTS GEKREGEN. HEB JE AL OPGESCHREVEN WAT HET SYSTEEM IS?
09.00	P:	WAT VOOR SYSTEEM BEDOEL JE?
10.00	T:	KIJK NOG EENS GOED NAAR DE HINTS EN VOER HINT 1 EN 2 UIT. ALS JE KLAAR BENT ZEG DAT DAN EVEN.
12.30	P:	IK BEN ERMEE KLAAR.
12.31	T:	WAT VOOR DING HEB JE GETEKEND?

Figuur 3 MUSPA – fase 1, protocol-fragment (alleen interactie-deel vermeld). 1e kolom: tijd sinds begin van de sessie; 2e kolom: T = leraar, P = leerling

Uit protocollen die we in fase 1 afnamen kwam bovendien steeds naar voren dat de docent te weinig informatie had om de leerling te kunnen begeleiden bij het oplossen van het probleem<sup>8</sup>. Met name het gebrek aan inzicht in de (aan)tekeningen was een belangrijke frustratie. Op grond hiervan werd een aantal interface-criteria geformuleerd waarvan de belangrijkste was dat de leerling zijn (aan)tekeningen niet meer naast maar op de terminal moest kunnen uitvoeren. Met de implementatie van deze interface belandden we in fase 2.

## 4.2 MUSPA – fase 2: De interface is geïmplementeerd

In deze fase bezit de computercoach nog geen onderwijskundige kennis maar de interface is gereed, zodat een onderwijsessie voor de leerling eruit ziet zoals dat in de uiteindelijke coach het geval is. De docent heeft nu twee terminals voor zich. Op de ene kan hij de grafische manipulaties van de leerling volgen, op de andere kan hij met de leerling in contact treden.



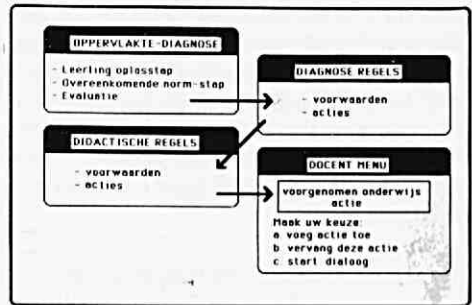
Figuur 4 MUSPA – fase 2, experimentele opstelling

Figuur 4 geeft een schematische weergave van deze situatie. Het scherm van de (grafische) leerling-terminal bevat linksboven een menu waarmee de leerling kan kiezen uit verschillende handelingstypen (tekst intikken, tekening maken, formules zien, etc.). Middenboven is een venster te zien waarin, afhankelijk van de menukeuze, diverse hulpmiddelen kunnen verschijnen (opgavetekst, rekenmachine, formule-venster, constanten-venster). Linksonder is de werkruimte voor de leerling (het eigenlijke 'kladblok'). Rechts daarvan bevinden zich de onderdelen (vaten, zuigers, kranen etc.), die de leerling kan gebruiken om tekeningen te maken.

Een typisch stuk interactieprotocol uit deze fase staat in Figuur 5. De manier van interventie vertoont een zekere consistentie. Er worden, op basis van het norm-spoor, steeds specifiekere hints gegeven. Eerst: *Zou je niet eerst even een analyse maken?* en later meer specifiek: *Ik bedoel ook een tekening*. Door implementatie van deze kennis ontstond een volledig<sup>9</sup> geïmplementeerde computercoach, zij het dat we ons bewust waren dat de kennis die aan diagnose en onderwijs-interventies (hintgenerator) ten grondslag lag buitengewoon summier van aard was. Hiermee belanden we in fase 3.

## 4.3 MUSPA – fase 3: de computercoach staat op wankel(e) eigen benen

In deze fase verandert de opdracht aan de menselijke docent. Hij moet nu vooral hardop denkende handelingen van de computercoach bekritisieren en als daarvoor aanleiding is de voorgenomen acties van de computercoach 'overrulen'. Teneinde de coach te kunnen kritisieren wordt het interne denkproces van de computercoach op de docent-terminal getoond. In Figuur 6 is dit schematisch weergegeven.



Figuur 6 MUSPA – fase 3, schematische weergave van het docent-terminal scherm

TIJD	BRON	INTERACTIE
02.09	P:	HIERTOE GEBRUIK IK DE EERSTE HOOFDWET
02.38	T:	ZOU JE NIET EERST EVEN EEN ANALYSE VAN HET PROBLEEM MAKEN?
03.33	P:	DE DRUK VERANDERT DE TEMPERAATUUR IS KONSTANT HET VOLUME VERANDERT
03.54	T:	IK BEDOEL OOK EEN TEKENING

Figuur 5 MUSPA – fase 2, protocol-fragment (alleen interactie-deel vermeld). 1e kolom: tijd sinds begin van de sessie; 2e kolom: T = leraar, P = leerling

Er zijn drie vensters waarop de werking van de drie componenten (interface, diagnose, onderwijs-expert) te zien is. In het vierde venster is zichtbaar met welk gedeelte van het norm-spoor de diagnose-expert het leerling-spoor probeert te vergelijken. De denkstappen van de computercoach komen overeen met het activeren van regels van het type: ALS (CONDITIE) DAN (ACTIE). Zowel de condities als de acties worden op het docent-scherm uitgeschreven.

Om het schema van Figuur 6 te verduidelij-

ken volgt hier een concreet voorbeeld. De situatie waaraan dit refereert is die van het protocol in Figuur 7, vlak vóór de regel op 17.41. De leerling lost de in de Inleiding van dit artikel gegeven opgave op, heeft daartoe een correcte tekening gemaakt, en is net klaar met het intikken van de tekst 'V1 is 2 liter'. In het 'oppervlakte-diagnose' venster geeft het systeem nu een weergave van de leerling-oplosstap (*type*: gegeven (probleemschets); *variabele*: V,1; *waarde*: 2,liter), van de overeenkomstige norm-oplosstap (*type*: gegeven (probleemschets); *variabele*: V,1; *waarde*: 2,liter), en een evaluatie (*getalwaarde*: ok; *eenheid* ok; *ontbrekende stappen*: probleemkenmerk (niet-adiabatisch<sup>10</sup>), probleemkenmerk (ideaalgas)). Het 'diagnose-regels' venster vertoont vervolgens de getriggerde diagnose regel, in dit geval 'ALS een leerling stappen overslaat EN de stap die de leerling wel doet valt niet buiten een gegeven procedurele aanwijzing, DAN is er sprake van een eenvoudige procedurele fout'. Het 'didactische regels' venster geeft het vervolg hierop, namelijk de didactische regel 'ALS een eenvoudige procedurele fout gemaakt is EN geen inhoudelijke fout gemaakt is, DAN geef een procedurele aanwijzing op één niveau specifiek dan de vorige aanwijzing'. Aangezien de vorige aanwijzing zeer globaal was (los de opgave op) is de nieuw gevonden aanwijzing ook nog vrij globaal: 'OK, maar je hebt het vraagstuk nog niet af. Je moet de probleemschets nog compleet maken'. Deze voorgenomen onderwijsactie wordt afgedrukt in het 'docent menu' venster.

Als de onderwijs-expert tot een actie heeft besloten, stopt het systeem en krijgt de docent de mogelijkheid om deze actie aan te vullen, te annuleren of te vervangen door een eigen actie (met behulp van menu-keuzen in het 'docent menu'). Het interactie-protocol in fase 3 kent daardoor nu drie deelnemers: de leerling, de menselijke docent en de computerdocent. Een voorbeeld is gegeven in Figuur 7, hierin komen naast uitingen van T (leraar) en P (leerling) nu ook uitingen van S (systeem) voor. Ter verduidelijking zijn aan het protocol verwijzigen toegevoegd naar de oplosprocedures die de leerling op een bepaald moment uitvoert (tussen vierkante haken).

Merk op dat de menselijke docent het noodzakelijk vindt om een positieve reinforcement te geven ('Goed ga maar verder'), en daarom

TIJD	BRON	INTERACTIE
00.19	S:	HALLO...HIER IS JE (COMPUTER/MENS)DOCENT. WE BEGINNEN!
01.51	S:	OPGAVE WORDT GEZOCHT. MOMENT GEDULD...
02.08	S:	HIER KOMT DE OPGAVE. GA JE GANG, EN VEEL SUKSES.
11.55	T:	[leerling maakt tekening met triviaal foutje] GOED GA MAAR VERDER <<SLUITEN>> [V1 is 2 liter]
17.41	S:	OK, MAAR JE HEBT HET VRAAGSTUK NOG NIET AF. JE MOET DE PROBLEEMSCHETS NOG COMPLEET MAKEN. [P1 is 120 kPa]
21.22	S:	OK, MAAR JE HEBT DE PROBLEEMSCHETS NOG NIET AF. JE MOET DE PROBLEEMKENMERKEN NOG DOEN. [T1 is T2]
26.09	T:	GA DOOR <<SLUITEN>> [gesloten systeem]
27.32	S:	JA, GA MAAR DOOR [V2 is 3 liter]
29.44	T:	GA DOOR <<SLUITEN>> [druk neemt af]
31.44	S:	JA, GA MAAR DOOR [leerling kiest formule: wet van Boyle]
35.29	T:	GELDT DEZE FORMULE VOOR ALLE SOORTEN GASSEN? <OVER>
37.13	P:	NEE, ALLEEN VOOR IDEALE GASSEN, DUS HIER GELDT IE WEL <OVER>
38.44	T:	KIJK NAAR DE OPGAVE EN VUL DE PROBLEEMSCHETS MET DIT SOORT KENMERKEN AAN <<SLUITEN>>

Figuur 7 MUSPA – fase 3, protocol-fragment (alleen interactie-deel vermeld). 1e kolom: tijd sinds begin van de sessie; 2e kolom: T = leraar, P = leerling, S = systeem; tussen [ ]: oplos-actie van de leerling (toegevoegd aan protocol)

de systeem-actie (die gebaseerd was op een triviaal foutje in de tekening) verwerpt. Ook blijkt dat de menselijke docent een (ingeklede) hint geeft omtrent wat er nu bedoeld wordt met de term 'probleemkenmerken'. Dit gebeurt niet meer op grond van het norm-spoor en betekent dat er extra kennis nodig is om dergelijke ingrepen te kunnen implementeren.

## 5 Toekomstverwachtingen en conclusies

Een kennisgebaseerd CAI-systeem is een omvangrijk en complex programma, en dat des te meer naarmate het onderwijsdomein omvangrijker en complexer (semantisch rijker) is. Het oplossen van natuurkundige problemen is, zowel qua subject als qua cognitieve vaardigheid, een semantisch rijk domein. Een goede computercoach voor dit domein, die ook nog een aanvaardbare interactiesnelheid heeft, lijkt de eerstkomende jaren nog niet realiseerbaar. Zelfs het nu bestaande allesbehalve complete prototype systeem kan alleen nog op grote, dure computers uitgevoerd worden. Maar de technische ontwikkelingen in compu-

terapparatuur gaan in een duizelingwekkend tempo en het is zeker niet denkbeeldig dat er over een jaar of tien goedkope micro-computers zijn met voldoende geheugencapaciteit en rekensnelheid voor dit soort programma's.

De twee volgende punten zijn meer fundamentele redenen waarom het huidige computercoach-systeem niet in de praktijk van het onderwijs bruikbaar is. Ten eerste is het systeem bedoeld voor gebruik in het onderzoek; de opbouw is daar ook naar, deze is bijv. meer gericht op flexibiliteit dan efficiëntie. Ten tweede, en van meer belang, weten we nog lang niet genoeg van (al dan niet correcte) cognitieve structuren en functies bij probleemoplossen in dit domein (leerling-model), van hoe deze af te leiden uit uitingen (diagnose), en welke feedback in welke situatie welk effect kan hebben bij welke leerlingen (didactische kennis). Om computergebaseerde onderwijs-systemen te maken, moeten we eerst weten hoe we moeten onderwijzen. En daarvoor is empirisch onderzoek nodig naar het handelen en denken van ervaren onderwijzers (onderwijs-experts) tijdens het geven van individueel onderwijs (Ohlsson, 1986).

De MUSPA-methodologie blijkt een waardevol instrument, niet alleen voor dit soort onderwijsresearch, maar meer in het algemeen voor de elicitering van kennis, vooral waar dialogen in het geding zijn. Ten eerste wordt de dialoog niet gestoord door het hardopdenken. Ten tweede veroorzaakt het steeds grotere aandeel van het programma een steeds beter geformaliseerde weergave van de gebruikssituatie, en daardoor een steeds beter *interpretatiemodel* (Wielinga & Breuker, 1984) met behulp waarvan nieuwe hardopdenk-protocollen zijn te duiden. Parallel daarmee wordt, ten derde, het steeds verder ontwikkelde model iedere keer weer bekritiseerd in de feitelijke gebruikscontext, en op grond daarvan verwijfd.

Deze procedure van steeds verdere verwijfing en toepassing van een model op grond van hardopdenk-protocollen vindt men ook bij het modelleren van andere cognitieve vaardigheden, en in het bijzonder van probleemoplossen (waartoe wij ook onderwijzen rekenen). De werkwijze is een equivalent van de zogenaamde 'empirische cyclus' (De Groot, 1970) zoals deze in de empirische wetenschappen wordt toegepast.

## Noten

1. Veeleer een brave beginner die precies volgens de regels de problemen oplost.
2. Thermodynamica moet in dit verband met een korreltje zout worden genomen. Het betreft een subset van dit domein, betrekking hebbende op de allereerste beginselen van de warmteleer.
3. Ideaal, in de zin van niets vergetend en consistent handelend.
4. Bovenop deze architectuur huist nog een component die de opgave selecteert of bepaalt dat de leerling uitgeleerd is.
5. Het idee dat er één norm zou zijn is natuurlijk onzin. Het 'denkspoor' van de computer bevat bijv. sequenties die omwisselbaar zijn. Meer in het algemeen gesteld: een normspoor geeft slechts één handelingsverloop, terwijl er meerdere correcte alternatieven kunnen zijn.
6. Echte intelligentie zou zich moeten uiten in de mogelijkheid om het gedrag van de leerling te REPRODUCEREN. Bijvoorbeeld door een deel van de eigen kennis te verminken. Dit leidt echter snel tot combinatorische explosies en is in het huidige project niet nagestreefd.
7. Aangezien de communicatie tussen de leerling en de docent via de computer verloopt, zijn bepaalde communicatie-methoden (zoals gezichtsuitdrukkingen) uitgesloten. De interface werkt in dat opzicht als een filter dat slechts een deel van de communicatie doorlaat.
8. Door het gebrek aan informatie werd de docent gedwongen de leerling uit te vragen. Deze vragen werkten soms onbedoeld als hint.
9. De diagnose kan men zich denken te bestaan uit twee lagen. Eerst een oppervlakkige diagnose die aangeeft welke kenniselementen wellicht niet of foutief bij de leerling aanwezig zijn. Daarna een diepere die deze diagnose probeert te duiden als een uiting van een bepaald mentaal model. Hoewel we uit parallel verloopend onderzoek een goed overzicht hadden gekregen over bestaande mentale modellen in dit domein (Stehouwer, Van Looy, Bierman & Van Someren, 1986) hebben we deze diepte-diagnose niet geïmplementeerd.
10. Niet-adiabatisch betekent: er is warmte-uitwisseling met de omgeving.

## Literatuur

- Anderson, J.R., Methodologies for studying human knowledge. *Behavioral and Brain Sciences* (nog te verschijnen). Abstract in: UUCP newsgroup mod.ai, Message-ID: <8701071550.AA01452...>, 1987.
- Ausubel, D.P., J.D. Novak & H. Hanesian, *Educational psychology, a cognitive view*. New York: Holt, Rinehart & Winston, 1978 (2nd ed).



- Bierman, D. J. & A. A. Anjewierden, *The use of a graphic scratchpad for students in ICAI*. New Orleans: Proceedings of the 27th ADCIS Conference (pp. 68-71), 1986.
- Bierman, D. J. & P. A. Kamsteeg, Het monitoren van het Gewenst Handelingsverloop met behulp van een 'intelligente' computercoach. In P. Verhagen & B. J. Wielinga, (Eds.), *Media in het onderwijs*. Lisse: Swets & en Zeitlinger, 1986.
- Breuker, J. A., J. J. Elshout, M. W. van Someren & B. J. Wielinga, Hardopdenken en protokolanalyse. *Tijdschrift voor Onderwijsresearch*, 1986, 11/5, 241-254.
- Jansweijer, W. N. H., A. J. J. Elshout & B. J. Wielinga, *The expertise of novice problem solvers*. Brighton: Proceedings of the European Conference on AI (pp. 576-585), 1986.
- Jansweijer, W. N. H., L. Konst, J. J. Elshout & B. J. Wielinga, PDP: *A protocol diagnostic program for solving problems in physics*. Paris: Proceedings of the European Conference on AI (pp. 278-280), 1982.
- Groot, A. D. de, *Methodologie*. Den Haag: Mouton, 1970.
- Kamsteeg, P. A., *Kennis van docenten bij individuele coaching*. Scriptie psychologische funktieleer (25.6.84.421). Amsterdam: Universiteit van Amsterdam, 1984.
- Kamsteeg, P. A. & D. J. Bierman, *Constraints on an expert system for use in ICAI*. Paper submitted for 3rd International Conference on AI and Education. Amsterdam: Universiteit van Amsterdam, 1987.
- Knoers, A. M. P., Instructiemethoden. In J. A. van Kemenade (Ed.), *Bijdragen uit de onderwijswetenschappen*. Alphen aan den Rijn: Samsom, 1973.
- Mettes, C. T. C. W. & A. Pilot, *Over het leren oplossen van natuurwetenschappelijke problemen*. OC-rapport 42. Enschede: Technische Hogeschool Twente, 1980.
- Mettes, C. T. C. W. & H. J. Roossink, *Teruggkoppelen bij het maken van vraagstukken*. OC-rapport 48. Enschede: Technische Hogeschool Twente, 1982.
- Ohlsson, S., Some principles of intelligent tutoring. *Instructional Science*, 1986, 14/3-4, 293-326.
- Reif, F., Theoretical and educational concerns with problem solving: bridging the gap with human cognitive engineering. In: D. T. Tuma & F. Reif (Ed.), *Problem solving and education, issues in teaching and research*. Hillsdale: Erlbaum, 1980.
- Shuell, T. J., Learning theory, instructional theory and education. In: R. E. Snow, P. A. Federico & W. E. Montague (Ed.), *Aptitude, learning and instruction, volume II: cognitive process analyses of learning and problem solving*. Hillsdale: Erlbaum, 1980.
- Stehouwer, M., J. van Looy, D. J. Bierman & M. W. van Someren, *Wat is er toch zo moeilijk aan thermodynamica*. Paper gepresenteerd op NIP-congres, Groningen, 1986.
- Wielinga, B. J. & J. A. Breuker, *Interpretation of verbal data for knowledge acquisition*. Pisa: Proceedings of the European Conference on AI (pp. 41-50), 1984.

### Curricula vitae

D. J. Bierman (1943) studeerde experimentele natuurkunde aan de U.v.A. en promoveerde hierin in 1972. Sinds 1972 werkzaam op het Psychologisch Laboratorium van de U.v.A. Doet onderzoek op het gebied van o.a. artificiële intelligentie en onderwijs.

P. A. Kamsteeg (1954) studeerde psychologie met specialisatie artificiële intelligentie, aan de U.v.A. Werkt sinds 1984 binnen de Vakgroep Psychonomie van de U.v.A. aan door de S.V.O. gesubsidieerde onderzoeksprojecten op het gebied van (intelligent) computerondersteund onderwijs.

Adres: Psychologisch Laboratorium U.v.A., Vakgroep Psychonomie, Weesperplein 8, 1018 XA Amsterdam

Manuscript aanvaard 11-6-'87

### Summary

Bierman, D. J. & P. A. Kamsteeg. 'A method for knowledge elicitation in individual coaching.' *Pedagogische Studiën*, 1987, 64, 335-343.

The aim of the research project 'A computercoach for thermodynamics' was to build a tutoring system around an existing computer model of problem solving in thermodynamics. The resulting prototype tutoring system consists of four modules: a domain expert, a diagnostic expert, a tutoring expert and a graphic interface.

Problems that rose during the project, partly due to the difficulty of eliciting specific didactic knowledge, led to the development of a method (MUSPA) of knowledge elicitation for use in an interactive computer system, using the developing system itself. A student works with the incomplete system while a human coach simulates the tasks the system cannot (yet) perform at another terminal. Both student and human coach think aloud and also their written interaction is recorded. The method consists of three stages: construction of the interface, construction of a minimal prototype, and refinement and extension of the prototype.