

AT MODELLERE TAVS VIDEN

Henning Boje Andersen

Den gængse opdeling mellem automatiske og kontrollerede processer gennemgås kort, idet det fremhæves, at flere moderne generelle modeller af menneskers kognitive præstationer stemmer overens med hensyn til den dobbelte arkitektur bag sondringen. Et bredt begreb om tavs viden fremstilles i relation til bestemmelser af eksplicit og implicit viden. Inden for dette brede begreb skelnes mellem (i) baggrundsviden og kontekstviden, samt mellem (ii) ikke-repræsentationel viden (perceptuelle og sensorisk-motoriske færdigheder) og den viden, der er indeholdt i den globale sammenhæng mellem vidensstrukturer. Artiklen slutter med nogle korte bemærkninger om mulighederne for at modellere tavs viden.

Indledning

Forskellige forfattere har givet forskellige karakteriseringer og bestemmelser af tavs viden (»tacit knowledge«). De klassiske bidrag kommer fra videnskabsfilosofi, dvs. fra Michael Polanyi (1958, 1966), Thomas Kuhn (1962, 1970), Norwood Russell Hanson (1958); fra lingvisten Noam Chomsky (1965, 1980a, 1980b, 1986); og fra filosofi - især Gilbert Ryle (1949), Ludwig Wittgenstein (1953, 1980) og Martin Heidegger (1927 og senere udgaver).

Det vil føre alt for vidt at inkludere en historisk oversigt over de begreber, som disse førende forfattere har diskuteret i forbindelse med tavs viden. Men for dog at give en antydning af bredden, vil jeg kort nævne et par af de illustrative eksempler, der har været givet på tavs viden.

Polyanyi bruger dels eksempler, der involverer mønstergenkendelse - fx. genkendelse af ansigter eller af eksemplarer af en kategori, og dels nævner han viden, der ligger bag menneskers udførelse af komplekse men tilsyneladende simple handlinger. Han beskriver fx. meget træffende, hvor lidt folk ved om, hvad de faktisk gør for at holde balancen på en cykel (1958, p. 49f). Kuhn beskriver i sin berømte og kolossalt indflydelsesrige karakteristik af normal eller paradigmatisk videnskab, hvordan videnskabsfolks indbyrdes forståelse hviler på udsagte og typisk uerkendte antagelser og normer, og han fremhæver, hvordan videnskabelige paradigmer afhænger af enighed om mønstergældende eksemplarer: »Shared exemplars serve cognitive functions commonly attributed to shared rules«. Chomsky bruger et begreb om tavs viden til at sammenfatte den ubevidste viden, som mennesker har om deres eget sprog: »we postulate unconscious {læs: tacit} knowledge of the

rules of the grammar if this postulation is empirically justified by the role it plays in explaining the facts of use and understanding and acquisition of language.« (se Hook, 1969, p. 155). Filosoffernes bidrag har været meget forskellige, og de trodser en kort opsummering (men se fx. den rolle de filosofiske refleksioner over tavs viden har i forbindelse med design og anvendelse af AI- og informationssystemer i Winograd & Flores, 1987; Dreyfus, 1979 og Dreyfus & Dreyfus, 1986).

I de følgende afsnit ridser jeg nogle forskellige karakteriseringer af tavs viden op, og jeg afslutter med et par principielle bemærkninger om modellering, herunder af tavs viden.

Delte arkitekturer

Menneskers kognitive udfoldelser afhænger af forskellige typer processer i hjernen/bevidstheden. Kognitive aktiviteter spænder vidt. Jens Rasmussen har foreslået en hensigtsmæssig og nu meget indflydelsesrig tredeling af kognitive aktiviteter. I denne model understreges i højere grad end i de fleste andre den centrale rolle som tavs viden spiller, herunder specielt de sensoriske og motoriske færdigheder - menneskers flydende, ubesværede og dynamiske kontrol af omgivelser og redskaber. (Se fx. Rasmussen, 1986, 1987; Vincente & Rasmussen, 1988 samt referencer i disse til flere arbejder; i Goodstein et al. 1988 findes introducerende oversigter). Vi finder her kognitive aktiviteter udspændt over et spektrum løbende fra »skill-based«, »rule-based« til »knowledge-based« aktiviteter; dvs. fra (a) mønstergenkendelse og kropslige færdigheder, over (b) rutinemæssig udførelse af stereotype opga-

Tavle 1:

Contrasting the properties of attentional and automatic control

Attentional control	Automatic control
Serial	Parallel
Slow	Rapid
Laborious	Effortless
Ressource-limited	No apparent limitations
Analytic	Intuitive
Computationally powerful	Operates by simple heuristics
Processes accessible to consciousness	Processes beyond the reach of awareness; only products available to consciousness
Essential for coping with novelty, but cannot be sustained for any length of time	Capable of handling routines and recurrences; but often ineffective in the face of change

(gengivet efter Reason, 1988)

ver, eksekvering af procedurer og udnyttelse af skematiske vidensstrukturer, til (c) opmærksomhedsbelastende, reflekteret indsats over for nye, krævende situationer, hvor stereotype løsninger er utilstrækkelige.

Rasmussens model har til fælles med en række generelle fremstillinger af kognitive præstationer - hvad J. Reason kalder »framework theories« (Reason, 1988), at den indeholder en principiel sondring mellem to kvalitative forskellige slags processer: De kontrollerede processer, der er sekventielle og normalt er bevidst tilgængelige, og de underbevidste, parallelle og analoge processer (jf. Rasmussen 1986, p. 76ff).

Denne grundlæggende sondring mellem *automatiske* og *kontrollerede* processer er der udbredt (men ikke universel) enighed om indenfor kognitionsforskning - omend det må tilføjes, at denne enighed naturligt nok er omvendt proportional med, hvor præcist de karakteristiske forskelle bag sondringen trækkes op.

De *automatiske* processer ligger bag menneskers omfattende repertoire af mønstergenkendelse og af ubesværede, letflydende færdighedsaktiviteter - hvilket ofte involverer dynamisk og kontinuert regulering og koordination af krop og omgivelser. Disse processer antages ofte at være implementeret i hjernen som massivt parallelle associative mekanismer. De automatiske processer er - i modsætning til deres produkt: fx. en genkendelsesakt, en handling eller bevægelse - ikke tilgængelige for aktørens bevidsthed, og de er tilsvarende ikke tilgængelige for direkte viljesafhængig kontrol. (Men de kan selvfølgelig styres indirekte derved, at en aktør sætter sig et nyt bevidst mål - fx. »der er vist is på vejen: nedsæt farten, brems ikke hårdt«, hvorved den automatiske kontrol af instrument (bil) får nye referenceværdier - jf. nedenfor). I dag er der intense bestræbelser på at modellere disse processer inden for den meget synlige og kraftigt ekspanderende forskningsretning, der kaldes »konnektionisme« (neurale netværk, parallelt distribueret procesring: pdp - se note 7), og processerne betegnes i den sammenhæng ofte som subsymbolske (jvf. fx. Smolensky 1988a, b).

Menneskers (og højerestående dyrs) evne til at koordinere sensorisk information og bevægelse - glidende, anticiperende og hyppigt i fase med dynamiske og hurtige forløb i omverdenen - kan muligvis slet ikke modelleres på kendte måder ved hjælp af associative net, men skal snarere forklares i termer af analoge modeller. (Se hertil Rasmussens beskrivelse af hvad han kalder den dynamiske verdensmodel, Rasmussen 1986, p. 89ff., samt en tidlig formulering i Rasmussen 1974. Se desuden Kohonen 1988 om antydninger af analoge repræsentationer i neurale net).

De *kontrollerede* processer ligger bag den for det meste bevidste aktivitet i den begrænsede arbejdshukommelse. De er typisk (måske uvægerligt) sekventielle, relativt langsomme og de er, hvad der er helt centralt, underlagt strikte ressourcebegrænsninger. Det kan diskuteres, om det er hensigtsmæs-

Tavle 2:

Some characteristics of Automatic and Control processes

Characteristic	Automatic processes	Control processes
Central capacity	Not required	Required
Control	Not complete	Complete
Indivisability	Wholistic	Fragmentized
Practice	Results in gradual improvement	Has little effect
Modification	Difficult	Easy
Serial-parallel dependence	Parallel Independent	Serial Dependent
Storage in LTM	Little or none	Large amounts
Performance level	High	Low, except when task is simple
Simplicity	Irrelevant	Irrelevant
Awareness	Low	High
Attention	Not required but may be called	Required
Effort	Little if any	Much

(gengivet efter Schneider, Dumais & Shiffrin, 1984).

sigt at stipulere, at de kontrollerede processer udføres langsomt og med besvær; for mennesker udfører jo et væld af rutinehandlinger, der, skønt de er bevidste, ikke kræver andet end momentvis monitorering fra opmærksomheden - et punkt hvis fremhævelse netop udmærker Rasmussens model. Desuden er de underlagt viljens herredømme (dvs. man kan eksempelvis vælge, at man nu vil gå fra andet til fjerde gear og ikke som ellers først til tredje; men man kan ikke vælge, at nu vil man ikke genkende sin nabo, og nu vil man ikke forstå, hvad der siges på modersmålet osv.). Se tavle 1 og 2 for to moderne opsummeringer af karakteristiske træk ved henholdsvis automatiske og kontrollerede processer.

Teorier eller modeller over kognitive præstationer, der indeholder denne principielle sondring mellem de to kvalitativt forskellige slags processer, kaldes undertiden dobbelt-arkitektur modeller (se fx. Reason, 1988). Med dette understreges, at opdelingen ikke er blot funktionel, men afspejler den kognitive arkitektur (som det udtrykkes med en term, der oprindeligt stammer fra datalogi - den bruges i lidt forskellige betydninger - se fx. Newell, 1982; Pylyshyn, 1984).

Dele af en taksonomi

Daniel Dennetts diskussion i artiklen »Styles of mental representations« (Dennett 1982/83) er et godt udgangspunkt for at etablere nogle nødvendige

og forhåbentligt nyttige sondringer indenfor de ret forskellige begreber om tavs viden, der findes. Dennetts hovedmål er at bestemme tavs viden over for eksplicit viden, og at skitsere hvordan eksplicit viden kan siges at afhænge af eksistensen af tavs viden. Undervejs defineres samtidig et gængs og simpelt begreb om, hvad han kalder implicit repræsenteret information. Dog skal det bemærkes, at Dennett strengt taget ikke taler om eksplicit eller implicit *viden*, men om forskellige måder hvorpå information kan være repræsenteret eller indeholdt i et system, herunder i menneskets hjerne/bevidsthed. Men om de følgende distinktioner formuleres i termer af viden eller i termer af information er næppe væsentligt, og de to udtryk her vil blive brugt i flæng.¹

Her er Dennetts forslag til en definition af eksplicit information:

Let us say that information is *explicitly* represented in a system if and only if there actually exists in the functionally relevant place in the system a physically structured object, a *formula* or *string* or *tokening* of some members of a system (or »language») of elements for which there is a semantic or interpretation, and a provision (a mechanism of some sort) for reading or parsing the formula. This definition of explicit information is exigent, but still leaves room for a wide variety of representation systems. They need not be linear, sequential, sentence-like systems, but might, for instance, be »map-reading systems« or »diagram interpreters«. (citeret fra Dennett, 1986, p. 216).

Det kan diskuteres, om det er hensigtsmæssigt at definere, sådan som Dennett gør, et repræsentationssystem (»language») så bredt, at det omfatter både (i) verbal information, som jo nødvendigvis er diskret, som har en rekursiv struktur, og som udviser en indholdsuaafhængig inferensstruktur; og (ii) repræsentationsformer som fx. er billed-lignende (ikoniske), eller analoge eller som ligger bag vore mentale modeller, dvs. de dynamiske og rumlige forestillinger der indgår i vores evne til at udvikle mentale scenarier over både fysiske og abstrakte forhold.

Imidlertid er det som nævnt først og fremmest Dennetts ærinde at definere eksplicit over for tavs viden (information) og at skitsere deres vekselvirkning - han sigter altså ikke efter at skelne mellem typer af eksplicit viden. Hvis vi vælger at udjævne forskellen mellem på den ene side sproglig, diskret information og på den anden side de forskellige former for billedlig information (analog, kontinuert, dynamisk osv.) vi er i besiddelse af, kan vi dog stadig bestemme en naturlig (men altså som nævnt uensartet) klasse af mentale tilstande og processer. Thi vi har at gøre med repræsentationer, der (i en eller anden fælles forstand, som man skylder senere at gøre præcis) har konstituentstruktur: Forestillingsbilleder, scenarier og dynamiske mentale modeller er systematisk sammensatte - ligesom sprog er det, men de er i modsætning til sprog næppe rekursive, og de udviser ingen inferensstruktur.

Hvis man tilslutter sig den grove skitse bag dobbelt-arkitektur modellen, er de eksplicitte repræsentationer at finde i arbejdshukommelsen og i den deklarative langtidshukommelse («paratviden»), og de er normalt, men ikke uvægerligt, tilgængelige for subjektets bevidsthed.²

Eftersom det meste af det, en person ved, er noget han slutter sig til ud fra anden viden han har, må vi ud over explicit viden (information) også indføre et begreb om implicit viden: dvs. viden eller information som potentielt kan genereres givet en passende lejlighed og givet normale inferentielle evner. F.eks. tror eller ved læseren sikkert, at tigre kan løbe hurtigere end pingviner, eller at der ikke er lavet maskiner, der kan sortere vittige vitser fra ikke-vittige. Men da disse tanker næppe har strejft læseren før, er disse antagelser i denne forstand implicite, indtil det øjeblik de produceres - fx. i en samtales forløb. Tilsvarende indfører Dennett et simpelt og nærliggende begreb om implicit viden eller information: »... for information to be represented *implicitly*, we shall mean that it is *implied* logically by something that is stored explicitly«. (ibid.).

Lad os først se på en mindre detalje ved dette begreb om implicit viden eller information. Implicit viden er alt, hvad der følger som konsekvenser af det man ved eksplicit, hedder det. Men der findes jo en ubegrænset stor (uendelig) mængde af konsekvenser, som vi enkeltvis kan drage ud fra det, vi eksplicit ved. Fx. kan jeg ved at studere og forstå aritmetikkens aksiomer siges at vide, at de er sande. Men hvad jeg ved eller antager må formentlig siges at udgøre en endelig mængde, hvorimod der følger uendeligt mange konsekvenser af aksiomerne (at 1001 er mindre end 2001 osv. osv.). Som Dennett påpeger, kan vi derfor strengt taget ikke sige, at implicit information i et system er potentielt eksplicit. I det mindste ikke hvis vi antager, at mængden af potentielt eksplicit viden er endelig. Dette er dog alt i alt et mindre væsentligt punkt i denne sammenhæng. (Men se fx. hertil artikler i samlingen Bogdan, 1986).

Hvad der er uafklaret og langt mere prekært ved begrebet om implicit viden og information, som her defineret, er to forhold, som nu skal ridses op i grove træk:

(A) Det er endnu et åbent og meget omfattende spørgsmål, hvordan vi skal opdele klassen af de forskellige inferentielle og associative transformationer, der er nødvendige for og tilstrækkelige til, at en person kommer frem til nye antagelser ud fra allerede explicit viden. Dette viser sig ved to sæt af problemer: Dels er det (A.1) et foreløbigt uafklaret spørgsmål, hvordan vi skal afgrænse egentligt inferentielle (logiske) transformationer fra andre typer: associative, ikoniske, analoge, dynamiske - jvf. hvad der anførtes ovenfor om det (u)hensigtsmæssige i at operere med én udifferentieret klasse af eksplicitte mentale tilstande og processer omfattende såvel de sproglige og diskrete som de ikke-diskrete, ikoniske eller analoge repræsentationer. Dvs. det vil formentlig være hensigtsmæssigt at udvide klassen af implicit information og viden til at omfatte det, der fremkommer som resultat af andre

processer end de i strikt forstand inferentielle. Dennes diskussion af »tacitly represented information« - jvf. det følgende - omfatter dog netop associative og »holistiske« forbindelser. Så i det omfang dette punkt kræver en afklaring af definitionen af implicit information, vil man kunne bygge videre på den præcisering af eksplicit information, der tilsvarende blev efterlyst ovenfor.

Men desuden (A.2) bør det nævnes, at klassen af inferentielle transformationer ikke i dag er velafgrænset. For det første er der den detalje, at det i nogen grad er vilkårligt, hvad man vælger at henregne under henholdsvis logiske konstanter og under ikke-logiske udtryk (hvad Tarski allerede for et halvt århundrede siden gjorde opmærksom på). Dette lille punkt er dog næppe en vanskelighed men kan betragtes som udgørende en familie af valgmuligheder. Men for det andet rejser der sig et sæt af åbne og dybe spørgsmål ved bestemmelsen af implicit information, så snart vi søger at præcisere, hvilke typer implikation, der bør komme på tale. For det er i dag ikke afklaret, hvordan vi bør opdele og karakterisere de forskellige former for ikke-klassiske og ikke-deduktive informationstransformationer, som mennesker foretager, og som man forsøger at modellere med henblik på at indlægge dem i AI-systemer. Lad os nævne i flæng: induktive, abduktive, temporale, kontrafaktiske og ikke-monotone inferensmønstre. I det hele taget er det et spændende og åbent problem at karakterisere den fleksible, hurtige og ubesværede måde, hvorpå mennesker håndterer usikre, vage og ufuldstændige informationer; og gør det med succes, idet vi jo for det meste kommer frem til slutninger, der i det store og hele må have gavnet overlevelse. Som Quine så prægnant har formuleret det i forbindelse med induktion: individer, der har for vane at tage fejl, når de inducerer, har en sørgelig men prisværdig tendens til at dø, inden de får afkom.

(B) Endelig bør nævnes, at der også inden for klassen af deduktive inferenser eksisterer et velkendt åbent spørgsmål om, hvordan vi begrænser disse. Det kan forekomme sært at søge at finde principper for begrænsninger af inferentielle transformationer. Men antag, at vi har som ambition ikke blot at karakterisere menneskers inferentielle kompetence (hvad der i sig selv beslaglægger en betragtelig forskningsindsats), men også deres performans, dvs. deres faktiske præstationer i realistiske kontekster. I så fald må vi selv sagt tage som udgangspunkt, at vi både har begrænsede kognitive (inferentielle) ressourcer, og at vi tilsyneladende har heuristikker for at prioritere disse. Dvs. vi må afdække principper for, hvor omfattende inferentielle ressourcer, vi under forskellige forudsætninger kan tilskrive mennesker (og lad os blot i starten negligere individuelle forskelle). Det kan lyde uskyldigt at antage generelt, at når en person ved, at $A \& B$, ved han også at A ; og tilsvarende, at når han både ved, at A og at A medfører B , så må han også vide at B . Men under sådanne antagelser følger, at man tilskriver aktører, hvad man har kaldt »logisk alvidenhed« (jvf. hvad vi anførte ovenfor om eksemplet med de uendeligt mange konsekvenser af aritmetikkens aksiomer). Proble-

met med at begrænse logisk alvidenhed består i at finde teoretisk begrundede og empirisk adækvate begrænsninger på menneskers inferentielle kapaciteter. Disse problemer har været diskuteret blandt logikere og filosoffer (se fx. hertil Hintikka, 1962; Boër & Lycan, 1986; Cherniak, 1987; se også Gärdenfors, 1987, for en fremstilling af berørte problemer), og vies stigende opmærksomhed blandt tilhængere af logik-baserede metoder i AI (se fx. Haas, 1986, samt kort oversigt over epistemisk logik i Genesereth and Nilsson, 1987, og flere relevante artikler i Halpern, 1986).

Det kræver understregning, at disse åbne spørgsmål, som hæfter ved etablering af modeller og empiriske resultater vedrørende menneskers inferentielle og inferenslignende kapaciteter, jo slet ikke medfører, at en definition som den, Dennett angiver for implicit information, er uanvendelig. Tværtimod - for den er kompakt, simpel og for et begrænset formål hensigtsmæssig. Ydermere giver den et frugtbart og sammenhængende perspektiv på en tilsyneladende inhomogen klasse af kognitive fænomener (informations-transformationer, inferenslignende mønstre) - en sammenhæng som kun viser sig på tværs af enkeltdiscipliners ret snævre fokus.³

Lad os efter disse kvalifikationer forberede lanceringen af det, Dennett beskriver som »tacitly represented information« eller tavs viden. Det vil gavne forståelsen, hvis man er opmærksom på, at han her samtidig leverer en delvis men generøs rehabilitering af Ryle. Rehabilitering, fordi Ryles »The Concept of Mind« (Ryle, 1949) på mange punkter forekommer uaktuel, som Dennett påpeger: Ryles fremstilling fremkalder billedet af en uerklæret behaviourist, og hans indvendinger mod cartesisk dualisme har i dag meget lidt brod i den form, de har. Især fordi den analytiske bevidsthedsfilosofi er grundigt ændret gnenem fremkomsten af funktionalisme, og ændret i en retning som Ryle næppe har kunnet have fantasi til at forestille sig. Mange antager i dag (filosoffer eksplicit, andre kognitionsforskere for det meste implicit eller underforstået) en materialistisk ontologi (»de eneste ting der eksisterer er materielle«), men samtidig anerkender man flere typer af egenskaber, relationer og irreducible beskrivelses- og forklaringsniveauer. Dvs. man søger at retfærdiggøre en metafysik, der tillader os at betragte bevidstheden (hjernen) på en og samme tid som en materiel substans og som et system af tilstande og processer, der har »emergente« egenskaber: tanker, formål osv. Diskussionen af denne »dobbelte standard« (kombinationen af entitetsmonisme og egenskabsdualisme eller -pluralisme) har i nær tilknytning til udviklingen i kognitionsforskning gjort bevidsthedsfilosofi til det langt mest aktive filosofiske felt siden ca. begyndelsen af 80'erne. Det er derfor tankevækkende at finde hos Ryle en diskussion af »knowing-how« og »knowing-that«, som på mange måder foregriber på en forfriskende anderledes måde moderne diskussioner om, hvad de ikke-symboliske processer består i. Man finder således hos Ryle en række synspunkter om færdigheder, der står i skarp kontrast til de symbol-processerende doktriner, der har hersket og stadig yder stor indflydelse inden for AI.

Mange af de argumenter, der i dag gengives i ret enslydende form mod de symbol-processerende systemers skrøbelighed og deres utilstrækkelighed som modeller for menneskelig kognition, har genklang i Ryles fremstilling. Lad os kort se på et par hovedpunkter i kritikken af disse systemer eller modeller.

Når vi taler om symbol-processerende systemer menes der her systemer, der er programmeret i sædvanlig forstand iflg. formelle regler; disse regler vedrører selvfølgelig udelukkende syntaktiske-strukturelle egenskaber - dvs. reglerne er totalt uafhængige af det indhold, som vi »udefra« tilskriver de strukturerede repræsentationer. (Disse klassiske modeller er beskrevet talrige steder; se fx. Haugeland, 1985, hvor der gives en klar fremstilling af, hvad han kalder GOFAI: Good Old-Fashioned Artificial Intelligence; se også Searle, 1980, for en ofte citeret kritik).

Disse systemer eller modeller er principielt mangelfulde som modeller over et bredt spektrum af kognitive aspekter, anføres det. For det første kan de efter alt at dømmes aldrig bruges til at eftergøre den utroligt hurtige og smidige mønstergenkendelse, mennesker uden videre er i stand til. Vi er i stand til at præstere, hvad der kræver en meget omfattende og raffineret behandling af information: vi kan fx. inden for få hundrede millisekunder genkende et ansigt eller et sted, vi ikke har set i årevis (at nogle af os nogle gange eller for det meste er langsommere er selvsagt irrelevant: spørgsmålet er, hvilke processer der ligger bag både optimale og retarderede præstationer). Men neuroner er umådeligt langsomme i forhold til dagens datamater - det anføres, at neurale signaler er måske 100.000 gange langsommere end datamaters). Vores evne til hurtigt, ubesværet og robust mønstergenkendelse baseret på disse vore langsomme neuroner har givet anledning til at opstille, hvad der er blevet kaldt »the 100-step program constraint« (Feldman, 1985):

Denne 100-trinsbegrænsning tager udgangspunkt i den nævnte kendsgerning, at vi er i stand til at udføre meget komplekse mønstergenkendelsesakter inden for brøkdele af et sekund, men, hedder det videre, de »regler« eller »algoritmer«, som for argumentets skyld kunne tænkes på mest økonomisk vis at karakterisere hjernens aktivitet under disse præstationer, må derfor skulle implementeres i programmer, der højst gennemløber nogle få hundrede trin (og ikke millioner og mere som det i virkeligheden synes at skulle kræve). Dvs. hvis vi tager højde for neuroners langsomme processering, er der faktisk kun tid til ca. 100 serielle trin fra det øjeblik, et ansigt præsenteres, til det er genkendt. Denne dramatisk skarpe begrænsning gennemtvinger hårde krav til de mulige typer af informationsprocesser, der kan komme på tale som instantieringer. Og det er utænkeligt at forestille sig, at disse i virkeligheden yderst komplicerede præstationer kan eftergøres af programmer, der effektuerer manipulation af syntaktisk strukturerede repræsentationer, hedder det. Parallisme er ingen udvej: for uanset hvor massivt parallelt man vil implementere sådanne algoritmer til mønstergenken-

delse, må selvfølgelig intet af de parallelle spor bruge mere end 100 trin.

Et andet men relateret angrebspunkt mod de klassiske symbol-processerende modeller kommer tættere på det begreb om tavs information, som Dennett lancerer. Antag at vi (for argumentets skyld) ser bort fra perception, sensori-motorisk koordination - kort sagt alt hvad der umiddelbart må indrømmes at indebære hurtig mønstergenkendelse - og lad os se på de kognitive præstationer, som man traditionelt har søgt at modellere inden for den symbol-processerende tradition: fx. problemløsning, ræsonneren og diagnose. Kan man tænke sig et informationssystem, hvis information udelukkende er repræsenteret eksplicit - i den rummelige betydning som er defineret ovenfor? Det kan man godt, for de eksisterer jo allerede, fx. i form af ekspertsystemer eller andre såkaldt regelbaserede systemer. (Det mest sofistikerede produkt af denne art er i dag formentlig Newells SOAR (Laird, Newell & Rosenbloom, 1987))

Men som det ofte er fremhævet (se især Winograd og Flores, 1986; Dreyfus og Dreyfus, 1986) findes der et grundlæggende problem ved denne art systemer: det er vanskeligt og tilsyneladende umuligt at gøre dem både fleksible og robuste over for den indflydelse, en varierende kontekst har på løsningen af et problem. Dette kan der tages højde for, når det domæne, de er designet til at blive placeret i, er begrænset. Fx. kan man komme langt med programmering af skak-maskiner og andre doomænisolerede systemer, idet man indskyder heuristikker, der resulterer i beskæring af søgerummet. Tilsvarende er der grund til at tro, at beslutningsstøtte- og ekspertsystemer vil kunne udvikles med held for relativt lukkede domæner, hvor opgaverne nok er komplicerede, men hvor enten kontekstvariationen er meget begrænset eller ingen indflydelse har på opgaveindhold og løsning.

Mennesker og mange dyr derimod navigerer i en foranderlig og delvis uforudsigelig omverden, og deres forfædres evne til at overleve har i ikke ringe grad skyldtes, at de har kunnet mobilisere, hvad de har lært i en fart tilpasset den konkrete situation. Da man ikke på forhånd kan gå ud fra, at nogen del af en persons omfattende common-sense viden vil være irrelevant for de opgaver, han møder i morgen, er der på forhånd ingen øvre grænse for, hvilke regler og »kendsgerninger« en vidensingeniør måtte lægge ind i et system, der skulle kunne løse opgaver under bredt varierende omstændigheder. (Bemærk at vi her taler om, hvad det indebærer at hævde, at kognitive præstationer består i at følge regler - dette vedrører ikke muligheden af evt. at designe intelligente systemer ud fra andre principper).

Lad os illustrere dette. Vi tænker os, at vi får leveret en robot - en af den nye type, hvor leverandøren har løst problemet med at simulere perception, sensori-motorisk koordination og i det hele taget hele den færdighedsbaserede aktivitet. Dvs. robotten er smidig og hurtig som en abe - den går graciøst uden om ting og kan flytte krystalglas og babyer uden katastrofer. Men den er, må vi tænke os, næsten uden eksplicit viden og må betegnes som dybt uvidende. Dog må vi for enkelheds skyld forestille os, at robotdesignerne

har formået at indlægge nogle meget intrikate inferensmekanismer i den; den kan fx. behandle usikker, ufuldstændig og fejlbarlig information, og den kan regne, deducere, inducere og abducere så hurtigt og sikkert som Sherlock Holmes. Denne robot, som altså er leveret i den fordomsfrie tabula rasa model, står nu parat til at modtage eksplicit viden. Den er kort sagt det ideelle slutprodukt i den udvikling, der startede med Simon og Newell's GPS: *The General Problem Solver* (Newell & Simon, 1963). Hvilken eksplicit viden vil det være nødvendigt at indlægge i dette umulige artefact, hvis det skulle udføre dagligdags opgaver, der kræver løs og hyppigt opdateret planlægning undervejs? Er der nogen som helst grænse for, hvad det kan være nødvendigt at vide for denne oprindeligt uvidende robot, hvis den eksempelvis skal ned i byen og handle for os, og måske på vejen hjem skal ind i banken for at overtale dem til at financiere denne satsning på »optimistisk softwareudvikling«? Næppe. Men selvom vi antager, at vi faktisk har denne information i eksplicit form, ja endog i et program, så kommer dertil den vanskelighed, at antallet af beregninger generelt vokser kraftigt med antallet af regler. Ja, det vokser eksponentielt, hvis ikke søgestrategier har indbygget meget snilde begrænsende heuristikker; hvilket må ske på bekostning af, at dele af vidensbasen afsnøres i en aktuel søgning - og dermed under konstant risiko for at en relevant og essentiel information bliver negligeret. I modsætning hertil bliver menneskers evne til at trække relevant episodisk viden frem jo næppe ringere i takt med væksten af episodiske erfaringer, og jo flere »regler«, de behersker (for nu fortsat at formulere det i denne vildledende »regel«-terminologi), des hurtigere - og netop ikke langsommere - bliver de til at anvende de relevante regler. Som det så træffende er blevet udtrykt: hvad vi har brug for er ikke så meget et begreb om indholds-adresserbar men snarere relevans-adresserbar hukommelse (Se Tienson, 1988, hos hvem man også finder citeret denne aktuelle men 250 år gamle observation af Hume: »One would think the whole intellectual world of ideas was at once subjected to our view, and that we did nothing but pick out such as were most proper for our purposes.« (*Treatise, I, I, vii*)).

Tankeeksempler af denne art er absurde, og den måske væsentligste kilde til absurditeten er to forhold. For det første er det angribeligt at antage, at den omfattende mængde af såvel common-sense som esoterisk viden mennesker besidder, kan formuleres som eksplicite regler - ikke fordi den er omfattende (selv om det indebærer måske et helt uoverstigeligt praktisk problem); men fordi den er sammenhængende. Og denne sammenhæng manifesterer sig i den kontekstafhængighed, som vore begreber og ideer udviser. Sprogfilosoffer har talt om »meningsholisme«; - de mener hermed (bl.a.), at vore individuelle antagelser om omverdenen kun har den mening de har, i kraft at være associativt og inferentielt forbundet med andre antagelser (se også Quine, 1953, hvor den klassiske fremstilling gives).

For det andet er eksemplet lige så absurd som science fiction film på underholdende vis kan være det, når de fremstiller viden og færdigheder, som

noget man kan indlægge: et menneskes viden og færdigheder leveret i én software-pakke.⁴ Vor viden og færdigheder er baseret på indlæring og induktion (men selvfølgelig også på at vi får mange ting at vide eksplicit og evt. præsenteret som generelle regler). Men det er i det mindste yderst tvivlsomt, om vi typisk eller altid indlærer via formulering af eksplicite (introspecterbare og formulerbare) hypoteser om læredomænet - et kontroversielt synspunkt, der er velkendt fra fx. Ericsson og Simon, 1984. Der er således en række robuste empiriske resultater (se fx. Berry, 1987; Hayes & Broadbent, 1988) vedrørende implicit indlæring; dvs. hvor læreprocessen ikke forløber via en eksplicit formulering af de regler, der kan beskrive domænet. Reglerne kan være korrekte som beskrivelse af vor kompetence adfærd (fx. vor evne til at sortere eksemplarer som frugter eller grøntsager). Men der er bedre forklaringer på vor evne end den, der hviler på antagelsen om internalisering af regler. I særdeleshed forklaringer baseret på henholdsvis konnektionistisk induktion (se note 7) og ideer om prototypisk repræsentation (se fx. Lakoff, 1987, kap. 1 og 2 for en bred og læsværdig oversigt over resultater og teorier om prototyper; se også hertil Larsen, 1984).

Lad os opsummere: vores omverden stiller os overfor det krav, at vi ofte hurtigt skal mobilisere, hvad vi ved og kan. Den bevidst tilgængelige viden er for det første dybt integreret med vor tavse viden: vores know-how, kropslige færdigheder (perception, sensori-motorisk kontrol), og for det andet er den både uhyre omfattende men også indholdsmæssigt forbundet gennem et utal af (etté)relevansforbindelser: associationer, analogier, inferenser m.m. Disse to momenter - integrationen med den tavse viden, og den tætte indre forbundethed - gør det uplausibelt at hævde, at den er repræsenteret som eksplicit, diskret information, eller at den kan simuleres i realtid på denne basis.

For fuldstændighedens skyld bør det nævnes, at der ingen hjælp er at hente i det begreb om implicit information, vi beskrev ovenfor. Thi nok kunne man forsøgsvis hævde, at mængden af eksplicit viden i virkeligheden er langt mindre, end vi her antyder. Dvs. mindre end vi har formodet, at de regelbaserede modeller må kræve. Men det skubber problemet om at mobilisere relevant viden over på de inferentielle ressourcer. For i så fald må det være menneskers implicite viden, der skal bære byrden. Men som Dennett anfører mod dette forslag, så kan implicit viden selvfølgelig først have en virkning (i adfærd eller over for andre mentale processer) på det tidspunkt, den er produceret (udledt). Den har ingen kausal rolle, så længe den er et facit i en mulig, men ikke-realiseret slutningskæde. Hvad vi har behov for er snarere et billede, iflg. hvilket den eksplicite viden kan mobiliseres og får konkret situationsafhængig betydning inden for et sammenhængende netværk af begreber og en ramme af kognitive færdigheder:

So, *implicit* depends on *explicit*. But in the sense of »tacit« I will use, it is the other way around: *explicit* depends on *tacit*. This is what Ryle was get-

ting at when he claimed that explicitly proving things (on black-boards and so forth) depended on an agent's having a lot of *know-how*, which could not itself be explained in terms of the explicit representation in the agent of any rules or recipes, because to be able to manipulate *those* rules and recipes there would have to be an inner agent with the know-how to handle those explicit items - and that would lead to an infinite regress. At the bottom, Ryle saw, there has to be a system that *merely* has the know-how. If it can be said to *represent* its know-how at all, it must represent it not explicitly and not implicitly - in the sense just defined - but tacitly. The know-how has to be built into the system in some fashion that does not require it to be represented (explicitly) in the system. (op.cit., p. 219).

Tavs viden eller information i denne forstand er altså primært defineret ud fra to forhold. For det første dens rolle i relation til eksplicit viden (denne vekselvirkning skitseres men diskuteres i ringe grad af Dennett, men vi forsøger nedenfor at skitsere nogle hovedtræk inden for rammen af de såkaldt dobbelt-arkitekturer). For det andet er tavs viden defineret negativt i forhold til begrebet om eksplicit repræsenteret viden. Dennett sammenligner med den måde, et instrument eller et simpelt artefact kan siges at indeholde og adlyde tavs information. Fx. indeholder en simpel lommeregner ingen datastrukturer, der svarer til aritmetikkens aksiomer. Men den »følger« disse love - fordi den er designet til at gøre det. Er disse love da - i denne foreløbigt løse forstand af begrebet - indeholdt som tavs information i lommeregneren? Dennett anfører, at »the calculator is a device with the dispositional competence to produce explicit answers to explicit questions . . . , but it does this without relying on any explicit representation within it«.

Men hermed er vi egentlig kommet frem til en kilden tvetydighed. For når vi overvejer, hvordan information og viden kan siges at være indeholdt eller repræsenteret i et meneskenskabt system, ser vi, at der rejser sig nogle principielt andre muligheder end de, vi har diskuteret ovenfor.⁵

I fremstillingen har vi imidlertid hidtil uden at gøre det klart opereret med et »absolut« eller ikke-relativt begreb om repræsentation og indhold. D. v. s. vi har talt om, hvad en person ved - uafhængigt af hvordan andre fortolker hans manifesterede adfærd, og han deres. Men der er et alternativt eller supplerende begreb om viden og information, iflg. hvilket betydningen af menneskers handlinger og kommunikative udfoldelser aldrig kan bestemmes isoleret, men altid må ses relativt til den sociale praksis, de indgår i.

I denne forstand af begrebet om »information eller viden, der er repræsenteret i x« kan man illustrere den nævnte relativitet ved at betragte sprog. For vi kan jo betragte begrebet om information indeholdt i et medium (et sprog, et intermediært system, en »intelligent message scheduler«).

Der er to forskellige måder, hvorpå information kan siges at være indeholdt i et kunstigt system. For det første kan informationen være indeholdt i systemet i kraft af designerens og brugernes intentioner: dvs. indholdet er

noget, der kun eksisterer relativt til designerens og brugernes formål og antagelser (herunder deres antagelser om hinandens formål og forudsætninger). I den udstrækning, at natursproglige systemer en gang passerer papegøje-stadiet (som vist er en retfærdig betegnelse for eksisterende natursproglige systemer, der har en meget tynd og alt for ufleksibel semantik), vil deres output være i stigende grad semantisk sammenhængende. Dvs. man *kan* gøre gældende, at det er vilkårligt, om man siger, at systemet forstår, hvad det modtager og afgiver, eller om man siger, at systemet er blot et medium. Men i jo højere grad der er indholdsmæssig sammenhæng og »common-sense« bag systemets svar, des mere tilbøjelige vil vi formentlig være til at tilskrive det mentale egenskaber. Og det vil sige, des mere tilbøjelig vil vi være til at indrømme, at der kan være information indeholdt eller repræsenteret i systemet uafhængigt af, hvad dets designere har villet, og hvad dets brugere ved og tror.

Forholdet mellem information der opfattes som repræsenteret uafhængigt af, henholdsvis relativt til, omgivende aktørers intentioner og antagelser er intrikat, og kan kun berøres overfladisk her (se Andersen 1988b for en lidt fyldigere fremstilling). Især inden for den retning af filosofien, der har været påvirket af den sene Wittgensteins ideer, er dette emne blevet grundigt drøftet, omend sjældent i relation til AI.⁶

Lad os opsummere: Vi har sagt, at der er to begreber om informationsindhold, og vi har antydnet, at de to begreber er tæt forbundne, men at de som led i en analyse kan og bør adskilles. De to begreber er:

Intra-subjektiv tavs viden: information der er tavst repræsenteret i et menneskes hjerne/bevidsthed eller i kunstigt system uafhængigt af andre aktørers viden, ønsker og intentioner.

og

Intersubjektiv tavs viden: information der er tavst repræsenteret i kommunikation mellem personer (evt. mellem systemer eller mennesker) og som per definition kun har det indhold, den har relativt til parternes gensidige fortolkning.

Baggrundsviden og kontekstviden

Vores beskrivelse oven for af det sammenhængende »holistiske« netværk af antagelser, formål, vidensstrukturer har som nævnt genklang i den fremstilling, mange filosoffer har givet. Et af de mest interessante bidrag til belysningen af tavs viden og berørte emner, der er fremkommet i nyere tid, er John Searles bog om intentionalitet (Searle, 1985). Searle skelner mellem, på den ene side, det han kalder »the Network« - hvad der i store træk svarer til det sammenhængende netværk, der er skitseret her - og på den anden side, hvad han kalder »the Background«. Dette sidste begreb, som i det store og hele

modsvares, hvad vi har kaldt tavs viden (og hvad Dennett kaldte »tacitly represented information«) beskriver Searle på en oplysende men skitsepræget måde: »The Background is a set of nonrepresentational mental capacities that enable all representing to take place« (jf. også Dreyfus, 1979).

I dette afsnit vil vi kort beskrive en sontring mellem to former for viden, der går på tværs af sondringen mellem tavs og explicit viden (og følgelig sondringen mellem »The Network« og »The Background« hos Searle).

Menneskers udvælgelse, udnyttelse og fortolkning af information i omverdenen afhænger i overvejende grad af, hvad de er i stand til at forvente, og hvad de faktisk forventer af omgivelserne; d.v.s. afhænger af hvilke vidensstrukturer (skemaer, »scripts«, »frames« o.s.v.) de har til rådighed og kan anvende med held over for strømmen af information. (Og i virkeligheden burde man tilsvarende tale om forventnings- frem for vidensstrukturer, idet de jo typisk tjener til at organisere en stor mængde information ud fra nogle få udløsende træk). Denne vores forhåndsviden kan med fordel opdeles i to slags, hvilket modsvarer to forskellige slags informationsbehov, personer kan have i krævende og uventede situationer. En persons *baggrundsviden*, som det vil blive kaldt her, omfatter dels alle hans kognitive kapaciteter, hans tavse viden sådan som den skitseredes ovenfor, og dels alt hvad han har lagret i sin langtidshukommelse (altså såvel episodisk som semantisk/generel viden). Men baggrundsviden er ikke tilstrækkelig til at klare omverdens krav. For udover denne viden kræver talrige situationer et lokalt her-og-nu kendskab, d.v.s. viden om den konkrete situations art og parametre. Hvad der kræves er, hvad man kan kalde *kontekstviden*. Denne sontring lader sig bedst illustrere med eksempler fra sproglig kommunikation, hvor det så tydeligt fremgår, at de to slags viden ikke altid optræder sammen. Lad os forestille os, at John finder et kassettebånd med nogle gamle optagelser på, heraf tilsyneladende nogle fra en eller anden lokalradio - måske er der et enkelt løsrevet brudstykke på båndet, der lyder »vejret i morgen vil blive tåget og koldt, men der loves opklaring hen under aften for vores region«. Men hvor? Hvornår? Som såkaldt kompetent sprogbruger vil John til et vist niveau forstå indholdet af denne isolerede ytring. Men han ved ikke, hvad den er om - d.v.s. han ved ikke, hvad dens konkrete indhold er. Denne manglende viden er, hvad vi her kalder kontekstviden. Den studeres i sammenhæng med kommunikation systematisk inden for situationsemantik (Barwise & Perry, 1983), men mange af de underliggende principper lader sig generalisere. Desuden bør det bemærkes, at situationsemantikere har fundet megen af deres inspiration i den »økologiske« Gibson-influerede litteratur.

Afsluttende bemærkninger om modellering

Når man diskuterer, hvad modeller og simuleringer er i forbindelse med kognitive fænomener, er det fornuftigt i første omgang at se bort fra, at det ofte

er *kvalitative* fænomener, der søges modelleret. For forholdet mellem domæne, modeller og simuleringer m.m. lader sig beskrive i generelle termer, uafhængigt af domænets natur. Jeg tror, at flg. ret simplificerede opdeling er nyttig, og at den hviler på antagelser, der er korrekte i de store træk. (Den ligner flere andre fremstillinger af disse forhold - se fx. indlæg i Brachman & Smith, 1980). Vi kan skelne mellem:

- A: forskningsdomænet med dets fænomener (fx. kognitive præstationer),
- B: hypoteser og teorier om disse fænomener,
- C: algoritmiske, måske approximative modeller til at afspejle teoriens og hypotesers indhold,
- D: simuleringer, dvs. kørsler på maskiner af de programmerede modeller med henblik på at studere
 - (1) egenskaber ved teorierne, eller
 - (2) at eksperimentere med modeller (dette er en form for empirisk matematik), eller
 - (3) at undersøge overensstemmelse mellem empiriske resultater fra domænet med resultater fra simuleringer.

Til punkt D og underpunkter kan bemærkes, at i dag er konnektionistisk forskning relativt teori-fattig og derfor hovedsagelig optaget af 1 og 2; men at den regelbaserede (symbol-processeringsskolen) især dyrker 1 og 3. Desuden kan påpeges, at der er et stort trin at overskride fra selv helt formaliserede teorier til implementerede algoritmer (fx. kan man blot tænke på, at mange formelle matematiske teorier ikke kan implementeres som algoritmer på datamater undtagen i en nødvendigvis tilnærmet, approximativ form).

Kvalitative fænomener er ting, der har et indhold, en mening; d.v.s. tanker og handlinger, og i afledt forstand sprog og andre medier (fx. datamater-som-medier), kulturelle frembringelser, ritualer m.m. Kvalitative fænomener lader sig kun i ringe grad formalisere (hvilket selvfølgelig ikke er det samme som at sige, at de kun kan diskuteres i upræcise termer). Alligevel kan der i nogle tilfælde opnås indsigt ved at prøve at formalisere og modellere nogle generelle træk ved dem, idet man jo ikke kan udelukke, at selv nok så bløde fænomener udviser invarianser. Omvendt ser man undertiden, at de modeller, man simulerer, udviser emergente egenskaber, der svarer til domænets bløde fænomener (se fx. hertil diskussionen i Smolensky, 1988b). Desuden giver moderne informationsteknologi os muligheden for at undersøge langt mere komplekse og finkornede modeller, end man tidligere har drømt om at udvikle, og specielt får man mulighed for at undersøge formelle modeller, hvis komplekse dynamiske adfærd viser sig at udvise emergente træk.⁷

Det er vigtigt at fastholde, at formalismers, formelle modellers og simuleringers rolle, ikke er at tvinge al indsigt på eksplicit form, men snarere at af-

prøve grænser for, hvilke typer begreber (notionelle strukturer) man p.t. kan eksperimentere med. Formalismer (og simulering) er nødvendige for den slags eksperimenter og afprøvninger. Men omvendt er det nødvendigt, at vi parallelt hermed fører en løbende (ikke-formaliseret!) diskussion om naturen af de kvalitative fænomener (tanker, formål, heuristikker, kognitive regulariteter og særheder), som vi ønsker at få indsigt i. Kun gennem en kvalificeret løbende debat om mål, metoder og forudsætninger for de aktuelle forsøg på præcisering og formalisering kan man undgå, at formalismerne dominerer.

Der er grund til at formode, at det integrerede netværk af antagelser, normer og prototypiske kategorier ikke lader sig modellere - fordi det af praktiske grunde er for omfattende, og fordi det af teoretiske grunde er sammenhængende («meningsholismen» - jf. ovenfor) og derfor kun kan beskæres med åbenlys risiko for, at forbindelser til relevant anden viden kastes over bord. Omvendt kan man foreløbigt formode, at den færdighedsbaserede kunnen (den tavse viden i denne forstand) lader sig modellere i det mindste stedvist - fordi den tilsyneladende er relativt modular. Dvs. de enkelte funktioner varetages tilsyneladende af relativt isolerede moduler - og specielt isolerede fra hvad personen ved. Hvis dette viser sig at være sandt, er situationen den måske ironiske, at jo tættere vi kommer på eksplicit viden, i jo højere grad vil det være umuligt at modellere den.⁸

NOTER

1. Det kan virke kejtet og egentlig selvmodsigende at tale om, at en person har »viden« om et eller andet, hvis det samtidig antages, at denne »viden« er fejlagtig. (Som bekendt definerede Platon viden som antagelser (doxa), der både er sande og justificerede). Det kan således virke som et brud på normal betydning at sige »John ved, at toget går kl. 8. Men det gør det ikke«. Men der er ikke nødvendigvis tale om en anden betydning af viden. For hvis den talende fx. nu synes, at John ikke bare tager fejl, men også er urimeligt skræsikker, når han påstår »Jeg ved, at toget etc.«, så kan den talende jo netop vælge denne måde at rapportere på - fx. for at opnå ironisk-retorisk effekt. I øvrigt vil nogen hævde, at termen »information« også har indbygget denne veridikale implication (dvs. »der foreligger information om, at p« skulle medføre, at det er sandt at p; ligesom »hun ved, at p« medfører, at p er sand). For begge termer (»viden«, »information«) gælder dog tilsyneladende, at det er kontekstbestemt, hvornår de bruges i den ukvalificerede veridikale betydning, og hvornår de bruges i den fejlbarlige betydning. Med en lille rest af modvilje følger jeg den nu gængse praksis i denne sammenhæng og taler om viden som noget, der kan være fejlagtig, (og jeg indser det nyttesløse i at foreslå, at man burde opgive at tale om »knowledge engineering« til fordel for fx. »belief engineering«).
2. Det vil føre alt for vidt at komme ind på den ret omfattende diskussion af og litteratur om ubevidste processer. William James argumenterede i sit klassiske værk *Principles of Psychology* («Can States of Mind be Unconscious?») mod antagelsen af ubevidste processer: »the unconscious is the sovereign means for believing, whatever one likes in psychology and of turning what might become a science into a tumb-

ling ground for whimsies«! Moderne kognitionsforskning må siges at være solidt involveret i at søge at forklare kognitive præstationer i de af James forbudte termer.

Se fx. Kihlstrom, 1987, Shevrin & Dickman, 1980, der indeholder omfattende oversigter; desuden Hollender, 1984, med følgende kommentarer). Inden for pragmatik er det ligeledes almindeligt at antage ubevidste inferenser (deduktive, abduktive; non-monotone etc.) baseret på formodninger om andre aktørers baggrunds- og kontekstviden. (Men se fx. den letlæste men omfattende diskussion i Sperber & Wilson, 1987).

3. Moderne kvalificeret kognitionsforskning bedrives typisk i miljøer, hvor de akademiske disciplingrænser er ophævede, men hvor deltagerne har en solid faglig baggrund i et grundfag og desuden er bekendte med visse af de andre fags grundtræk og resultater. Dette forskningssociologiske forhold udgør den nødvendige baggrund for, at man kan levere den brede forskningsindsats, der skal til at forklare noget så måske usammenligneligt kompliceret som kognitive præstationer. Når vi eksempelvis prøver at skabe en taksonomi for begreber om tavs viden og dens komplementer er det naturligt, at man på dette stadie må vælge mellem enten (i) at give forståelige, kompakte men stormaskede definitioner (der udjævner mange distinktioner hidrørende fra forskellige fag) eller (ii) at give finmaskede definitioner, der respekterer de åbne spørgsmåls mange muligheder, men som til gengæld er uoverskuelige med deres mange garderinger, og derfor let kan vise sig mindre nyttige. På et tidligt stadie i et forskningsfelt må man typisk stille sig tilfreds med en upræcis definition; som sådan er den en fremtidsveksel, og dens værdi vil først fremgå den dag, den enten er opgivet som ufrugbar eller bliver indløst enten gennem en etableret enighed om mønstergældende eksemplarer. En af T. Kuhns store fortjenester var at fremhæve, hvor illusorisk det er at tro, at i videnskab må alt være veldefineret i klassisk analytisk forstand.
4. I filmen *Bladerunner* optræder robotter eller »replicas«, i hvem viden indsættes eller indhældes som væske i en beholder. En del af det fascinerende ved at høre og se fx. replica-pigen, der i filmen overtager en afdød, rigtig piges samtlige færdigheder og barndomserindringer, er formentlig følelsen af det fornuftsudfordrende. (»Her, tag denne pille, så kan du spille klaver«).
5. Dette er en tvetydighed, som ikke kommer frem i Dennetts fremstilling. Tværtimod bidrager han til sammenblandingen af (eller tilsløringen af sondringen mellem) det, vi kalder absolut og relativ information (se hans analogi med det implicitte og virkelige budskab i en spions explicitte budskab, som blot er en dækhistorie - p. 217). Dette er naturligt ud fra Dennetts instrumentalistiske grundstandpunkt om tilskrivning af tanker, viden og formål til datamater, dyr og mennesker: »instrumentalistisk« i den forstand, at spørgsmålet om det er meningsfuldt og sandt eller falsk at sige fx., at en termostat ved og vil noget som helst, ikke har noget »objektivt« svar. Iflg. Dennett er det udelukkende et spørgsmål om nytte! »Hvis det er nemmere at sætte sig ind i og fortælle andre, hvordan systemet virker (termostaten, programmet/systemet), når man bruger mentalistiske termer (»ved«, »tror«, »prøver« o.s.v.), så er det både meningsfuldt og praktisk at gøre det. Og så er der i øvrigt mulighed for, at en praksis opstår iflg. der bliver regler for, hvornår sådanne tilskrivninger regnes som sande eller falske.
6. Traditionen hidrørende fra især Wittgenstein har sat sig interessante spor i debatten om Human-Computer Interaction. Nogle principielle, filosofiske betragtninger findes fx. hos Johannesen (1988), Harré (1988). Konsekvenser for design diskuteres flere steder i Rasmussen (1986). I Berry (1987) diskuteres nogle praktiske problemer med at fremdrage tavs viden, medens vidensrepræsentation desuden berøres i Collins (1986) vidtrækkende essays. Den omfattende men indtil for ganske nylig isolerede og noget indavlsprægede filosofiske debat om, hvad det vil sige at følge en regel - er det noget en person gør i kraft af, at reglen i en eller anden forstand er repræ-

senteret i hans hovede, eller er det i kraft af, at andre kan se og acceptere, at han følger reglen - har i nyere tid især været præget af diskussionen i Kripke (1982); se også hertil samlingen Holtzman & Leich (1981).

7. Det vil være påfaldende - selv i en så forholdsvis generel fremstilling som i denne artikel - at diskutere muligheder for at modellere tavs viden uden at nævne konnektionisme. Emnet er dog så omfattende, at jeg har valgt at indskrænke mig til at sammenfatte nogle hovedtræk i denne lange oversigtsnote; og i øvrigt findes der efterhånden mange introducerede diskussioner af konnektionisme (jf. referencer nedenfor), ligesom dette bind indeholder artikler om emnet (se O. Bernsen og K. Plunkett, dette bind).

Konnektionisme («connectionism») er det velvalgte navn, som J. Feldman har givet til en bred og løst defineret klasse af modeller, der i de seneste år har tiltrukket stor opmærksomhed, betragtelige forskningsmidler, og en del kontrovers (disse modeller kaldes også pdp («parallel distributed processing») modeller eller neurale netværk). Denne forskningsretning viderefører (nogle siger: genopliver) studier, der startede i 50'erne, men som gennemgik en stille periode i ca. 15 år, indtil den kraftige aktivitet satte ind fra begyndelsen/midten af 80'erne.

Konnektionistiske modeller (systemer) er netværk af mange (evt. millionvis) af meget simple men »massivt« forbundne processorer. En populær type af netværk indeholder typisk »forneden« et lag af receptorprocessorer (som tænkes at modtage signaler fra netværkets omgivelser) og »foroven« et lag af output-processorer, der repræsenterer netværkets respons. Udover input processorerne, der modtager eksterne signaler, modtager hver processor signaler fra de evt. mange naboprocessorer, hvormed den er forbundet, og den beregner en aktiveringsværdi iflg. en eller anden given simpel algoritme; denne værdi sendes så videre til forbundne processorer. Processorerne påvirker således direkte forbundne processoreres aktiveringsværdi moduleret gennem den specifikke »vægt«, der indeholdes i hver forbindelse. Vægtene ganges med aktiveringsværdien, og resultatet indgår som input med negativ (inhibitorisk) eller positiv (excitatorisk) værdi til den modtagende processor. Når indgåede signaler til input-lagets receptorprocessorer er propageret op gennem nettet til respons-processorerne bliver (i en typisk familie af modeller) »fejll« algoritmisk propageret baglæns gennem nettet som justeringer af vægtene. Ved gennemløb af mange propageringer (op gennem systemet ved præsentation af eksterne signaler, d.v.s. præsentation af et læreeksemplars træk indkodet til receptor-processorerne og påfølgende justering oppe-fra-og-ned af vægte) kan et system opnå en stabil tilstand vis-a-vis læresættet (d.v.s. en tilstand, hvor præsentation af et eksemplars træk til receptorer og følgende »opad« aktivering ikke medfører justering af vægte).

I et konnektionistisk system er »information om de underliggende regler«, der hersker for et læresæt af eksemplarer, ikke indeholdt i nogen som helst eksplicit repræsentation; men den er indeholdt i fordelt form i de selvprogrammerede vægte, der gradvist er etableret, efter at et netværk evt. har opnået stabilitet (eller har nået et forud fastsat mål af tilfredsstillende respons over et sæt af nye eksemplarer fra domænet).

Det moderne hovedværk inden for konnektionisme er Rumelhart & McClelland (1986) - især Don Normans afsluttende kapt. i bd. 2 samt de fire første kapt. i bd. 1 er gode introducerende oversigter. En aggressiv men principiel kritik af konnektionistiske modellers rolle som modeller af kognitive fænomener findes i Fodor & Pylyshyn (1988), medens et afbalanceret og i visse henseender imødekommende svar gives i Smolensky (1988a,b). Bl.a. Broadbent (1985) og Anderson (1987) giver udtryk for en ofte fremført traditionel skepsis overfor konnektionistiske modeller: de kan tjene som implementeringer men ikke til at belyse kognition »at the computational/algorithmic level«. Smolenskys grundige fremstilling (1988a) belyser flere tvetydigheder og opklarer udbredte misforståelser; men flere af de grundlæg-

gende distinktioner har endnu ikke fået en blot tilnærmelsesvis endelig form. Se også Pedersen (1988) og Rasmussen (1986) for en skitsering af andre mulige alternativer til traditionelle »symbol-behandlings-approaches«.

Konflikten mellem konnektionisme og traditionel diskret symbol-baseret modellering trækkes undertiden hårdt op: »AI is in a state of paradigmatic mess with little agreement on the nature of intelligence or the proper methodology for studying it . . . connectionism challenges some of the fundamental assumptions of on which most work in symbolic AI has been based . . .« (Goel, 1988, p. 28); men - og det ville heller ikke kunne forventes - den påståede modsætning bliver der ikke gjort forsøg på at trække principielt op i de underholdende og tit lidt selvspottende kongresberetninger: ». . . Expert Systems have had their days (seven years, in fact) and it was on the cards that another fad would take over soon« (Forsyth, 1988). En forskningsekstern kilde til modsætning er meget naturligt konkurrence om fælles bevilgningssourcer: »The reason why the US AI community (academic and commercial) has taken up the neuralnet model so enthusiastically is straightforward. It is primarily because the Department of Defence has decided that neuralnet computing is a high-priority strategic technology. As an example, the UCLA AI Lab has recently started ten new projects concerned with neural networks while seven (Schank-style) symbolic AI projects are due to be terminated shortly. This switch didn't come from inside the university, it happened as a result of strong prompting from DARPA [Defence Advanced Researched Project Agency] and other funding bodies. But why should US government agencies be pushing AI research into a particular direction? To put it crudely, they see it as a defence against the Red menace and the Yellow peril at the same time« (ibid.).

Når tilhængere af konnektionisme undertiden fremstiller denne type modeller som alternativer til mere traditionelle former for kognitiv modellering, kan det således have såvel en (forsknings)politisk som en saglig baggrund. Den nu ombremmede entusiasme bag udviklingen af ekspertsystemer har efterladt en ophobning af uindfriede forventninger; disse har nu fundet en ny holdeplads.

En lidt anden (men i virkeligheden tæt beslægtet) grund til de til tider opskruede forventninger til konnektionisme er måske at finde i det forhold, vi berører i hovedteksten: nemlig at symbol-baserede modeller er så iøjnefaldende mangelfulde til at forklare, hvordan tavs viden manifesterer sig i færdighedsbaserede præstationer og i menneskers smidige udnyttelse af relevant common sense og esoterisk) viden i hurtigt skiftende situationer. Om konnektionistiske modeller vil vise sig at bidrage til belysning heraf, er det altfor tidligt at afgøre.

8. Flere personer har påvirket mine forestillinger om tavs viden. Især skylder jeg at takke Jens Ramussen og Stig Andur Pedersen for mange diskussioner. Jeg er også taknemmelig for nogle inspirerende bemærkninger, som Berndt Brehmer knyttede til et foredrag, jeg holdt i Uppsala i foråret 1988 om emnet.

REFERENCER

- ANDERSEN, H.B., ed. (1988a): *Quasi-Rational Thinking*, Forskningscenter Risø, Risø-M-2752 (Dec. 1988).
- ANDERSEN, H.B. (1988b): »Types of Tacit Knowledge«, Forskningscenter Risø, Risø M-2753 (Dec. 1988).
- ANDERSON, J.R. (1983): *The Architecture of Cognition*, Cambridge, MA: Harvard Univ. Press.
- ANDERSON, J.R. (1987): »Methodologies for Studying Human Knowledge«, *Behavioral & Brain Sciences*, 10, 3.

- BARWISE, J. & PERRY, J. (1983): *Situations and Attitudes*, MIT Press.
- BERRY, D. (1987): »The Problem of Implicit Knowledge«, *Expert Systems*, 4, 5.
- BODEN, M. (1988): *Computer Models of Mind*. Cambridge Univ. Press.
- BOGDAN, R.J., ed. (1986): *Belief*, Oxford.
- BOËR, S.E. & LYCAN, W.G. (1986): *Knowing Who*. MIT Press, Cambridge, Mass.
- BRACHMAN, R.J. & SMITH, B. eds. (1980): *Special Issue on Knowledge Representation*, *Sigart Newsletter*, 70, Feb. 1980.
- BROADBENT, D. (1985): »A Question of Levels: Comment of McClelland and Rumelhart«, *J. Exper. Psychol. General*, 114, 2, 189-92.
- BROADBENT, D.E., FITZGERALD, P. & BROADBENT, M.H.P. (1986): »Implicit and Explicit Knowledge in the Control of Complex Systems«, *Brit. J. Psych.*, 77, 33-50.
- CHOMSKY, N. (1965): *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Mass.
- CHOMSKY, N. (1980a): »Rules and Representations«, *Behavioral and Brain Sciences*, 3, 1-15.
- CHOMSKY, N. (1980b): *Rules and Representations*. Columbia Univ. press, New York.
- CHOMSKY, N. (1986): *Knowledge of Language: Its Nature, Origin and Use*. Praeger, New York.
- COLLINS, H.M. (1985): *Changing Order: Replication and Induction in Scientific Practice*, Sage, Beverly Hills & London.
- DENNETT, D. (1982/83): »Styles of mental representation«, *Proc. Arist. Soc.*, vol LXXXIII. Genoptrykt i Dennett (1987).
- DENNETT, D. (1987): *The Intentional Stance*. MIT Press, Cambridge, Mass.
- DREYFUS, H. & DREYFUS, S. (1986): *Mind Over Machine*. Blackwell, Oxford.
- DREYFUS, H. (1979): *What Computers Can't Do* (2nd rev. ed.), Harper & Row, New York.
- ERICSSON, A. & SIMON, H. (1984): *Protocol Analysis*, MIT Press, Cambr. Mass.
- FAHLMANN, S.E. & HINTON, G.E. (1987): »Connectionist Architectures for Artificial Intelligence«, *Computer*, January.
- FELDMAN, J. (1985): »Connectionist Models and their Applications: Introduction«, *Cognitive Science*, 9, 1-2.
- FODOR, J. & PYLYSHYN, Z. (1988): »Connectionism and Cognitive Architecture: A Critical Analysis«, *Cognition*, 28, 2-71.
- FORSYTH, R. (1988): »AI: East/West Conference Report«, *AISB (Artificial Intelligence & the Simulation of Behaviour) Quarterly*, Autumn 1988, no. 66.
- GÄRDENFORS, P. (1988): *Knowledge in Flux*, MIT Press, Cambr. Mass.
- GOEL, A. (1988): »A Report on the AAAI Symposium on Parallel Models of Intelligence«, *Sigart Newsletter*, No. 106, Oct. 1988.
- GOODSTEIN, L.P., ANDERSEN, H. & OLSEN, S.E., eds. (1988): *Tasks, Errors and Mental Models*. Taylor & Francis, London.
- HAAS, A.R. (1986): »A Syntactic Theory of Belief and Action«, *Artificial Intelligence*, 28, 245-92.
- HALPERN, J. Y. (1986): *Theoretical Aspects of Reasoning about Knowledge: Proceedings of the 1986 Conference*, Morgan Kaufman, Los Altos, Calif.
- HANSON, N.R. (1958): *Patterns of Discovery*, Cambr. Univ. Press.
- HARRÉ, R. (1988): »Wittgenstein and Artificial Intelligence«, *Philosophical Psychology*, 1, 1, 105-15.
- HAUGELAND, J. (1985): *Artificial Intelligence: The Very Idea*. MIT Press, Cambr. Mass.
- HAYES, N. & BROADBENT, D.E. (1988): »Two Modes of Learning for Interactive Tasks«, *Cognition*, 28, 249-76.
- HEIDEGGER, M. (1927/1967): *Sein und Zeit*, opr. udg. i *Jahrbuch f. Philos. u. phänom. Forschung*, 11. oplag, Niemeyer, Tübingen 1967 og senere.

- HINTIKKA, J. (1962): *Knowledge and Belief*, Cornell University Press, Ithaca, N.Y.
- HOLLENDER, D. (1986): »Semantic Activation without Conscious Identification in Dichotic Listening, Parafoveal Vision and Visual Masking: A Survey and Appraisal«, *Behavioral & Brain Sci.* 9, 1-66.
- HOLTZMAN, S. & LEICH, C. (1981): *Wittgenstein: To follow a Rule*, Routledge & Kegan Paul, London.
- HOOK, S. ed. (1969): *Language and Philosophy*, New York Univ. Press.
- JANLERT, L.-E. (1988): »Pictorial Knowledge Representation«, *Proc. 8th European Conf. on Artificial Intelligence*, Pitman Publ., London 149-51.
- JOHANNESSEN, K. (1988): »Tankar om tyst kunnskap«, *Dialoger Nr. 6: Tyst Kunnskap*, (Arbetslivscentrum, Box 5606, Stockholm).
- JOHNSON-LAIRD, P.N. (1983): *Mental Models*, Cambridge Univ. Press.
- KIHLSTROM, J.F. (1987): »The Cognitive Unconscious«, *Science*, vol. 237 (18, Sept.).
- KOHONEN, R. (1988): »An Introduction to Neural Computing«, *Neural Networks*, 1, 3-16.
- KRIPKE, S. (1982): *Wittgenstein on Rules and Private Language*, Blackwell, Oxford.
- KUHN, T. (1962/1970): *The Structure of Scientific Revolutions*, 2nd exp. ed., Chicago Univ. Press. Dansk udg.: *Videnskabens Revolutioner*, Fremad, Kbh. 1973.
- LAKOFF, G. (1987): *Women, Fire and Dangerous Things*, Chicago Univ. Press.
- LAIRD, J.E., NEWELL, A. & ROSENBLOOM, P.S. (1988): »SOAR: An Architecture for General Intelligence«, *Artificial Intelligence*, 33, 1, 1-64.
- LARSEN, S.F. (1984): »Kognitionens Logikker«, *Psyche & Logos*, 5, 2, 221-42.
- MCCLELLAND, J. & RUMELHART, D., eds. (1986): *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, bd. I-II. MIT Press, Cambridge, Mass.
- MILLER, A.I. (1986): *Imagery in Scientific Thought Creating 20th Century Physics*, Birkhäuser, Boston, Basel, Stuttgart.
- NEWELL, A. (1982): »The Knowledge Level«, *Artificial Intelligence*, 18, 81-132.
- NEWELL, A. & SIMON, H. (1963): »GPS - A Program That Simulates Human Computers and Thought«, i: Feigenbaum, E. & Feldman, J. (eds.): *Computers and Thought*, McGraw Hill, New York.
- NORMANN, D. (1986): »Reflections on Cognition and Parallel Distributed Processing«, i: McClelland & Rumelhart 1986, bd. 2.
- PEDERSEN, S.A. (1988): »Modelling the Background of Knowledge and Belief«, i Goodstein et al., 1988.
- POLANYI, M. (1966): *The Tacit Dimension*, Doubleday, New York.
- POLANYI, M. (1958): *Personal Knowledge*, Routledge & Kegan Paul, London.
- QUINE, W.V.O. (1953): »Two Dogmas of Empiricism«, i Quine, W.V.O.: *From a Logical Point of View*, Harper, New York.
- RASMUSSEN, J. (1974): »The Human Data Processor as a System Component: Bits and Pieces of a Model«, Forskningscenter Risø, R-8-74 (Juni, 1974).
- RASMUSSEN, J. (1986): *On Information Processing and Human-Machine Interaction*, Elsevier, Amsterdam.
- RASMUSSEN, J. (1987): »Mental Models and the Control of Actions in Complex Environments«, Forskningscenter Risø, Risø-M-2656 (Jul, 1987).
- REASON, J. (1988): »Framework Models of Human Performance and Error: A Consumer Guide«, i L.P. Goodstein et al. (1988).
- RUMELHART, D.E., HINTON, G.E. & WILLIAMS, R.J. (1986): »Learning Representations by Back-Propagating Errors«, *Nature*, 323, 9 Oct.
- SCHNEIDER, W., DUAMAS, S.T. & SHIFFRIN, R.M. (1984): »Automatic and Control Processing and Attention«, i Parasuraman, R. & Davies, D.R., eds.: *Varieties of Attention*, Academic Press, Orlando, Florida.
- SEARLE, J. (1983): *Intentionality*. Cambridge Univ. Press.

- SHEVRIN, H. & DICKMAN, S. (1980): »The Psychological Unconscious«, *American Psychologist*, 35, 5, 421-34.
- SMOLENSKY, R. (1988b): »The Constituent Structure of Connectionist Mental States: A Reply to Fodor and Pylyshyn«, *The Southern J. of Philos.*, 26, *Supplement: Connectionism and the Philosophy of Mind*.
- SPERBER, D. & WILSON, D. (1986): *Relevance*, Blackwell, Oxford.
- STABLER, E. (1983): »How are Grammars Represented?«, *Behavioral and Brain Sciences*, 6, 391-422.
- TIENSON, J. (1988): »Introduction to Connectionism«, *The Southern J. of Philos.*, 26, *Supplement: Connectionism, and the Philosophy of Mind*.
- VICENTE, K. & RASMUSSEN, J. (1988): »A Theoretical Framework for Ecological Interface Design«, Forskningscenter Risø, Risø-M-2736 (Aug. 1988).
- WINOGRAD, R. & FLORES, F. (1986): *Understanding Computers and Cognition*, Ablex, New Jersey.
- WITTGENSTEIN, L. (1953): *Philosophical Investigations*, Blackwell, Oxford.
- WITTGENSTEIN, L. (1980): *Remarks on Philosophical Psychology*, Blackwell, Oxford.