

KOGNITIONS- OG
KUNSTIG INTELLIGENS-FORSKNING
MELLEM LOGIK OG PSYKOLOGI:
ET SEMIOTISK SYNSPUNKT

Michael May

Artiklen præsenterer et »meta-semiotisk« perspektiv på kognitionsforskningen (CS), ifølge hvilken man hverken bør søge grundlaget for CS i den psykologiske udforskning af menneskelig intelligens eller i konstruktionen af en »kunstig intelligens« (AI), men derimod i udforskningen af de mulige udformninger af kognitive processer og formationer. Teoretisk AI vil understøtte en sådan forskning ved at analysere egenskaberne for datamatiske modeller, der kan simulere »det kognitive«.

CS bør interessere sig specielt for biologisk implementeret intelligens, men det bør ske på baggrund af et teoretisk arbejde med forskellige niveauer af »tvangsbindinger« (logiske, fysiske, biologiske etc. constraints), ud fra hvilke vi kan erkende de forskellige naturlige og kunstige implementeringer af intelligens som realiseringer af abstrakte muligheder, der knytter sig til tegn, redskaber og modeller. Metasemiotikken studerer de almene egenskaber ved tegn (og ved betydning og kommunikation), redskaber, herunder maskiner) og modeller (fysiske og mentale modeller).

CS står over for tre fundamentale problemer: 1) at overvinde modstillingen af psykologiske og logiske opfattelser af kognition, 2) at reformulere den fænomenale verdens status inden for teorier om kognition og 3) at reformulere konflikten mellem det klassiske AI-paradigme og det konnektionistiske paradigme (»neurale netværk«) som et index på et reelt teoretisk problem, - nemlig spørgsmålet om forholdet mellem symbolske og sub-symbolske processer og om integrationen af diskrete og kontinuerte modeller af disse kognitive processer.

I forbindelse med institueringen af »informationspsykologi« og »informatisk psykologi ved de psykologiske institutter i henholdsvis København og Århus, kan det være relevant at få en diskussion om psykologiens forhold til kognitionsforskningen (i det følgende forkortet CS efter »Cognitive Science«) og »kunstig intelligens«-forskningen (i det følgende forkortet AI efter »Artificial Intelligence«). Mit bidrag hertil er det generelle *semiotiske* synspunkt, som fremlægges i artiklen - dels med henblik på den overordnede opfattelse af AI- og CS-forskning og dels med henblik på semantik, som et særligt område for denne forskning, hvor semiotikken kan blive relevant.

Forsøgene på at formulere en *kalkulatorisk semantik*, det vil sige at formulere betydnings-strukturer og -processer som et »beregnbart« problem, udgør ikke kun et teoretisk centralt problem i AI og CS, men er også et centralt problem for mange praktiske AI-projekter, der involverer simulation af naturligt sprog. Jeg kan dog ikke i denne artikel komme nærmere ind på de forskellige tilgange til semantik men må nøjes med at beskrive en række almene problemner omkring afgrænsningen af AI og CS. Hvad jeg forstår ved semiotik vil fremgå i løbet af artiklen.

Den frankensteinianske opfattelse af AI

Undertiden møder man i AI-litteraturen, hvad jeg vil kalde en »frankensteiniansk« opfattelse af målet med AI-forskningen: at AI skulle dreje sig om at »konstruere en person eller, mere beskedent, et dyr«¹. Inden for sprogets område bliver målet for AI da at konstruere tænkende og talende maskiner, at konstruere systemer der »forstår« naturligt sprog etc. Denne opfattelse beror efter min mening dels på en manglende skelnen mellem teoretisk og anvendt AI og dels på en uheldig forenkling af forskningsmetodologien i AI.

Mens anvendt AI fx. omfatter konstruktionen af ekspertsystemer og industrirobotter til specifikke formål (og altså ikke sigter mod konstruktionen af »generelle robotter«), så må teoretisk AI dreje sig om den teoretiske konstruktion af et erkendelsesobjekt (men spørgsmålet er hvilket). Forskningsmetodologien i AI er knyttet til computerteknologien på den måde, at man i AI kan udføre simulationer ved at konstruere programmer, der kan repræsentere den model, som man ønsker at undersøge eksperimentelt, og derefter køre programmerne på en datamaskine. Modeller, der er udformet således, at de i princippet kan repræsenteres i et programmeringssprog (men ikke nødvendigvis et af de aktuelt foreliggende), og som i princippet kan køres på en datamaskine (men ikke nødvendigvis med en af de aktuelt foreliggende maskinarkitekturer, kaldes *datamatiske modeller*.

Inden for sprogets område vil målet for teoretisk AI være knyttet til evalueringen af datamatiske modeller af forskellige sproglige funktioner, sådan som det sker i *datalingvistikken*, mens målet for anvendt AI inden for dette område kan være, at konstruere systemer til maskinel oversættelse af tekster eller at konstruere brugergrænseflader til vidensbaserede systemer, som kan understøtte dialoger i naturligt sprog. Sådanne *dialogsystemer* vil imidlertid ikke »forstå« det, der bliver sagt, selv om man i litteraturen ofte betegner dette område som »natural language understanding« eller »natural language comprehension«. Maskinelle dialog-systemer kan understøtte en dialog i naturligt sprog, fordi de kan *simulere* visse almene sprogfunktioner i kraft af, at systemets program repræsenterer en datamaskin model af disse funktioner. Systemets måde at simulere almene sprogfunktioner på behøver ikke at »ligne« den måde, hvorpå mennesker faktisk udfører disse funktioner, men det

er ikke desto mindre på grund af sådanne slutninger, at nogle AI-forskere vil hævde, at dialog-systemer »har sprogforståelse«, og at man erkender menneskers »sprogevne« ved at konstruere AI-systemer, der simulerer den. Disse slutninger hænger til dels sammen med en sammenblanding af anvendt og teoretisk AI og med en forenkling af forskningsmetodologien i AI. Men hvis det på den anden side er muligt for AI-forskningen faktisk at konstruere systemer, der kommunikerer relevant i naturligt sprog, omend det blot er partielt og simulatorisk, så må det vel få konsekvenser for vores opfattelse af sprog og sprogfunktioner i almindelighed; men hvilke? Lad os først se nærmere på, hvad man kan forstå ved AI som en videnskab.

AI som proto-videnskab

I det første nummer (1986) af det engelske tidsskrift »Artificial Intelligence Review«, tager John A. Campbell udgangspunkt i vanskeligheden ved at klassificere AI og hævder, at AI bedst kan karakteriseres som en *protovidenskab*², hvis struktur og grundlag endnu ikke er udformet. Der kan dog skelnes mellem tre niveauer af AI-forskning. På det laveste niveau, der nærmest må karakteriseres som rent datalogisk, drejer AI sig om udviklingen af specielle maskinarkitekturer, programmeringssprog og programmeringsteknikker, der kan understøtte AI-eksperimenter og anvendelser. På det mellem-liggende niveau drejer AI sig om udviklingen af *heuristikker* (dvs. »tommel-finger-regler«) eller, hvis det er muligt, *algoritmer* (dvs. »effektive procedurer«) til at løse specifikke AI-problemer (for så vidt håbet om at simulere en »generel intelligens« er opgivet). På det højeste niveau drejer AI sig om konstruktionen af *datamatiske modeller*³, - det vil sige modeller, der kan undersøges eksperimentelt ved at blive implementeret som programmer, der da kan »køres« reelt på en datamaskine eller imaginært som et tankeeksperiment.

AI er ifølge Campbell en protovidenskab på samme måde som botanikken før Linné: den er på jagt efter en almen ramme for beskrivelser, der vil muliggøre systematiske abstraktioner og almene teorier. Han fremhæver selv to af de abstraktioner, der må indgå i enhver fremtidig behandling af AI som videnskab: heuristikbegrebet og modelbegrebet. Netop den systematiske beskrivelse af heuristikker har hidtil manglet i AI-litteraturen, der har været præget af ad hoc-løsninger på specifikke problemer. AI-eksperimenter og anvendelser har ofte ikke udnyttet tidligere erfaringer eller allerede eksisterende viden, men i stedet startet fra grunden med at sammenflikke en løsning til lejligheden.

Alan Bundy har iværksat en katalogisering af »AI-redskaber« (heuristikker, algoritmer, repræsentationsformalismer, programmeringsteknikker etc.) - netop begrundet i et ønske om at etablere *en kerne af fælles teknikker*, samt overhovedet at artikulere AI som et forskningsområde med en *fælles metodologi og fælles kriterier for evaluering* af forskningsresultater⁴. Etable-

ringen af en sådan kerne har ikke været noget let arbejde og det er kun partielt udført. Bundy påpeger, hvordan resultater i AI ofte beskrives i en »se mor, ingen hænder«-stil, hvor implementeringen af modeller i programmer kun er ufuldstændigt dokumenteret (hvis den overhovedet er gennemført og ikke blot postuleret mulig); og selv når implementeringen er dokumenteret, er de anvendte teknikker sjældent abstraheret fra programmerne og formuleret som generelle teknikker. Da AI endnu ikke er et begrebsligt kohærent forskningsområde, er der heller ikke etableret en *fælles terminologi*, og det gør det endvidere vanskeligt at afgøre, om beskrivelser af resultater, modeller og teknikker er beskrivelser af noget nyt, eller om der er tale om allerede kendte resultater, modeller og teknikker beskrevet med en ny terminologi.

Adfærds- og kognitions-orienterede AI-definitioner

Et aspekt af AI som protovidenskab tydeliggøres, hvis man spørger, hvad de forskellige datamatiske modeller i AI er modeller af. Forskellige traditionelle områder af AI vil give delvist forskellige svar.

Bundys katalogisering dækker så forskellige områder som: automatisk bevisførelse, automatisk inferens, automatisk programmering, datamodeller, ekspertsystemer, logikprogrammering, maskinarkitekturer, maskinel indlæring, maskinelt syn, mønstergenkendelse og billedbehandling, naturlig sprogsimulering, planlægning, problemløsning, programmeringssprog, robotik, spilstrategier, søgestrategier, talegenkendelse og -syntetisering, vidensrepræsentation. Spørgsmålet er så, hvad der er fælles, hvis ellers AI har noget konsistent erkendelsesobjekt. De fleste AI-fremstillinger kredser her omkring termen »intelligent adfærd«, men på helt forskellige måder. Det er væsentligt her at se på nogle af disse forsøg på at definere AI, idet de er afgørende for opfattelsen af forholdet mellem AI og CS, og dermed for psykologiens forhold til dem begge.

Ifølge Avron Barr & Edward A. Feigenbaum (1981) er *symbolsystemer* og *informationsbehandling* de fundamentale begreber i AI, og de definerer AI som den *del af datalogien*, der er optaget af »at designe intelligente computersystemer«. Ved intelligente systemer forstås de »systemer, som udviser de egenskaber, vi forbinder med intelligent i menneskelig *adfærd*«.

Denne definition er orienteret mod anvendt AI (: design af intelligente computersystemer) og er endvidere adfærdsorienteret, idet »intelligent« er en beskrivelse af adfærdsformer og ikke af de maskinelle eller menneskelige aktører, der udfører dem⁵.

Elaine Rich betragter også symbolsystemer som et fundamentalt begreb. Ifølge Rich er AI baseret på *hypotesen om fysiske symbolsystemer*, - formuleret af Allen Newell⁶ som det fundamentale fælles grundlag for AI & CS, - der hævder, at de processer, der indgår i produktionen af intelligent adfærd, kan *simuleres* af et system af fysiske symboler og mekanismer, der konstruerer

symbolstrukturer ved at operere på symbolerne. Computeren er det *redskab*, hvormed denne manipulation af symboler og konstruktion af symbolstrukturer kan realiseres. Disse symbolstrukturer *repræsenterer viden* (for brugeren af dette redskab), - dels generel viden om et problemområde (domænevden) og dels specifik viden om, hvordan problemer kan løses i dette område (heuristisk viden). Elaine Rich definerer AI som »studiet af de måder, hvorpå computere kan bringes til at udføre *kognitive opgaver*«⁷.

I modsætning til Barr & Feigenbaum, der ser intelligens som en adfærdsmæssig kvalitet, - det vil sige som en beskrivelse af en handlingsemæssig performans (maskinen »udviser intelligent adfærd«), implicerer Rich intelligens som en kognitiv kvalitet, det vil sige som en beskrivelse af en kognitiv performans (maskinen »udfører kognitive opgaver«). Et spørgsmål til Richs definition af AI vil helt oplagt være, hvad der kvalificerer en computers datamatiske »opgaver« som værende *kognitive*.

Skellet mellem den adfærds- og kognitions-orienterede definition af AI har mange teoretiske og praktiske konsekvenser. Fra en ekstrem behavioristisk position, vil man kunne hævde, at et AI-system kun skal vurderes på sin evne til at *efterligne* en menneskelig intelligent adfærd, uanset hvordan det finder sted. Systemet selv kan da opfattes som en »black-box«, der blot skal »designes« på en adækvat måde for at producere efterligningen. Denne opfattelse forekommer ikke kun inden for den adfærdsorienterede AI-forskning og inden for det hidtidigt fremherskende *PSS-paradigme* (:»Physical Symbol Systems«) for CS, men ses endnu mere rendyrket inden for det konkurrerende *PDP-paradigme* (:»Parallel Distributed Processing«, mere eller mindre ækvivalent med forskningen i »konnektionisme« og »neurale netværk«). Den ekstremt behavioristiske position har ikke så stor videnskabelig interesse, fordi den hverken giver anledning til teoridannelse i forhold til de fænomener (den menneskelige »intelligente adfærd«), der efterlignes, eller i forhold til et konstrueret overordnet erkendelsesobjekt (såsom »intelligens« eller »kognition« hos dyr, mennesker og maskiner). PDP-eksperimenter kan være relevante for erkendelsen af, hvilken *form for underliggende arkitektur*, der mest effektivt kan »køre« de *matematiske modeller* af formodet *sub-symboliske processer* som fx. visuel perception og de *datamatiske modeller* af *symbolske processer* som fx. analyse (»parsing«) af syntaktiske og semantiske strukturer i naturligt sprog, og vil dermed kunne bidrage til *datamatisk neurovidenskab*. Derimod vil PDP-eksperimenter formodentlig ikke bidrage til erkendelsen af de symbolske processer, blot ved at disse processer efterlignes i deres effekter (som fx. et maskinelt »neuralt netværk«, der trænes til at mønstergenkende orddele på dansk ud fra eksempler⁸), men uden at nogen teori om dét, der efterlignes, er *repræsenteret* i simulationen. I modsætning hertil vil PSS-eksperimenter, i det mindste inden for den kognitionsorienterede AI-forskning, altid implicere en (i bedste fald teoretisk begrundet) model, der er repræsenteret i det program, der simulerer de symbolske processer, som eksperimentet drejer sig om - selv om det ofte kniber med at

opretholde skellet mellem *teorier, modeller og programmer* i AI-litteraturen.

Når vi taler om det alternative paradigme til PSS er det relevant med W.A. Phillips⁹ at skelne mere præcist mellem (a) den teoretiske forskning i *neurale netværks* abstrakte egenskaber, der formuleres i matematiske teorier, (b) *datamatisk neurovidenskab* der analyserer de komputationelle egenskaber ved specifikke neuronale strukturer, (c) *PDP-modeller* af kognitive processer, der udgør en særlig form for »kognitive modellering«, dvs. en form for psykologistisk AI, og (d) *kognitiv neurovidenskab*, der arbejder med det funktionelle neuronale grundlag for de forskellige kognitive moduler, men ikke nødvendigvis på basis af datamatiske modeller.

Fra spillende automater til ekspertsystemer

Interessen for den ekstremt behavioristiske position beror nok også på selve fascinationen af den kunstige efterligning og det automatiske liv, som vi kender det fra automatiske figurer og dukker, der har en lang historie fra antikens »talende« tempelstatuer¹⁰, over 1700-tallets kunstige fugle til vore dages talende dukker. Blot noget tilnærmelsesvis ligner naturligt sprog, hvadenten det kommer fra en legetøjsdukke eller fra en computerskærm, bliver vi nærmest påtvunget en forestilling om, at der er et budskab heri, - vi bliver *suggereret til at tro, at der er nogen, der kommunikerer noget til os*. Det interessante fænomen heri er selve denne suggestive virkning ved de sproglige udtryk og den forestilling, den giver anledning til¹¹, hvorimod den tekniske »indmad« bag dette blændværk ikke nødvendigvis har nogen videnskabelig interesse. Som John A. Campbell påpeger, vil en konsekvent behavioristisk position da også skulle medregne von Kempelen's skakspillende »automat« fra 1769 under AI-systemer, selv om maskinen bestod af en dværg skjult i en kasse¹².

Spil er et godt eksempel på, hvordan AI-miljøet i praksis har forladt behaviorismen. De første succes-historier i AI var udviklingen af skak- og damspillende automater i 1950-erne. Det første velfungerende program til dam (»checkers«) blev udviklet af Arthur Samuel, og det første skakprogram af Alex Bernstein. Begge arbejdede for IBM, da disse programmer vakte opmærksomhed, hvilket var til stor forlegenhed for IBM, der på dette tidspunkt markedsførte den ny teknologi som »elektroniske regnemaskiner« og gjorde alt for at distancere sig fra forestillinger om »tænkende maskiner«. IBM's aktionærer var dels nervøse for, at spillende automater ville give IBM et useriøst ry og dels for, at deres regnemaskiner skulle blive associeret med science fiction litteraturens truende fantasier om robotter, der overtager magten¹³. Men på grund af spil-programmernes relative succes, blev spillende automater anset for en central problemstilling for AI i 50-erne. Allen Newell, J.C. Shaw og Herbert A. Simon skrev i 1958, at »hvis man kunne ud-tænke en vellykket skakmaskine, ville man øjensynlig have gennemskuet

kærnen af den menneskelige forstands anstrengelser¹⁴«. Spil som dam og skak blev prototypiske eksempler på menneskelig problemløsning, som AI-systemerne måtte eftergøre.

I 70-erne er man blevet mere forbeholden over for de spillende automaters placering i AI-projektet¹⁵. Skakspillende automater blev sat i kommerciel produktion og kunne nu købes som almindeligt legetøj på linje med andre computerspil. Skak er et eksempel på, hvordan den interne afgrænsning af AI inden for AI-miljøet, såvel som »humanistiske« forskeres eksterne afgrænsning imod AI-projektet, hele tiden flytter sig relativt til, hvad der faktisk bliver realiseret teknisk og kommercielt.

Uden at komme nærmere ind på udviklingen af spillende automater eller på forsøgene på at realisere en generel intelligens (formuleret som generel problemløsning), vil jeg med Elaine Rich¹⁶ fremhæve en konklusion, som AI-forskningen måtte drage af sine erfaringer i 60-erne. De ringe resultater med at få AI-systemer til at udføre generel problemløsning på basis af generelle algoritmer havde demonstreret: (a) nødvendigheden af at begrænse og kontrollere den søgning, der må indgå i problemløsning, gennem en *heuristisk viden*, - såvel som (b) nødvendigheden af at repræsentere *domænespecifik viden* og generel »*common sense*«-viden i et AI-system, som skal udføre problemløsning inden for et givet domæne (vidensområde).

Det centrale for AI bliver da opfindelsen af *heuristikker* til at begrænse og kontrollere søgningen, opfindelsen af adækvate former for *vidensrepræsentation* og opfindelsen af *interferensmekanismer* til manipulation af disse repræsentationer, hvorimod *søgestrategier i sig selv og rent algoritmisk beregning ikke udgør en del af AI*, men af den almindelige datalogi. Tilbagekuende kan man hævde, at populariteten af spil i den første AI-forskning i 50-erne skyldtes, at *spillende automater som problematik tilsyneladende tillod en udelukkelse af vidensrepræsentations-problematikken*¹⁷. Til gengæld kom denne problematik til at definere kærnen i AI-projektet i 70-erne, hvor *ekspertsystemer* kom til at udgøre det prototypiske eksempel på et AI-system.

Jeg har nu formuleret en begrundelse for, at AI-miljøet i praksis har forladt en behavioristisk opfattelse af AI, og at spillende automater ikke længere ses som prototypiske AI-systemer, nemlig at man praktisk og teoretisk stødte mod problemer omkring *repræsentation af viden; en problematik, der kræver et skift til en mere kognitions-orienteret opfattelse af AI-projektet*. Idet repræsentationsproblemer bliver centrale, er det ikke længere muligt at opretholde en *black box*-opfattelse af AI-systemet, som et system, der blot skal efterligne en menneskelig adfærd, dvs. producere en adfærd, der blot skal »ser intelligent ud«; det er nødvendigt at interessere sig for, *hvordan* viden kan repræsenteres og manipuleres datamatisk, således at de datamatiske processer kan udformes på en intelligent måde.

Hermed bliver modsigelsen inden for AI-projektet mere tydelig: drejer det sig egentlig om *at efterligne en menneskelig intelligens*, hvadenten det nu er som adfærd eller som tænken?

Den anvendte AI's forhold til psykologien: efterligning af menneskelig adfærd og kognition?

I forarbejdet med at konstruere en dam-spillende automat måtte Arthur Samuel konsultere litteraturen om dam, men opdagede, at den enten bestod af abstrakte overvejelser fjernt fra spillets praksis eller af optegnelser af særligt mesterlige spil uden abstraktion af spillets heuristiske principper. Han søgte derfor hjælp hos kendte damspillere, men uden særligt held. En dammester skulle have arbejdet på projektet i et halvt år, men som Samuel fortæller¹⁸: »Vi kom ingen vegne. Jo mere han prøvede at analysere sine egne tankeprocesser, jo mere forvirret blev han, og til sidst forlod han projektet og sagde, at han virkelig var bange for, at han var blevet en dårligere damspiller i forsøget på at tænke over det.«

Samuel var blevet overbevist om, at det ikke var formålet med AI-projektet at efterligne menneskelig tænken. Hans ønske om at få mere at vide om »problemløsningen« hos damspillere var ikke rettet mod en erkendelse af, *hvilken metode mennesker bruger for at løse problemet*, men mod erkendelsen af, *hvori det virkelige problem bestod*¹⁹. Forskellen mellem hjernen og en computer var, ifølge Samuel, for stor til, at man skulle forsøge at efterligne den måde, hvorpå mennesker udfører en given opgave. I forsøgene på at konstruere skakautomater løb man ind i helt parallelle problemer: dataløjerne havde ikke gjort sig klart, hvad det var, de ville opnå ved at interviewe skakspillere om deres tænkning. Drejede det sig om at efterligne denne tænkning, sådan som den var udformet hos nogle skakmestre, eller drejede det sig om at finde ud af, hvilke objektive problemsituationer der opstår i skakspil?

Denne uklarhed gentog sig i 70-erne i forbindelse med konstruktionen af ekspertsystemer. Lad os sige, at det drejer sig om et medicinsk beslutningsstøttesystem. Er »vidensingeniørens« opgave nu at få ekspliciteret nogle lægelige eksperters viden inden for et givet domæne og at fremanalysere deres daglige beslutningsprocedurer og det heuristiske grundlag og den implicite viden bag disse, i håb om at få et AI-system til at simulere *denne* beslutningsproces? Drejer det sig med andre ord om at simulere en ekspert på en *psykologisk adækvat måde*? Eller er »vidensingeniørens« opgave tværtimod at konstruere et system, der kan repræsentere den lægelige viden inden for et givet domæne på en *logisk sammenhængende måde*, således at det alene skal drage logiske slutninger af denne viden, - *uanset om en sådan logisk slutningsproces ikke svarer til den psykiske beslutningsproces*? Mig forekommer det, at et beslutningsstøttesystem kan være en støtte for beslutninger, netop derved at systemet ikke efterligner (»mimer«, »spejler«) den psykiske (eller sociale) beslutningsproces, men giver en hjælp til at overskue de logiske konsekvenser af en stor mængde data og muliggør en simuleret fremskrivning af konsekvenser af hypotetiske beslutninger. Vi ser her problematikken omkring *en psykologisk versus en logisk position* inden for AI, - ikke blot som en ren te-

oretisk problematik, men også som et praktisk problem, der dukker op inden for den anvendte AI.

For at kunne drage nogle konklusioner om AI-forskningens mål og dens forhold til psykologien, er det nødvendigt først at vende tilbage til det »frankensteinianske« udgangspunkt, og se nærmere på den særstilling, som jeg mener robotikken og beslægtede AI-områder har i forhold til teoretisk AI.

Robotikkens særstilling

Den anvendte AI omfatter naturligvis ikke kun spil (som en nedre grænse for AI) og eksperter-systemer, men også områder som *robottik* og de hermed beslægtede områder: maskinsyn, mønster- og billedgenkendelse, talegenkendelse og -syntetisering. Alle disse robotteknikker har en paradoksal placering i forhold til den teoretiske bestemmelse af AI-projektet. Robottik er en veletableret ingeniørvidenskab, der i sit teoretiske indhold snarere er knyttet til *kybernetikken*²⁰ end til AI. Robotikken medtages stadig i nogle introduktioner til AI, som et egentligt AI-område²¹, men teknologisk, teoretisk og videnskabssociologisk tror jeg, at robottik vil blive mere og mere adskilt som et selvstændigt område, og at det vil tage de beslægtede områder med sig, dvs. maskinsyn (eller robotsyn), mønster og billedgenkendelse²² samt genkendelse og syntetisering af tale.

Teknologisk set er robottik en form for anvendt kybernetik på linje med automatisk processtyring og finder sin anvendelse i industrien. Når det samtidig er muligt, at se robottik som en form for anvendt AI, beror det alene på, at man kan hævde, at der indgår AI-teknikker i den datamatiske styring af industrirobotter. De egentlige AI-problemer, der kommer tættest på robottik, angår ikke konstruktionen af robotter, men de almene problemer omkring generering og koordinering af planer for handlinger, hvor disse handlinger er givet som diskrete repræsentationer. Men i robottik såvel som i anvendt kybernetik drejer det sig ikke om *diskrete handlingsrepræsentationer*, men om *kontinueret kontrol af bevægelser*. Og dette er som teoretisk problematik et spørgsmål om kybernetik og klassisk mekanik.

På samme måde vil jeg afgrænse AI fra konstruktionen af systemer til maskinsyn og systemer til *genkendelse og syntetisering af tale*. Denne sidste problematik må ikke forveksles med den centrale AI-problematik, der drejer sig om simulering af naturligt sprog via konstruktionen af systemer til *analyse og generering af sprog*. Den første konstruktionsopgave implicerer i vid udstrækning teoretiske problemer knyttet til matematisk akustik og matematisk beregning (snarere end datamatiske modeller), men i arbejdet med at konstruere sådanne systemer vil man støde møde den *kategorielle perception af fonologiske forskelle*, som et alvorligt teoretisk problem, der angår alle overgange fra *analoge* til *digitale* former for analyse og repræsentation af tale²³.

Der er altså noget, der radikalt adskiller robotik og beslægtede områder fra det AI-projekt, jeg forsøger at indkredse som det centrale. Men hvordan præcisere denne særstilling, som jeg insisterer på at tildele robotikken?

Lad os tage et typisk robotproblem: perceptionen af fysiske objekter af forskellig form og størrelse og koordinationen af robotens bevægelser til fx. at gribe og flytte disse objekter. Dette vil som ingeniørmæssig problematik involvere en lang række beregningsmæssige problemer fx. af differentialgeometrisk art. Sådanne beregninger kan styre robotens bevægelser direkte via forskellige feedbackmekanismer, uden at roboten udfører nogen logiske slutninger inden for en datamatisk model af disse objekter, og uden at den nødvendigvis selv konstruerer en plan for sine bevægelser (objektflutningerne kan fx. være et led i en på forhånd fastlagt sekvens). Den egentlige AI-problematik opstår først, hvis roboten konstrueres således, at den ikke blot udfører matematiske (numeriske) beregninger, men: (a) rekonstruerer datamatiske modeller, der kan repræsentere en viden om verden, (b) drager logiske slutninger ud fra den repræsenterede viden, (c) konstruerer planer for sine handlinger.

Disse punkter er langt fra uproblematisk i sine implicite forudsætninger, men de fremhæver, hvilke typer af opgaver AI-forskere stiller sig. Om et AI-system, der faktisk udfører disse opgaver, er udformet som en robot eller ej, er teoretisk set irrelevant! Terry Winograd designede i sin berømte disputats fra MIT, *Understanding natural Language* (1972), et system (SHRDLU), der kunne modtage, analysere og udføre simple ordrer og forespørgsler om en lukket modelverden, der bestod af en robotarm og forskellige legetøjsklodser. SHRDLU-systemet *simulerer* operationen af robotarmen, der kan manipulere med klodserne, så deres indbyrdes placering ændres. Når Winograd kunne nøjes med at simulere operationen beror det netop på, at selve dens udførelse er irrelevant for de opgaver, som han stillede sig ved at konstruere systemet²⁴: *at benytte datamatiske modeller til at etablere en eksperimentel erfaring om hvilke former for viden, der er nødvendig, for at »forstå« naturligt sprog, - i første omgang i forhold til en lukket og stærkt begrænset modelverden.*

Det er måske ikke helt så enkelt at afgrænse robotik fra teoretisk AI, som fremstillingen ovenfor giver indtryk af. I 80-erne er der således opstået en ny disciplin, *datamatisk geometri* (»computational geometry«)²⁵, der søger en systematisk beskrivelse af de algoritmer og heuristikker, der knytter sig til datamatiske modeller for bevægelse i rummet (i forhold til objekter etc.) og konstruktionen af planer herfor, hvorved robotik i nogen grad kan løsrides fra de kybernetiske og kontrolteoretiske problemer og formuleres som en ren datalogisk disciplin. Der kan da være tale om en disciplin, der udgør en del af teoretisk AI, hvis man indfører et skel mellem:

- (1) datamatiske modeller, der simulerer *sub-symboliske processer*, og
- (2) datamatiske modeller, der simulerer *symboliske processer*,

og man endvidere anerkender, at (1) også indgår i AI-projektet (og ikke blot i PDP-modeller inden for CS), selv om kun (2) giver anledning til en egentlig

videnspræsentationsproblematik.

Når der ikke uden videre er tale om repræsentation af viden ved simulering af sub-symboliske processer, så beror det på, at denne simulering kan finde sted som en ren *numerisk beregning*. Således benytter den klassiske mekanik differentiaalligninger til at beskrive legemers bevægelser i rummet. Disse ligninger repræsenterer for fysikeren en viden om de almene bevægelseslove og konkret en viden om et bestemt mekanisk system. Hvis et computerprogram beregner bevægelserne i det konkrete mekaniske system, udfører det en *analog simulation* af systemet. Systemet er derimod ikke *digitalt repræsenteret* i programmet på en sådan måde, at man kan tale om videnrepræsentation. Dette ville kræve, at den fysiske viden om systemet, blev formuleret på en anden måde, som ville sætte programmet i stand til at drage *kvalitative slutninger* om systemets bevægelser, snarere end at beregne dem. Sådanne formuleringer af fysisk viden er under udvikling i *kvalitativ proces-teori*²⁶ og i »*naiv fysik*«-forskningsprogrammet, der søger at formalisere hverdagslig »common-sense«-viden om *virtuelle begivenheder* i den fysiske verden (som fx. de mulige måder væsker kan bevæge sig på i forhold til flader og rum)²⁷. I forhold til disse AI-teorier kan robotsystemer komme til at fungere som en eksperimentel udfordring: robottik kan benyttes til at konfrontere AI-modellerne med en *aktuel verden*²⁸. Men der er da tale om en AI-forskningsmetodologi snarere end om robottik som anvendt AI. Det er således stadigvæk muligt at argumentere for, at implementeringen af AI-systemer i en eller anden form for robotversion er irrelevant for det teoretiske indhold af datamatiske modeller inden for AI, og det er derfor misvisende, når Eugene Charniak & Drew McDermott definerer det endelige mål for AI som »at bygge en person, eller mere beskedent, et dyr«.

Resumerende kan vi fastslå, at AI både kan defineres adfærdorienteret, som simulering af intelligent adfærd, og kognitivt, som simulering af »kognitive opgaver«, - men i alle tilfælde er der tale om en *simulering* ved hjælp af *symbolsystemer, heuristikker og datamatiske modeller*. Foruden disse grundbegreber er *vidensrepræsentation* blevet den centrale problematik for den del af AI, der arbejder med *simulering af symbolske former for kognition*. AI omfatter også *simulering af sub-symboliske former for kognition*, sådan som områderne robottik og maskinsyn demonstrerer. Men der er et problem heri, nemlig at ren matematisk analyse og rent algoritmiske beregninger ikke bliver til AI, blot fordi de udføres af en computer. Enhver brug af computere ville i så fald være anvendt AI. Hvis AI overhovedet kan formuleres som en disciplin med *et konsistent teoretisk indhold*, må vi søge en afgrænsning af AI-projektet, der ikke er over-inklusiv, med henblik på at søge at fundere *en kerne af centrale begreber og teoretiske hypoteser*. Dette indebærer også en afgrænsning af teoretisk AI fra anvendt AI (eksperysystemer, robottik etc.). En sådan kerne er først ved at blive etableret i AI-miljøet, der hidtil har været præget af manglende konsensus om forskningsresultater. Historisk er AI endnu en før-videnskabelig disciplin (»protovidenskab« i Campbell's termi-

nologi). Et afgørende led i en videre afklaring af AI-projektet vil være diskussionen af dets forhold til psykologien og kognitions-forskningen (CS). Denne diskussion kan vi komme på sporet af ved at spørge, hvad det er, AI-forskerne erkender noget om, når de konstruerer datamatiske modeller og simulerer symbolske og sub-symbolske processer.

Hjernen som en computer: en misrepræsenteret problematik

I Eugene Charniak & Drew McDermott's introduktion til AI hedder det, at »kunstig intelligens er studiet af mentale evner (mental faculties) gennem brugen af datamatiske modeller«²⁹. Men hvorfor skulle brugen af datamatiske modeller fortælle os noget om »mentale evner«? Charniak & McDermott hævder selv, at deres definition hviler på *en antagelse om, at hjerner fungerer »ligesom« computere*. Men denne antagelse er, som de selv fremhæver, tilsyneladende forkert: hjernens funktion kan i almindelighed ikke adskilles fra organisationen af det levende materiale («wetware»), der gennem millioner af års evolution har udviklet mange andre egenskaber, end de egenskaber som computerens »hardware« er bundet til (på et tilfældigt tidspunkt i den kulturhistoriske udvikling af denne teknologi). Der er en omvendning på spil her, som er typisk for AI-forskningen: natur- og kulturhistorisk ville den rimelige antagelse om forholdet mellem hjernen og computeren jo være, at det omvendt er AI-miljøet, der forsøger at udvikle en teknologi, der på visse begrænsede områder ser ud til at fungere lige som en hjerne. Hjernevarer her trods alt først! På den anden side har hjernens funktion været relativt ukendt, mens computeren er et velbeskrevet objekt, så i erkendelsen af hjernens funktion, har man søgt at elaborere *computermetaforen* til at udgøre en egentlig *begrundet analogi*, der kan danne udgangspunkt for teoretiske hypoteser om hjernens funktion. Men vi taler i så fald om datamatisk neurovidenskab og ikke om AI!

Charniak & McDermott opretholder omvendingen, men modificerer deres antagelse til den fundamentale arbejdshypotese, at »det som hjernen gør, på et eller andet niveau, kan anskues som en form for kalkulation (:computation)³⁰«. Dette er *den funktionalistiske antagelse*, - men i en bestemt udformning. I den udstrækning denne antagelse er korrekt vil det, ifølge Charniak & McDermott, være frugtbart at konstruere datamatiske modeller af forskellige »mentale evner«, som vi forsøgsvis vil fortolke som faktiske psykiske kompetencer, sådan som de beskrives i psykologien. Det kan nu se ud som om, at AI dermed bliver en del af psykologien, men dette afviser de. Den eksperimentelle psykologi søger empirisk at bekræfte eller afkræfte sine hypoteser om *hvordan hjernen og »det psykiske« fungerer*, hvorimod AI er tilfreds med, *at et program virker*, og at det blot *benytter én ud af mange mulige måder, hvorpå en given »intelligent adfærd« kan produceres*³¹. Heraf må man slutte, at et AI-program blot *simulerer* den psykiske kompetence (den

»mentale evne«), som antages at ligge bag en manifesteret »intelligent adfærd« hos mennesker eller andre dyr. Hermed bliver det imidlertid også uklart, hvordan datamatiske modeller forholder sig til psykologisk empiri. Den eksperimentelle psykologi vil øjensynlig ikke kunne falsificere AI-modeller af en given kompetence, for så vidt disse modeller ikke er bundet af den måde, hvorpå den tilsvarende psykiske kompetence er udformet på hos fx. mennesker. Omvendt vil AI-eksperimenter måske kunne falsificere psykologiske modeller af psykiske kompetencer, for så vidt disse modeller er udformet som datamatiske modeller, hvis »funktionalitet« da kan undersøges eksperimentelt på en datamaskine. Imidlertid er sådanne eksperimenter jo begrænset af og bundet til de foreliggende maskinarkitekturer og programmeringsprog; så blot fordi en bestemt implementering (en bestemt fysisk realisering) af den datamatiske model ikke virker, kan man ikke slutte, at modellen ikke er en adækvat model for den givne psykiske kompetence. For at gennemføre en sådan slutning, måtte man enten vise, at enhver implementering af modellen vil fejle, dvs. at den er *logisk umulig at realisere* (fordi den er inkonsistent), eller vise, at modellen er *uforenelig med det, vi ved om hjernens funktion*. Det sidste ville være en opgave for *neurovidenskaberne*, men forudsætter formodentlig en viden om *koblingen mellem et kalkulatorisk niveau og et neurofysiologisk niveau af hjernens virkomsghed*, som vi endnu ikke har, men som det er målet for den datamatiske neurovidenskab og for den teoretiske forskning i neurale netværk at tilvejebringe. Det første kunne være en del af AI-projektet, for så vidt dette ikke begrænses af og bindes til *udforskningen af de psykiske kompetencer*, sådan som de er realiseret hos mennesker eller andre dyr, men til *udforskningen af de logisk mulige udformninger af »mentale« kompetencer i en mere abstrakt forstand*.

Phillipe J. Hayes har søgt at bestemme enheden og forskellen mellem AI og psykologi ved at hævde, at de begge søger »at forklare menneskelig rationel adfærd« men blot har forskellig forskningsmetodologi. AI søger den samme erkendelse som den eksperimentelle kognitionspsykologi, og udgør også en empirisk videnskab, men den kræver ikke en detaljeret overensstemmelse med psykologisk empiri, men søger at teste forklaringer og teorier gennem kørsler af programmer³². Phillippe Hayes fremhæver, at AI kun kan godkende teorier og forklaringer, der er datamatiske. Han undgår at sammenblande teorier og programmer, men den rent metodologiske adskillelse af psykologi og AI er helt utilstrækkelig. AI har efter min mening slet ikke det samme erkendelsesobjekt som den eksperimentelle psykologi, hvilket jeg igen vil fremhæve ved at hævde, at teoretisk AI drejer sig om *udforskningen af de logisk mulige udformninger af abstrakte »mentale« kompetencer gennem konstruktionen af datamatiske modeller, der simulerer sådanne kompetencer*. Med formuleringen »logisk mulige udformninger« vil jeg fremhæve, at AI-modeller ikke nødvendigvis er underlagt de tvangsbindinger (constraints), der gælder for »det psykiske apparat«, sådan som det faktisk er udformet hos mennesker eller andre dyr ifølge den viden, vi har fra etologien,

neurofysiologien, kognitionspsykologien etc. Jeg vil foreslå, at det tværtimod drejer sig om *at abstrahere fra denne viden for at kunne analysere de almene* (»universelle«) *begrænsninger for, at systemer af enhver art kan danne billeder, danne betydninger, danne logiske slutninger etc., udveksle disse indbyrdes* (:»kommunikere«), *handle i forhold til en verden etc.* Med denne formulering vil jeg søge at beskrive et fælles grundlag for AI og CS, der må ses som *meta-semiotisk*. Det specielle ved AI er den eksperimentelle metodologi i forhold til denne almene problematik, dvs. konstruktionen af datamatiske modeller af sådanne »mentale« fænomener. Et eksempel fra den mere spekulative AI-forskning kan illustrere, hvad jeg mener med »universelle begrænsninger«. I en artikel, hvis tema nærmest må betegnes som »exo-semiotik«³³, søger Marvin Minsky at deducere sig frem til en række universelle begrænsninger for levende systemer af enhver art, - ud fra den ide, at »alle intelligente problemløsere i sidste ende er underlagt de samme tvangsbindinger, - begrænsninger på rum, tid og stof. For at kunne udvikle effektive måder at omgås med disse tvangsbindinger på, må de have måder at repræsentere de situationer de møder, og de må have processer, der kan manipulere disse repræsentationer.«³⁴

Lad os vende tilbage til Charniak & McDermott's position. De tager ikke konsekvensen af deres egen skelnen mellem de *logisk mulige* og de *psykisk realiserede* udformninger af de kompetencer, der kan generere eller producere »intelligens«: for hvis AI udforsker de logiske mulige modeller heraf, der desuden er datamatiske, hvorfor skulle dette så samtidig være et studium af de »mentale evner« forstået som *psykiske* kompetencer? AI må have et erkendelsesobjekt, der helt eller delvist er forskelligt fra psykologiens.

Denne inkonsekvens og den teoretiske uklarhed, som den forårsager, beror på den særlige udformning af *funktionalismen*, som Charniak & McDermott (og mange andre) opererer med, nemlig den fysiologiske reduktionisme. Termen »funktionalisme« betegner ikke det samme inden for fx. antropologi og psykologi, men set i relation til kognitionsforskning kan man ifølge Ned Block opregne tre forskellige betydninger: (a) funktionalisme som en forskningsstrategi, der søger en bestemt type af forklaringer, hvor et systems egenskaber forklares ud fra delenes funktion og indbyrdes organisation; (b) funktionalisme som en teoretisk antagelse om at det psykiske kan forklares »mekanistisk«, som forskellige former for beregninger, der udføres på repræsentationer; (c) funktionalisme som en metafysisk antagelse om at psykiske tilstande er funktionelle tilstande³⁵.

Funktionalismen som teoretisk antagelse kan siges at give videnskaber som psykologi og lingvistik et teoretisk råderum. Funktionalismen reducerer hverken det psykiske til bevidsthedsindhold eller til fysiologiske processer i hjernen, men opretter *et »mellemliggende niveau«* for beskrivelse og forklaring. Det er derfor en utidig amputering af kognitionsvidenskaben at udforme funktionalismen som en form for fysiologisk reduktionisme, der søger at »udtømme« det kognitive i analogien mellem hjernen og computeren.

Funktionalismen er en antagelse om det psykiske og ikke om hjernen³⁶. Hjernen som en computer er derfor en misrepræsenteret problematik, når den fremføres som det generelle grundlag for CS og AI, men ikke når den formuleres som et specifikt videnskabeligt problem, som det sker i datamatisk neurovidenskab.

Tegn, redskaber og modeller

Jeg vil nu gå tilbage til analogien hos Alan Bundy mellem AI og botanikken på Linnés tid. Bundys katalogiseringsarbejde i forhold til AI-redskaber synes virkelig at være en art botanisering i AI-miljøet, hvor de forskellige foreliggende redskaber først må beskrives side om side, før man kan gøre sig håb om at give mere strukturelle beskrivelser af de forskellige mulige redskabers funktionalitet og indre forbundethed.

Etableringen af biologien som videnskab måtte gennem en tilsvarende overgang fra botanisk og zoologisk klassifikation af de *eksisterende* individer til den biologiske beskrivelse af de *mulige* enheders (typer af celler, organer, individer, økologiske nicher) funktionalitet og forbundethed³⁷. Biologi som teoretisk videnskab blev først etableret, idet man kunne bevæge sig fra klassifikationen af konkrete individer og deres rent formelle forbundethed ifølge en taxonomi til opstillingen af hypoteser om livets væsentlige egenskaber. Disse er i første omgang blevet formuleret som *funktionelle egenskaber* på forskellige niveauer (bl.a. som fysiologi og evolutionsteori for celler, organer, individer og økologiske nicher). Efterhånden som biologien er blevet solidt etableret som videnskab, fremstår dens mest almene erkendelser som en *teoretisk biologi*, der blandt andet søger at formulere en art *topologisk forbundethed* mellem form og funktion af de forskellige *moduler*, hvorved livet *består*: på tværs af forskellige arter og på tværs af forskellige typer af celler, organer etc., beror livets funktionalitet på det forhold, at forskellige fysiske former og strukturer, kan *realisere forskellige »abstrakte« maskiner og redskaber*³⁸. På dette punkt viser det sig nu, at den teoretiske biologi og den AI-inspirerede kognitionsforskning (CS) har delvist sammenfaldende forskningsprojekter! Dette er grunden til, at jeg her har skitseret denne analogi mellem klassifikation og teoretisering af AI-»redskaber« og naturens »redskaber«, selv om vi endnu ikke ved, hvad den betyder.

Analogien har imidlertid længe haft praktisk betydning for ingeniørvidenskaberne, der ofte har søgt at udnytte naturens former og problemløsninger (som fx. i Leonardo da Vincis berømte studier af fuglevinger og mulige flyvemaskiner). I vort århundrede er *bionikken* blevet formuleret som det sammenlignende ingeniørvidenskabelige studium af designprincipper, optimalitetsprincipper o.lign. i levende og maskinelle systemer³⁹. Nu hvor man inden for flere videnskaber søger at formulere rammerne og grundlaget for en mere almen kognitionsforskning, er det værd at være opmærksom på, hvordan problematikkerne inden for bionik er systematisk forbundne med be-

grebet om »optimal design« af organiske former og funktioner inden for *teoretisk biologi*⁴⁰ og optimalitets-begreber inden for kybernetik og systemteori.

Bionik og AI deler mange problemstillinger, men her skal jeg blot fremhæve spørgsmålet om efterligning. Jævnfør Leonardo da Vincis fuglestudier, så lykkedes det jo først at konstruere flyvende maskiner, idet man ophørte med at efterligne fuglenes vinger. Noget tilsvarende har gjort sig gældende i forsøget på at udvikle »tænkende« maskiner - det gælder ikke om at efterligne tænkning i dens psykologiske udformning eller i dens neurologiske grundlag, men om at simulere de *gyldige slutninger*, der kan tænkes om et eller andet, dvs. det drejer sig om *logik*. Dette er, hvad man vil hævde fra en logicistisk position, hvor kognitionsforskning (CS) og AI ikke synes at have meget med hinanden at gøre. Men bionik-eksemplet fortæller os noget andet. Selv om det er en væsentlig pointe, at flyvemaskiner ikke er konstrueret således, at de efterligner fuglens måde at flyve på, så er flyvemaskiner og fugle *udformninger af en almen mulighed*, en kunnen flyve, der beror på de samme aerodynamiske principper. *Kognitionsforskningen bør epistemologisk placeres på et tilsvarende niveau for tænkningens almene mulighed*, selv om den også må have en særlig forpligtelse til at erkende den »biologisk implementerede intelligens«⁴¹ i sine psykologiske udformninger og i sit neuronale grundlag.

Idéen om de organiske former og funktioner som realiseringer af abstrakte redskaber og maskiner leder på flere måder frem til kognitionsforskningen. Ikke blot af epistemologiske grunde men også af natur- og kulturhistoriske grunde. Det er vigtigt også at fremhæve *kulturhistorien* som en væsentlig kilde til kognitionsvidenskabelige problemer, idet der øjensynlig foregår en kamp om rækkevidden af den tværvideenskabelige tilgang til CS: på den ene side sker der en naturvidenskabelig opstramning omkring kærnediscipliner som filosofisk logik, lingvistik, neurovidenskab og eksperimentel psykologi og på den anden side forsøger enkelte forskere at holde CS-forskningsprogrammet åbent over for problemstillinger inden for antropologi, semiotik, sociologi etc.⁴². Jeg gør mig ikke til talsmand for, at CS er en humanistisk disciplin, men for at det er for tidligt at lukke problemstillingen i CS omkring logiske, psykologiske og neurologiske problemstillinger.

Den kognitive virksomhed hos mennesker er en virksomhed, der (a) kulturhistorisk har frembragt konkrete redskaber (herunder maskiner) og modeller, der godt nok er historiske »opfindelser«, men som også er en form for »genopdagelse« af abstrakte redskabs- og model-egenskaber, som naturen allerede har realiseret i de organiske formers funktionalitet, - og (b) er sammenvævet med den kommunikative virksomhed, der naturhistorisk beror på dannelsen af særskilte *tegn-funktioner* (der er afledt af tidligere former for funktionalitet ved mekanismer som fx. »ritualisering« og »mimicry«⁴³) og på den videre udvikling af egentlige sproglige funktioner.

Metasemiotik og »naiv fysik«

Hermed aftegner der sig muligheden af en almen analyse af tegn, redskaber og modeller i deres indre forbundethed. En sådan analyse vil jeg betegne som *meta-semiotisk*⁴⁴, idet *semiotik* er den almene teori for *betydningsdannelse og kommunikation* (men formuleret som en almen tegnteori)⁴⁵. Det er min hypotese her, at meta-semiotik kan tjene som en overordnet begrebsdannelse, der vil kunne etablere en dialog mellem de forskellige (mere snævert definerede) paradigmer i kognitionsforskningen.

Det er naturen, der har *frembragt* de forskellige former for abstrakt funktionalitet, der knytter sig til tegn, redskaber og modeller, og *realiseret* dem (fysisk-biologisk) i de levende systemer, længe inden vi i kulturhistorien »opfinder« sprogtegn, redskaber og modeller som noget, der kan *eksternaliseres*⁴⁶ som selvstændiggjorte konkrete funktioner og objekter og dermed gøres til genstand for *brug*. Kulturhistoriske opfindelser som fx. kruset eller kniven udgør i en vis forstand *genopdagelser af forskellige former for nyttighed* (»brugsværdi«), hvis funktionelle grundlag allerede eksisterede i naturen. Når vi har kunnet »finde på« at *fremstille* sådanne *genstande* beror det på, at vi har kunnet *forestille* os de *sagsforhold*, der knytter sig til deres *brug* og dermed til deres form og funktion. Objekternes generelle fremstillelighed og forestillelighed hænger på denne måde snævert sammen på et dybere (ontologisk) niveau⁴⁷, hvor objekternes *tingslighed* beror på deres topologiske egenskaber. Krusets tingslighed kan partielt beskrives som en »kunken rumme« og knivens tingslighed partielt som en »kunken skære«, hvor disse to egenskaber dels beror på noget rent topologisk og dels på verdens fysiske beskaffenhed, som gør, at der findes faseforskelle mellem fast og flydende (så faste hule genstande kan rumme væsker) og forskelle i relativ hårdhed (så hårde skarpe genstande kan skære i relativt blødere genstande).

Man er ved at genopdage denne type af problemstillinger inden for AI-forskningen, idet man støder mod vanskelighederne ved at simulere forståelse og generering af naturligt sprog. Mange af disse vanskeligheder beror på vores vanskelighed ved at eksplicittere og formalisere vores umiddelbare forståelse af semantiske sammenhænge og vores generelle erfaring af og viden om verden, som er forbundet hermed.

I det såkaldte *naiv-fysik*-forskningsprogram søger Patrick Hayes, Jerry Hobbs og andre at eksplicitte og formalisere »common sense«-viden om (især) den fysiske verden⁴⁸, men der kan i dette projekt være en indre modsigelse mellem *en psykologisk og en logistisk opfattelse af AI*. Er formålet med naiv fysik at lave en *kognitiv modellering* af menneskers (mis-)repræsentationer af fysiske forhold i dagligdagen sammenlignet med den teoretiske fysik? Dette er den psykologistiske position inden for »cognitive modeling«⁴⁹, der falder godt i tråd med psykologiens oprindelige bestemmelse som en disciplin, der skulle søge af (bort-)forklare, hvorfor verden ikke »opleves« af erfaringens subjekt således som den »virkelig er« ifølge den ra-

tionelle erkendelse, dvs. den fysiske beskrivelse af verden⁵⁰. Eller er formålet med naiv fysik at rekonstruere de *regionale ontologier* for forskellige områder af den *fænomenale fysik*, der både ligger til grund for den hverdagslige og den videnskabelige beskrivelse? Dette er den logicistiske (og den meta-semiotiske) position⁵¹.

En logicistisk opfattelse af AI og en meta-semiotisk opfattelse af CS vil godt kunne støtte hinanden i et fælles forskningsprogram, som måske vil kunne føre kognitionsforskningen ud over logik/psykologi-modsigelsen. Semiotik er i traditionen fra Peirce (den analytiske semiotik) netop ækvivalent med logik i en udvidet forstand, en tegn-logik, og semiotikken hos Peirce er på en måde allerede en almen kognitionsvidenskab.

Kognitionsforskningen står over for tre store udfordringer. Den ene er *overvindelsen af logik/psykologi-modsigelsen*, - både i forhold til formuleringen af en epistemologisk adækvat position for CS og i forhold til den fundamentale vanskelighed med denne modsigelse inden for semantikken⁵². Den anden er *reformuleringen af den fænomenale verdens status* inden for videnskabelige teorier om kognition, således at de »naive teorier« om den fysiske og den sociale verden ikke søges bortelimeret, men tværtimod integreret som en væsentlig del af CS⁵³. Og den tredje afgørende udfordring for CS er ophævelsen af paradigme-striden mellem PSS- og PDP-paradigmet for i stedet at se denne modsætning som et index på en reel problematik, der angår digitale og analoge former for repræsentation i kognitive processer. Udfordringen her vil teoretisk set bestå i *formuleringen af mulige overgange mellem diskrete og kontinuerte modeller af symbolske og sub-symbolske processer*⁵⁴. De seneste udviklinger inden for »kognitiv semantik« (George Lakoff, Ronald Langacker, Leonard Talmy, Eve Sweetser o.a.) og inden for den »modale semiotik« (Jean Petitot, Per Aage Brandt) hvor betydningens formdannelse søges funderet i en dynamisk topologi⁵⁵, kan på alle tre punkter bidrage til kognitionsforskningen og til funderingen af dens meta-semiotiske grundlag, men dette ligger uden for rammerne af denne artikel.

NOTER

1. Eugene Charniak & Drew McDermott: *Introduction to Artificial Intelligence* (1985, Addison-Wesley), p. 7.
2. John A. Campbell: »On artificial intelligence«, *Artificial Intelligence Review* (1986), Vol. 1, No. 1, p. 3-9. Og jvf. John A. Campbell: »Principles of Artificial Intelligence«, i: Masoud Yazdani (e.): *Artificial Intelligence: Principles and applications* (1986, Chapman and Hall).
3. Datamatisk er den danske betegnelse for det engelske »computational«, der egentlig betyder »beregnet«. Etymologisk kommer »computational« fra latinsk computare, der betyder »tælle«, »beregne«; afledt fra putare, der betyder »beskære«, »rense«, »korrigere«. Datalogi er den danske betegnelse for det engelske »computer science«.

4. Alan Bundy (ed.): *Catalogue of Artificial Intelligence Tools* (1986, Springer), 2. rev.ed.
5. Avron Barr & Edward A. Feigenbaum (eds.): *Handbook of Artificial Intelligence*, Vol. 1 (1981, William Kaufmann), p.xi, 3.
6. Allen Newell: »Physical Symbol Systems«, *Cognitive Science* (1980), vol. 4, p. 135-183.
7. Elaine Rich: »Artificial Intelligence«, p. 9, i: S.C. Shapiro (ed.): *Encyclopedia of Artificial Intelligence* (1987, Wiley).
8. Et sådant »neuralt netværk« er konstrueret herhjemme af Søren Brunak og Benny Lautrup og præsenteret ved DAIS-mødet på RUC d. 6.5.1988.
9. W.A. Phillips, oplæg ved forskerkursus i »Filosofiske perspektiver i konnektionisme« (Odense 1988).
10. Jvf. Richard L. Gregory: *Mind in Science* (1981), Kap. 2.
11. Forestillingen om, at der er nogen (: et subjekt), som kommunikerer noget (: en viden) til os, så snart noget sprogligt (noget der er »lige som« sprog eller slet og ret noget, som vi kan opfatte som *tegn*) bliver manifesteret, så at sige »præsenteret«, for os udefra, er en fundamental problematik i såvel semiotikken som i psykoanalysen.
Det er i en sådan sammenhæng, at man må fortolke Joseph Weizenbaum's berømte erfaringer med sit ELIZA-program, der indgød mange brugere en illusion om, at de faktisk talte med en forstående »psykiater«, som de endog helst ville »tale med« i enrum, på trods af, at de vidste, at ELIZA var et maskinelt implementeret program. Jvf. Joseph Weizenbaum: *Computer Power and Human Reason* (1976), (1984, Penguin Books).
12. Jvf. John A. Campbell: »Three uncertainties of AI«, p. 250 i: Masoud Yazdani & Ajit Narayanan (eds.): *Artificial Intelligence: human effects* (1984, Ellis Horwood). Baron von Kempelen's automat er afbildet uden på *Psyke & Logos* 1984, N. 2 (Tema: Menneske og maskine).
13. Jvf. Pamela McCorduck: *Machines Who Think. A Personal Inquiry into the History and Prospects of Artificial Intelligence* (1979, W.H. Freeman), Ch. 7.
14. A. Newell, J.C. Shaw, H.A. Simon: »Chess Playing Programs and the Problem of Complexity«, *IBM Journal of Research and Development* 1958, Vol. 2, N. 4. Her cit. fra P. McCorduck 61979), op.cit., p. 160.
15. Jvf. J.A. Campbell: »Three uncertainties of AI«, i: Masoud Yazdani o.a. (1984), op.cit.
16. Jvf. Elaine Rich: »Artificial Intelligence«, i: S.C. Shapiro (ed.): *Encyclopedia of Artificial Intelligence* (1987, Wiley); For en mere positiv vurdering af GPS se Pamela McCorduck (1979), op.cit., Ch. 10.
17. Jvf. Elaine Rich (1987), op.cit., p. 9-10.
18. Pamela McCorduck (1979), op.cit., p. 152.
19. Pamela McCorduck (1979), op.cit., p. 152.
20. Kybernetikken blev oprindeligt formuleret af Norbert Wiener i *Cybernetics: Control and Communication in the Animal and the Machine* (1948).
21. Således hos Patrick Henry Winston: *Artificial Intelligence*, 2.ed. (1984, Addison-Wesley), Ch. 10.
22. Datamatisk billedbehandling, computergrafik og lignende må anses for en del af den almindelige datalogi - hverken en del af AI eller af robotikken.
23. Jvf. Stephen Isard: »Levels of representation in computer speech syntheses and recognition«, in: Masoud Yazdani (ed.): *Artificial Intelligence. Principles and applications* (1986, Chapman and Hall). En teori om den kategorielle perception af fonologiske forskelle (baseret på den matematiske katastrofeteori som formuleret af René Thom) er fremsat af Jean Petitot-Cocorda i *Les catastrophes de la parole* (1985, Maloine).

24. Jvf. Terry Winograd: »A Procedural Model of Language Understanding«, i: R. Schank & K. Colby (eds.): *Computer Models of Thought and Language* (1973, W.H. Freeman); Terry Winograd: »What does it mean to understand language?«, *Cognitive Science* 1980, Vol. 4, p. 209-241.
25. Jvf. Chee-Keng Yap: »Preface: Special Issue on Robotics«, *Algorithmica* 1987, Vol. 2, N. 4, p. 363-365.
26. Jvf. Kenneth D. Forbus: »Qualitative process theory«, *Artificial intelligence*, 1984, N. 24, p. 85-168.
27. Jvf. P.J. Hayes: »The Second Naive Physics Manifesto« og »Naive Physics I: Ontology for Liquids«, in: Jerry R. Hobbs & Robert C. Moore (eds.): *Formal Theories of the Commonsense World* (1985, Ablex).
28. Jvf. J. Michael Brady: »Artificial Intelligence and robotics« (1985), i: Masoud Yazdani: *Artificial Intelligence. Principles and applications* (1986, Chapman and Hall), Ch. 7.
29. Eugene Charniak & Drew McDermott: *Introduction to Artificial Intelligence* (1985, Addison-Wesley), p. 6.
30. Det er vigtigt for den begrebslige præcisering at holde det åbent, hvilken *form for beregning*, der er tale om. Det er derfor også vigtigt ikke at lukke for problemstillingen ved at give en fortolkende oversættelse af »computation« som fx. informationsbehandling, symbolbehandling eller databehandling. Jeg foretrækker derfor at oversætte »computation« bogstaveligt som beregning eller *kalkulation* undtagen i sammenhænge, hvor der findes en hævdevunden term som fx. datalingsvistik for »computational linguistics«.
31. Jfr. Eugene Charniak & Drew McDermott: *Introduction to Artificial Intelligence* (1985, Addison-Wesley), p. 6-7.
32. Jvf. Phillip J. Hayes: »On the difference between psychology and AI«, i: Masoud Yazdani & Ajit Narayanan (eds.): *Artificial Intelligence: human effects* (1984, Ellis Horwood).
33. Dvs. Semiotik for liv i universet og for kommunikation med liv uden for jorden.
34. Jvf. Marvin Minsky: »Communication with alien intelligence«, *Byte* 1985 (Tema: Artificial Intelligence), Vol. 10, N. 4.
35. Jvf. Ned Block: »Introduction: What is Functionalism?«, i: Ned Block (ed.): *Readings in the Philosophy of Psychology* (1980, Methuen), Vol. 1.
36. Dette er i den filosofiske debat omkring CS især blevet fremhævet af Hillary Putnam og Jerry A. Fodor, der begge afviser at identificere typer af »mentale tilstande« med fysiske tilstande i hjernen. Jvf. Ned Block (1980), op.cit. og Ned Block & Jerry A. Fodor: »What psychological states are not«, *The Philosophical Review* 1972, Vol. 81, N. 2 (også i J.A. Fodor: *Representations. Philosophical Essays on the Foundations of Cognitive Science* (1982, Harvester Press), ch. 3.
37. Jvf. René Thoms fremstilling af disse overgange fra empiriske og klassificerende beskrivelser af givne former (som fx. Linnés beskrivelser af planternes empiriske morfologi), over reduktionistiske årsagsforklaringer og funktionelle beskrivelser (som fx. genetikken fra Mendel til molekylærbiologien) til strukturelle beskrivelser af formernes »dynamik« (som fx. i C.H. Waddingtons teoretiske embryologi) i videnskaber som fysik, biologi og lingvistik. Se »La linguistique, discipline morphologique exemplaire«, *Critique* (1974), N. 322.
38. Et eksempel herpå giver teorien for »cellulære automater«, hvis udgangspunkt er von Neumanns overvejelser (i 1950-erne) over muligheden af teoretisk at beskrive selvreproduktionen som en egenskab ved abstrakte maskiner (automater).
39. For en mere præcis definition, se Horst Heynert: *Grundlagen der Bionik* (1976).
40. Optimalitetsprincipper i naturens »design« blev først formuleret som teoretisk problematik af D'Arcy Thompson i *On Growth and Form* (1917) og Nicolas Rashevsky i *Mathematical Biophysics* (1938). Begge disse foregangsmænd inden for den teore-

tiske biologi søgte at fundere en teori om organiske former på fysiske principper. Den teoretiske biologi blev etableret som et selvstændigt forskningsområde i 70-erne med værker som C.H. Waddington (ed.): *Towards a Theoretical Biology* Vol. I-IV (1968-1972, Edinburgh U.P.) og Robert Rosen (ed.): *Foundations of Mathematical Biology* Vol. I-III (1972-1973, Academic P.).

41. Ifølge FAST-rapporten til EF om »Cognitive Science in Europe« (1986) ses opgaven for CS netop som studiet af »biologisk implementeret intelligens« og dermed som sideordnet med den »kunstigt« implementerede intelligens (AI). Herved er der en fare for, at CS fortaber sig i en ren psykologistisk position og overser muligheden for at udforme forskningsprogrammet inden for rammerne af spørgsmålet om tænkningens almene mulighed (hvor rammerne for CS bliver sammenfaldende med meta-semiotikkens forskningsprogram).
42. Fremhævet i rapporten »Kognitionsforskning i Danmark« fra det Hummanistiske Forskningsråds FTU-udvalg (1988), genoptrykt i B. Karpatschof (ed.): *Artikelsamling til informationspsykologiens videnskabsteori* (1988, Psyk. Lab.).
Opstramningen kan fx. aflæses i en lærebog som Neil A. Stillings o.a. (eds.): *Cognitive Science* (1987, MIT), mens forsøget på at holde CS åbent kan aflæses i værker som Howard Gardner: *The Mind's New Science: a History of the Cognitive Revolution* (1985) og Richard Gregory: *Mind in Science. A History of Explanation in psychology and Physics* (1981).
43. Ved *ritualisering* forstås en evolutionær proces, hvorved en allerede eksisterende instinkuel aktivitet (hos fugle fx. redebygning og yngelpleje) bliver »frigjort« fra sin oprindelige kontekst og sine oprindelige udløsningsmekanismer og »typificeret« i sin form og intensitet, således at *aktiviteten omfunktioneres som et abstrakt redskab* (hos fugle fx. som et redskab for synkronisering af parringsaktiviteten). Ritualisering er således et begreb for den evolutionære *afledning af naturlige tegn* og opkomsten af kommunikativ adfærd. *Mimicry* betegner en anden form for afledning af naturlige tegn, hvor en allerede etableret aktivitet eller organisk form hos én art efterlignes af en anden art, som da *lokalt bliver en model af denne aktivitet eller form*. *Mimicry* opstår især omkring social kommunikation mellem arter (som ved advarselskrig i en fuglekoloni) eller som simulation og dissimulation i jagt/byttedyr-systemer. Ritualisering blev først beskrevet af Julian Huxley, men teorien er udviklet af Niko Tinbergen og Desmond Morris i 1950-erne, jvf. D. Morris: *Patterns of Reproductive Behaviour* (1970).
44. Som her formuleret er meta-semiotikken inspireret af James Bunnns spredte overvejelser over tegn, redskaber og modeller i J.-H. Bunn: *The Dimensionality of Sign, Tools and Models* (1981, Indiana U.P.). Den mere systematiske forestilling om meta-semiotik stammer fra den strukturelle semiotik (Hjelmslev).
45. En generel lettilgængelig introduktion til semiotikken er Umberto Eco: *A Theory of Semiotics* (1977, Macmillan). Semiotikken har to forskningstraditioner, den *analytiske semiotik* funderet af C.S. Peirce og den *strukturelle semiotik* funderet af F. de Saussure, L. Hjelmslev og A.J. Greimas. En analyse af problematikkerne i de to traditioner er Herman Parret: *Semiotics and Pragmatics. An evaluative comparison of conceptual frameworks* (1983, John Benjamins). Greimas-skolen har formuleret sine begreber i en Dictionaire i to bind, A.J. Greimas & J. Courtès (eds.): *Sémiotique* 1, (1979, Hachette) & 2 (1986, Hachette), hvoraf første bind er oversat til dansk på forlaget Basilisk.
46. I virksomhedsteoretisk forstand, som udviklet af Benny Karpatschof, jvf. »Grænser for automatisering«, *Psyke & logos*, 1984, N. 2.
47. Jvf. i denne sammenhæng Martin Heideggers overvejelser omkring *Herstellen/Vorstellen* og *Ding/Gegenstand/Sache* i artiklen »Das Ding«, i: *Vorträge und Aufsätze* (1954).

48. Jvf. Jerry R. Hobbs & Robert C. Moore (eds.): *Formal Theories of the Commonsense World* (1985, Ablex).
49. Se fx. Dedre Gentner & Albert Stevens (eds.): *Mental Models* (1983, LEA).
50. Det er således psykologiens udgangspunkt i psykofysikken må forstås, jvf. Georges Canguilhem klassiske artikel »Qu'est-ce que la psychologie?« (1958), *Cahiers pour l'Analyse* 1966, N. 1. Engelsk oversættelse - »What is psychology?«, *Ideology & Consciousness* 1980, N. 7.
51. Jvf. Jerry Hobbs o.a.: »Commonsense metaphysics and lexical semantics«, *Computational Linguistics* 1987, Vol. 13, N. 3-4.
52. Jvf. fx. Barbara Hall Partee: »Semantics - Mathematics or Psychology?«, i: R. Bäuerle, U. Egli & A. von Stechow (eds.): *Semantics from Different Point of View*.
53. Jvf. fx. Andy Clarks kritik af forsøgene på at etablere en »neurofilosofisk« eliminerings af de »folkepsykologiske« begreber. A. Clark: »From Folk Psychology to Naive Psychology«, *Cognitive Science* 1987, vol. 11, p. 139-154.
54. Som det også blev fremhævet af Stig Andur Pedersen i »Formelle modeller af menneskelig kommunikation«, i: T. Söderqvist (ed.): *Informationssamfundet* (1985, Philosophia).
55. Jvf. P. Brandt: *La charpente modale du sens* (thèse de doctorat, Sorbonne 1985) (1988, Romansk Inst., Århus).