

Tankevækkende netværk mellem biologi, datalogi og fysik

Claus Emmeche

Udforskningen af neurale netværk rejser også spørgsmål, som ligger mellem videnskab og filosofi. At de endnu ikke har klare svar, gør dem ikke mindre interessante, og de vil blive stillet med voksende styrke, efterhånden som forskningen i neurale netværk afkaster nye resultater. Spørgsmålene drejer sig bl.a. om, hvilken art videnskabelig genstand, man udforsker. Firkantet sagt: er det 'biologi og psykologi med fysiske metoder' eller er det en ny og mere generel fysik om komplekse, dynamiske systemer, der blot har virkelige nervesystemer som et eksempel? Hvor store dele af menneskets kognition og psyke falder indenfor en sådan generel teoris gyldighedsområde? Diskussionen drejer sig samtidig om de fundamentale begreber, forskellige systemer må beskrives med.

Fundamentale systemtyper

Hvilke grundlæggende typer systemer realiserer naturen? Biologer har den intuition, at organismer er noget særligt, og at selv gode ideer om 'selvorganiserende dissipative strukturer' eller 'komplekse dynamiske systemer med kollektiv adfærd' ikke indfanger alle unikke træk ved organismer som almen systemtype. Det kan være træk som selvreproduktion og dannelse af højere biologisk organisation. Eller det kan være principper, som ellers er ukendte i fysiske systemer, f.eks princippet om evolution via naturlig selektion af genotyper, som udgør en molekylær 'kodeskrift' for dannelse af de fænotyper, selektionen direkte virker på.

Omvendt har fysikere ofte den intuition, at hjernen og dens adfærd, for ikke at sige psyken, netop er emergente effekter af fundamentale fysiske processer, og at man blot mangler en tilstrækkelig god fysisk teori om sådanne systemer. Endelig taler både fysikere og neurobiologer om 'neural computation', som er et udtryk, de færreste dataloger bruger, fordi det ikke svarer til den beregningsform, man kender fra de almindelige computere, som den teoretiske datalogi af idag bygger på.

Neurale netværk er altså et perspektivrigt forskningsfelt, hvor fysiske, matematiske, neurobiologiske og datalogiske metoder yder væsentlige bidrag til at stille nye spørgsmål om beregningens natur. Kernen i den grundvidenskabelige interesse (set med en biologs øjne) er på langt sigt at give en teori for hjernens funktion. Studiet af neurale netværk er en opgave, hvis betydning for biologien potentielt er lige så afgørende som beskrivelsen af arvematerialets organisation, der med molekylærbiologien førte til en revolution i forståelsen af det levende. Men vi må huske, at der kan være afgørende aspekter ved biologiske neuroners signalbehandling, som teorier om neurale

netværk ikke klarlægger, og som må forstås til bunds før en samlet teori for f.eks psykiske funktioner hos dyr og mennesker kan gives.

Analog beregning

Neurale netværk er praktiske, fordi de anviser metoder til at løse komplicerede beregningsopgaver og nye veje til at forstå principperne for, hvad naturlig realiseret beregning og informationsbehandling er, måske også i mere abstrakt forstand. Dette har ikke blot biologisk og psykologisk interesse. Den datalogiske teori for beregning har hidtil været forbundet med begreber som algoritmer, programmerbarhed samt ideen om beregning forstået som diskontinuerte tilstandsovergange, hvor en maskine 'læser' og 'skriver' elementære symboler ifølge nogle regler. Se box: Turingmaskinen. Har vi med de neurale netværk en ny, grundlæggende model for (analog) beregning?

Fysik og ingeniørvidenskab har jo længe benyttet sig af analoge beregningsmetoder. Det er velkendt, at man ved at konstruere bestemte elektriske kredsløb kan simulere en bevægelsesform, som beskrives af et system af koblede differentialligninger, der kan være vanskelige at løse analytisk, men som lader sig 'beregne' analogt v.h.a. sådanne opstillinger. De analoge kredsløb mangler naturligvis de digitale computers universelle programmerbarhed, de er dedikerede til at løse lige præcis den konstruerede opgave, men ikke desto mindre er de altså et klassisk eksempel på det, man kunne kalde 'fysisk realiseret beregning'. Men hvad er så det eksakte forhold mellem fysisk beregning, der realiseres i analoge kredsløb, f.eks i de neurochips som nu produceres, og så det formelt definerede beregningsbegreb man kender fra datalogien?

Dette spørgsmål fører vidt omkring, men lad mig her nøjes med én tangent. Set ud fra ét, måske naivt, perspektiv 'beregner' et fysisk system jo aldrig noget som helst i sig selv, uanset om beregningen sker analogt eller digitalt: Systemet følger blot de fundamentale naturlove i sin udvikling af de fysiske tilstande, der karakteriserer systemet. Ligeså lidt som månen 'beregner' sin næste stedposition ud fra Newtons love, men blot følger disse love, ligeså lidt 'beregner' en chip resultatet af en additionsproces. Også den følger lovmæssigt en adfærd som beskrives af rent fysiske ligninger. At chippen beregner er således en afledet, eller 'tilskrevet' egenskab: Det er chippens særlige konstruktion og dens lovbundne adfærd, som gør, at den systematisk kan behandle signaler, der på et abstrakt niveau har en fortolkning som tal. Ikke sådan at forstå, at der skal være nogen menneskelig fortolker involveret i selve processen – så var fidusen ved comput-

ere forsvundet – men forstået sådan, at processen skal ske netop systematisk, hvilket bl.a. vil sige med næsten nul støj. Dette er forudsætningen for at mekaniseringen af beregningsprocessen overhovedet kan finde sted: Tilskrivningen af betydning til input og output skal kunne ske pålideligt og systematisk.

Turingmaskinen

Turingmaskinen (TM) er en teoretisk model af en computer, konstrueret i 1936 af den engelske matematiker Alan Turing. Turing forestillede sig et apparat, som bestod af en boks og en lang strimle. Strimlen er inddelt i felter, der hver kan rumme et symbol. Boksen har et antal tilstande, som kan ændres ifølge en tabel for tilstandsovergange. Boksen kan – idet den følger regler af typen: "Hvis symbol = x og tilstand = y, gør da operation = p." – bevæge sig hen over strimlen og udføre operationer af typen: Ændre (skrive) symbol på strimlen, ændre egen tilstand, bevæge sig et antal skridt frem eller tilbage og (evt) stoppe. Afhængig af reglernes art (programmeringen), kan den udføre forskellige opgaver.

Ud fra dette tankeeksperiment kunne Turing vise, at der eksisterer en *universel* maskine (UTM), der kan simulere alle konkrete TM'er. En konkret TM svarer til programmet i den universelle. En UTM svarer igen til en generel programmerbar digital computer. Turings tese er, at alt hvad der er effektivt beregneligt, kan beregnes på en UTM.

TM'en er en formalisering af den intuitive ide om effektiv beregning. Den er også en antropomorf maskine: En talteoretiker, der rummer reglerne for matematiske operationer, som han udfører på en endimensional tavle.

Der er i artiklens sammenhæng 3 tag-hjembudskaber om TM.

1. Det er alment anerkendt, at hvis et problem kan beregnes (dvs. løses på en TM), så kan det løses på en digital computer.
2. Det er også anerkendt at ikke alle problemer er Turing beregnbare.
3. Det er et åbent spørgsmål om dette koncept for generel beregning indfanger alle aspekter af beregning, forstået som informationsbehandling; og om det overhovedet er frugtbart i forståelsen af neurale netværk.

Intentionelle systemer

Ovenstående udgør i al fald, hvad vores almindelige forståelse af størrelser som naturlove, fysiske genstande, tal, beregning, og betydning lægger op til. Vi kan skelne mellem ting, som har visse iboende egenskaber, og ting der blot ser ud 'som om' de har dem under en given beskrivelse.

Månens bevægelse ser ud 'som om' den beregner sin bane efter en newtonsk algoritme, men vi vil vel helst ikke tro at den bruger beregning eller behandler information for at bevæge sig. Månen bevæger sig blot i overensstemmelse med de love, der er det fysiske rum iboende, uanset om vi tilskriver det noget eller ej. På samme måde udvikler den analoge simulation af systemet af differential-ligninger sig som den gør p.g.a. sin iboende tendens til at følge bestemte naturlove, uanset om vi tilskriver systemet nogen beregnende adfærd eller ej. Det vil sige, dets 'beregning' er – under denne naive synsvinkel – afledet eller tilskrevet. Tilskrevet af hvem? Selvfølgelig af mennesker, som i kraft af de menneskelige evner til abstrakt tænkning, meningsfuld kommunikation, bevidste forestillinger o.l. kan håndtere fysiske eksemplarer (tokener) af bestemte symbolske typer (f.eks matematiske symboler) på en sådan måde, at de på systematisk vis kan benyttes til beregning. Ethvert menneske kan lære fysik, ikke enhver abe. Disse evner er knyttet til det, filosoffer har benævnt *intentionalitet*, forstået som det at kunne se noget (f.eks noget kridt på en tavle) som betydende eller repræsenterende noget andet (f.eks et begreb, et tal, osv.). Uden intentionalitet kunne vi ikke se, høre, eller tænke på konkrete ting som noget, der betyder noget andet, nemlig generelle, abstrakte og begrebslige størrelser. Uden intentionalitet kunne vi heller ikke bruge en PC'er, fordi den som enhver anden computer af klassisk type netop er et *fortolket, automatisk, formelt* system. – Fortolket: heri ligger det intentionelle. Automatisk: heri ligger det fysiske (computerens hardware). Formelt: heri ligger det datalogiske i klassisk forstand. Så snart vi (naivt) taler om computere, forudsætter vi derfor intentionelle systemer. Vi er nu fremme ved en foreløbig konklusion og et spørgsmål:

At neurale netværk er eller repræsenterer en vision om en ny type computere betyder måske, at Turing-konceptet for beregning som noget datalogisk fundamentalt skal revideres og udvides. Her vil relationen mellem teoretisk fysik og datalogi skulle gennemtænkes påny. Et helt andet, og endnu åbent spørgsmål er, om de meget komplicerede neurale netværk, som engang i fremtiden evt. måtte nærme sig den menneskelige tænkning i kapacitet, også vil have deres 'intentionelle' aspekt iboende sig. For vi har endnu ikke sagt hvoraf denne intentionalitet kommer: om den i sig selv kan forstås som en emergent effekt af tilstrækkelig komplekse neurale informationssystemers vekselvirkninger; om det kræver større lighed med netop hjernens neurale informationsbehandling, eller om der skal andre ting til.

Hvis hjernen er en neural analog computer, så forklarer dette ikke, *hvordan* den har den egenskab, at kunne tilskrive noget betydning, fordi hverken en analog eller en digital computer i sig selv (ifølge det naive synspunkt ovenfor) havde 'beregning' eller 'betydning' som iboende egenskaber. Det har kun levende organismer med nervesystem og især de som tilmed har et sprog. Måske skal vi så også på dette punkt gøre op med den naive snusfornufts skelnen mellem de egenskaber, tingene kan

have 'i sig selv', og de egenskaber, som tilskrives dem. Risikoen er, at vi i så fald nødsages til at se temmelig mange genstande i verden som analoge computere, måske måner inklusive.

Beskrivelsen af hjernen som et biofysisk system, der realiserer beregninger, synes altså at løbe ind i problemer med beregningsbegrebet og med – en temmelig grundlæggende egenskab ved hjernen – dens intentionalitet. Dette problem gælder såvel neurale netværk, som traditionel kunstig intelligens. Det er fornuftigt at forskningen for det meste ikke lader sig hæmme af eksistensen af begrebslige problemer i selve dens fundament. På den anden side er det uhensigtsmæssigt hvis man lader som om de slet ikke findes.

Vi kan med de neurale netværk ane konturerne af en egentlig teori for 'intelligent' og 'adaptiv' informationsbehandling i biologiske systemer. Måske vil den endda være formuleret som et særligt fælde af fysikkens teorier om komplekse, dynamiske systemer. Det er umiddelbart ønskeligt for den teoretiske biologi, som har været al for isoleret fra fysikken. Der er netop brug for flere "tankevækkende" netværk mellem folk fra hidtil adskilte discipliner. Disse netværk vil især intensiveres hvis det

viser sig, at biologisk beregning – det som dyr gør, når de sanser objekter i omgivelserne og handler derefter – kan forstås ud fra teorier om ikke-lineær dynamik, statistisk mekanik, faseovergange, m.v. Meget tyder på, at netop disse teorier kan indfange noget af den kollektive adfærd, som er resultat af myriader af interaktioner mellem nerveceller i hjernen. Om en sådan biofysisk teori kan indfange alle aspekter af det, hjerner gør, når de fungerer i organismer, er måske mere tvivlsomt.



Claus Emmeche, biolog, lic.scient., gæst ved CONNECT, Niels Bohr Institutet og Center for Kognitionsforskning, Roskilde Universitetscenter.

Edwards High Vacuum Turbomolekularpumper

- ★ 4 basistyper
- ★ High Power DC-motor
- ★ 40 flangevarianter

- ★ Ny EXT E økonomiversion
- ★ 17 modeller
- ★ Lavt støj- og vibrationsniveau

EDWARDS

Edwards High Vacuum er en af verdens førende producenter af vakuum-pumper og -udstyr. Den nye EXT E version gør det nu muligt at anskaffe en Turbomolekular Vakuumpumpe til en fornuftig pris. Kan køles enten med luft eller vand, og kan styres af den normale EXC350 og EXC500 Turbo-kontroller. Edwards High Vacuum er også Dry Pumps, Diff. Pumps, Vakuum måleinstrumenter, Frysetøringsanlæg og På-dampningsanlæg.

Ring for yderligere oplysninger!

Marielundvej 36
2730 Herlev
4291 75 11

Buch & Holm A/S