

# Kunstig intelligens

Forestillingen om, hvordan man skaber kunstig intelligens, har forandret sig over tiden. Fra at tro, at hjernen kan erstattes af en fantastisk regnemaskine, og til at se intelligens som en myretue af samarbejdende enheder.

## Stærk og svag kunstig intelligens

Kunstig intelligens (KI) er maskiner, der er i stand til at overveje, lære og tage beslutninger på samme niveau som et menneske. Kunstig intelligens er en direkte oversættelse af det engelske artificial intelligence, og ofte benyttes den engelske forkortelse "AI" i stedet for den danske "KI".

Man skelner mellem to forskellige programmer i kunstig intelligens-forskningen, nemlig stærk og svag KI. Svag KI er systemer, der blot er i stand til at efterligne visse træk ved menneskelig tænkning, mens stærk KI er systemer, der virkelig har intelligens af samme karakter og dybde som vi mennesker.

Skellet blev indført af den amerikanske filosof John R. Searle. Searle har ikke nogen indvendinger mod svag KI, men han mener, det principielt er umuligt at skabe kunstig intelligens af den stærke form. En computer fungerer rent syntaktisk; den manipulerer med symboler efter strengt formelle regler, men disse syntaktiske operationer alene er ikke nok til at give computere forståelse af den betydning eller mening, symbolerne har. Af den grund vil en computer aldrig kunne forstå noget på samme niveau som et menneske.

## Historisk baggrund for kunstig intelligens

Forskningen i KI er nært forbundet med de computere, der blev opfundet i begyndelsen af 1940'erne. Selve udtrykket kunstig intelligens stammer fra John McCarthy, der første gang benyttede det i 1955 i en ansøgning om midler til en konference i Dartmouth året efter. På denne konference deltog en række af de forskere, der i de følgende år gik forrest i KI-bevægelsen, bl.a. Allen Newell og Herbert Simon, der i 1969 grundlagde KI-laboratoriet på Carnegie Mellon-universitetet, samt Marwin Minsky og John McCarthy, der sammen grundlagde den indflydelsesrige KI-forskning på Michigan Institute of Technology. Dartmouth-konferencen var den første af sin slags og bliver normalt betragtet som dén historiske begivenhed, der gjorde kunstig intelligens til et selvstændigt forskningsområde.

## Tænkning som kalkule

Drømmen om at mekanisere den menneskelige tænkning har dog rødder langt tilbage i historien. Det første forsøg på at finde love for tænkning ses hos den græske filosof Aristoteles (384-322 f.Kr.), der grundlagde den formelle logik. I formel logik undersøger man argumenter ved at betragte deres logiske form. Man kigger ikke på de enkelte begrebs betydning, men kun på den overordnede logiske sammenhæng, der dannes af ord som "og", "eller", "alle" m.v. Aristoteles' logik var forholdsvis begrænset, men i 1800- og 1900-tallet blev den formelle logik væsentligt udviklet og forbedret, bl.a. udviklede George Boole (1815-1864) den binære logik, der udgør hjertet i computere.

En række filosoffer har udbygget ideen bag den formelle logik ved at forestille sig et særligt "tankesprog", hvor alle tanker kan udtrykkes som symboler, og tænkning foregå som en ren mekanisk procedure, hvor symbolerne håndteres efter bestemte formelle regler eller "tanelove". Ideen ses i en tidlig form hos W.G. Leibniz (1646-1716), og kommer til udtryk i forskellig form i de følgende århundrede. For KI-forskningen har navnlig Jerry Fodors *The Language of Thought* (1975) haft stor indflydelse.

I årene omkring år 1900 blev ideen om at reducere tænkning til en mekanisk udregning gennemført med en vis succes inden for matematikken. Anført af den tyske matematiker David Hilbert lykkedes det at samle al matematik i et enkelt system. En væsentlig del af menneskets tænkning, nemlig matematikken, kunne herefter betragtes som mekanisk håndtering med symboler efter formelle regler. Man er dog i dag blevet opmærksom på en række alvorlige problemer i denne såkaldt formalistiske opfattelse af matematikken.

### **Beregningsmaskiner**

Den første mekaniske regnemaskine, der kunne håndtere alle fire regnearter (plus, minus, gange og dividere) var den såkaldte pascalina, der blev udviklet af Blaise Pascal i 1645. I de følgende århundreder skete en voldsom udvikling af forskellige mekaniske regneapparater. Brugen var dog begrænset, for de ikke var programmerbare og kunne derfor kun løse præcis de opgaver, de var skabt til at løse. I 1837 fremsatte Charles Babbage imidlertid ideen om en programmerbar maskine, dvs. en maskine, der med den samme fysiske hardware kunne løse forskellige opgaver alt efter, hvilket program den var udstyret med. Det lykkedes dog aldrig Babbages at bygge sin såkaldte Analytical Engine, til dels fordi dens design var meget kompliceret. Først med fremkomsten af den elektroniske computer i begyndelsen af 1940'erne lykkedes det at lave programmerbare beregningsmaskiner.

Med kombinationen af programmerbare computere og opfattelsen af, at tænkning består af en mekanisk beregning, var jorden beredt for de første forsøg på at skabe tænkende maskiner.

### **Paradigmer inden for kunstig intelligens**

Forsøgene på at skabe kunstig intelligens kan i grove træk inddeles i tre forskellige forskningsprogrammer eller paradigmer, der har hver sin opfattelse af, hvad tænkning er, og af, hvordan intelligens kan skabes ad kunstig vej: 1) det klassiske symbolparadigme, 2) konnektionisme og 3) embodied KI. Af disse tre var det klassiske symbolparadigme dominerende fra 1950'erne til midt i 80'erne, hvorefter de andre to paradigmer fik indflydelse. I dag er et helt fjerde paradigme, der fokuserer på selvorganisering, så småt begyndt at tegne sig (for en nærmere beskrivelse af paradigmerne, se følgende).

### **Det klassiske symbolparadigme**

Den centrale tese i det klassiske symbolparadigme er, at al tænkning, inklusiv menneskers tænkning, er mekaniske operationer i et formelt system. Et formelt system er uafhængigt af den fysiske form, det realiseres i, og derfor er der i princippet ikke nogen hindring for, at computere foretager de samme formelle operationer som menneskehjernen - og dermed tænker på samme måde som et menneske. Det kræver blot, at computeren har tilstrækkelig regnekraft og er udstyret med det rigtige program.

Det klassiske symbolparadigme havde i starten stor succes. Det lykkedes hurtigt at skabe systemer, der kunne spille kryds og bolle og skak på et rimeligt niveau, og i 1967 præsenterede trio'en Allen Newell, Herbert Simon og J. Cliff Shaw programmet General Problem Solver, der var i stand til at bevise logiske og matematiske sætninger. Sammenligninger med studerende antydede, at programmet oven i købet gjorde det på nogenlunde samme måde som vi mennesker. Nøglen til disse tidlige systemer var heuristisk søgning, dvs. procedurer, hvor computeren ved hjælp af tommelfingerregler og intelligente vurderinger finder en god løsning blandt en gruppe mulige løsninger uden at afprøve samtlige mulige løsninger.

Det er aldrig for alvor lykkedes at overføre succeserne fra skak og matematik til mere realistiske problemer, f.eks. til robotter, der befinder sig i den virkelige verden. Med det såkaldte rammeproblem, der blev opdaget af John McCarthy og Patrick Hayes i 1969, peges der på et principielt problem for den klassiske tilgang. Problemet går i korte træk ud på, at enhver handling i den virkelige verden har umådelig mange

konsekvenser. Hvis man flytter en kop, flytter man ikke blot koppen, man hvirvler også luftmolekyler rundt, ændrer trykfordelingen på bordpladen osv. Derfor er der stor forskel på at planlægge og forstå konsekvenser af handlinger i den virkelige verden frem for at planlægge i et spil skak, hvor hver handling har et trods alt begrænset antal konsekvenser.

Det klassiske symbolparadigmes rent formelle tilgang har desuden vist sig at være dårlig til naturlige sprog (som dansk og engelsk). Systemerne betragter ordene som symboler, der flyttes rundt efter bestemte regler som brikkerne i et spil skak. Men menneskers sprogbrug kan næppe forklares af den slags snævre regler. Visse teoretikere hævder, at de klassiske systemer er principielt begrænsede, idet menneskets sprogbrug bliver påvirket af vores kropslighed og omgang med tingene omkring os.

### **Det konnektionistiske paradigme**

Den grundlæggende idé i konnektionismen er at lave systemer, der på visse punkter efterligner menneskehjernens biologiske opbygning. Hjernen består af et stort antal nerveceller, kaldet neuroner. Hver enkelt neuron kan i isolation kun foretage meget simple beregninger, men neuronerne er sammensat i et enormt netværk, der gør det muligt for dem at samarbejde og på den måde foretage meget komplekse beregninger. Opgaven for forskningen i kunstig intelligens er derfor, ifølge konnektionisterne, dobbelt. For det første må man forstå, hvordan netværk af beregningsenheder fungerer, og for det andet må man prøve at lave kunstige netværk, der kan efterligne nogle af de funktioner, hjernens biologiske netværk har.

### **Kroppen i kunstig intelligens**

Ifølge embodiment-filosofien er menneskets intelligens knyttet tæt sammen med vores fysiske eksistens. Man kan derfor ikke forstå (eller kopiere) vores evne til abstrakt tænkning, medmindre man først har forstået vores helt grundlæggende evne til at spille sammen med det omgivende miljø.

Hypotesen er i høj grad udviklet inden for den såkaldte kognitive semantik, ikke mindst af psykologen George Lakoff og filosofen Mark Johnson, der med en række konkrete studier søger at vise, at en stor del af menneskets sprog er bygget op om metaforer hentet i basale erfaringer. For eksempel er almindelige udtryk som "tiden løber", "priserne stiger" og "han virker ubalanceret" alle metaforiske, hvis man kigger efter. Dette viser, ifølge Lakoff og Johnson m.fl., at helt basale erfaringer skaber orden på vores verdensopfattelse og giver os evnen til abstrakt tænkning.

Embodiment-filosofien understøttes dog også af andre studier af, hvordan mennesker i praksis tænker. I en artikel fra 1994 påviser David Kirsh og Paul Maglio således, at mennesker ikke altid løser beregningsopgaver gennem abstrakt tænkning, men i visse situationer vælger at udføre bestemte handlinger i stedet. Kirsh og Maglio undersøgte brugere af computerspillet Tetris, men man kan f.eks. forestille sig, hvordan man drejer en puslespilsbrik for at se, om den passer, frem for at lukke øjnene og beregne. Med andre ord løser man beregningsopgaven ved at handle og sanse.

### **Embodiment i KI - autonome agenter**

Inden for den kunstige intelligens slog embodiment-filosofien igennem i midten af 80'erne i et paradigme, der på engelsk går under navne som embodied-AI, nouveau-AI og autonomous agents research. Den nye kunstige intelligens er karakteriseret ved en "nedefra-og-op-tilgang". Man tager typisk udgangspunkt i robotter, og arbejder først og fremmest på at sætte dem i stand til at klare sig på egen hånd i et omskifteligt og dynamisk miljø. Først når dette er på plads, forsøger man at bygge mere avancerede egenskaber oven på de basale. På den måde afspejler man, hvordan menneskets intelligens ifølge

embodiment-filosofien er opbygget. De nye, selvstændige, men typisk meget simple robotter går under navnet autonome agenter.

De første forsøg med simple robotter af autonom agent-typen blev foretaget af William G. Walter i 1940'erne. Disse forsøg blev dog stort set ignoreret, og først i midten af 1980'erne, da det klassiske symbolparadigme i manges øjne havde spillet fallit, var folk parate til at prøve noget nyt. Gennembruddet kom med Rodney Brooks' subsumption arkitektur, der gjorde det muligt at lave robotter, der med en vis robusthed overfor den virkelige verdens uforudsigelighed og dynamik var i stand til at løse opgaver som at indsamle tomme coladåser i et hektisk kontormiljø.

De nye robotter har i begyndelsen af det nye årtusind fået et kommercielt gennembrud, fordi robotternes evne til at begå sig i et omskifteligt og uforudsigeligt miljø har gjort dem brugbare inden for både industri, underholdnings- og plejesektoren. Det er dog aldrig lykkedes at udvikle de autonome agenter ud over et meget basalt intelligensniveau.

### **Selvstrukturerende systemer**

Et system kan kaldes selvstrukturerende, hvis der af sig selv opstår en overordnet orden eller struktur ved, at systemets enkelte dele følger bestemte regler. Man ser det ofte i den biologiske verden, f.eks. i myretuer, hvis struktur ikke opstår gennem central styring og planlægning, men ved at de enkelte myrer følger nogle bestemte, simple regler.

Selvstrukturering har længe været et tema i forskningen i kunstig intelligens. I KI-forskningen kan kunstige neurale netværk ses som et tidligt eksempel på brug af selvstrukturering, men det er navnlig i forbindelse med det øgede fokus på biologisk realistiske systemer i den nye KI, at man mere målrettet er begyndt at bruge selvstrukturering. Princippet er bl.a. blevet benyttet af belgieren Luc Steels, der i 90'erne fik autonome agenter til at udvikle et fælles sprog ved hjælp af simple selvstruktureringsmekanismer. Princippet benyttes også af Adaptronics Group på Syddansk Universitet, der bl.a. udvikler robotter, sammensat af mindre enheder. Robotternes overordnede adfærd og form opstår som følge af enhedernes samarbejde efter simple regler.

Man kan endnu ikke tale om et samlet paradigme eller forskningsprogram om selvstrukturering, men den begyndende brug vidner om et fundamentalt skift i opfattelsen af, hvordan kunstig intelligens skabes. Forskerne ser ikke længere sig selv som urmagere, der skal designe og kontrollere systemerne ned i mindste detalje. I stedet gælder det om at skabe de rigtige betingelser, så den overordnede kompleksitet kan vokse frem af sig selv.