

Kunstig intelligens – død eller levende?

Thomas Bolander

Adjunkt ved Informatik og Matematisk Modellering (IMM), DTU



Kunstig intelligens – død eller levende?

Termen **kunstig intelligens** (artificial intelligence, **AI**) betegner “the science and engineering of making intelligent machines, especially intelligent computer programs” (John McCarthy).

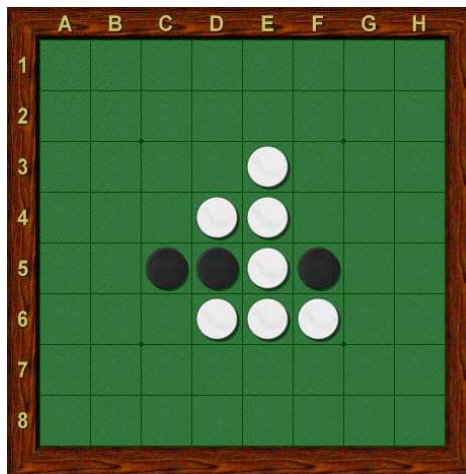
Det eneste problematiske ord her er ‘intelligent’. Hvad vil det egentlig sige at en robot eller et stykke software er intelligent?

Hvad er intelligens?

Intelligens er et meget vanskeligt begreb. 52 amerikanske psykologer underskrev i 1994 en (forsøgsvis) definition som starter sådan her:

“Intelligence is a very general mental capability that, among other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience.”

Vore dages AI-systemer kan i en vis begrænset udstrækning alle disse ting formodentlig på nær det at forstå komplekse idéer. Det er dog svært at svare entydigt på, da ‘forståelse’ igen et meget vanskeligt begreb.



AI og forståelse

I AI's filosofi skelner man mellem stærk og svag AI.

Stærk AI henviser til AI som reelt kan *forstå* og være *selvbevidst*.

Svag AI henviser til AI som er orienteret mod specifik problemløsning eller ræsonering uden krav om et højt niveau af menneskelignende kognitive evner.

Det findes en overvældende mængde af filosofiske argumenter både for og (mest) imod den teoretiske eksistens af stærk AI. Stærk AI er dog under alle omstændigheder ikke en realitet i dag.

Turing-testen

Alan Turing (ham som brød tyskernes Enigma-kode under anden verdenskrig) fandt i 1950 på den såkaldte **Turing-test**, en test til at checke om en maskine udviser stærk AI.

I en Turing-test er en menneskelig dommer i skriftlig dialog med henholdsvis et andet menneske og en maskine (en computer). Hvis dommeren ikke kan fortælle hvem der hvem har maskinen bestået testen (det forudsættes at både mennesket og maskinen forsøger at virke menneskelige).

Hvert år konkurrerer natursprogs-AI'er mod hinanden om **Loebner-prisen**. Prisen går til den AI som klarer en Turing-test bedst muligt. Vinderen får 2.000\$. Hvis maskinen klarer Turing-testen fuldstændigt får vinderen 25.000\$. De 25.000\$ er endnu aldrig blevet udbetalt.

Turing-testen

De sidste par år har Jabberwacky (jabberwacky.com) vundet Loebner-prisen. Her er en dialog jeg havde med den igår:

Jabberwacky: Have you read any books lately?

Thomas: Yes.

Jabberwacky: What books have you read?

Thomas: "I am a strange loop".

Jabberwacky: Yes you are.

Thomas: Is that supposed to be funny?

Jabberwacky: Yes.

Thomas: OK, accepted, it was kind of funny.

Jabberwacky: No. Who are you trying to convince?

Thomas: You.

Jabberwacky: Tell me something funny.

Thomas: The best joke I know is in Danish.

Jabberwacky: That's a poem I guess, tell me another.

Thomas: Why don't you tell me one.

Jabberwacky: I don't know any?

Thomas: How can you not know any jokes?

Jabberwacky: No, not that, what about star wars?

Thomas: I don't like star wars.

Jabberwacky: Why don't you like star wars?

Thomas: I find it boring.

Jabberwacky: How very British of me. I'm sorry. Are you going out tonight?

Thomas: No, I'm preparing a talk for tomorrow.

Jabberwacky: Preparing to become what?

...

Stærk AI

Det er vist rimeligt at konkludere at stærk AI—i hvert fald hvad angår AI som kan bestå Turing-testen—stadig er lysår ude i horisten. Relevansen af Turing-testen kan dog også diskuteres: Målet med forskning i flyveteknik er jo heller ikke at lave maskiner som kan flyve præcist som duer, således at selv duerne ikke kan kende forskel!

Men under alle omstændigheder: Det er meget tvivlsomt om der i dag findes nogen AI-systemer som kan siges at udvise reel forståelse og selvindsigt.

De fleste forskere stiller sig derfor tilfredse med svag AI:

Computerprogrammer og robotter som kan udføre en begrænset grad af ræsonering og problemløsning med et praktisk anvendeligt udbytte.

AI's storhed og fald

I begyndelsen af “de glade 60'ere” var forventningerne til AI ekstremt høje. Komplexiteten af den menneskelige intelligens blev stærkt undervurderet, og de fleste forventede at maskiner med menneskelignende intelligens var lige om hjørnet:

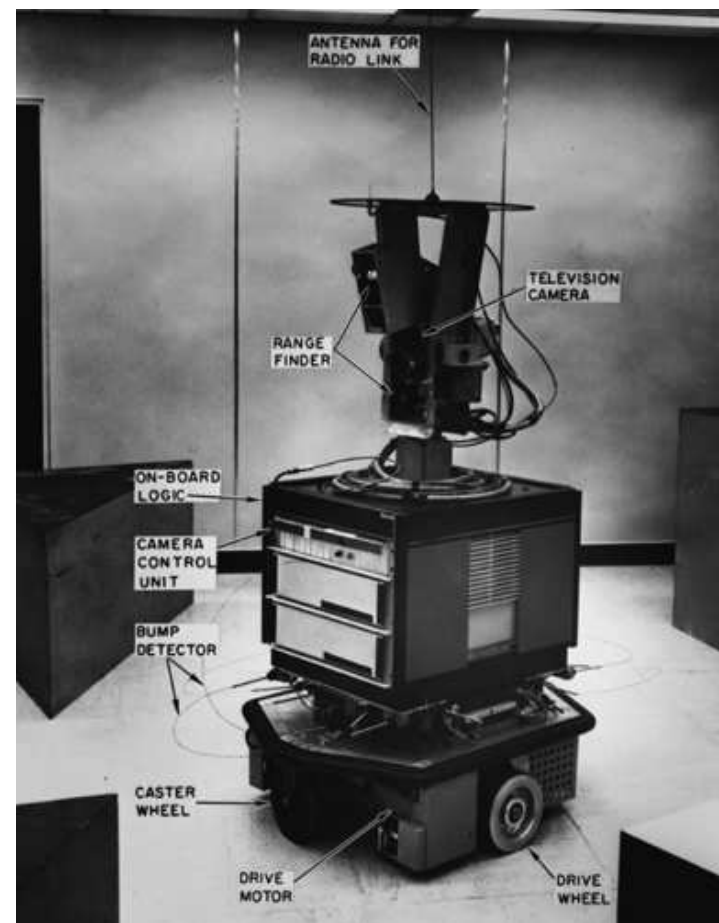
“It is not my aim to surprise or shock you—but the simplest way I can summarize is to say that there are now in the world machines that think, that learn and that create. Moreover, their ability to do these things is going to increase rapidly until—in a visible future—the range of problems they can handle will be coextensive with the range to which the human mind has been applied” (Herbert Simon, 1957).

Denne ‘visible future’ trækker øjensynligt ud...

AI's storhed og fald

Den grænseløse optimisme i begyndelsen af 60'erne skyldes også at man faktisk den gang på meget kort tid fik lavet AI-systemer som kunne løse et væld af problemer som på den tid blev opfattet at kræve 'reel intelligens':

- skak-spillende computere som kunne slå dygtige amatører;
- computere som kunne bevise geometriske teoremer bedre end de fleste matematikstuderende;
- computere som kunne løse geometriske analogi-problemer i stil med IQ-tests;
- computere som kunne løse problemregningsopgaver på folkeskole-niveau;
- det første 'elektroniske menneske', robotten Shakey.



AI's storhed og fald

I dag vil vi dog næppe opfatte en skakcomputer som blot slavisk regner 8 træk frem og vælger det træk som maksimerer en simpel matematisk evalueringfunktion som 'intelligent'. Den slags computerprogrammer hører dog stadig ind under det område som kaldes kunstig intelligens, selvom det for mange er lidt skuffende fordi det forekommer mere 'kunstigt' end 'intelligent'.

AI's storhed og fald

De store succeser i begyndelsen af 60'erne satte ambitionsniveauet endnu højere op, og der blev investeret enorme summer i AI-forskning. Der skulle dog vise sig at være tale om en 'AI-boble'.

Et af de ambitiøse projekter i USA var maskin-oversættelse af russiske videnskabelige tekster til engelsk i kølvandet på russernes opsendelse af Sputnik. Systemerne var dog blot simple syntaks-transformatorer hvilket blandt andet gav anledning til følgende underholdende oversættelse fra engelsk via russisk tilbage til engelsk:

Original: "The spirit is willing but the flesh is weak".

Oversættelse: "The vodka is good but the meat is rotten".

AI's storhed og fald

I slutningen af 60'erne begyndte udviklingen af AI at gå meget langsommere, og man blev klar over at reelt tænkende maskiner måske alligevel ikke var lige rundt om hjørnet—og at sådanne maskiner måske slet ikke kunne eksistere.

Skuffelsen var stor, og interessen i at støtte AI-forskning gik pludselig stærkt ned ad bakke. I 1973 besluttede den britiske regering at sløjfe al forskning i AI på nær på to universiteter. AI'en blev af mange opfattet som død.

AI's storhed og fald

Det bratte fald af AI'en skyldes givetvis to ting:

- De oprindelige naive og urealistiske forventninger til lethed i at modellere og simulere menneskelignede intelligens.
- Det generelle fænomen indenfor ethvert nyt videnskabeligt fagområde, at når først den lavthængende frugt er plukket går alting pludselig meget langsommere (uden at det dog behøver at betyde at området er dødt).

Nyt liv til AI

I slutningen af 80'erne får AI nyt liv igen, men belært af erfaringen holder de flest sig nu til den svage AI og at 'tage hvad man kan få' fremfor at have urealistiske forventninger. Det resulterer i en række nye succeser for AI indenfor begrænsede domæner:

- USA's forsvar bruger et planlægningssystem til at håndtere logistikken i forbindelse med Golf-krigen i 1991, resulterende i en besparelse større end den samlede mængde af penge benyttet til AI-forskning i perioden 1950-1991.
- I 1994 kører en computerstyret bil over 1000 km i tæt trafik i Frankrig.
- I 1997 slår IBM's skakcomputer *Deep Blue* verdensmesteren Gary Kasparov.
- I 2001 vinder en computer en konkurrence i aktiehandel.



Intelligens eller ej?

Selvom de computere der kan lave logistik, spille skak, køre i bil og handle med aktier (og tjene på det!) nok ikke er hvad vi i dag vil kalde 'tænkende maskiner' er det dog computere som løser problemer vi tidligere opfattede som krævende 'ægte intelligens'.

Så hvad forskningen i AI nok primært gør er at rykke vores grænse for hvad vi opfatter som 'ægte intelligens': I takt med at computere erobrer flere og flere af menneskets enemærker må vi revurdere vores opfattelse af hvad der er egentlig intelligens.

Babbage's analytiske maskine—en slags programmerbar regnemaskine—beskrevet af Charles Babbage i 1840'erne blev også af bl.a. matematikeren Ada Lovelace beskrevet som en slags intelligent maskine, mens ingen i dag ville opfatte en programmerbar lommeregner som intelligent.

Intelligens eller ej?

I virkeligheden er det måske sådan at intelligens snarere handler om den *måde* hvorpå vi som mennesker udfører ræssonering end det som vi faktisk er i stand til at opnå med denne ræssonering.

Menneskets intelligens er fleksibel og dynamisk på en måde som ingen computer er, men det vi på et ethvert givet tidspunkt er i stand til at opnå med vores intelligens er muligvis noget som alligevel kan simuleres på en passende avanceret og kløgtigt programmeret computer.

Men vores evne til at være selvbevidste, selvobserverende og nærmest transcendere os selv er stadig et stort mysterium, og noget som ingen computer er blot tilnærmelsesvis i nærheden af. For ikke at tale om den fri vilje...

Det symbolske paradigme

Selvom man ikke nødvendigvis tror på eksistensen af stærk AI er det stadig den menneskelige hjernes processer som er inspirationen bag de fleste landvindinger indenfor AI. Der er to hovedparadigmer indenfor AI, det symbolske paradigme og det sub-symbolske paradigme.

- **Det symbolske paradigme.** Her forsøger man direkte at simulere det højeste niveau af kognitive processer i hjernen, det vil sige bevidst, sprogligt forankret, ræssonering. AI-systemerne udstyres med et 'kunstsprog' (et formelt sprog) som de benytter til at repræsentere viden om verden og til at ræsonere omkring verden. Metoderne er primært baseret på formel logik, et område som oprindeligt blev udviklet i begyndelsen af det 20. århundrede med henblik på at forstå og sikre matematikkens grundlag.

Det subsymbolske paradigme

- **Det sub-symbolske paradigme.** Her forsøger man modsat at simulere de lavere niveauer af processor i hjernen: de biologiske processer. Det gøre f.eks. med (kunstige) neurale netværk og evolutionære algoritmer.

Systemerne indenfor de to paradigmer har forskellige styrker og svagheder:

Det symbolske paradigme er godt til at lægge planer i komplekse domæner, mens det det sub-symbolske paradigme er godt til mønstergenkendelse og oplæring af rutine-opgaver.

I det symbolske paradigme har man meget kontrol over sit system, kan matematisk bevise egenskaber ved det m.m. I det subsymbolske paradigme må man i langt højere grad bygge på empiri.

Symbolsk versus subsymbolsk

Ofte konstrueres systemer som kombinerer de to paradigmer i f.eks. robot-køretøjer som både skal planlægge ruter og rækkefølger af handlinger (symbolsk) og skal kunne navigere ved hjælp af robotsyn (sub-symbolsk).

AI i dag

AI-metoder benyttes i dag blandt andet til:

- Computerspil—styring af non-player characters.
- Søgmaskiner som f.eks. Google.
- Ekspertsystemer—f.eks. medicinske diagnosticeringssystemer.
- Biologisk sekvensanalyse.
- Planlægningssystemer til f.eks. rejseplanlægning.
- E-mail spamfiltre.
- Autonome køretøjer, f.eks. støvsugere og græsslåmaskiner.

Konklusion

Kunstig intelligens er et stærkt fagfelt i rivende udvikling og med stort potentiale, selvom det man konstruerer ikke nødvendigvis fortjener etiketten 'intelligent'. Men i takt med udviklingen af den kunstige intelligens får vi også gradvist en bedre forståelse af hvad intelligens så reelt er, idet vi—om ikke andet—få den afgrænset nedefra.

<http://www.imm.dtu.dk/~tb>