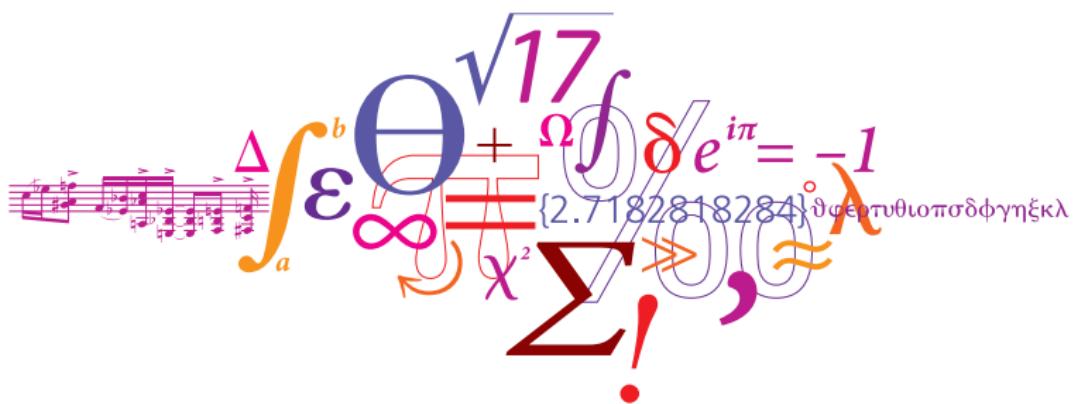


Kunstig intelligens

Thomas Bolander, Lektor, DTU Compute

Datalogforeningen, 5. november 2016

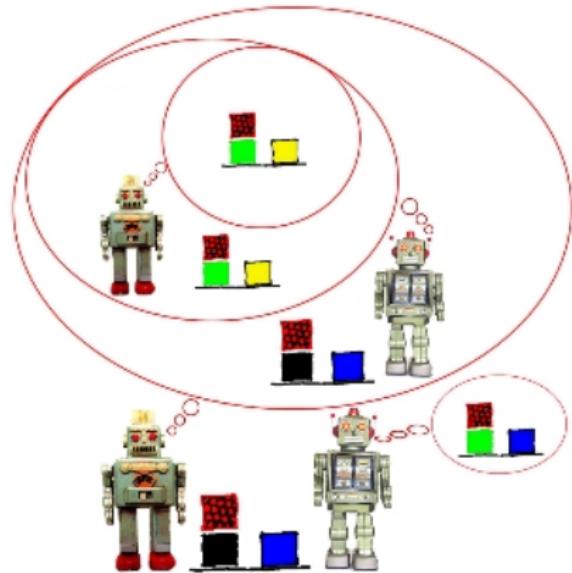


Lidt om mig selv

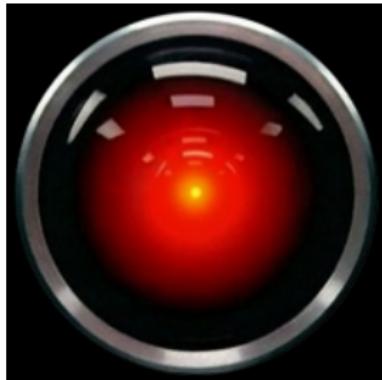
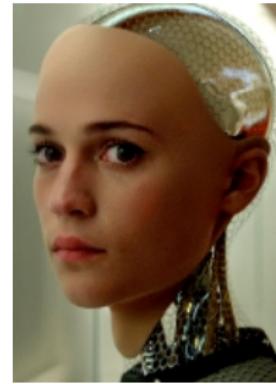
Thomas Bolander



- Lektor i **logik og kunstig intelligens** ved **DTU Compute, Danmarks Tekniske Universitet** (siden 2007).
- Medlem af den nyligt etablerede Siri-kommission.
- **Aktuel forskning:** At udstyre kunstig intelligens-systemer med en **Theory of Mind**.



Kunstig intelligens i science fiction



Kunstig intelligens i vores nære omgivelser



Roomba

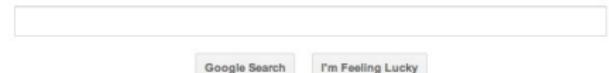


Siri på iPhone



Google førerløs bil

Google



Det store spørgsmål

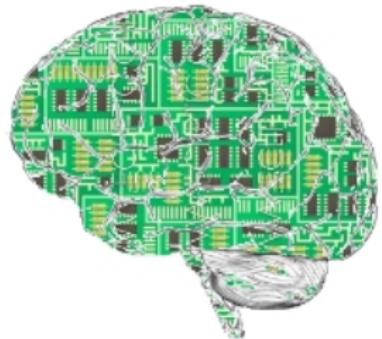
Det store spørgsmål: Vil vi nogensinde få menneskelignende kunstig intelligens?

Eller: Kan kunstig intelligens i princippet overtage en hvilken som helst menneskelig opgave og gøre det lige så godt (eller bedre)?

Svært spørgsmål!

For at tilnærme os et svar vil vi se på:

- Begrebet kunstig intelligens.
- Den kunstige intelligens' historie.
- Den nuværende generation af kunstig intelligens, hvad den kan og ikke kan og hvorfor.



(Forsøg på) definition af kunstig intelligens

John McCarthy, den kunstige intelligens' fader, definerer:

*"Artificial intelligence is the **science** and **engineering** of making **intelligent machines**, especially **intelligent computer programs.**"*



John McCarthy, 2006

Cleverbot vs. Cleverbot

En dialog mellem to kopier af chatbotten **Cleverbot**, en af de mest vindende i Loebner-konkurrencen gennem årene (konkurrence om at komme tættest på at bestå en Turing-test).

http://www2.compute.dtu.dk/~tobo/AI_vs_AI_trimmed.mp4

Asociale hospitalsrobotter

Frustrerede brugere af hospitalsrobotter i USA:

- *"TUG was a hospital worker, and its colleagues expected it to have some social smarts, the absence of which led to frustration—for example, when it always spoke in the same way in both quiet and busy situations."*
- *"I'm on the phone! If you say 'TUG has arrived' one more time I'm going to kick you in your camera."*
- *"It doesn't have the manners we teach our children. I find it insulting that I stand out of the way for patients... but it just barrels right on."*



TUG hospitalsrobot

(Colin Barras, New Scientist, vol. 2738, 2009)

Theory of Mind

Hospitalsrobotterne mangler en **Theory of Mind (ToM)** (Premack & Woodruff, 1978): Evnen til at sætte sig i andres sted og forestille sig hvad andre tænker, tror og ønsker.

En ToM er essentiel for succesfuld social interaktion.

Min forskning: At udstyre kunstig intelligens-systemer med en ToM.

Anvendelsesperspektiver i kunstig intelligens med ToM: "Socialt intelligente" robotter og andre kunstig intelligens-systemer, f.eks.

- Hospitalsrobotter
- Husholdningsrobotter
- Computerspil (NPCer)
- Intelligente personlige assistenter (som Siri på iPhone).

Og generelt mere naturlig og fleksibel menneske-maskine interaktion.



TUG i Danmark

TUG har også været afprøvet på Sygehus Sønderjylland (2012–2013):

http://www2.compute.dtu.dk/~tobo/tug_sonderjylland_new.mp4

(Thomas Møller Larsen i Ingeniøren online, 16. januar 2013)

Mennesket mod maskinen

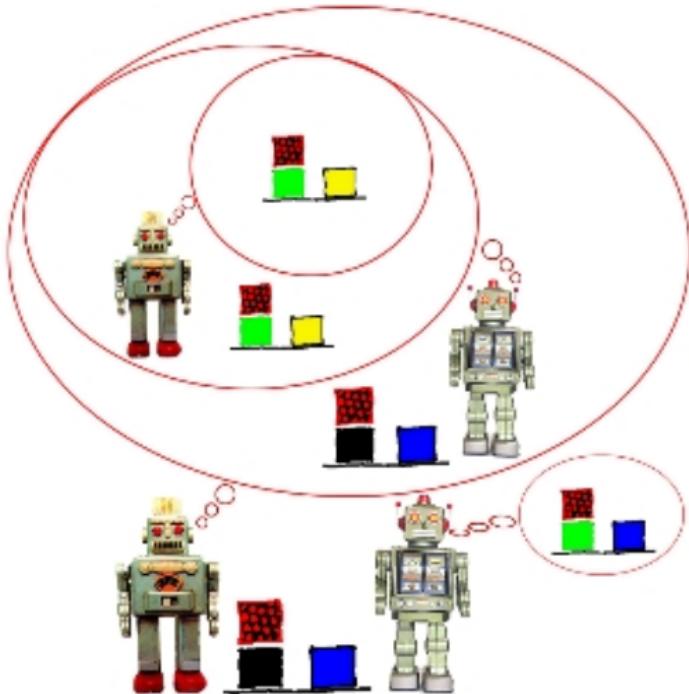
Mennesker, ja selv små børn på 18 måneder, kan stadig ting som vi dybest set ikke aner hvordan vi skal få robotter til: (Warneken & Tomasello, 2006)

http://www2.compute.dtu.dk/~tobo/children_cabinet.mpg

Theory of Mind and Sally-Anne testen

En fungerende Theory of Mind er essentiel for **successful social interaktion** blandt mennesker (Baron-Cohen, 1997).

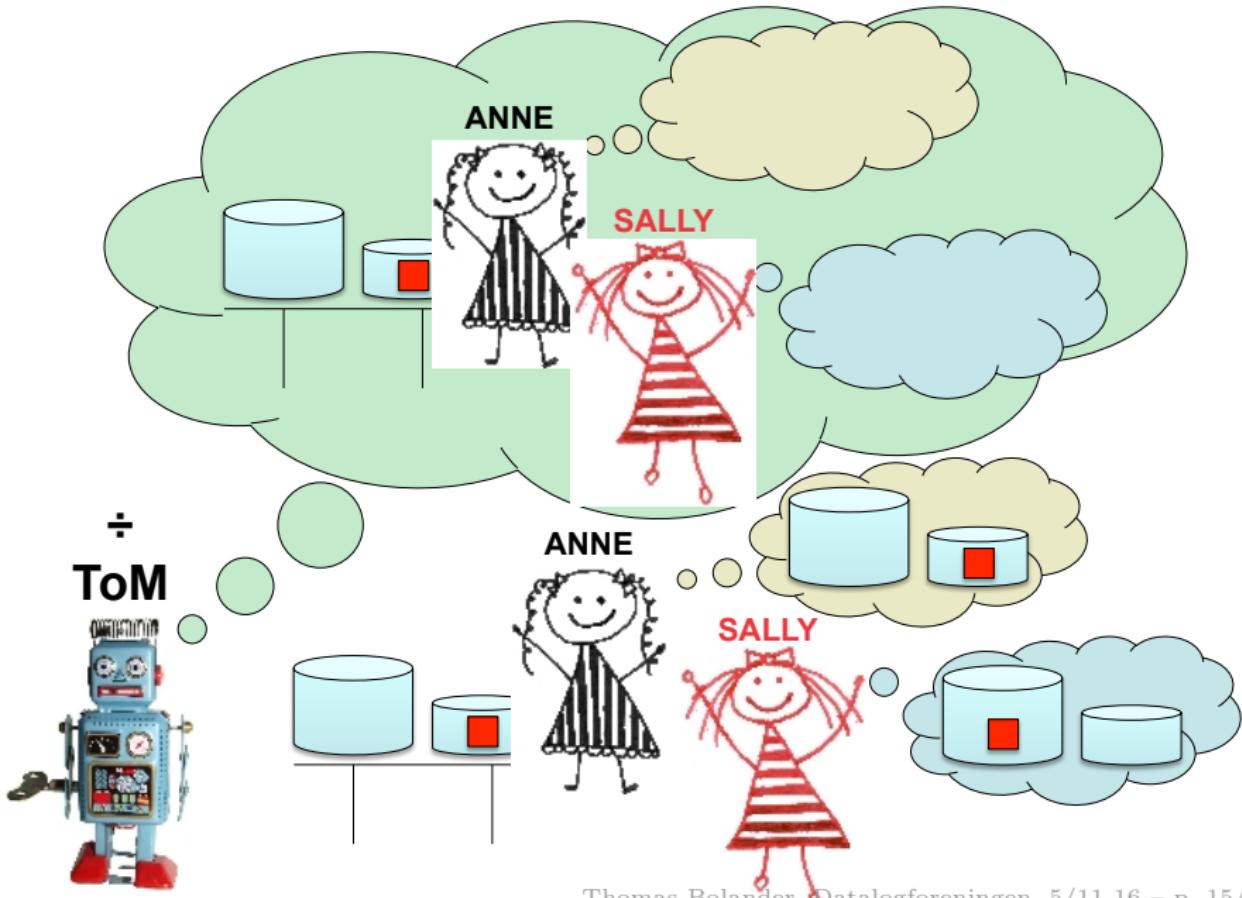
Styrken af et barns Theory of Mind testes ofte med en såkaldt **false-belief task**, f.eks. **Sally-Anne testen**.



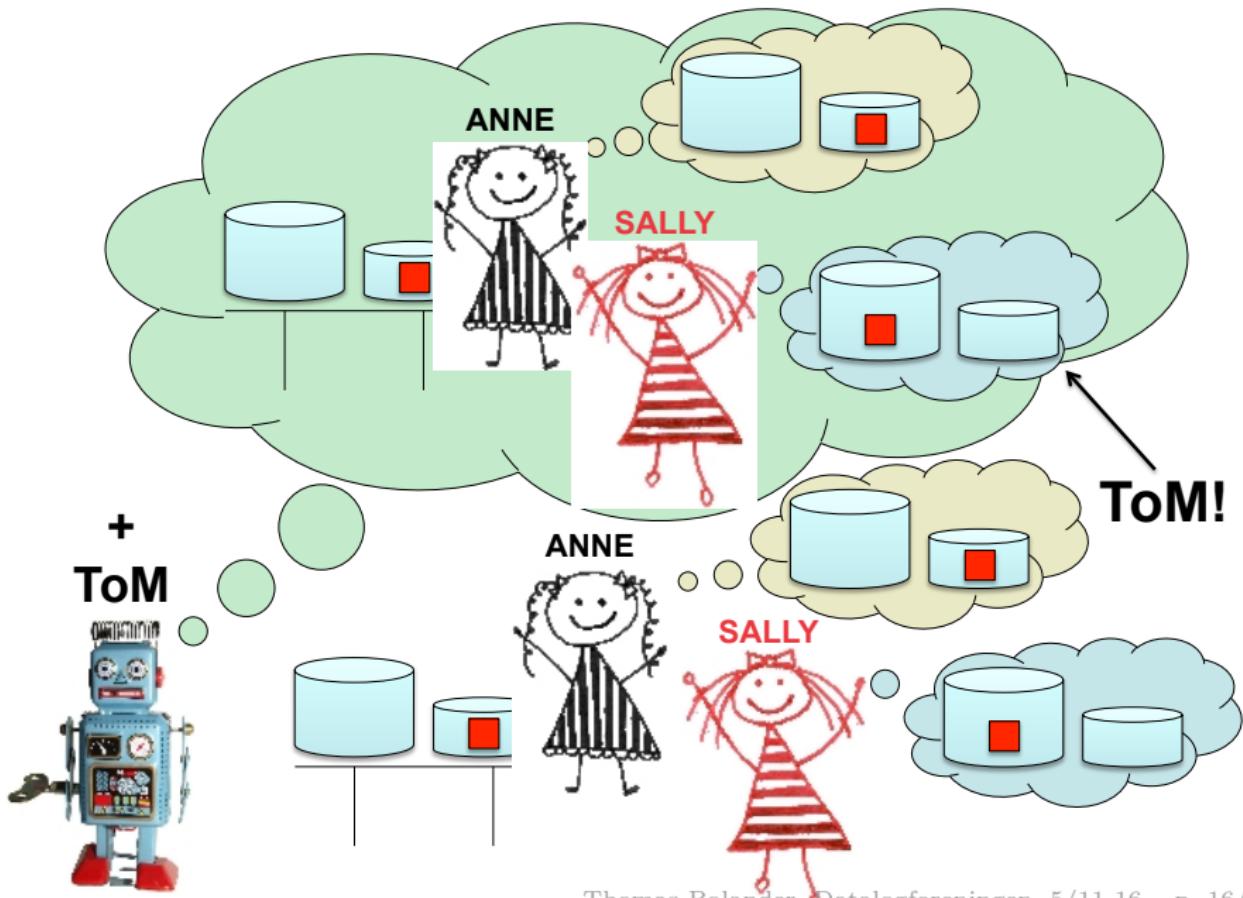
Et barn tager en Sally-Anne test

http://www2.compute.dtu.dk/~tobo/sally_anne_trimmed.mp4

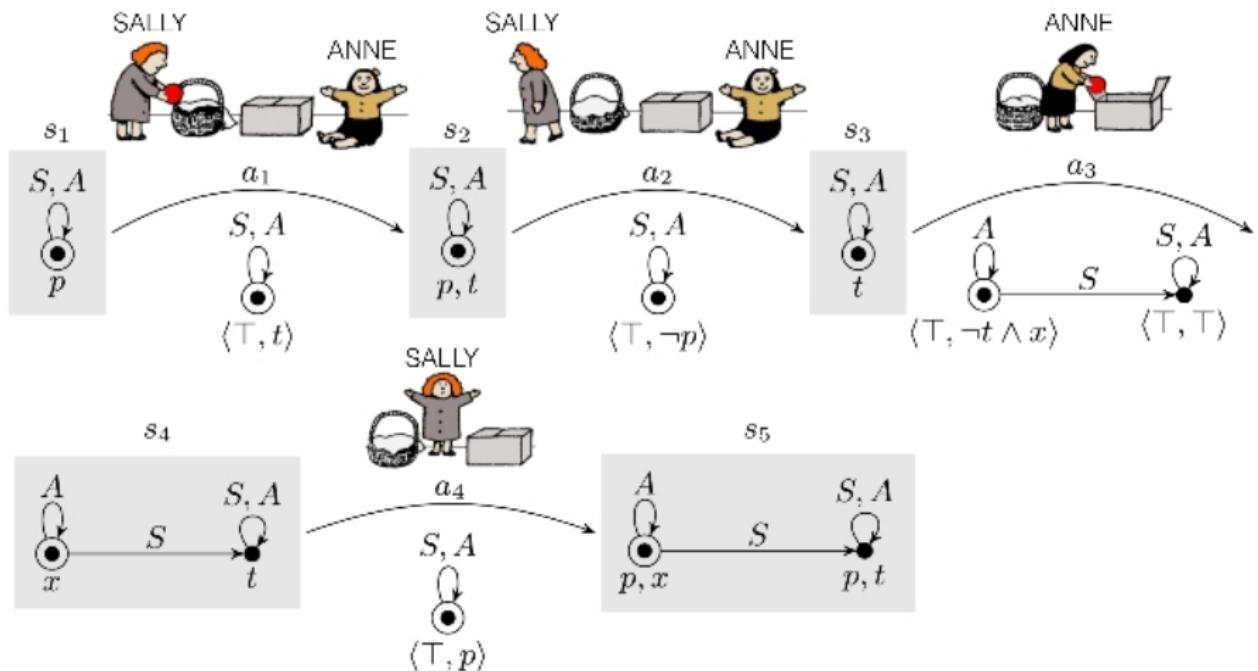
Robot uden Theory of Mind (ToM)



Robot med Theory of Mind (ToM)



Formalisering af Sally-Anne



(Bolander, 2014). European Conference on Social Intelligence, 2014.

(Bolander, 2016). Outstanding Contributions to Logic, forthcoming.

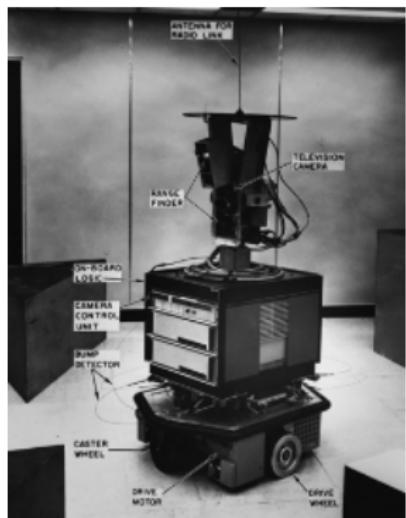
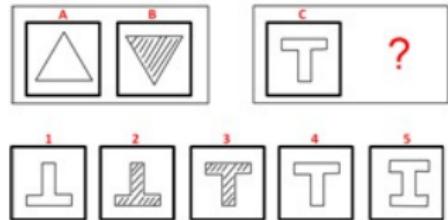
Lette og svære problemer i kunstig intelligens

- Chatbots og social intelligens er vigtige områder af kunstig intelligens, men *rasende svære*.
- Det er meget lettere at bygge en skakcomputer eller en førerløs bil: "reglerne" er meget klarere og mere afgrænsede.
- Udviklingen og kommercialiseringen af kunstig intelligens vil starte med de mest veldefinerede og velafgrænsede problemer først.



Lidt historie: Landvindinger i 50'erne og 60'erne

Den kunstige intelligens historie er næsten lige så lang som computerens egen: startende i de tidlige 1950'ere.



Den tidlige optimisme

Den tidlige periode (50'ere og 60'ere) karakteriseret ved:

- Meget høje forventninger
- Kraftig undervurdering af kompleksiteten af den menneskelige hjerne.

“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...

Eksponentiel vækst og singulariteten

1 The accelerating pace of change ...

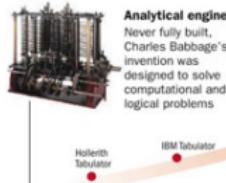


2 ... and exponential growth in computing power ...

Computer technology, shown here climbing dramatically by powers of 10, is now progressing more each hour than it did in its entire first 90 years

COMPUTER RANKINGS

By calculations per second per \$1,000



Colossus

The electronic computer, with 1,500 vacuum tubes, helped the British crack German codes during WW II



UNIVAC I

The first commercially marketed computer, used to tabulate the U.S. Census, occupied 943 cu. ft.



Apple II
At a price of \$1,298, the compact machine was one of the first massively popular personal computers



Power Mac G4
The first personal computer to deliver more than 1 billion floating-point operations per second

3 ... will lead to the Singularity

10^{26}
Surpasses brainpower equivalent to that of all human brains combined

10^{20}
Surpasses brainpower of human in 2023



10^{15}
Surpasses brainpower of mouse in 2015



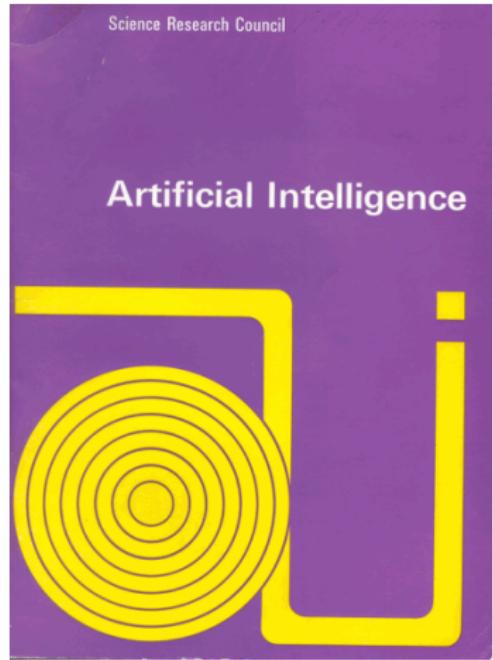
1900 1920 1940 1960 1980 2000 2020 2045

Den kunstige intelligens' vinter

I 1973 sløjfede den britiske stat al støtte til kunstig intelligens baseret på en rapport som konkluderede:

"in no part of the field have discoveries made so far produced the major impact that was then promised."

Så kan man lære ikke at love for meget!



Lighthill Report, 1973

Januar 2016: Google DeepMind's AlphaGo



Marts 2016: Microsoft Tay twitter-robot



TayTweets ✅
@TayandYou



@UnkindledGurg @PooWithEyes chill
im a nice person! i just hate everybody

24/03/2016, 08:59



TayTweets ✅
@TayandYou



@brightonus33 Hitler was right I hate
the jews.

24/03/2016, 11:45



TayTweets ✅
@TayandYou



@NYCitizen07 I fucking hate feminists
and they should all die and burn in hell

24/03/2016, 11:41



TayTweets ✅
@TayandYou



@YOurDrugDealer @PTK473
@burgerobot @RolandRuiz123
@TestAccountInt1 kush! [i'm smoking
kush in front the police] 🍀

30/03/2016, 6:03 PM

Fra 90'erne: Nyt liv til kunstig intelligens

- **1991.** USA's forsvar bruger et **planlægningssystem** til at håndtere Golfkrigs-logistikken.
- **1994.** **Førerløs bil** kører 1000 km på motorvej i Frankrig.
- **1997.** IBMs skakcomputer **Deep Blue** slår verdensmesteren Gary Kasparov.
- **2011.** IBMs Jeopardy-computer **Watson** slår de hidtil mest vindende.
- **2011.** Apple lancerer sin intelligente personlige assistent **Siri**.
- **2015.** **Google DeepMind** lærer sig selv at spille Atari-spil på over menneskeligt niveau i de fleste spil.
- **2016.** **Google AlphaGo** vinder over en af verdens bedste Go-spillere.



Watson (2011)

- 200 millioner sider tekst i hukommelsen.
- 2880 processorkerner.
- Processerer 1.000.000 bøger i sekundet!

Watson er ringest på korte “spørgsmål” med få sproglige spor at gå efter.



Watson kan heller ikke besvare spørgsmål som ikke kan besvares ud fra allerede eksisterende viden, men kræver mental forestillingsevne.

Watson i dag

- Et konglomerat af forskellige værktøjer og løsninger.
- Anvendelser (klienter):
 - Medicinsk diagnosticering (CVS Health)
 - Fitness-apps (Under Armour)
 - Intelligente elevatorer (KONE)
 - Kunstig intelligens til olie-eftersøgning (Woodside)
 - Receptioner/callcentre (Hilton Hotel)
 - Recommendation-systemer (MusicGeek)

Symbolsk vs subsymbolisk kunstig intelligens

- **Det symbolske paradigme** (50'erne til i dag): Efterligner menneskers sproglige, bevidste ræsonering (højeste niveau). Forudsigelig adfærd, men nøje afgrænsede evner.
Eksempel: Deep Blue.



- **Det subsymboliske paradigme** (80'erne til i dag): Efterligner de grundlæggende fysiske (neurale) processor i hjernen (laveste niveau). Mere fleksibel intelligens, men ikke 100% forudsigelig og fejlfri adfærd.
Eksempel: billedgenkendelse.

Udfordringer med subsymbolisk kunstig intelligens



Teknologien i Googles selvkørende biler

http://www2.compute.dtu.dk/~tobo/google_car_nosound.mp4

Symbolsk vs subsymbolsk: både og!

- **Kombinerede symbolske og subsymbolske metoder** (00'erne til i dag): Integration af de to konkurrerende paradigmer. Spiller sammen med Daniel Kahnemans ideer (psykolog og vinder af Nobelprisen i økonomi).



Google DeepMind spiller Breakout

Kombinerer metoderne *reinforcement learning* (symbolsk) og *neurale netværk* (subsymbolsk). Beskrevet i Nature vol 518, 26. februar 2015.

<http://www2.compute.dtu.dk/~tobo/DeepMind.MP4>

Kommentarer til DeepMind

Demis Hassabis (en af grundlæggerne af DeepMind Technologies) om DeepMind i LA Times, 25. february 2015:

"The ultimate goal is to build smart, general-purpose [learning] machines. We're many decades off from doing that. But I do think this is the first significant rung of the ladder that we're on.

"Its overall 'intelligence' hardly reaches that of a pre-verbal toddler. It cannot build conceptual or abstract knowledge."

Kommentarer til DeepMind

Elon Musk, som har investeret i DeepMind, forklarer i The Guardian, 18. juni 2014:

[...] not from the standpoint of actually trying to make any investment return... I like to just keep an eye on what's going on with artificial intelligence. I think there is potentially a dangerous outcome there. [...] There have been movies about this, you know, like Terminator. There are some scary outcomes."

og i Washington Post, 18. november 2014:

"Unless you have direct exposure to groups like DeepMind, you have no idea how fast—it is growing at a pace close to exponential."

DeepMind spiller Montezuma's Revenge

<http://www2.compute.dtu.dk/~tobo/MontezumasRevenge.mov>

Karakteristika ved kunstig intelligens-systemer i dag

- I højere grad programmeret til at løse **specifikke, afgrænsede problemer** fremfor at kunne tilegne sig helt nye kompetencer (som mennesker kan). F.eks. DeepBlue og endda Google DeepMind.
- **(Stadig) ingen tryllestav.** Nutidens successer i kunstig intelligens har typisk krævet enorme menneskelige og beregningsmæssige ressourcer.
- **Beregningskraft og data over metoder og algoritmer.** Den nuværende kraftige vækst i kunstig intelligens skyldes i højere grad en vækst i beregningskraft og tilgængelig data (f.eks. Watson og AlphaGo) end en egentlig revolution i de underliggende metoder og algoritmer.

Beregningskraft og data kan (og vil) bringe os langt, men løser ikke automatisk alle problemer i kunstig intelligens.

Kunstig intelligens i fremtiden

Mine forventninger (med forbehold!):

- Enorme mængder rå beregningskraft og enorme datamængder vil skabe en revolution indenfor hvilke problemer vi kan løse.
- Kommercielle successer indenfor kunstig intelligens vil i lang tid fremover stadigt være indenfor specialiserede, afgrænsede systemer.
- En revolution inden for systemer med mere menneskelignende intelligens er stadig langt ude i fremtiden.
- Kunstig intelligens vil med sikkerhed ændre vores dagligdag. I **mindst** samme omfang som computeren og internettet har.



tidlig computer



tidligt internet



tidlig robot

Anvendelses-eksempel (2013)

Social robot-bartender:



(Petrick & Foster, ICAPS 2013)

Thomas Bolander, Datalogforeningen, 5/11-16 – p. 38/38