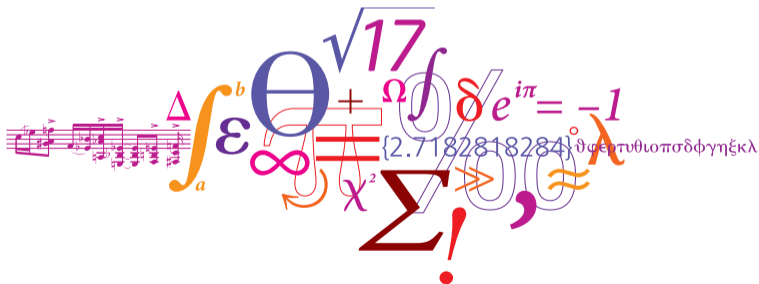


Socialt intelligente robotter (så meget som de nu kan blive det)

Thomas Bolander, DTU Compute, Danmarks Tekniske Universitet

16. september 2019





Vagn Lundsgaard Hansen



Mai Gehrke



Holger Bech Nielsen



Douglas Hofstadter



Maria, Audra, Jonas

H. C. Ørsted: Altings sammenhæng— og tranformationer mellem repræsentationsrum



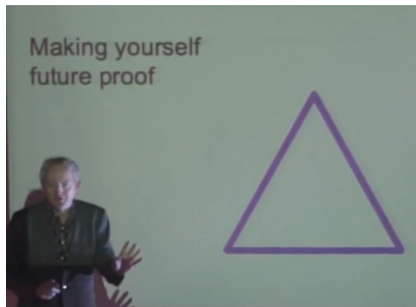
[http://www2.compute.dtu.dk/~tobo/
klangfigurerer.mp4](http://www2.compute.dtu.dk/~tobo/klangfigurerer.mp4)

Hvordan undgår vi at
fremtidens robotter
og computere bliver
sociale analfabeter?

De 3 sværeste problemer i kunstig intelligens



Carl Frey, 20. april 2017
*Udd.- og Forskningsministeriets
Årsmøde 2017, Kolding*



Toby Walsh, 18. marts 2017
Science & Cocktails, København

Begge har **social intelligens** blandt de 3 menneskelige evner som er sværest at simulere på computere og i robotter.

Social intelligens: Evnen til at forstå andre og den sociale kontekst, og dermed at interagere succesfuldt med andre.

Social kunstig intelligens: Hvorfor?

- Fleksibel og naturlig interaktion med mennesker.
- Forklarlighed: AI-systemer der kan gøre sig forståelige for mennesker (skabe tillid).

Eksempel på nødvendigheden af social intelligens. Kritiske brugere af hospitalsrobotter i USA.

- *"I'm on the phone! If you say 'TUG has arrived' one more time I'm going to kick you in your camera."*

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



TUG hospitalsrobot

Social intelligens: Hvad er det?

Et kognitionspsykologisk eksperiment med en dreng på 18 måneder. Han har ikke fået nogen instruktioner.

(Warneken & Tomasello, 2006)

http://www2.compute.dtu.dk/~tobo/children_cabinet_trimmed.mov

Social intelligens: Hvad er det?

Et kognitionspsykologisk eksperiment med en dreng på 18 måneder. Han har ikke fået nogen instruktioner.

(Warneken & Tomasello, 2006)

Drengen ser ud til at have evnen til at **sætte sig i den voksnes sted** og forstå hvad han ønsker at opnå.

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

Theory of Mind er essentiel for menneskers sociale intelligens.

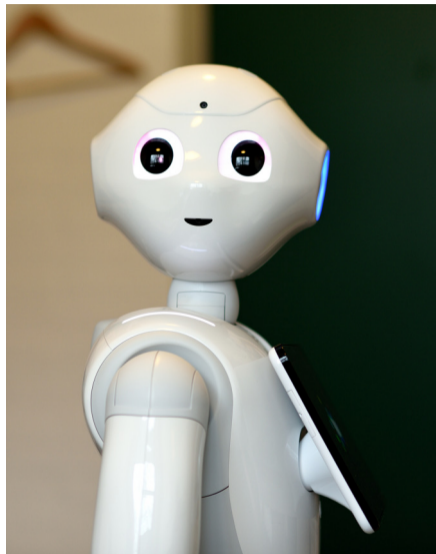
(Baron-Cohen, 1997)



Et barn tager Sally-Anne testen

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

Demo af vores Pepper-robot R2DTU



To hovedparadigmer i kunstig intelligens – og i R2DTU

Det symbolske paradigme (1950–): Simulerer menneskelig symbolsk, sproglig, bevidst ræsonnering. Søgning, planlægning, logisk ræsonnering. **På robotten:** symbolsk forankring (DTU), Theory of Mind (DTU), automatiseret planlægning (DTU), implicit koordinering (DTU), sprogforståelse (DTU).

↑
regler og ræsonnementer

mønstre og perception

↓
Det subsymbolske paradigme (1980–): Simulerer de fundamentale fysiske processer i den menneskelige hjerne. Neurale netværk. **På robotten:** genkendelse af bevægelse (CMU), tale (DTU), objekter (U Michigan), ansigter (CMU).

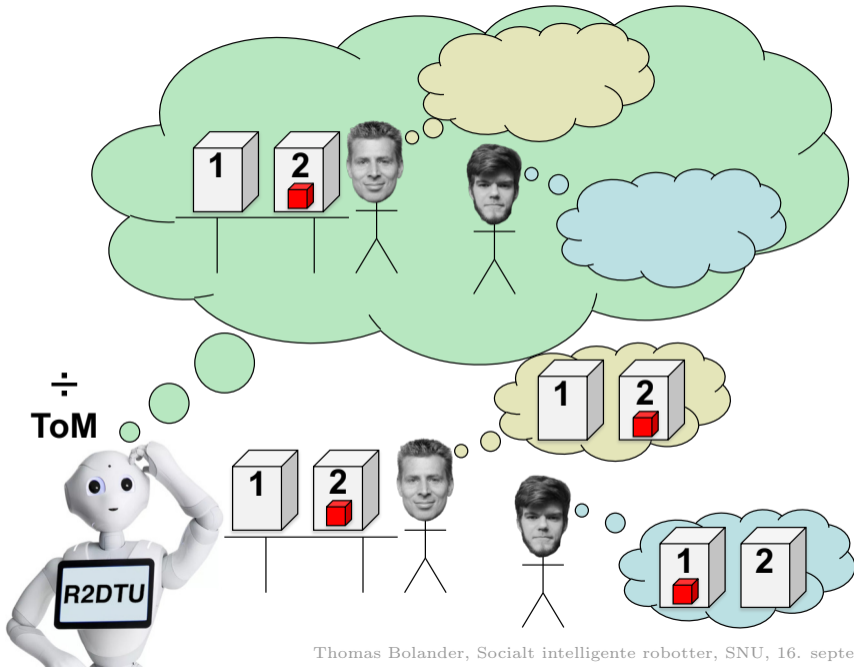


Dronningens nytårstale

Så er det nytårsaften. Vi har alle forskellige mål i livet. Men det danske nytår, der starter i dag, er en tid til at fejre, hvem vi er. Det er tid til at gøre noget andet. Det er tid til at give tilbage, at tage noget tilbage, at skabe noget nyt. Det er tid til at tænke på, hvad der er vigtigt i livet. Vi ønsker ikke bare at bevare det, vi allerede har.

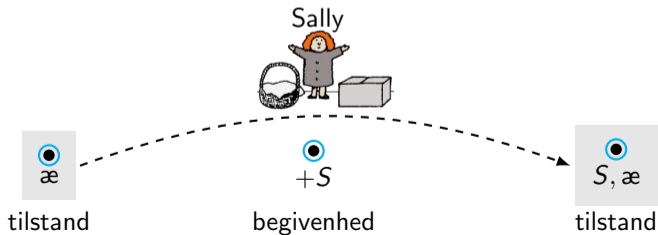
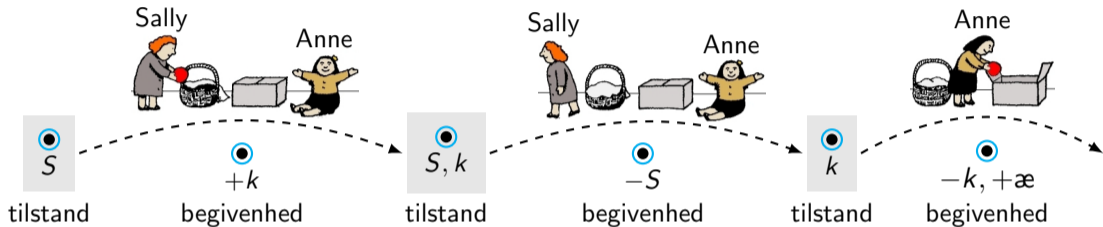
Nogle vil hellere tilbringe tid med en, de elsker, end med en gruppe venner. Måske vil familien gerne spille et spil sammen eller lege tagfat. For et par år siden startede Danmark en sport, der er meget gammel, men stadig meget vigtig. Det er en lille turnering kaldet "Svenskninge". Den finder sted hvert år. Her er tidsplanen: lørdag 15.-16. januar 20:00-19:00 og søndag 17.-18. marts 18:00-19:00. Jeg ønsker alle et godt nytår, og håber I forbliver sunde og raske! Hvis I har spørgsmål eller kommentarer, så kontakt mig endelig.

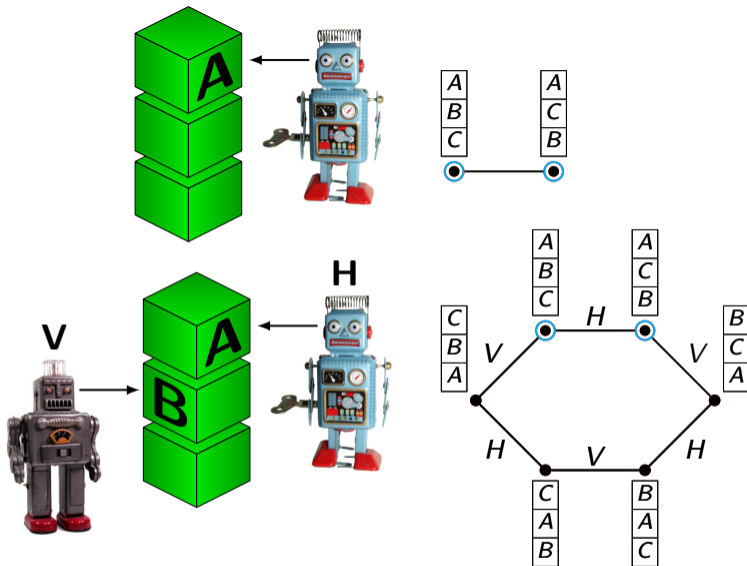
GUD BEVARE DANMARK



Sally-Anne uden en Theory of Mind

S: Sally. A: Anne. k: kurv. æ: æske.

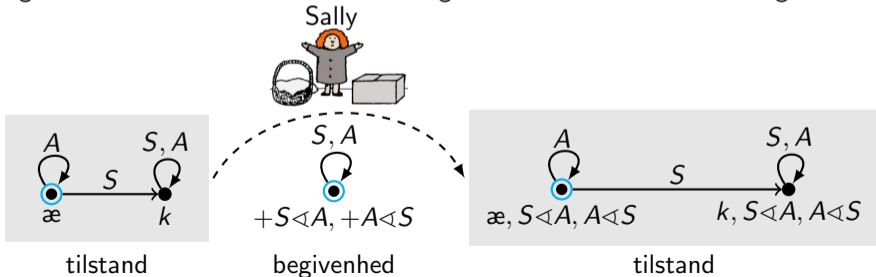
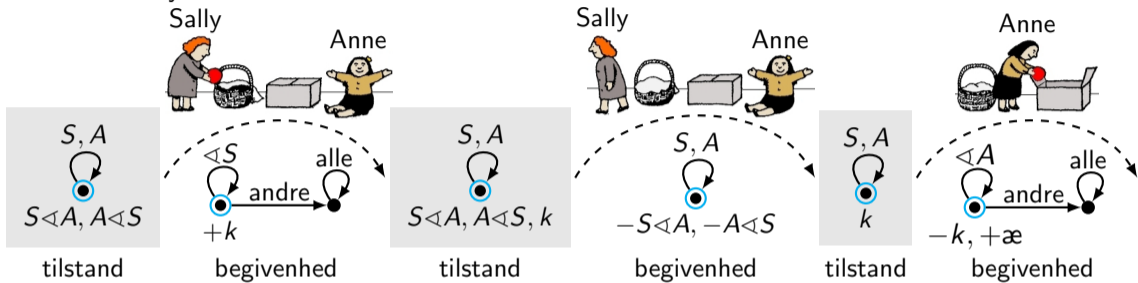




Modellerne til højre kaldes **epistemiske tilstande**. Punkterne er de **mulige tilstande**, kanterne angiver **usknelighed** mellem tilstande. Den blå robots verdensbillede er beskrevet ved punkterne markeret \bullet .

Sally-Anne med en Theory of Mind

S: Sally. A: Anne. k: kurv. æ: æske.

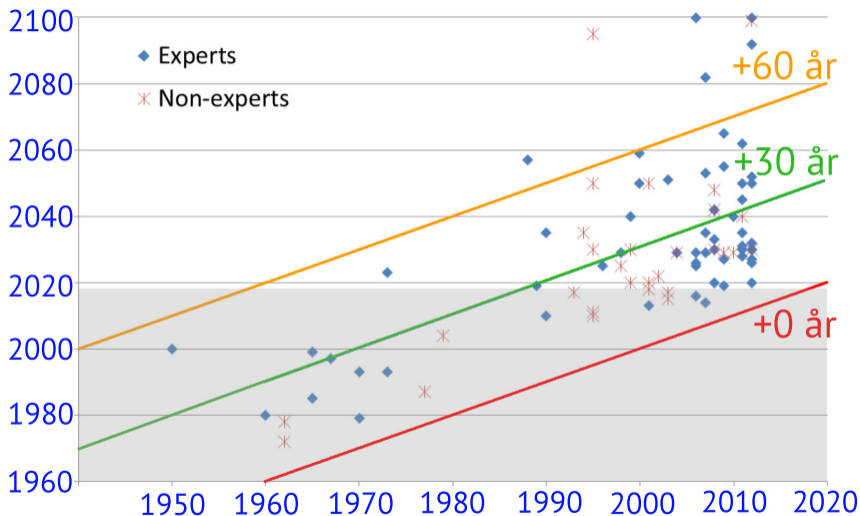


(Bolander: *Seeing is Believing*. Chapter in *Outstanding Contributions to Logic*, Springer 2018)

Fremtiden for menneskelig og kunstig intelligens?



Hvornår kunstig intelligens på menneskeligt niveau?



(Armstrong & Sotala: How We're Predicting AI—or Failing To. Beyond Artificial Intelligence, Springer, 2015) med farvede linjer og gråt område tilføjet af mig.