

Kognitionsvetenskaplig introduktionskurs

Föreläsning 4:
Fysiska symbolsystemhypotesen,
Language of thought,
Kinesiska rummet

1

Vad är tänkande?

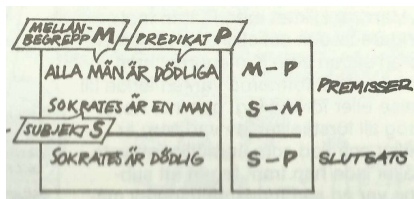
Enligt kognitionsvetenskapen: ett slags informationsberäkningsprocess

- **Tänkande som symbolisk beräkning (algoritmer, GOFAI)**
- Tänkande som sub-symbolisk beräkning (neurala nät)
- Tänkande som en situerad och distribuerad process
- Tänkande som en förkroppsligad process

2

Symbolisk beräkning

- använder symboler (representationer)
- som man manipuleras formellt genom formella procedurer (regler/algoritmer)
- Aristotelisk syllogism som exempel på symbolisk beräkning



3

3

Symbolisk beräkning

- I **logiskt giltiga** resonemang följer slutsatsen med logisk nödvändighet ur premisserna
- Alltså: om premisserna i ett logiskt giltigt resonemang är sanna så är slutsatsen också sann
- Det är formen på argumentet som avgör giltigheten – ej innehållet
- Är följande logiskt giltiga resonemang?

Om P så Q
inte Q

inte P

Om inte P så inte Q
 P

 Q

4

4

Symbolisk beräkning

- Är följande ett giltigt resonemang?
Om P (äter grönsaker) så Q (får efterrätt)
inte Q (får efterrätt)

inte P (äter grönsaker)
- Ja!
- Är följande ett giltigt resonemang?
Om inte P (äter grönsaker) så inte Q (får efterrätt)
 P (äter grönsaker)

 Q (får efterrätt)
- Nej!

5

5

Symbolisk beräkning

- Något i världen ska beräknas
- Representation i ett symbolsystem sätts upp, t ex A, B, C
- Beräkning sker på symbolerna (utan hänsyn till deras innehåll)
- Resultatet (symbolerna) tolkas/avläses i relation till världen
- Skiljer på Syntax (form) och Semantik (innehåll)!


6

6

Exempel

Hur skulle ni göra en symbolisk beräkning av följande?

Om det är onsdag idag, så är Sam på jobbet.
 Det är onsdag idag.
 Alltså är Sam på jobbet.






7

7


Representationer och symbolers medium

- En symbol har alltid en fysisk existens, men denna är mediumoberoende, så länge det finns metoder för "avläsning"
- Exempel

Mugg

01101101 01110101 01100111 01100111




8

8

Fysiska symbolsystemshypotesen

- "Ett fysiskt symbolsystem uppfyller nödvändiga och tillräckliga villkor för generell intelligens" (Simon & Newell, 1975)
 - Nödvändigt villkor innebär att om något är intelligent så är det också ett FSS
 - Tillräckligt villkor innebär att om något är ett FSS så är det också intelligent




13

13

Karakteristika hos FSS

1. Symboler är fysiska mönster
2. Symbolerna kan kombineras till komplexa symbolstrukturer
3. Det finns processer för att manipulera komplexa symbolstrukturer
4. Processerna för att generera och transformera komplexa symbolstrukturer kan representeras av symboler



14

14

1. Symboler är fysiska mönster


INSTRUKTION:

Nuvarande tillstånd	Nuvarande symbol	Ny symbol	Förflyttning	Nytt tillstånd
A	1	1	1R	A

A

A

1 1 1 + 1 1 # # # # #



15

15

2. Symbolerna kan kombineras till komplexa symbolstrukturer

INSTRUKTION:

Nuvarande tillstånd	Nuvarande symbol	Ny symbol	Förflyttning	Nytt tillstånd
A	1	1	1R	A


A

A

1 1 1

+ 1 1

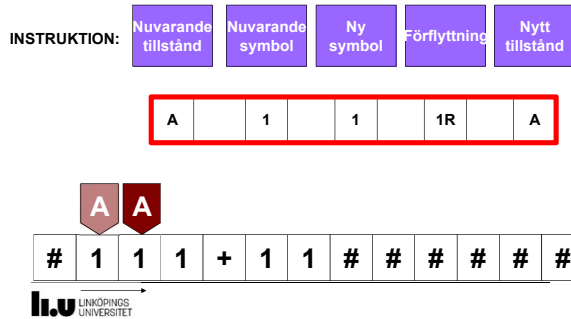
#



16

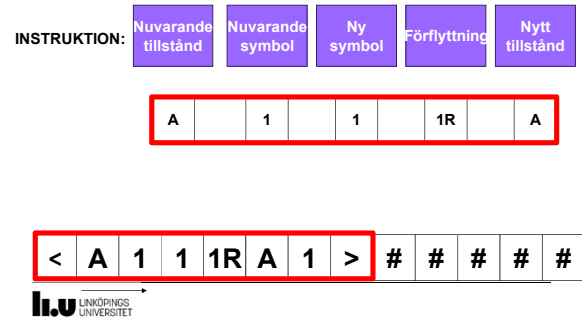
16

3. Det finns processer för att manipulera komplexa symbolstrukturer



17

4. Processerna för att bearbeta komplexa symbolstrukturer kan representeras av symboler



18

Dator som FSS

- Symboler 1 eller 0
 - Fysiska symboler elektriska/logiska kretsar som är på (1) eller av (0)
- Tecken kan kombineras till ord, strängar eller uttryck
 - 011001010110111000100000011011010111010101110011
e n m u s
- Algoritmer bestående av enkla entydiga instruktioner t ex:
 - konkatenering av strängar
 - problemlösning som sökning
- Datorprogram är skrivna i högnivå programmeringsspråk
 - 01101001011001100010000001101101001111000110001
i f m > i

li.u LINKÖPINGS UNIVERSITET

19

Language of thought hypothesis (Fodor, 1975)

- Tänkande sker i ett mentalt språk, *Mentalese*.
- Mentalese skiljer sig i vissa avseenden från naturligt språk (ej introspektiv tillgång, avsaknad av prosodi, m.m.).
- Meningar uttryckta i Mentalese är fysiska symbolstrukturer som manipuleras enligt syntaktiska regler (ett slags grammatik). Alltså: tanken har, precis som språket, **kombinatorisk syntax**.
- Förändringar i Mentalese (dvs. på en fysisk/formell/syntaktisk nivå i symbolsystemet) påverkar systemets *propositionella attityder* (dvs. på en semantisk/meningsfull nivå), och *vice versa*. Alltså: tanken har, precis som språket, **kompositionell semantik**.
- Detta gör det möjligt att förklara "causation by content".

li.u LINKÖPINGS UNIVERSITET

20

Propositionella attityder och Intentional Realism

- Beliefs och Desires är exempel på psykologiska tillstånd
- De kallas ofta för propositionella attityder eftersom de uttrycker
 - attityder (tror att ..., önskar att ..., osv.) till
 - propositioner (det är soligt idag, jag kommer att klara tentan, osv.)
- Mycket mänskligt beteende kan förklaras i termer av propositionella attityder.

li.u LINKÖPINGS UNIVERSITET

21

Propositionella attityder och Intentional Realism

Sinnesintryck (t ex ljud och ljusvågor) omvandlas successivt via informations-bearbetning till olika representationer som resulterar i handling

"Hjälp" ljudvågor och hjärnaktivering (fysiska tillstånd) ->

Jag tror att någon behöver hjälp (Belief – Uppfattning av situation) ->

Jag vill hjälpa (Desire – Önskning) ->

Jag ska springa mot ljudet (Intention – Avsikt) ->

Mina ben rör sig i riktning mot ljudet (handling)

li.u LINKÖPINGS UNIVERSITET

22

Causation by content

- Hur kan psykologiska tillstånd som Beliefs och Desires orsaka beteende?
 - Jämför
 - den fysiska handlingen läkaren som slår med hammaren mot knät utför och som reflexmässigt får benet att röra sig, med
 - den viljestyrda benrörelsen hos den som önskar gå på lina och ska ta första steget
 - I det senare fallet är det **innehållet** i önskan (desire) som **orsakar** beteendet

23

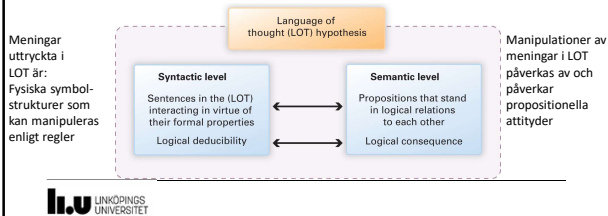
Causation by content, forts

- Omformulering av problemet – Hur kan hjärnan bearbeta information då den endast bearbetar formella egenskaper utan hänsyn till semantiska egenskaper?
- Svar: På samma sätt som datorer!
 - Datorer bearbetar mekaniskt sekvenser av 1:or och 0:or utan minsta kunskap om vad dessa representerar
 - Genom att följa regler som baserats på avsikten med representationen sker dock bearbetningen så att de semantiska egenskaperna respekteras/bevaras

24

Language of Thought

1. Det sker kausal interaktion mellan olika fysiska tillstånd
2. De fysiska tillstånden har en meningslik struktur som avgör hur de interagerar med varandra (Syntaktisk nivå)
3. Kausala relationer mellan meningar respekterar de logiska/rationella relationerna mellan innehållet dessa meningar representerar (Semantisk nivå)



25

Language of Thought

- Förklarar hur FSS kan fungera i mänskligt medvetande
 - Mentala representationer (symboler) realiserar i ett **formellt mentalt språk**
 - Ej talat naturligt språk som svenska eller engelska, universellt och medfött
 - Ej tillgängligt via introspektion
 - Mentala processer (dvs. tänkande) är **kausala interaktioner** mellan dessa strukturer
 - Slutledningar kan göras utifrån regler som endast beror på strukturen/syntax
 - Eftersom det mentala språket är ett formellt språk så följer regler semantiska restriktioner: Om premisser är sanna måste också slutsatser vara sanna

27

Kinesiska rummet (Searle, 1980)

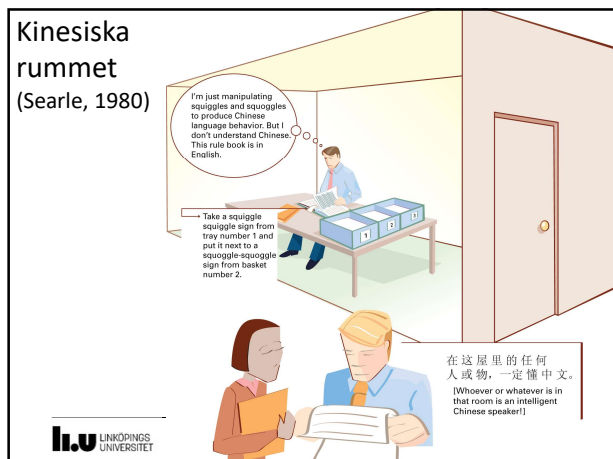
- Argument vars huvudinslag är ett tankeexperiment (en filosofisk metod)
- Argument mot tänkande som symbolberäkning (t.ex. FSSH och LoT): "the appropriately programmed computer with the right inputs and outputs would ... have a mind in exactly the same sense human beings have minds" (Searle, 1980)
- Argument för att symbolmanipulation är inte **tillräckligt** för generell intelligens, säger inte att symbolmanipulation inte är del i tänkande.

29

Kinesiska rummet (Searle, 1980)

- Antag att AI-forskning har lyckats skapa en dator som betar sig som att den förstår kinesiska. Datorn tar kinesiska som input och, genom att följa instruktioner i ett datorprogram, ger kinesiska som output.
- Anta även att datorn är så övertygande att den klarar av Turing-testet: den övertygar en kinesisk talare att programmet själv är en levande kinesisk talare. Datorn svarar lämpligt på alla frågor den får.
- Frågan som Searle vill besvara är: har datorn *bokstavligen* förståelse för kinesiska? Eller *simulerar* den bara förmågan att förstå kinesiska?

31



32

Kinesiska rummet – motargument

- **Systemsvaret**
 - Personen utför flera komplicerade deluppgifter: känner igen kinesiska tecken, slår upp i regelbok, sätter ihop svar
 - Systemet som helhet, dvs personen och regelboken, förstår kinesiska
- **Robotsvaret**
 - Behöver interagera med omgivning och andra människor för att kunna förstå

li.u LINKÖPINGS UNIVERSITET

33

Symbol grounding problem

- Kinesiska rummet illustrerar ett vidare problem – gapet mellan formell bearbetning av symboler och genuin tanke och förståelse
- Men hur blir symboler meningsfulla (även i det mänskliga fallet)?
 - ord (lingvistiska symboler) är meningsfulla eftersom vi lär oss en betydelse när vi lär oss att använda dem
 - men tankar verka ha betydelse i sig själva – vi har kört fast
- Möjliga vägar
 - Kognition är inte informationsbearbetning
 - Kognition är icke-symboliska informationsbearbetning

li.u LINKÖPINGS UNIVERSITET

34