

INNEHÅLL

Hallå där, kompis!	9
Medveten mot alla odds	13
Ur led är tiden	22
Rummet växer	36
Närvaro	49
Tänkande materia I	62
Exorcisterna	75
Experimentalisterna	92
Den (o)fria viljan	115
Spegelbild av en främling	129
Guds avbilds avbild	151
Evolution	170
Zombien får en själ	194
Tänkande materia II	210
Epilog	231
Tack	237
Källförteckning	238

Till Miriam, Naima och Amala
– de tre nornor som spunnit de bästa och vackraste
trådarna i min livsväv.

Hallå där, kompis!

PROFESSOR HENRIK SCHÄRFE öppnar dörren till sitt arbetsrum på Aalborgs universitet för att släppa in mig. Jag har knappt hunnit ta ett par steg in i rummet förrän jag förnimmer den där typiska närvarokänslan när man upptäcker att *hoppсан, där sitter det visst någon!* Alldeles riktigt – därinne sitter samme Henrik Schärfe som just nu håller på att stänga dörren bakom oss.

Jag förstår att den sittande är Schärfes mekaniska tvilling Geminoid-DK. Porträttlikheten är perfekt, huden av silikon ser levande ut, och roboten är lika välvårdad och välklädd som sin förlaga. Jag kan se den göra de där nästan omärkliga tyngdpunktsförskjutningar som alla människor gör. Och om man fokuserar blicken på skjortbrösten kan man ana en långsam andhämtning. Geminoid-DK är just nu aktiverad med en automatisk rörelseslinga som är tillräckligt lång och komplicerad så att inte känslan av upprepning smyger sig på en i första taget. Roboten sitter vid skrivbordet. Schärfe tar plats vid bordsändan och själv sätter jag mig mittemot roboten. Jag samtalar med Schärfe, kastar ner en och annan anteckning men har svårt att koncentrera mig. Jag sneglar på roboten som inte alltid tycks

lyssna på intervjun. Då och då vänder den på huvudet och tittar förstrött ut genom fönstret. Den böjer sig framåt, lutar sig bakåt och så – plötsligt – gör den en snabb huvudvridning, stirrar rakt på mig och läpparna kröks i ett svagt, *menande* leende. Ja, vad menar den? Jag känner mig egendomligt generad och drabbas av en flashback.

På 1970-talet när AI, artificiell intelligens, var nytt och mycket i ropet, var jag i USA och samlade material till min bok *Ingenjörstrupper i hjärnan*. Jag besökte Kenneth Mark Colby, en f d elitspelare i schack, på ett laboratorium i Los Angeles. Colby hade utvecklat datorprogrammet PARRY (som står för paranoid). Programmet simulerade beteendet hos en psykiatrisk patient som plågades av fixa idéer; syftet var just att testa en teori om hur paranoida vanföreställningar uppstår. Jag kommunicerade via ett tangentbord på engelska med PARRY och presenterade mig som hans nya psykiater. PARRY älskade att få tala om sina problem och var mycket positiv till sin nya doktor – ända tills jag satte hårt mot hårt och påstod att han fantiserade. Då blev han rasande och kallade mig för en kvacksalvare som borde byta jobb. Jag var faktiskt tacksam för att han inte kunde komma störtande livs levande ut ur datorskärmen med mord i blicken. Då säger Colby plötsligt till mig:

– Vi har faktiskt lurat dig... Vi har arrangerat den här intervjun med en verklig patient!

För bråkdelen av en sekund trodde jag honom och greps av panik. Elektronisk paranoia är smittsam. Här sitter jag nu och nonchalerar Geminoid-DK på samma sätt som jag inte riktigt tog PARRY på allvar. Med en diffus känsla av skuld. Eller skäms jag för att jag inte vet hur jag ska förhålla mig till en docka?

Det är just det som är problemet. Sociala robotar tycks likom människor utstråla *närvaro*. Och tvetydigheten i denna

närvaro är vad Henrik Schärfe kallar *blended presence*, en sammanblandad närvaro. Han har redan sett symtomen hos mig.

– Är du medveten om att du ibland kallar geminoiden för »den«, ibland för »han«? Det beror på i vilken situation eller vilket sammanhang du tänker på den. Dess mänskliga drag gör den till »han«; när du talar om den i tekniskt avseende blir det »den«. Du ser kanske att det här rummet är riggat med videokameror som inte är riktade mot roboten utan mot dig. Liksom våra försökspersoner vet du att du möter en robot. Men beroende på situationen rör du dig hela tiden fram och tillbaka på en skala mellan det mänskliga och det maskinella. Det är det som är *blended presence* och som vi undersöker med hjälp av geminoiden, säger Henrik Schärfe.

Det finns en hög gullighetsfaktor hos många robotar – tänk bara på knubbiga robotsälar och andra mekaniska sällskapsdjur som prövats inom åldringsvården eller på de troskyldiga och rörande tappra R2-D2 och C-3PO i Stjärnornas krig. Men när robotar liknar människor kan de väcka helt andra känslor. Robotikprofessorn Masahiro Mori har myntat begreppet *the uncanny valley*, den kusliga dalgången. Om man tänker sig ett diagram där höjdaxeln står för det familjära, för igenkänning, och längdaxeln för graden av mänsklig form hos roboten får man en stigande kurva av ökad igenkänning i takt med växande mänsklig likhet. Men när likheten blir *för* stor drabbas vi av alienation och t o m skräck. Kurvan gör en djupdykning ner i den kusliga dalgången. Först när vi har vant oss vid robotens mänsklighet har vi nått dalgångens andra vägg och känner ånyo ett stigande förtroende.

Och visst kan Geminoid-DK då och då kännas kuslig – framför allt det där menande leendet. Det utstrålar ett slags hemligt samförstånd som jag inte bett om. Jag försöker komma under-

fund med mina egna motstridiga känslor. Varför är jag generad? Henrik Schärfe menar att jag upplever ett slags skändning av »den mänskliga formen« och reagerar på det. Nej, det där låter för abstrakt för att tillfredsställa mig. Snarare finns nyckeln till min känsla i den där upplevelsen i Kenneth Mark Colbys laboratorium i Los Angeles. Vad är det roboten säger när den ler? Den säger:

– Hallå där, kompis! Jag är en maskin. Är inte du det också?

Ja, vad är en människa? En serie hävstänger och en kemisk motor. En maskin. Förvisso är det så – och ändå vet vi alla att detta är en grovt förenklad och missvisande beskrivning. Det är något som fattas.



Henrik Schärfe och Geminoid-ДК.

FOTO: Gorm Olesen

Medveten mot alla odds

GEMINOID-DK är en fjärrstyrd docka men den dag kan rändas när en intelligent och självständigt tänkande maskin kommer att fråga mig om jag också är en maskin och om vi inte kommer från samma verkstad. Jag skulle kunna svara på två sätt. Skillnaden mellan oss, skulle jag kunna säga, är att jag har en själ och det har inte du. Men en logisk maskin kräver omedelbart att jag bevisar riktigheten i detta påstående. Det kan jag inte.

Den andra försvarslinjen är att hävda att jag har en rik inre värld, något som jag naturligtvis inte heller kan leda i bevis. Jag kan dock argumentera för den på ett sätt som kanske inte logiska maskiner men väl logiska människor kan ta till sig. Alltså: jag upplever mening och närvaro på ett sätt du, robot, inte gör. Du är inte ens ett *du*, d v s inte ens en person. Du är ett *det*. Det är ett rimligt antagande att andra människor precis som jag är medvetna och har upplevelser. Jag känner andra som jag känner mig själv. Men jag kan inte känna en maskin på samma sätt som jag känner mig själv. Hos oss människor finns det en inre värld av tät närvaro. Upplevelserna har en tvingande omedelbarhet: himlens svindlande blåhet, doften av krossade smultron, den döende klangen från ett pianoackord, ilningen längs ryggraden men också den pulserande smärtan i en infekterad tandrot. Om

du är en robot kan du bete dig *som om* du upplevde saker och ting men då simulerar du bara något mänskligt. Du är en zombie.

Missförstå mig inte. Jag talar inte om filmdukens zombier, om arméer av halvt förmultnade döingar som sömngångaraktigt marscherar upp ur sina gravar och som är onaturligt starka och suktar efter människokött. Alla zombier har nämligen inte kyrkogårdsjord under naglarna. I de filosofiska salongerna – om man nu kan tala om sådana på 2000-talet – finns det en elegant variant. För zombier figurerar då och då i de filosofiska tankeexperimenten när man diskuterar identitetsproblem eller hur medvetande och kropp förhåller sig till varandra. Det oroväckande med den filosofiska zombien är att det inte är säkert att man kan skilja den från alla oss andra som är övertygade om att vi *inte* är zombier. Den filosofiska zombien är rumsren, välvårdad och synnerligen belevad – men också till synes spontan. Om den snubblar i trappan och slår sig på smalbenet kommer den att bete sig som alla andra: skrika till och gnida det ömmande stället. Men i själva verket har den inte känt någon smärta, för den har inget medvetande utan reagerar mekaniskt med ett inprogrammerat smärtbeteende. Den äter inte för att den är hungrig utan för att den fysiologiska apparaten beordrar den att inta näring. Tro den inte om den säger att maten är delikat; det är bara en inlärd verbal respons.

Kognitionsfilosofen David Chalmers är en av dem som anser att tankeexperiment med zombier är meningsfulla. Anta att du kunde göra en exakt kopia av dig själv genom att foga ihop en sådan atom för atom. Skulle denna kopia vara identisk med originalet? Om det är fallet har man förutsatt att kopian är medveten – att den inte är en zombie. En kopia av David Chalmers skulle rentav som honom skriva filosofiska essäer om medvetandets natur och kanske hävda att David Chalmers

skulle kunna vara en zombie. Detta är nämligen precis vad David Chalmers gör. Han utesluter inte möjligheten att en kopia som överensstämmer med originalet atom för atom ändå skulle kunna vara en zombie. Medvetandet är således enligt Chalmers någonting mer och utöver den materiella strukturen.

Ändå betraktar sig Chalmers som naturalist och materialist. Han är känd för sin distinktion mellan det »lätta« och det »svåra« problemet när det gäller medvetandets natur. Det »lätta« problemet är att hitta korrelat mellan neuron och medvetenhet – att se vilka delar av hjärnan som är aktiva när vi blir medvetna och om denna aktivitet kan tänkas ha speciella kännetecken. Det svåra problemet är att begripa hur en geléartad klump av celler, elektriska impulser och kemiska substanser kan skapa något så till synes väsensskilt från all materia som medvetenhet. Ett annat sätt att uttrycka Chalmers ståndpunkt är att man inte kan reducera subjektiva upplevelser till hjärnprocesser. De är lika fundamentala till sin natur som tid och rum.

Många underkänner dock tankeexperiment med zombier. Vi har ju inte empiriskt kunna påvisa att det finns något sådant som en zombie. Vi kan inte ens föreställa oss hur en zombie skulle bete sig utan en inre värld. Skulle den verkligen inte kunna särskiljas från en normal människa? Om inte, ja, då är det ofruktbart att ens ta upp ämnet, menar skeptikerna. Vi har inte lärt oss något nytt bara för att zombier skulle kunna vara möjliga i *logisk* mening.

Ett av skälen till att zombier är så svårsmälta är säkert att sådana resonemang i sin förlängning får absurda och skrämmande konsekvenser. Det är ett gammalt filosofiskt problem att vi egentligen inte vet vad som rör sig i andra människors medvetanden. Hur kan jag vara säker på att *de andra* inte är zombier? Det finns ingen som helst möjlighet för mig att fastställa detta.