

November 2012 Årgång 33 Nr 4

Filosofisk tidskrift

- ANDERS BERGLUND
3 Kunde Descartes gud ha skapat allt på sådant sätt
att han inte hade existerat?
- LARS BERGSTRÖM
8 Analytiska satser
- TORBJÖRN TÄNNSJÖ
21 Nazismens ideologi och judeutrotningen
- TOM ERIKSSON
29 Alan Turings två teser 1950
- ANDERS LUNDBERG
43 Medvetandefilosofins förbistring
- 54 RECENSION
Lars Bergström om *Filosoferna* av Svante Nordin
- 58 NOTISER

Filosofisk tidskrift

Utges av Stiftelsen Filosofisk tidskrift och Stiftelsen Bokförlaget Thales

Redaktör och ansvarig utgivare: Lars Bergström

Filosofisk tidskrift utkommer med fyra nummer per år

Prenumeration per år inom landet: 200 kr (inkl. moms); utom landet: 300 kr

Lösnr: 55 kr (inkl. moms)

För prenumerationsärenden kontakta:

Nätverkstan ekonomitjänst, Box 31120, 400 32 Göteborg

Telefon 031-743 99 05 Fax 031-743 99 06

ekonomitjanst@natverkstan.net

Förlagsadress: Bokförlaget Thales, Box 50034, 104 05 Stockholm

info@bokforlagetthales.se www.bokforlagetthales.se

Filosofisk tidskrift utges med stöd från Statens kulturråd

Copyright © Respektive författare 2012

Grafisk form: Ulf Jacobsen

Satt med Indigo Antiqua

Tryck: Carlshamn Tryck & Media AB, Karlshamn 2012

ISSN 0348-7482

Kunde Descartes gud ha skapat allt på ett sådant sätt att han inte hade existerat?

1. INLEDNING

En av Descartes centrala metafysiska teser är att Gud med en enda fri viljeakt har skapat allt som existerar förutom sig själv (CSMI, 200–201; CSMII, 293–94; CSMK, 25, 193).¹ Här skall jag försöka att besvara frågan huruvida Descartes teser om Guds natur är förenliga med att Gud kunde ha skapat allt på ett sådant sätt att Gud själv inte hade existerat.² Vid första anblicken ger Descartes olika teser om Guds natur stöd åt olika svar på denna fråga, varför Descartes teser om Guds natur skulle vara oförenliga med varandra. Betrakta två av Descartes andra centrala teser om Guds natur: tesen att Gud har gränslös makt och tesen att Gud existerar med nödvändighet (CSMK, 235, 363; CSMII, 31, 83). Tesen att Gud har gränslös makt tycks per definition medföra att Gud kunde ha skapat allt på vilket sätt som helst, inklusive på ett sådant sätt att han själv inte hade existerat. Tesen att Gud existerar med nödvändighet tycks dock blockera just denna slutsats. Om Gud existerar med nödvändighet kunde han inte ha skapat allt på ett sådant sätt att han inte hade existerat.

2. BREVEN TILL MERSENNE OCH MESLAND

Descartes är inte den förste som har försvarat tesen att Gud har gränslös makt, men han skiljer sig från sina föregångare i ett viktigt avseende. Där till exempel Thomas av Aquino håller för sant att Gud har gränslös makt

¹I texten citeras band I–II av *The Philosophical Writings of Descartes* via bandnummer som CSM. Band III av samma verk citeras som CSMK. Descartes håller för sant att det hör till Guds väsen att alltid existera (CSMII, 79). – Det finns alltså ingen tidpunkt då Gud har skapats, och börjat existera (jfr Williams 1978, 157–58). Ingen substans, inklusive Gud, har alltså skapat Gud eftersom Gud inte är skapad. Descartes menar dock att Gud mäktar existera utifrån sig själv, eftersom Gud har gränslös makt. Härigenom är Gud, i en mening, sin egen orsak (CSMII, 79–80).

²Frågan jag behandlar här uppstod på ett fredagsseminarium i filosofi vid Umeå universitet. Jag vill tacka Emma Beckman, Sten Lindström, Bertil Strömberg och Inge-Bert Täljedal för deras synpunkter på en tidigare version av denna uppsats.

inom ramarna för det logiskt möjliga så menar Descartes att Guds makt inte ens är begränsad till det logiskt möjliga (Thomas av Aquino, bok I, art. 84; CSMK, 235).

Hos Descartes är positionen att Guds makt inte är begränsad till det logiskt möjliga i huvudsak en konsekvens av tesen att Gud har skapat allt som existerar med en fri viljeakt. Låt oss kalla de påståenden som Gud har gjort till nödvändiga sanningar för *aktuellt nödvändiga sanningar*. Några av Descartes exempel på aktuellt nödvändiga sanningar är ”Triangelns vinkelsumma är lika med summan av två räta vinklar”, ”Två logiskt oförenliga påståenden kan inte samtidigt vara sanna” och den sanning som Descartes menar att alla andra sanningar har som sanningsgrund, ”Gud existerar” (CSMK, 235, 24–25).³ Eftersom Gud har skapat allt med en fri viljeakt, menar Descartes, så implementerade han inte de aktuellt nödvändiga sanningarna med nödvändighet. För att Gud skall ha varit *betingad* att implementera de aktuellt nödvändiga sanningarna, så måste de relevanta betingelserna ha existerat oberoende av Guds viljeakt, vilket är oförenligt med Descartes tes att Gud har skapat allt som existerar med en enda viljeakt. Alltså var Gud inte betingad att skapa allt som existerar enligt Descartes, utan kunde ha skapat allt på ett annorlunda sätt. Gud kunde till exempel ha gjort det falskt att alla cirkelns radier är av samma längd, han kunde ha låtit bli att skapa en värld, och han kunde ha gjort det falskt att logiskt oförenliga påståenden inte kan vara sanna samtidigt (CSMK, 25, 235).

Breven till Mersenne och Mesland är problematiska eftersom Descartes både håller för sant att de aktuellt nödvändiga sanningarna är nödvändigt sanna och att Gud kunde ha gjort dessa sanningar falska. Om Gud kunde ha gjort de aktuellt nödvändiga sanningarna falska så förefaller de aktuellt nödvändigt sanningarna att vara kontingent sanna, och Descartes framstår som motsägelsefull. Vi måste behandla detta problem eftersom ”Gud existerar” är en aktuellt nödvändig sanning. I enlighet med det ovanstående kunde Gud ha gjort även *denna* aktuellt nödvändiga sanning falsk.

Anledningen till att Descartes påståenden i breven till Mersenne och Mesland framstår som motsägelsefulla är att vi inte har tagit hänsyn till att Descartes behandlar två olika aspekter av Guds makt i sina teser om Guds natur. Där Descartes hävdar att Gud kunde ha gjort de aktuellt nödvändiga sanningarna falska betraktar han den makt som Gud har i kraft av sin fria vilja *per se*. Det är i kraft av denna fria vilja som Gud har implementerat de aktuellt nödvändiga sanningarna *som* aktu-

³Descartes hävdar att ”Gud existerar” är den mest *eviga* av alla möjliga sanningar, men att denna sanning också är en aktuellt nödvändig sanning följer av att Descartes håller för sant att Gud existerar med nödvändighet (CSMI, 197, 306).

ellt nödvändiga sanningar. Oberoende av den fria viljeakt med vilken Gud har skapat allt som existerar, vilket inkluderar implementeringen av de aktuellt nödvändiga sanningarna, har de aktuellt nödvändiga sanningarna inte *någon* modal status (jfr Alanen 1988, 189). Där Descartes å andra sidan hävdar att de aktuellt nödvändiga sanningarna är nödvändigt sanna betraktar han de aktuellt nödvändiga sanningarna som implementerade.

3. KONVERSATIONEN MED BURMAN

Så långt är Descartes modala metafysik överskådlig. Konversationen med Frans Burman 1648 leder dock till oklarheter. Vid den punkt där oklarheterna uppstår behandlar konversationen Descartes tes att det endast existerar en akt hos Gud, en akt med vilken Gud samtidigt får kunskap om allt, implementerar sin vilja och skapar allt (CSMI, 201; CSMK, 347). Burman hävdar att det är tänkbart att Guds beslut att skapa allt som existerar hade kunnat vara annorlunda, och att Gud därför också kunde ha skapat allt på ett annorlunda sätt. Burmans påstående tycks vara helt analogt med Descartes egna påståenden i breven till Mersenne och Mesland, men Descartes invänder ändå att det ur en metafysisk synvinkel är omöjligt att tänka sig att Guds viljeakt att skapa allt som existerar kunde ha varit annorlunda. I realiteten kunde Guds viljeakt inte ha varit annorlunda menar Descartes, eftersom Guds viljeakt inte är skild från Gud själv. Guds existens är inte primär i förhållande till Guds faktiska viljeakt, och han kunde inte ha existerat utan denna viljeakt (CSMK, 347–48).

I breven till Mersenne och Mesland skiljer Descartes tydligt på vad som är möjligt relativt den makt Gud har i kraft av sin fria vilja *per se* och vad som är möjligt relativt Guds faktiska viljeakt. Denna skillnad framträder inte lika klart i konversationen med Burman. Den bästa tolkningen av konversationen med Burman är att Descartes försöker driva Burman till att avstå från hypotetiska funderingar kring vad Gud kunde ha gjort i kraft av sin fria vilja *per se*, och *uteslutande* betrakta Guds faktiska vilja *som* implementerad. Denna tolkning motiveras i synnerhet av Descartes varnande anmärkning att det ligger bortom vår förmåga att besvara metafysiska frågor om Guds natur och handlingar, en anmärkning som påminner om argumentationen i brevet till Mesland. Fastän Gud till exempel kunde ha gjort det falskt att logiskt oförenliga påståenden inte kan vara sanna samtidigt skall vi inte försöka förstå hur detta kan vara sant, menar Descartes i brevet till Mesland. Att försöka begrippliggöra detta är enligt Descartes ett fruktlöst projekt, eftersom Gud har skapat vårt tänkande på ett sådant sätt att vi bara kan förstå sådant som

är möjligt relativt de *aktuella* modala sanningarna. Med andra ord, vi kan bara förstå sådant som är möjligt relativt Guds *faktiska* viljeakt (jfr CSMK, 235).

4. VARFÖR GUD INTE KUNDE HA SKAPAT ALLT PÅ ETT SÅDANT SÄTT ATT HAN SJÄLV INTE HADE EXISTERAT

I enlighet med Descartes modala metafysik gäller alltså följande med avseende på Guds egen existens. Att Gud existerar är en central aktuellt nödvändig sanning, men samtidigt är en av Descartes centrala teser om Guds natur att Guds makt inte kan ha några gränser. Om aktuellt nödvändiga sanningar om Gud i varje avseende delar den modala status som andra aktuellt nödvändiga sanningar har, så leder vår diskussion till slutsatsen att Gud också kunde ha gjort aktuellt nödvändiga sanningar om sig själv falska. På samma sätt och i samma mening som Gud till exempel kunde ha gjort det falskt att logiskt oförenliga påståenden inte kan vara sanna samtidigt kunde Gud i sådant fall ha skapat allt på ett sådant sätt att han inte hade existerat, och den aktuellt nödvändiga sanningen ”Gud existerar” hade i sådant fall varit falsk.

Även om den rimligaste tolkningen av Descartes *bör* vara att ”Gud existerar” och andra aktuellt nödvändiga sanningar om Gud har en speciell status så skulle det vara godtyckligt att utan ytterligare stöd för detta göra antagandet att aktuellt nödvändiga sanningar om Gud är uteslutna från de aktuellt nödvändiga sanningar som Gud kunde ha gjort falska.⁴ Därför skall vi inte göra något sådant antagande. För att kunna besvara frågan huruvida Gud kunde ha skapat allt på ett sådant sätt att han inte hade existerat skall jag i stället utgå ifrån Descartes tes att Gud saknar förmågan att beröva sig själv sin egen existens.

Först och främst, inga ting kan existera och inga sanningar kan gälla med mindre än att Gud vill att detta skall vara fallet. Denna premis stöds av Descartes tes att det inte kan finnas något subsisterande ting, någon lag eller någon sanning som inte är beroende av Gud (CSMII, 293–94), och av hans tes att Gud är den ”universella” eller ”totala” orsaken till allt och att inget kan vara fallet med mindre än att Gud vill det (CSMK, 272, 358). Följaktligen kan ingen annan substans vilja (att ting skall existera eller att sanningar skall gälla) infrias med mindre än att denna substans vilja sammanfaller med Guds vilja. För det andra, Gud saknar förmågan att beröva sig själv sin egen existens. (Det vore en ofullkomlighet hos Gud om han var kapabel att beröva sig själv sin egen existens.) Denna premis är hämtad från ett brev från Descartes

⁴Jämför Wilson (1978, 124). Se även Curley (1984, sektion VII).

till en okänd adressat (se CSMK, 212, 212n). Det följer härmed att ingen substans kunde ha berövat Gud hans existens, och därmed att Gud inte kunde ha skapat allt på ett sådant sätt att han inte hade existerat.

Om detta argument är rimligt så har det en viktig konsekvens. Om Gud inte kunde ha skapat allt på ett sådant sätt att han inte hade existerat, så är Descartes teser om Guds natur inbördes oförenliga. Guds makt kan inte vara gränslös samtidigt som han inte kunde ha skapat allt på ett sådant sätt att han inte hade existerat. Vi måste därför antingen förkasta det argument jag presenterat eller den tolkning av Descartes enligt vilken han förbehållslöst håller för sant att Guds makt är gränslös.

LITTERATUR

- Alanen, Lilli. 1988. Descartes, Omnipotence and Kinds of Modality. I *Doing Philosophy Historically*, red. Peter H. Hare, s. 182–96. Buffalo: Prometheus Books.
- Cottingham, John, Robert Stoothoff, och Dugald Murdoch, red. 1984 och 1985. *The Philosophical Writings of Descartes*, band I–II. Cambridge: Cambridge University Press.
- Cottingham, John, Robert Stoothoff, Dugald Murdoch och Anthony Kenny, red. 1990. *The Philosophical Writings of Descartes*, band III. Cambridge: Cambridge University Press.
- Curley, Edwin M. 1984. Descartes on the Creation of the Eternal Truths. *The Philosophical Review* 93, nr 4 (Oktober), s. 569–97.
- Thomas av Aquino. *Summa Contra Gentiles*.
- Williams, Bernard. 1978. *Descartes: The Project of Pure Enquiry*. Hassocks: The Harvester Press.
- Wilson, Margaret Dauler. 1978. *Descartes*. London: Routledge & Kegan Paul.

Distinktionen mellan analytiska och syntetiska sanningar spelade en viktig roll för Kant och ungefär samma distinktion förekom också tidigare, t.ex. hos Leibniz och Hume. Den har senare ifrågasatts och diskuterats av många filosofer, särskilt sedan W. V. Quine publicerade sin klassiska artikel ”Two Dogmas of Empiricism”.¹ Men någon vidare enighet har inte uppnåtts. Somliga har blivit övertygade av Quines argument – eller åtminstone av vissa av hans argument – men andra tycks mena att de är dåliga och att Quines skepsis mot idén om analytiska satser i större eller mindre grad är obefogad.

Analytiska satser antas ofta vara detsamma som ”begreppsliga sanningar”, dvs. språkliga satser som på något sätt är, eller kan inses vara, sanna i kraft enbart av sin betydelse (inklusive betydelsen av de ord de innehåller).² Satser som är falska i kraft av sin betydelse kan eventuellt också kallas analytiska, men i så fall hellre ”analytiskt falska”; vanligen betyder dock ”analytisk” detsamma som ”analytiskt sann”. Sanna och falska satser som varken är analytiskt sanna eller analytiskt falska kallas ”syntetiska”.

Teorier, hypoteser och observationsatser inom empiriska vetenskaper antas i allmänhet vara syntetiska, även om vissa grundläggande principer eventuellt kan vara analytiska. Men många anser att de språkliga satser som förekommer inom matematik, logik och filosofi (ofta) är analytiska. Det kan man dock inte vara säker på, bland annat därför att det fortfarande råder en hel del oklarhet och oenighet om vad en analytisk sats är för något.

Språkvetenskap eller lingvistik brukar antagligen uppfattas som en

¹Quines artikel framfördes som föredrag i december 1950 – vid ett möte med American Philosophical Association i Toronto – och publicerades först 1951. Detta är kanske den mest uppmärksammade artikeln i analytisk filosofi efter andra världskriget. Här håller jag mig till en något reviderad version som kom 1953. Ytterligare en version kom 1961.

²Quine opererar med flera olika betydelser av ”analytisk”, men i inledningen till ”Two Dogmas” säger han att analytiska sanningar är de som är ”grounded in meanings independently of matters of fact” (1953a, s. 20).

empirisk vetenskap, som handlar om faktiskt förekommande språk och språkligt beteende, och de satser och teorier som förekommer i språkvetenskap är därför antagligen till stor del syntetiska. Men hur är det med satser av typen ”Satsen s är analytisk”, där s är en faktiskt förekommande sats? Är sådana satser analytiska eller syntetiska?

Det mest närliggande är väl att säga att de är syntetiska. Att en språklig sats s har en viss betydelse eller mening är väl något som är sant eller falskt beroende på hur språket faktiskt används. Det är alltså sant eller falskt beroende på hur verkligheten är beskaffad. Huruvida satsen s är analytisk är då också beroende av hur verkligheten är beskaffad. Satsen ”Satsen s är analytisk” kan alltså inte vara sann ”enbart i kraft av sin betydelse”.

Men å andra sidan kunde man kanske säga att om s är analytisk, så är den ju sann enbart i kraft av sin betydelse – och med tanke på vad ”analytisk” betyder så måste väl i så fall satsen ”Satsen s är analytisk” som helhet också vara sann enbart i kraft av sin betydelse, och alltså analytisk. Ett vanligt men trivialt exempel på en analytisk sats är ”Ungkarlar är ogifta”.³ Om man vet vad den betyder och dessutom vet vad ”Satsen ... är analytisk” betyder, så kan man väl direkt inse att även satsen ”Satsen ’Ungkarlar är ogifta’ är analytisk” är analytisk?

Hur *avgör* man om en sats i t.ex. svenska språket är analytisk? Är det en fråga för språkvetare eller filosofer? Jag antar att det är filosofer, snarare än språkvetare, som mest intresserat sig för analytiska satser, men kanske är de mindre kompetenta än språkvetare att avgöra vilka satser som är analytiska. Eller också kanske det beror på vilka satser det är fråga om. Språkvetare är kanske bättre på att klassificera satser som ”Alla ungar är ogifta” och ”All konst är skapad av människor”, medan filosofer är bättre på att avgöra om satser som ” $5 + 7 = 12$ ” eller ”Ingenting kan uppstå ur intet” är analytiska. Men hur är det med ”Alla kuber har tolv kanter” och ”Alla människor är dödliga”?⁴

Det är fortfarande mycket som är oklart med analyticitet. Låt oss se lite närmare på vad Quine har att säga om saken.

1. QUINE 1946

Quines mest kända uppsats om analyticitet presenterades alltså 1950. Den föregicks av många och långa diskussioner med filosofer som Rudolf Carnap, Alonzo Church, Nelson Goodman, Morton White och Alfred

³ Detta är en sats som även Quine är beredd att klassificera som analytisk; se t.ex. Quine 1991 (2008), s. 395–97.

⁴ Om det skulle finnas någon varelse som liknar människor i alla avseenden, utom i det att han eller hon inte är dödlig, så skulle vi kanske anse att det trots allt inte är en människa – om ordet ”människa” har sin vanliga mening.

Tarski.⁵ I december 1946 höll Quine också ett föredrag om analyticitet vid universitetet i Pennsylvania. Texten till detta föredrag publicerades aldrig under hans livstid, men den finns nu utgiven i en postum samlingsvolym.⁶

Här utgår Quine från att en analytisk sats är en sann sats, vars sanning följer av ordens betydelser. Detta tolkar han i sin tur som att en sats är analytisk om den är resultatet av att ersätta ett uttryck, i en logiskt sann sats, med ett uttryck som har samma betydelse (samma mening).⁷ Här finns då två problem. För det första, vad är *logisk sanning*? För det andra, vad är *samma betydelse* (dvs. synonymi)?

När det gäller logisk sanning finns huvudsakligen två möjligheter. Antingen innefattar det enbart satslogik och predikatlogik, eller också dessutom mängdlära. Quine själv vill av vissa skäl låta logiken omfatta endast sats- och predikatlogik. Men han noterar att detta kan innebära att matematiken inte blir analytisk, vilket många anser eller hoppas att den är.⁸

Som en första approximation tänker sig sedan Quine att *synonymi* kan förklaras på följande sätt: ”två uttryck har samma betydelse för [personen] x vid [tidpunkten] t om substitution av det ena för det andra, i varje sats som godtas av x vid t , ger en sats som godtas av x vid t ”.⁹

Här kan man notera att Quine tänker sig att betydelsen hos ett uttryck kan variera från person till person och från tidpunkt till tidpunkt. Det verkar rimligt – folk tolkar ibland satser olika – men då kan en sats också vara analytisk för en person men inte för en annan, och vid en tidpunkt men inte vid en annan. Även detta verkar rimligt, men det brukar näs-

⁵Se Quine 1991, s. 391–93.

⁶Quine 1946 (2008).

⁷Quine formulerar det så här: ”A statement is analytic if it can be got by putting one expression for another with the same meaning in a (logical principle, or) statement which is logically true” (1946, s. 25). Samma mening, synonymi, avser här och i fortsättningen *kognitiv* synonymi; stilistiska skillnader spelar ingen roll. Se även Quine 1981, s. 48–51.

⁸De logiska positiviserna ansåg t.ex. att matematiska satser är analytiska. Quines skäl för att låta logiken omfatta endast sats- och predikatlogik behöver vi inte här gå in på.

⁹Min översättning av ”two expressions have the same meaning for x at t if substitution of the one for the other, in any statement believed by x at t , yields a statement believed by x at t ” (Quine 1946, s. 34). – Man kan notera att detta förslag endast anger ett *tillräckligt* villkor för synonymi (för x vid t). Men det är kanske rimligt, ty två uttryck kan eventuellt vara synonyma utan att uppfylla detta villkor. Ordet ”bror” är t.ex. kanske synonymt med ”manligt syskon”, men om vi byter ut ”bror” mot ”manligt syskon” i ”Han är min bror” får vi en sats som inte ens är grammatiskt välformad, nämligen ”Han är min manligt syskon”. Se även nedan.

tan aldrig uppmärksammas. Och kanske vet vi alldeles för lite om vad betydelse och synonymi är för att vi ska kunna ha någon klar uppfattning om detta.

Quine har flera invändningar mot sitt eget approximativa förslag.

(1) Han menar bland annat att det att en person ”godtar” (eng. ”believes”) en sats vid en tidpunkt ska förstås som en *disposition* att bete sig på ett visst sätt – men han anser att Carnaps problem med att analysera dispositioner ännu inte är löst. Denna invändning är nog inte så drabbande; numera har man kommit längre med analysen av dispositioner och Quine själv har i senare skrifter varit fullt villig att hänvisa till sådana.¹⁰

(2) Ett annat problem är homonymer. Ordet ”sund” kan t.ex. ibland betyda detsamma som ”ett relativt smalt vattenområde mellan två landområden” och ibland detsamma som ”frisk”. Vi kan inte utgå från att detta är två olika ord, menar Quine, ty de skiljer sig bara i betydelse och betydelselighet är ju just det vi ska definiera. Ett och samma ord blir därför synonymt med två olika uttryck, som inte sinsemellan är synonyma. Quine anser att detta är en mycket allvarlig komplikation.¹¹

Men är det så allvarligt? Den föreslagna definitionen av synonymi tycks kunna fungera i många fall och den ger ju faktiskt bara ett *tillräckligt* villkor för synonymi. Att det finns andra fall av synonymi, t.ex. de som inbegriper homonymer, gör att definitionen inte är heltäckande, men det är kanske inte så viktigt i detta sammanhang. För övrigt antyder Quine själv i en senare uppsats att problemet med homonymi och flertydighet kan bemästras om man fokuserar på synonymi mellan längre uttryck, dvs. hela satser eller sekvenser av satser.¹²

(3) Ytterligare en invändning, som framförs av Quine, är att två uttryck kan *visas* ha samma betydelse trots att detta inte är *uppenbart* – och då kanske de inte uppfyller kriteriet på synonymi. Hans exempel är (en variant av) ”4” och ”262144”. Men har dessa två uttryck verkligen samma betydelse? De betecknar samma tal, men att två uttryck betecknar detsamma medför inte att de också har samma betydelse eller samma mening; detta är något som Quine själv brukar betona.¹³ Exempelvis menar han att ”9” och ”antalet planeter” betecknar samma tal, men har olika mening.¹⁴ Själv är jag beredd att gå med på detta – även om jag inte

¹⁰Se t.ex. Quine 1995, s. 21 och 76.

¹¹Quine 1946, s. 34. Den nämns, i förbigående, även i Quine 1953a, s. 32.

¹²Se Quine 1953b, s. 58.

¹³T.ex. i Quine 1953a, s. 21. Där säger han: ”Confusion of meaning with extension, in the case of general terms, is less common than confusion of meaning with naming in the case of singular terms”.

¹⁴Numera betecknar de inte ens samma tal, eftersom man nu har infört en definition av ”planet” som medför att det endast finns *åtta* planeter (kring solen).

riktigt vet vad mening/betydelse är för något – men varför kan man då inte också säga att uttrycken "4"⁹ och "262144"⁹ har olika mening?

Någon skulle kanske vilja säga att det är *nödvändigt* att "4"⁹ och "262144"⁹ betecknar samma tal, medan det *inte* är nödvändigt att "9"⁹ och "antalet planeter"⁹ betecknar samma tal. Men det stämmer knappast. Det är väl inte nödvändigt att "4"⁹ och "262144"⁹ betecknar samma tal; det hänger ju på vilka tecken vi råkar använda för att beteckna tal. Däremot är det förstås nödvändigt att det tal vi nu faktiskt betecknar med uttrycken "4"⁹ och "262144"⁹ är identiskt med sig självt. Men varför skulle detta medföra att de två uttrycken har samma mening?

Kort sagt: jag tycker inte att Quine 1946 egentligen har anfört några riktigt avgörande invändningar mot sitt eget approximativa förslag till definition av synonymi. Kanske var det därför han själv aldrig publicerade detta föredrag.

Men i föredraget har han också ett annat förslag till vad analyticitet är, ett förslag som inte bygger på någon synonymirelation. Om man inte kan definiera analyticitet i termer av synonymi, så kan man kanske göra tvärtom: definiera synonymi i termer av analyticitet.¹⁵ Men vad är i så fall kriteriet på analyticitet? Quine föreslår här att det har att göra med graden av vår motvilja mot att förkasta en sats som falsk, när vi ställs inför motstridande evidens. Och han menar att detta antyder att analyticitet – och därmed synonymi – blir en fråga om grader.¹⁶

Att analyticitet har att göra med vår ovillighet att förkasta en sats kan vara rimligt. När våra uppfattningar inte hänger ihop eller är motsäggelsefulla, så att något måste revideras, har vi normalt flera alternativ att välja mellan.¹⁷ Men vilka är egentligen "vi" här? Kanske tänker sig Quine ett "socialt" kriterium, så att medlemmarna av en språkgemenskap – eller kanske en elit inom den – på något sätt ska komma överens om om *hur* analytiska olika satser ska vara. Alternativt kunde man tänka sig att graden av ovilja kan variera från person till person, och från tidpunkt till tidpunkt, så att analyticitet i denna mening bör relativiseras till person och tidpunkt (analogt med Quines approximativa definition av synonymi). En sats kunde eventuellt sedan sägas vara analytisk, utan relativisering, om den är analytisk för alla eller de flesta personer inom en språkgemenskap vid alla eller de flesta tidpunkter.

Quines förslag behöver, såvitt jag kan se, inte heller leda till att analyti-

¹⁵Se Quine 1946, s. 35: "two expressions are the same in meaning if, if you put the one for another in a statement *s* to form a statement *s'*, the conditional 'If *s*, then *s'*' is analytic."

¹⁶Jfr Quine 1946, s. 35. Att synonymi är något som bör antas förekomma i olika grader säger Quine också i Quine 1953b, s. 63.

¹⁷Se även Quine 1953a, s. 42–43.

citet blir en fråga om grader. Alternativt kunde man väl säga att analytiska satser är de som vi i *maximal grad* – eller i en grad över en viss nivå, som vi då måste komma överens om – är ovilliga att förkasta eller revidera.

En fördel med Quines förslag att gå direkt på analyticitet, utan att först definiera synonymi, är för övrigt att vissa satser kan tyckas vara sanna i kraft av sin betydelse, trots att de inte kan omvandlas till logiska sanningar genom att uttryck ersätts av synonymer. Ett exempel kunde vara ”Om Anna är äldre än Bo och Bo är äldre än Cecilia, så är Anna äldre än Cecilia”. Men Quine lägger ingen större vikt vid detta förslag, han nämner det mest i förbigående.

2. QUINE 1950

Quines argument 1946 mot analyticitet – eller mot distinktionen mellan analytiska och syntetiska satser – var såvitt jag kan se inte särskilt övertygande. Blev de bättre i den berömda ”Two Dogmas of Empiricism”?¹⁸

Här diskuterar han en definition av synonymi som i vissa avseenden liknar det approximativa kriteriet från 1946, nämligen att ”synonymi mellan två språkliga uttryck består i att de är utbytbara i alla kontexter utan att sanningsvärdet förändras; utbytbarhet, med Leibniz uttryck, *salva veritate*”.¹⁹

I ett extensionellt språk – ett predikatlogiskt språk med individvariabler, predikat, kvantifikation och sanningsfunktioner – räcker inte detta för synonymi; i ett extensionellt språk är nämligen uttryck utbytbara *salva veritate* om de har samma extension. De behöver inte dessutom ha samma intension (mening). I ett intensionellt språk som även innehåller modala begrepp som ”nödvändigt” räcker däremot kriteriet för synonymi. Men begreppet nödvändighet är i lika hög grad i behov av förklaring som begreppet analyticitet.²⁰ Begreppen analyticitet, synonymi, mening och nödvändighet hänger intimt ihop; kan vi förklara analyticitet eller nödvändighet, så kan vi också förklara betydelse/mening och synonymi.

¹⁸Timothy Williamson anser för sin del att ”Quine’s arguments are generally found much less compelling than they once appeared” (Williamson 2007, s. 50). Men det är tveksamt om Williamson verkligen har förstått Quines argument.

¹⁹Quine 1953a, s. 27; min översättning av: ”the synonymy of two linguistic forms consists simply in their interchangeability in all contexts without change of truth value; interchangeability, in Leibniz’s phrase, *salva veritate*”.

²⁰Quine 1953a, s. 31: ”if a language contains an intensional adverb ‘necessarily’ in the sense lately noted, or other particles to the same effect, then interchangeability *salva veritate* in such a language does afford a sufficient condition of cognitive synonymy; but such a language is intelligible only if the notion of analyticity is already clearly understood in advance”.

Men är detta ett bra argument mot analyticitet? Varför kan man t.ex. inte helt enkelt acceptera begreppet nödvändighet? Vi kan då säga att satsen "Alla ungarlar är ogifta" är analytisk om det är nödvändigt att alla ungarlar är ogifta. Och mer generellt: en sats s är analytisk om, och endast om, det är nödvändigt att s är sann.

Alternativt kunde man ta fasta på ett annat förslag i "Two Dogmas", nämligen att "en sats helt enkelt är analytisk när den är synonym med en logiskt sann sats".²¹ Varför räcker inte det? Varför är begreppet synonymi så problematiskt?

Det är inte helt lätt att se, i "Two Dogmas", varför Quine inte accepterar intensionella begrepp som mening, synonymi, nödvändighet och analyticitet. 1991 säger han: "Jag tror inte att jag någonsin verkligen har förstått något som jag inte kan uttrycka i extensionellt språk."²² Men varför det?

Man kan hålla med om att det ibland är svårt att veta vad man talar om i ett intensionellt språk. Men det kan också vara svårt att veta vad som sägs i ett extensionellt språk; där hänger det ju på om man förstår de predikat som används. Och det tvåställiga predikatet "är synonym med" och det enställiga predikatet "är analytisk" kan förekomma i ett språk som är extensionellt.

I själva verket kan man nog säga att det avgörande för Quine är att sådana predikat som "mening", "synonymi" och "analytisk" inte har någon klar *empirisk* innebörd; det finns inget objektiva kriterium i termer av intersubjektivt observerbara förhållanden, som kan användas för att avgöra om uttryck är synonyma eller om satser är analytiska respektive syntetiska. På ett ställe säger han: "En ordboksförfattare är en empirisk vetenskapsman, vars uppgift är att redogöra för redan föreliggande fakta; och om han förklarar 'ungkarl' som 'ogift man' så är det för att han tror att det råder en relation av synonymi mellan dessa uttryck, implicit i allmänt eller föredraget språkbruk som existerar redan innan hans eget arbete. Begreppet synonymi som förutsätts här kräver fortfarande en förklaring, rimligen i termer som refererar till språkligt beteende."²³ Och lite senare påpekar han att "någon skiljelinje mellan analytiska och syn-

²¹Min översättning av: "a statement may be described as analytic simply when it is synonymous with a logically true statement" (Quine 1953a, s. 38).

²²Min översättning av: "I doubt that I have ever fully understood anything that I could not explain in extensional language" (s. 500).

²³Min översättning av: "The lexicographer is an empirical scientist, whose business is the recording of antecedent facts; and if he glosses 'bachelor' as 'unmarried man' it is because of his belief that there is a relation of synonymy between these forms, implicit in general or preferred usage prior to his own work. The notion of synonymy presupposed here has still to be clarified, presumably in terms relating to linguistic behavior" (Quine 1953a, s. 24).

retiska satser har helt enkelt inte dragits. Att det finns en sådan distinktion är en oempiristisk dogm bland empirister, en metafysisk trossats".²⁴

I boken *Word and Object* 1960 betonas samma sak. Quine säger där att flera av hans kritiker har tyckt att han har alldeles för höga krav på klarhet när det gäller att karakterisera analyticitet och synonymi, men han menar att allt han kräver är en grov karakteristik i termer av dispositioner till verbalt beteende.²⁵ Boken inleds också med några meningar – de tre första i förordet – som ger en rätt övertygande bakgrund till detta krav: "Språk är en social konst. När vi lär oss det är vi tvungna att förlita oss uteslutande på intersubjektivt tillgängliga tecken på vad man ska säga och när. Därför finns det ingen grund för att jämföra språkliga betydelser, utom i termer av människors dispositioner att öppet reagera på socialt observerbara stimuleringar."²⁶

Alla fakta om språkliga uttrycks betydelser och synonymiförhållanden måste alltså vara förankrade i människors språkliga beteenden i sociala sammanhang. Någon "mening" utöver detta finns inte och kan inte finnas, eftersom språklig mening är något som skapas genom språkanvändning.

3. EMPIRISKA KRITERIER

Men är det då verkligen omöjligt att formulera empiriska kriterier för synonymi och analyticitet? Har Quine övertygande visat det?

Här måste man först skilja mellan olika slags satser. Quine skiljer mellan *tillfällessatser* ("occasion sentences") och *stående satser* ("standing sentences"). De förra har sanningsvärden som kan variera från ett tillfälle till ett annat (t.ex. "Det regnar" och "Han är glad idag"), de senare har samma sanningsvärde under längre tid (t.ex. "Quine var glad en eftermiddag på hösten 1993", " $5 + 7 = 12$ " och "Alla människor är dödliga").

Quine anser att man kan definiera synonymi för tillfällessatser på

²⁴Min översättning av: "a boundary between analytic and synthetic statement simply has not been drawn. That there is such a distinction to be drawn at all is an unempirical dogma of empiricists, a metaphysical article of faith" (Quine 1953a, s. 37).

²⁵"I ask no more, after all, than a rough characterization in terms of dispositions to verbal behavior" (Quine 1960, s. 207). Och i "Two Dogmas in Retrospect" säger Quine: "Repudiation of the first dogma, analyticity, is insistence on empirical criteria for semantic concepts" (Quine 1991, s. 397).

²⁶Min översättning av: "Language is a social art. In acquiring it we have to depend entirely on intersubjectively available cues as to what to say and when. Hence there is no justification for collating linguistic meanings, unless in terms of men's dispositions to respond overtly to socially observable stimulations" (Quine 1960, s. ix).

följande sätt: ”Att två tillfällessatser har samma betydelse för en given talare vid ett givet stadium i hans utveckling består i att han då har en disposition [...] att ge samma bedömning (bejaka, förneka, avstå) av de båda satserna vid varje tillfälle.”²⁷

För stående satser, och det Quine kallar *eviga* satser – dvs. sådana som alltid har samma sanningsvärde – blir det svårare. Men en bra bit kan man komma genom att ta fasta på uttryck som är synonyma (dvs. utbytbara) i tillfällessatser; sådana uttryck kan antas vara synonyma även i stående och eviga satser.

Men såvitt jag förstår kan man på detta sätt endast hitta kriterier på synonymi och analyticitet *relativt till personer och tidpunkter*. Det ger ingen grund för att tala om synonymi och analyticitet i absolut mening, utom i de fall där det (nästan) inte blir några skillnader mellan olika personer och tidpunkter. Och sådana fall är nog förhållandevis sällsynta. Dock är det klassiska exemplet ”Alla ungarlar är ogifta” kanske analytisk för nästan alla, nästan alltid.²⁸

Detsamma gäller, som jag redan har sagt, om vi tar fasta på Quines förslag från 1946 att analytiska satser är de som man i maximal grad är ovillig att förkasta eller revidera. Även sådan analyticitet kan rimligen variera från person till person och från tidpunkt till tidpunkt. Men däremot kan den väl, åtminstone i princip, klara Quines krav på empirisk avgörbarhet. Frågan blir ju vilka dispositioner att revidera vissa andra dispositioner en person har vid en given tidpunkt.

Quines har ett annat förslag från 1974, som han också tar upp 1991; en sats s är analytisk för en person x (för vilken s ingår i modersmålet) om x fick lära sig att s är sann som ett led i att x lärde sig vissa ord som ingår i satsen.²⁹ Även här är det alltså fråga om ett relativiserat begrepp, även om det kanske finns vissa satser som är analytiska i denna mening för de allra flesta. I alla händelser verkar empirikravet i princip kunna uppfyllas även här.

Men även om vi kan hitta stående satser som är analytiska eller synonyma i *absolut* mening, dvs. för varje person vid varje tidpunkt, så är det inte säkert att de *intersubjektivt* har ”samma” betydelse för alla personer vid någon tidpunkt. Och det är svårt att se att man skulle kunna ange något empiriskt kriterium på sådan intersubjektivtydelselikheter. Quine för sin del tror förstås inte på något sådant – utom i fråga om observationssatser, där han anser att intersubjektiv synonymi inte är något problem. Mer ge-

²⁷Quine 1995, s. 76, min översättning. Se även Quine 1981, s. 49–51.

²⁸Men själv har jag ifrågasatt att det skulle råda synonymi mellan ”ungkarl” och ”ogift man” och att ”Alla ungarlar är ogifta” är analytisk (se *Filosofisk tidskrift* 2005 nr 1 och 2005 nr 4).

²⁹Jfr Quine 1991, s. 395–96 och 1993, s. 64.

nerellt kommer man här in på hans tes om översättningens obestämdhet.³⁰

Somliga filosofer anser att ord och satser har en absolut betydelse i ett språk och att betydelser alltså inte är något som varierar från en person till en annan.³¹ Ordboksförfattare beskriver ju också ords betydelser, utan att relativisera till person och tid, och de får väl antas ha empirisk evidens för sina beskrivningar. Man kunde tycka att hypoteser om absolut synonymi då måste kunna baseras på sådan evidens. Men påståenden om synonymi förekommer inte särskilt ofta i ordböcker. Snarare beskrivs olika användningar under olika typiska omständigheter, utan att man upprätthåller någon klar skillnad mellan information om ord och information om det orden refererar till.³²

Sanningen är nog att det är filosofer, snarare än språkvetare, som har intresserat sig för synonymi. Och den synonymi det då är fråga om är inte någon *fullständig* betydelselikheter i alla avseenden, så att även t.ex. emotiva och stilistiska nyanser är precis desamma. Filosofer intresserar sig nästan uteslutande för det som brukar kallas *kognitiv* synonymi, dvs. betydelselikheter i de avseenden som spelar roll för satsers sanningsvärden.³³

Den norske filosofen Arne Næss har föreslagit ett kriterium för kognitiv likvärdhet (synonymi) mellan satser som säger att två satser är kognitivt likvärda för en person i en situation om personen *inte kan tänka sig* möjligheten av att godta det som den ena satsen uttrycker för honom i situationen och samtidigt förkasta det som den andra satsen uttrycker för honom i situationen.³⁴ Analogt kunde man föreslå att en sats *s* är analytisk för *x* vid

³⁰Jfr t.ex. Quine 1995, s. 81–82.

³¹Timothy Williamson säger t.ex. att "each individual uses words as words of a public language; their meanings are constitutively determined not individually but socially, through the spectrum of linguistic activity across the community as a whole. The social determination of meaning requires nothing like an exact match in use between different individuals; it requires only enough connection in use between them to form a social practice" (s. 98).

³²Jfr Quine 1995, s. 83. Quine säger också följande: "Though the word 'meaning' is ubiquitous in lexicography, no capital is made of a relation of sameness of meaning" (ibid.).

³³Jfr Quine: "we are not concerned here with synonymy in the sense of complete identity in psychological associations or poetic quality; indeed no two expressions are synonymous in such a sense. We are concerned only with what may be called *cognitive synonymy*" (1953a, s. 28).

³⁴Jfr Næss 1981, s. 10–13. Næss betonar att detta endast gäller satser som "kan tänkas vara godtagbara under vissa sakförhållanden och icke godtagbara under andra sakförhållanden". Quine hänvisar i en fotnot till Næss och hans "experimentation on analyticity intuitions", se Quine 1960, s. 67. Næss bok har på svenska titeln "Empirisk semantik", vilket antyder att han åtminstone delvis är ute efter just det som Quine efterlyser.

t om, och endast om, x inte kan tänka sig möjligheten av att förkasta det som s uttrycker för x vid t .

Man kan ”undersöka empiriskt” om en sats är analytisk för en person vid en tidpunkt i denna mening; det är ju bara att fråga personen, eventuellt i samband med att man ger en utförlig instruktion om hur frågan ska uppfattas. Men personens uppfattnings- och föreställningsförmåga bör rimligen inte spela någon roll, det avgörande ska väl vara om det *finns* några möjliga omständigheter, en ”möjlig värld”, i vilken satsen, som den tolkas av x vid t , skulle vara falsk. Om det inte finns någon sådan värld, dvs. om satsen, som den tolkas av x vid t , är *nödvändigt* sann, så är den analytisk. Men härmed är vi tillbaka till Quines problem. Hur avgör man empiriskt om satsen, som den tolkas av x vid t , är nödvändigt sann? Att x själv på förfrågan *säger* att den är det – eller att den inte är det – kan inte gärna vara avgörande.

4. TILLBAKA TILL 1946

När det gällde tillfällighetssatser tänkte sig ju Quine att en person x vid t kan ha en disposition att bejaka satsen ”vid varje tillfälle”. Kunde man inte på liknande sätt säga att en sats är analytisk för x vid t om, och endast om, x vid t har en disposition att bejaka satsen under alla möjliga förhållanden? I själva verket är man väl då tillbaka till Quines förslag från 1946 att analytiska satser är de som ”vi” – eller, som här, en person vid en tidpunkt – i maximal grad är ovilliga att förkasta eller revidera.³⁵ Varför inte hålla fast vid detta? I princip kan man väl undersöka empiriskt om en sats är analytisk i denna mening; det är ju en fråga om att undersöka våra dispositioner till språkligt beteende.³⁶

Någon skulle kanske vilja invända att analytiska satser ju ska vara sanna i kraft av sin *betydelse* och satsers betydelse tycks inte direkt spela någon roll i Quines förslag 1946. Men det är ingen bra invändning. Satsers betydelse bestäms enligt Quine av våra dispositioner till språkligt beteende och det är ju också sådana dispositioner som bestämmer analyticitet enligt förslaget från 1946. De olika partiella förklaringar av analyticitet som Quine har föreslagit i olika skrifter passar också bra in som specialfall under förslaget från 1946. Jag menar alltså att lösningen på Quines problem fanns redan då.

³⁵Förslaget har också en viss likhet med Williamsons begrepp ”modal-analytisk”: en sats är analytisk i denna mening om den har en betydelse som är sådan att det är nödvändigt att varje sats med den betydelsen är sann i varje kontext (jfr Williamson 2007, s. 60–63).

³⁶Sådana dispositioner är enligt Quine fysikaliska egenskaper som vi normalt identifierar genom vissa symptom (se Quine 1995, s. 21).

5. ÄR ANALYTICITET INTRESSANT?

Varför har då inte Quine själv hållit fast vid sitt förslag från 1946? Eller: om han har hållit fast vid det, varför har han inte sagt det? Varför nämns det inte i ”Two dogmas” och senare skrifter? Kanske beror det på att analyticitet i denna mening – låt oss kalla det ”Q-analyticitet” – inte är av något större filosofiskt intresse.³⁷

Att en sats är Q-analytisk för oss vid en tidpunkt är fullt förenligt med att den inte är Q-analytisk för oss vid andra tidpunkter. Olika satser kan också vara Q-analytiska för olika personer. En sats kan vara Q-analytisk trots att den inte är sann och trots att vi inte har några goda grunder att tro att den är sann. Detta slags analyticitet tycks alltså varken vara kunskapsteoretiskt eller metafysiskt intressant.

Man kan också notera att en person mycket väl kan vilja förkasta en sats som är, eller hittills har varit, Q-analytisk för de allra flesta, kanske även för personen själv. Detta överensstämmer med Quines tes i ”Two Dogmas” att ”ingen sats är immun mot revision”.³⁸ Det stämmer också bra med Williamsons tes att det inte finns några satser, inte ens s.k. logiska sanningar, som inte kan förnekas av en person, utan att denne missförstår språket.³⁹ Det passar också bra ihop med Williamsons tes att vad som idag framstår som de bästa exemplen på analytiska eller begreppsliga sanningar kan i morgon ifrågasättas, på begripliga grunder, av någon uppfinningsrik tänkare.⁴⁰

Men Williamson tror – antagligen i motsats till Quine – att vi i stor utsträckning har samma intuitioner om vilka satser som är, respektive inte är, analytiska.⁴¹ Det stämmer nog inte, utom möjligen för mycket triviala exempel. Mer intressanta satser som t.ex. ”Alla människor är dödliga”, ”All konst är skapad av människor”, ”Man kan inte låtsas vara vaken”, ”Det finns oändligt många primtal”, ”Jag tänker, alltså existerar

³⁷I den sista bok han publicerade säger Quine ingenting alls om analyticitet – utom för en väldigt speciell sorts satser, s.k. observations-kategorialer; se Quine 1995, s. 45.

³⁸Quine 1953a, s. 43: ”no statement is immune to revision”. Se även 1991, s. 396–97.

³⁹Williamson definierar här ett begrepp analyticitet – som han kallar ”epistemologiskt” – så här: ”a sentence *s* is analytic just in case, necessarily, whoever understands *s* assentes to *s*” (2007, s. 73–74). Han menar att inga satser är analytiska i denna mening. Han argumenterar för följande tes: ”No given argument or statement is immune from rejection by a linguistically competent speaker” (ibid., s. 97).

⁴⁰”What strike us today as the best candidates for analytic or conceptual truth some innovative thinker may call into question tomorrow for intelligible reasons” (Williamson 2007, s. 126).

⁴¹Jfr Williamson 2007, s. 52.

jag”, ”Ingenting är orsak till sig självt”, ”Kunskap är sann och välgrundad tro”, ”Elefanter är större än myror”, ”Personer kan tänka”, ”Allt som har utsträckning har också färg”, ”Gud är allsmäktig”, ”Bergsbestigare är modiga”, ”Allt som är nödvändigt är möjligt”, ”Rationella aktörer maximerar förväntat värde”, ”Kuber har tolv kanter”, ”Bildning kräver naturvetenskapliga kunskaper” och ”Äktenskap är en relation mellan man och kvinna” är antagligen analytiska för somliga personer (vid vissa tidpunkter), men inte för andra.

Hur som helst har nog Quine hela tiden haft rätt i att det inte finns någon filosofiskt intressant form av allmän analyticitet eller ”begreppslik sanning”.

LITTERATUR

- Næss, Arne. 1981. *Empirisk semantik*. Solna: Esselte studium.
- Quine, W. V. 1946. ”On the Notion of an Analytic Statement”. I W. V. Quine. *Confessions of a Confirmed Extensionalist and Other Essays*, utg. av D. Føllesdal och D. B. Quine, s. 24–35. Cambridge, Mass.: Harvard University Press, 2008.
- . 1953a. ”Two Dogmas of Empiricism”. I *From A Logical Point of View*, s. 20–46. Cambridge, Mass.: Harvard University Press. Sv. övers. ”Två av empirismens dogmer”. I *Filosofin genom tiderna: 1900-talet; Före 1950*, andra upplagan, utg. av K. Marc-Wogau, S. Carlshamre och M. Gustafsson, s. 481–501. Stockholm: Thales, 2010.
- . 1953b. ”The Problem of Meaning in Linguistics”. I *From A Logical Point of View*, s. 47–64. Cambridge, Mass.: Harvard University Press.
- . 1960. *Word and Object*. Cambridge, Mass.: The MIT Press.
- . 1981. ”Use and Its Place in Meaning”. I *Theories and Things*, s. 43–54. Cambridge, Mass.: Harvard University Press.
- . 1991. Two Dogmas in Retrospect, omtryckt i W. V. Quine. *Confessions of a Confirmed Extensionalist and Other Essays*, utg. av D. Føllesdal och D. B. Quine, s. 390–400. Cambridge, Mass.: Harvard University Press, 2008.
- . 1993. *Strävan efter sanning*. Nora: Nya Doxa.
- . 1995. *From Stimulus to Science*. Cambridge, Mass.: Harvard University Press.
- Williamson, Timothy. 2007. *The Philosophy of Philosophy*. Oxford: Blackwell.

Nazismens ideologi och judeutrotningen

1. INLEDNING

Harald Ofstads bok, *Vårt förakt för svaghet* har utkommit i en ny tryckning på förlaget Karneval, med ett intressant förord av min kollega i Uppsala, Folke Tersman. Harald Ofstad var min lärare och en av mina företrädare som professor i praktisk filosofi vid Stockholms universitet (Lars Bergström innehade professuren mellan oss). Ofstads bok var tämligen nyutgiven (1971, på norska), då jag började läsa filosofi. Jag läste den i svensk översättning 1972. Jag läste ivrigt, den bjöd på ett oväntat möte med tankar, som få på den tiden tog på allvar. Jag läste den ändå med besvikelse. Jag minns att jag tänkte på Peer Gynt, som skalar sin lök. Han river av lager efter lager (komponent efter komponent i den nazistiska ideologin, i Ofstads framställning), men han finner ingen riktig kärna. Annat än, i Ofstads fall, den mycket allmänna tanken om förakt för svaghet. Visst är detta med förakt för svaghet ett spännande komplex, men problemet med det är ju att det är så allmänmänskligt. Det är inte något som kan särskilja den nazistiska ideologin.

Jag läser hans bok på nytt 2012 och får samma känsla. Det är spännande att ta del av referaten från Hitler. Jag tvingas gå till *Mein Kampf* för att själv läsa. Men *Vårt förakt för svaghet* ger inte svar på de frågor jag vill ha besvarade. Boken ställer dem knappt ens en gång. Den är osystematisk och svårtillgänglig. Den provocerar emellertid till eget tankearbete. Här ska jag ge två exempel på reflexioner, som fötts ur läsningen av den.

Det är framför allt två frågor man som läsare önskar att Ofstad hade formulerat klart och besvarat. Den ena rör den nazistiska ideologin i relation till andra politiska ideologier. Är det rimligt att hävda att det finns något sådant som en unik nazistisk ideologi, om man därmed menar en plan för hur samhället ska utvecklas och något slags motivering för en sådan plan? Jag ställer själv frågan och lutar åt att svaret är nej.

Den andra frågan rör det nazistiska försöket att systematiskt utrota judarna. Blir det begripligt i ljuset av den nazistiska ideologin? Svaret på den frågan måste också bli, nej. Detta kan synas trivialt om det inte finns någon särskild nazistisk ideologi, men jag vill försöka att driva en

mera långtgående tes. Jag tycker inte att det över huvud taget finns något förnuftigt stöd för detta folkmord i någon komponent i den nazistiska ”mixeln” av olika ideologier.

2. FINNS DET EN NAZISTISK IDEOLOGI?

Låt oss granska olika komponenter i den nazistiska ideologin, för att försöka se om de var och en är unika för nazismen. Ta först nationalismen. Ofstad gör olika distinktioner här, men inte de väsentliga, tycks det mig. För det första bör vi skilja mellan vad man kunde kalla en liberal och en konservativ nationalism. Nazismen tycks mig vara ett klart men föga originellt uttryck av den konservativa nationalismen.

Den liberala nationalismen kan ses som en universell princip. Varje nation ska ha sin egen stat. Det var en idé som ofta kombinerades med olika tankar om demokrati, civilisation och utveckling, i rörelser riktade mot konservativa imperier under andra hälften av 1800-talet. I vår tids multikulturella verklighet tycks den ha nått vägs ände. Det innebär att den har övertagits av rörelser som Sverigedemokraterna.

Den konservativa nationalismen är annorlunda. Den antar ingen principiell form. Den är snarare en ren attityd av gillande, riktad mot den egna nationen.

Typiskt för den konservativa nationalismen är också att man ser det man gillar, den egna nationen, det samhälle man tillhör, som en organism. Ja, man är beredd att gå längre. Man tillskriver nationen eller samhället en personlighet och en vilja. Den samtida konservativa tänkaren Roger Scruton har uttryckt saken så här:

Samhället är något för mer än en stum organism. Det har en personlighet och en vilja. Dess historia, institutioner och kultur är bärare av mänskliga värden – kort och gott, det är ett ändamål i sig självt lika väl som ett medel. (*The Meaning of Conservatism*, s. 23, min övers.)

Den tanken får ingen ordentlig behandling hos Ofstad. Han pratar om en orientering mot symboler (snarare än konkreta individer), sentimentalitet osv., men får därmed inte fast den konservativa nationalismens kärna. Visst kan en konservativ nationalist vara upptagen av olika symboler, men det hindrar inte att han kan vara intresserad också av individers väl och ve. Hitler uttrycker ofta ett sådant intresse, riktat mot ”riktiga” tyskar. Men det väsentliga är i sista hand att det går väl för nationen. Och det väsentliga är att nationen ses som en organism med en egen vilja. Det är därmed möjligt att det måste gå illa för vissa medlemmar av nationen då nationen triumferar.

Detta är emellertid konservativt allmängods. Och då det rör sig om

en attityd, snarare än en universaliserbar åsikt, är det möjligt för den konservativa nationalisten att gilla (också i sentimentala ordalag) den egna nationen, utan att bry sig ett dugg om hur det går för andra nationer. Hon vill t.ex. leva och dö i Norden, punkt slut. För hundra år sedan var denna hållning vanlig. Idag betraktar vi den nog lite till mans med överseende.

Nazismen intog också en hållning som kan beskrivas som anti-rationalistisk. Man motsatte sig tanken om att politik går ut på att rationellt formulera problem, granska olika alternativ, för att efter en förnuftig kalkyl välja det som tycks ge det bästa förväntade utfallet (bedömt utifrån den egna ideologin). Man hyllade i stället tanken om det intuitiva ledarskapet, utövat som ett slags konst. Detta är emellertid också det konservativt allmängods. Framför allt har Michael Oakeshott i vår tid artikulerat det tydligt (långt tydligare än Hitler förmådde).

Med anti-rationalismen, och tilltron till ledarskapet som en konst för den som hade gåvan, följde också en stark anti-demokratisk, anti-parlamentarisk inställning. En sådan slutsats drar inte alla konservativa idag, men i det tidiga 1900-talet var också detta allmängods. Ja, t.o.m. en tänkare och författare som Thomas Mann kunde före första världskriget (i *Betrachtungen eines Unpolitischen*) hylla denna inställning, då kultur ställdes mot civilisation, elitstyre mot demokrati och nationalism mot kosmopolitanism. Först efter det första världskrigets katastrof kunde Mann satirisera denna hållning. Den får representaras av Naphta i *Bergtagen*. Men porträttet av Settembrini är inte mindre skruvat, och man kan inte ens här vara riktigt säker på var författarens sympatier ligger. De klarnade nog i själva verket först i och med Thomas Manns påtvingade exil.

Nazismen hade som bekant inget eget ekonomiskt program, att jämföra t.ex. med socialisters strävan att förstatliga avgörande produktionsmedel. Man accepterade den rådande ekonomiska ordningen. Detta var ju också av strategisk betydelse. Det möjliggjorde maktövertagandet. Hade Hitler försökt välta kapitalismen över ända hade han knappast nått de politiska framgångar han gjorde.

Hans syn på samhället som en organism, med specifika uppgifter för olika individer, gjorde det möjligt för honom att under en tid flirta med arbetarrörelsen. Flirten fick emellertid inget riktigt programmatiskt uttryck och den avslutades i och med uppgörelsen med SA.

Man kunde tänka att åtminstone rasläran var något för nazismen specifikt och särskiljande. Jag tror inte att så är fallet. Här behövs emellertid flera klagöranden, för att vi ska förstå hur nazisterna tänkte.

3. RASLÄRAN

För det första var Hitler anhängare av en gängse föreställning om att den mänskliga arten (rasen) stod inför ett biologiskt hot. Välfärdsstatens inrättningar medförde att människor, som i en tidigare normal miljö, hade gått under, överlevde. De kunde också reproducera sig. Resultatet blev en försämring av vad vi idag skulle kalla den mänskliga genpoolen. Om 1000 år, förmodar Hitler, kan mänskligheten ha gått under, om inget radikalt görs.

Var denna uppfattning specifik och särskiljande för nazismen? Nej, den var allmängods. De enda som ifrågasatte den var i själva verket vissa kristna grupper, som inte trodde på evolutionsläran över huvud taget, och kommunister, som trodde på evolutionsläran i dess av Lysenko modifierade form (med en stark tro på vad vi idag skulle kalla epigenetik). Jag tror man måste medge att de som trodde på detta hot var förnuftigare än de som, på grund av sina religiösa eller ideologiska skyggglappar, förnekade det. Vi vet idag att hotet inte var reellt, men det vet vi på grund av senare rön om hur recessiva anlag fungerar i ett evolutionärt sammanhang. Slutsatsen som drogs i många länder, också i Tyskland, var att den mänskliga genpoolen måste försvaras mot detta hot. Åtgärden var sterilisering av vad man uppfattade som icke livsvärda individer. Sådan sterilisering förekom som bekant inte bara i Tyskland utan också i t.ex. USA och Sverige.

Hitler var också anhängare av en raslära med innebörden att det fanns olika mänskliga raser, vissa överlägsna andra i framför allt intellektuell och skapande hänseende. Högst stod vad han omtalade som den ariska rasen. Den stod t.ex. över svarta människor, men också över asiater. Om en arisk och en svart individ fick barn, förmodade Hitler, skulle de representera ett genomsnitt av föräldrarnas intellektuella kapacitet. Här tycks han ha tänkt ungefär i analogi med hur vi urskiljer olika hundraser, framavlade, med mycket bestämda och unika karaktäristika.

Inte heller denna föreställning var specifik för nazisterna. Den föranledde tanken att man inte ska blanda olika raser. Då skulle vi snart förlora resultatet av årstusenden av avel. Det vore orätt av hänsyn till den ”bästa” rasen och, indirekt, även av hänsyn till det allmänna bästa. Andra raser kunde nämligen, enligt Hitler – t.ex. japaner – bära en kultur vidare, åtminstone under en period, men de kunde inte själva skapa och utveckla en kultur. Detta var förbehållet medlemmar av den ariska rasen, som alltså behövdes för mänsklighetens fortsatta utveckling.

Hitler drog här samma slutsats som man drog i Sydafrika eller i den amerikanska södern. Raserna borde hållas åtskilda. I annat fall skulle alltså förädlingsvinsten med den ariska arten snart gå till spillo. Han tog miste också på den här punkten, men han var på intet sätt originell i sitt

misstag. Också här fanns viss "evidens" för hypotesen. Försök gjordes att pröva den. Ett exempel var den tyske docenten i anatomi, Eugen Fischer. Han var fascinerad av Mendels genetik, som just återupptäckts. Han anställde 1908 en studie i nuvarande Namibia. Han undersökte svarta och vita och individer av blandras och fann att de vita var intelligentast, de svarta dummast och de av blandras (bastarderna) föll någonstans mitt emellan de renodlade rasgrupperna. Resultatet publicerades i boken *Die Rehoboter Bastarde und das Bastardisierungsproblem beim Menschen* (Jena: Gustav Fischer, 1913). Just så tänkte också Hitler om saken.

Vad ska man då säga om antisemitismen? Den är inte att ta miste på i *Mein Kampf*. Den tar sig våldsamma och tjatiga uttryck, men knappast några som är specifika för just nazisterna. Antisemitismen var spridd och Hitler tar den till sig med hull och hår. Han hävdar att judarna är en ras, inte en religion. Problemet med dem är emellertid inte i första hand att de skulle hota den ariska rasens renhet (genom rasblandning). Hotet är mera omedelbart. Judarna utgör som en grupp ett hot mot den tyska nationen, menar han, ja mot all mänsklig civilisation. Här kopplar han för övrigt ihop sin antisemitism med en annan och lika starkt antipati som han hyser, anti-kommunismen. I själva verket är det som om de tu i hans föreställningsvärld blir ett. Judar ses som bolsjeviker och omvänt. Hans konspiratoriska tänkande går här över alla gränser. Ändå är detta, måste vi nog konstatera, en trist del av tidens allmängods.

4. JUDEUTROTNINGEN

Jag har hävdad att det inte finns någon specifik nazistisk ideologi. Vad vi finner är en mix av traditionella uppfattningar, rätt och förslaget nyttjad av ledande nazister i kampen om den politiska makten. Nazisterna var särskilt hänsynslösa i sin framfart, men det är ju något annat än att företräda ett särskilt politiskt program. Hur är det då med min andra huvudfråga, som rör judeutrotningen? Fanns det stöd för den i nazisternas politiska program och tänkande (t.ex. i *Mein Kampf*). Jag tror man måste svara nej på den frågan.

I den mån judarna sågs som en ras, vilken var underlägsen den ariska, och därmed kunde utgöra ett hot mot den ariska rasen, om rasblandning tilläts, kunde detta problem enkelt hanteras genom förbud mot sådana förbindelser. Ett sådant förbud utfördes ju också i och med Nürnberglagarna.

Det bör dock noteras att det finns en vacklan hos Hitler. Trodde verkligen Hitler själv på att det fanns något sådant som en judisk ras? Troligen inte. Han förnekar det i sitt "testamente". Han skriver där:

Vi använder för enkelhetens skull begreppet ”judisk ras”. I verkligheten och ur genetisk synpunkt finns ingenting sådant som ”den judiska rasen”. Den finns emellertid ett samfund på vilket begreppet faktiskt kan användas och vars existens judar över hela världen medvetet sluter sig till . . . det är denna människogrupp som vi ger namnet ”den judiska rasen”.

Han tvekade emellertid inte att underblåsa myten om en judisk konspiration. Judarna hade tillskansat sig viktiga positioner i samhället, vilka de avsåg att bruka för destruktiva syften. De utgjorde därmed ett hot mot den tyska nationen. Kanske trodde han själv att detta var en realitet. Det var en spridd uppfattning, och att spela på den var den hur som helst ett sätt att vinna politisk framgång.

En sådan föreställning, om nazisterna tog den på allvar, skulle kunna motivera att judar fråntogs sina positioner i samhället, för att göras ofarliga. Det var ju också vad man gjorde, med början i ”kristallnatten”. Denna föreställning skulle kunna motivera att man fråntog judar deras egendomar och tvingade dem i exil. Det kunde motivera samarbete med den sionistiska rörelsen, i syfte att bli av med landets judar. Det var också något som förekom.

Allt detta visade sig vara gångbart politiskt. Vid universiteten var det inte svårt att finna personer, som gärna övertog de avskedade judarnas tjänster. Inom de kapitalägande grupperna var det inte svårt att finna tjänstvilliga, som övertog de judiska rörelserna. Och med genetiker, som med ämnets nestor, Alfred Ploetz, var det inte svårt att knyta en allians. Också kritiskt tänkande genetiker fann det lätt att hålla sig till en ovetenskaplig antisemitism (vilken också främjade deras karriärer) om de i utbyte från staten fick vad de uppfattade som en sund eugenisk politik.

Kanske kunde en sådan hållning, där judarna utmålades som ett hot mot den nationella säkerheten, också motivera att man placerade dem i koncentrationsläger? Det var ju gängse behandling av människor, som ansågs ”samhällsfarliga”. I Sverige placerades kommunister och vänster-radikala i läger som Storsien.

Men kunde dessa föreställningar göra tanken om ett systematiskt massmord på judar begriplig? Jag tror inte det. Det är vanligt att citera Hitler ur *Mein Kampf* då han skriver:

Om man någon gång vid krigets början och under dess lopp tagit femtontusen av de hebreiska folkförförarna under en likadan giftgasbehandling, som hundratusentals präktiga tyska arbetare ur alla grupper och yrken måste utstå i fält, hade krigets miljoner fallna icke dött förgäves. Tvärtom hade det kanske kunnat rädda till livet en miljon tyskar av stort värde för framtiden, om dessa femton-tusen undanröjts i rätt tid.

Men nog är det en grov övertolkning att se detta som ett programmatiskt ställningstagande för ett systematiskt dödande av alla judar. Något sådant ställningstagande görs inte i *Mein Kampf*. Och något sådant tycks inte heller nödvändigt, av immanenta ideologiska skäl. Det är nog att förhindra rasblandning och att frånta judarna deras ställning i samhället.

Hur ska man i så fall förstå det systematiska försöket att döda alla judar, vilket ju drevs också på ett sätt som brukar anses ha hämmat den tyska krigsförmågan?

5. IDEALISM ELLER TILLFÄLLIGHETER?

Jag har undvikit en komponent i den nazistiska ideologin, som Ofstad lägger stor vikt vid: idealismen. Den är inte unik för nazismen, men klart är att den tog unikt starka uttryck i den nazistiska maktutövningen. Idealismen ska förstås som en beredskap att offra enskilda individer (sig själv och andra) för en högre sak. Denna hållning idoliseras av Hitler. Vi är nog alla beredda att hemfalla åt idealism. Vi är beredda att hylla ”hjältar” som offrar sig för det allmänna bästa. Problemet med idealismen är att det går att ta miste. Om offren sker för en ond sak, så är de ju helt förgäves! Jämför den norska motståndsrörelsen med Anders Behring Breivik!

Kan idealismen på något vis förklara det nazistiska folkmordet? Kan den institutionaliserade idealismen, i form av stränga krav också på likgiltighet inför offren, i kombination med en förvriden uppfattning om hur världen bör utvecklas, förklara det nazistiska folkmordet?

Jag tror att just institutionaliseringen av likgiltigheten för offren ger en viktig del av förklaringen. Här fanns varken utrymme för sentimentalitet eller hatiskhet. Men den institutionaliserade likgiltigheten för offren räcker inte ensam för att förklara det systematiska folkmordet. Om vi tänker oss att judeutrotningen verkligen var kontraproduktiv från nazistisk utgångspunkt, så behövs någon ytterligare förklaring. Varför fick den ske?

Det ligger nära till hands att här peka på (skrämmande) enkla institutionella förhållanden. De gjorde att man ”osökt” kunde gå från internering till utrotning av en hel folkgrupp.

Det politiska systemet byggde på att en enda man skulle fatta alla beslut. Det gällde den högsta nivån, men det gällde också alla lägre nivåer i hierarkin. Hitler är utförlig på den punkten i *Mein Kampf*. På varje nivå i hierarkin gällde också principen om absolut lydnad inför den högre instansen.

Det brukar antas att Hitler aldrig gav någon uttrycklig order om judeutrotningen. Han låter sina underlydande förstå att ”problemet” måste få en ”slutgiltig lösning”. Vilket var problemet? Tja, en viktig del av det utgjordes väl av själva förhållandet att man hade en stor grupp människor internerade. Vad skulle man göra med dem?

Himmler försöker efter bästa förmåga gissa sig till Hitlers avsikter (i skräck för att missförstå dem), konkretisera dem i olika beslut. Höss tar över i Auschwitz, ”uppfinner” bruket av cyankalium. Den nödvändiga utrustningen finns på plats. Gaskamrar kamoufleras som duschrum. Krematorierna, som ursprungligen introducerats för att bränna de infekterade liken av lägerfångar, vilka duktat under för epidemier, finns på plats (en i sig lärorik historia som jag inte här kan gå in närmare på). Hanteringen går som på räls. Alla gör vad som förväntas av dem. Problemet får sin lösning. Ingen är beredd att ta ansvar för något beslut. Den situationistiska psykologin, exemplifierad av Milgrams försök, förklarar resten. Idel banal ondska, som Hannah Arendt uttryckt saken.

6. SLUTSATS

Om det inte fanns någon unik nazistisk ideologi, om massmordet på judar saknar förnuftigt stöd i det hopkok på idéer, som brukar omnämnas som nazismens ideologi, om det kan förstås som resultatet av en samhällsorganisation, som lät det hända, vilken slutsats ska vi i så fall dra för framtiden? Vi bör undvika det slags hierarkiska samhällsorganisationer, som underlättar känslan av ansvarsfrihet, och vi bör låta bli att sätta människor vi ogillar i koncentrationsläger. Det kan låta platt, men det är vad som omedelbart slår mig.

1. INLEDNING

Alan Turing (1912–1954) är en av 1900-talets stora tänkare. (Den som vill få en överblick över Turings storhet kan få ut mycket av samlingsverket Teuscher 2010.) Turing gjorde sina största insatser inom logik och matematik. Hans teori om s.k. Turing-maskiner är viktig inom logiken, och bildar en abstrakt grundval för vetenskapen om datorer. Under andra världskriget var han en av de viktigaste som var med och löste tyskarnas ”Enigma-kod” för kryptering, vilket bidrog till krigets utgång. Han borde alltså ha blivit betraktad som krigshjälte i England efter kriget, men så blev det inte. I stället blev han på grund av sin homosexualitet så skandalöst behandlad av landets rättsmaskineri att han tog livet av sig. Detta är en historia som bör berättas. Men nu skall jag ta upp en annan sak.

1950 publicerade han i tidskriften *Mind* en artikel som efterhand kom att bli en av de mest citerade filosofiska artiklarna överhuvudtaget. Den hette ”Computing Machinery and Intelligence”. Själv anser jag att denna artikel är viktig på ett negativt sätt. Den har bidragit till nutida oförståelse för medvetandets problem och det genom att formulera en falsk teori som mest är en tillämpning på datorerna av den s.k. behavioristiska traditionens synsätt. Men genom att vara en särskilt pregnant artikulering av detta synsätt är det en teori som är värd att titta närmare på.

Den *behavioristiska traditionen* inom psykologi och filosofi utmärks av en tendens att i någon bemärkelse identifiera ett psyke med ett visst beteende. Detta har preciserats i olika riktningar vilket gett en uppsättning olika behaviorismer. Traditionens historia började på 1910-talet inom ämnet psykologi. Ett ställningstagande till Turings artikel måste ta itu med den gamla stridsfrågan om behaviorismen, och det skall jag göra här via Turings sätt att formulera problemet.

Man kan alltså göra en ”*minimaltolkning*” av Turings idéer i artikeln: Att det bara är en variant av *behaviorism* tillämpad på datorer. Den tolkningen håller sig närmast texten. Men man kan också göra en ”*maximaltolkning*” som anknyter till nutida idéer – som brukar uppträda med namnet ”*funktionalism*”.

Jag skall här inte gå in på detaljerna i artikeln utan plocka fram det som jag uppfattar som det väsentliga i den. Jag skall alltså inte lusläsa och analysera Turings formuleringar utan gå in på de *problem* han tar upp. Därför skall jag omformulera och generalisera hans tankar.

2. TURINGS PROBLEM

Turings artikel med dess spekulationer om datorers möjligheter, inledde den forskning som senare har kommit att kallas ”artificiell intelligens” (förkortat AI). Hans grundproblem var frågan ”Kan man genom lämplig programmering skapa ett mänskligt psyke i en dator?”, och hans svar var ”Ja!”. Frågan gällde inte vad som var tekniskt möjligt att åstadkomma 1950, utan vad som är *möjligt i princip*. Man kunde uppfatta det som ett problem som kunde lösas i en framtid på sciencefiction-avstånd från, inte bara 1950, utan också från vår nutid.

Jag måste nu förklara vad jag menar med ordet ”psyke”. Jag använder det här som en översättning av engelskans ”mind”, som är ett bra ord. I engelska kan man göra en distinktion mellan ”mind” och ”consciousness” – och det är inte så lätt att göra något motsvarande i svenska språket. Det senare engelska ordet betyder *medvetande* på svenska, men för det förra har vi inget etablerat ord i svenska. Jag skall alltså använda ”psyke”. Det avser allt som vi syftar på med psykologiska termer, rubb som stubb, både sådant som har med medvetandet att göra och sådant som inte har det.

Att jag använder ett särskilt ord, ”psyke”, för allt detta innebär inte att jag betraktar psyket som ett särskilt objekt – eller en ”entitet” som filosoferna säger. Våra psykologiska termer, både i vardagligt språk och i vetenskaplig psykologi, fungerar på mycket olikartade sätt i språket, och anger tillsammans ett begreppsligt heterogent område. Så är emellertid *inte* fallet med begreppet ”medvetande”, vilket jag skall återkomma till.

Man kan undra om Turing verkligen avsåg att ställa sin fråga om *allt* i psyket eller om han begränsade sig till en del av det, nämligen till *intellektet*. I hans överskrift ingick ordet ”intelligence” och i början av texten ställer han frågan om en maskin *kan tänka*. Det låter som om det skulle handla om intellektet. Men hans begrepp om ”tänkande” framstår i den följande texten som så generell att det verkar överskrida intellektet. Hur som helst: den allmänna frågan ovan, alltså den om psyket i allmänhet, är den *mest intressanta*. Jag skall hålla mig till den.

Bakom Turings fråga ovan finns en annan mer generell fråga: ”Kan man på konstgjord väg i en artefakt, en ”maskin”, skapa ett (ev. mänskligt) psyke?” Vi kan kalla detta ”*syntesproblemet*”.

Man kan hos Turing tolka in två förslag till lösning av syntesproblemet. *Minimaltolkningen* är det *behavioristiska* förslaget: maskinen kon-

strueras så att den *beter sig som om den hade* ett psyke. Då *har* den också – med begreppslik nödvändighet enligt Turing – ett psyke. Psyket sitter så att säga i *beteendet*. Denna tolkning håller sig nära texten.

Så har vi maximaltolkningen, som kan kallas ”*mjukvaraförslaget*”. Innebörden är: det är genom *programmeringen* av en dator som man kan syntetisera ett psyke. Psyket sitter så att säga i detta *program*, i *mjukvaran*. Olika program kan ge upphov till samma beteende, men enligt den här idén finns psyket i något unikt, eller åtminstone i någon snäv mängd av, program. Denna idé som numera kallas (vissa former av) ”funktionalism” eller ”stark AI” finns inte uttryckligen formulerad i texten, men kan kanske läsas in i den.

Om denna tolkning skall vara intressant, bör den kunna tillämpas också på oss människor, och då måste den ge en innebörd åt begreppet ”mjukvara” som kan tillämpas på våra hjärnor.

Båda de här förslagen har till konsekvens att *hårdvaran*, själva de ”hårda” tekniska detaljerna i maskinens uppbyggnad, är *begreppsligt oväsentlig*, vilket i sin tur betyder att psyket av båda uppfattas som en *abstrakt egenskap* hos hårdvaran. Detta är det väsentliga med båda.

Vad är det då som är intressant med Turings fråga och svar? De är ju intressanta för dem som är intresserade av datorer, och det finns många som har det intresset. Men de är också intressanta på ett annat sätt, nämligen som en uppfattning om *oss själva*. Turing framställde en teori om ”psykets natur i allmänhet, och om hans teori är riktig får det stora konsekvenser för vår uppfattning om oss själva.

För att se att denna teori är filosofiskt viktigt skall jag ge ännu en formulering av den problemställningen som är inblandad. Man kan se den som ett *variationsproblem* i filosofen Edmund Husserls mening: hur mycket kan man ändra på, variera, en människa utan att hon förlorar sitt psyke? Husserl varierade en företeelse för att bestämma dess *väsen* och kallade det ”eidetisk variation”. Han ville se hur mycket man kan variera något utan att det förlorar sitt väsen. Så kan man också se det i det här fallet. Turings ståndpunkt om oss människor, torde ha varit att man utan att psykets väsen går förlorat kan variera vår mänskliga hjärnas ”våtvara” till godtycklig annan hårdvara som har en viss abstrakt egenskap gemensam med vår hjärna

Nå, hade Turing rätt? Mitt eget – och många andras – svar är: *Nej!*

3. TURINGS TESER

Vi kan betrakta närmare hur Turing resonerade. Man kan urskilja två teser i hans resonemang. Den ena kan sägas vara tesen (ATT): Det är möjligt *att* en dator genom lämplig programmering kan bete sig så likt

en människa, att den kan *lura* en mänsklig observatör till att tro att den *är* en människa. När en dator lyckas med att luras på detta vis, sägs den ha klarat *Turings test*.

Den andra är tesen (OM): *Om* en dator klarar Turings test, så *har* den ett *mänskligt psyke*. Dvs. om en dator klarar testet så har man löst syntesproblemet.

Man får förgäves leta efter exakt dessa formuleringar i Turings text. Ändå går hela hans artikel ut på att hävda dessa teser. Teserna är ofta dåligt åtskiljda i debatten om Turings artikel, men de leder till mycket olikartade filosofiska problem.

Turing tänkte sig att vissa villkor måste vara uppfyllda för att en dator skulle kunna luras på det där viset: observatören, eller ”testaren” som jag skall säga, får inte ha möjlighet att se datorn i sitt skåp (eller vad det var 1950). Datorn måste vara dold för testaren. Kontakten mellan testare och dator får ske genom utbyte av textmeddelanden, dvs. genom det vi nu kallar *chattande*. En dator klarar alltså Turings test om den kan chatta som en människa.¹

Man kan lätt se att denna Turings chattversion av problemen inte är något vidare bra. Den framstår som en onödig begränsning av det filosofiska problemet. Testobjektets beteenderepertoar blir för liten och testsituationen är alltför laboratorieartad. Redan den okroppsliga synen på psyket ger problem. Genom Turings chatt-test kan man avgöra t.ex. om testobjektet behärskar ett språk, kan skriva dikter och spela schack, men inte om den kan spela fiol, skala en potatis eller ens gå på två ben.

Än värre blir det om man kommer in på medvetandet, testobjektets upplevelser. Antag att objektet under chattandets gång blir ”het på gröten” som man säger och inleder ett veritabelt sex-chattande med testaren. Om det då visar sig att testobjektet är en dator som chattar på detta vis, kan man då dra slutsatsen att den *har* heta lustar? Ja, enligt OM kan man det.² Men det verkar inte rimligt att på fullt allvar anta att datorn har heta lustar. Det enda rimliga är att endast säga att datorn i detta fall *kan chatta som om den hade* det. (Om en filosofisk vän av ordning nu påpekar att jag borde ge ett *argument* för denna åsikt, så kan jag svara att jag längre fram skall göra det.)

¹Turing la på ytterligare villkor på testet. Han tänkte sig att tre parter skulle vara inblandade, och att testet skulle bli mer formellt bestämt som ett slags spel, ”the imitation game”. Detta finner jag vara överflödigt och jag skall inte gå in på det.

²Man kanske kan invända att Turing aldrig avsåg att man skulle kunna dra en sådan slutsats eftersom han skulle ha begränsat sin tes till *intellektet* i någon vid bemärkelse. Jag kan svara, som ovan, att det filosofiska problemet blir mycket mer intressant om tesen OM syftar på ett fullblodspsyke, ty då avser Turings teori att säga mycket mer om oss människor.

När vi möter en människa, möter vi en kropp med mycket omfattande beteendepertoar i en komplex situation. Den vi möter har ett kroppsspråk, ansiktsuttryck, och en röst. Och vi ingår båda i någon social situation med alla de relationer en sådan kan innefatta. Filosofen Sören Stenlund och sociologen Johan Asplund har, på olika vis, påpekat dessa sociala omständigheter som en kritik mot Turing.³

Turings idéer var redan sciencefictionartade 1950 och är det än idag. Ingen dator klarar Turings test idag, dvs. om chattandet får bre ut sig obehindrat. Varför då begränsa science fiction till chattsituationen och inte i stället dra på för fullt? Låt oss göra det senare!

Vi *generaliserar* alltså testet och de båda teserna till en *människoliknande robot i sociala situationer*. Roboten har en kropp som ser precis ut som en människas och den har ett avancerat människoliknande beteende. Roboten *klarar* det generaliserade testet om den *lurar* människorna i sociala situationer att den *är* en människa. Vi får tänka oss att den lurar på ett avancerat sätt. Den lurar även i nära relationer med människor. Och om den t.ex. lägger sig på schäsongen hos en psykoanalytiker och börjar associera fritt, så blir analytikern lurad att där ligger en människa.

Vidare: Enligt den generaliserade tesen ATT* är det genom lämplig programmering av datorn i en datorstyrd robot möjligt *att* få roboten att klara det generaliserade testet. Enligt den generaliserade tesen OM* *har* en datorstyrd robot ett mänskligt psyke *om* den klarar det generaliserade testet.

Det är så här man bör formulera Turings problem. På detta vis ställs de filosofiska problemen på sin spets. Vilka problem? Jo, följande.

För det första hävdar ATT* att en dator kan bete sig precis som en människa, dvs. att en dator kan styra en robot så att denna betar sig precis som en människa. Problemet kring denna tes handlar alltså om hur avancerat beteende som är möjligt för en dator.

Och vi kan se att OM* är en specialisering av följande generella *behavioristiska* tes till datorfallet (P): Om en varelse, levande eller konstgjord, *beter sig som om* den har ett (mänskligt) psyke, så *har* den ett sådant.

Det stora krusket för behavioristiska teorier är medvetandet. Krusket för P är alltså följdtesen (M): Om en varelse *beter sig som om* den har ett medvetande, så *har* den ett medvetande.

Den *behavioristiska traditionen* inom psykologi och filosofi utmärks, kan man säga, av att anta P och oftast även P:s omvändning. Många psykologer och filosofer har ända från början gett invändningar mot dessa idéer. Det hade alltså länge funnits en debatt om dessa saker före 1950.

Turing hävdade alltså inget nytt med sin tes OM* (eller OM). Turing

³Stenlund 1990, Asplund 2002.

ger emellertid i texten ett slags motivering för dessa teser via en motive-
ring för M, nämligen genom att betrakta M som ”a polite convention”
– en *artig konvention*. Turing tänkte sig att följande är byggt på en *kon-*
vention: när vi möter en varelse som beter sig som om den har ett med-
vetande *bemöter vi den som en medveten varelse* – med allt vad det innebär,
dvs. så gör vi i vanliga fall. Det är förvisso sant (och mycket rimligt) att
vi i de flesta fall bemöter människor, hundar och katter och många andra
djur på detta vis.⁴ Frågan är dock om detta kan betraktas som byggt på en
konvention. Har vi inte goda skäl att anta att dessa varelses *är* medvetna?
Alltså ett antagande om fakta och inte en konvention. En del nutida filo-
sofer skulle dock hålla med Turing i denna sak. Så kommer emellertid det
stora problemet: Turing hävdar att vi kan (bör?) *utvidga* denna antagna
konvention *till en dator* som klarar hans test. Men varför skulle man vara
artig mot en dator på detta vis? Det tycks ju hänga på om datorn verk-
ligen *är* medveten eller ej. Det framstår som fullt rimligt att vara artig
mot människor och djur men oartig mot datorer. Turing går alls inte in
på den saken. En naturlig tolkning är att han förutsatte det som borde
ha bevisats, dvs. han förutsatte den behavioristiska ståndpunkten M och
tillämpade den på datorerna.

Jag skall strax ta upp problemen kring OM* på ett annat sätt.⁵ En sak
som man dock direkt kan se är att om man kan *förkasta* OM*, och det
skall jag visa att man kan, så behöver man inte ta ställning till ATT*. Ty
även om en dator skulle klara testet, så följer när OM* förkastats ingen-
ting om datorns eventuella psyke.

4. ALGORITMER

Jag måste säga något om ATT*. Jag skall först säga något om vad en *dator*
är. Jag skall betrakta datorn mycket allmänt, nämligen som ett system
vars uppförande bestäms av att den *programmerats* att följa någon *algo-*
ritm. Det sistnämnda begreppet ovan kan exemplifieras av de räknesätt
för heltal som vi lär oss tidigt i skolan, dvs. addition, subtraktion, osv.
De bestäms av *regler*. Med en regel kan man ta itu med en mängd olika
fall. Regeln för addition skall tala om hur man kan räkna ut $1 + 1$, $1 + 2$,
 $2 + 3$, osv. Additionsregeln täcker alltså *oändligt* många enskilda fall, och
så är det också för de tre andra räknesätten och många andra algoritmer.

⁴Det är ju välkänt att det också finns hemska undantag då människor och
djur blir behandlade som om de saknade känslor och förmåelser. Alltså
koncentrationsläger, djurfabriker, osv.

⁵Jag skall i denna artikel säga en del om tesen ATT*, dock kan jag av
utrymmesskäl inte säga så mycket som den förtjänar. Denna tes är i och för sig
mycket intressant och förtjänar en egen artikel.

Redan på dessa elementära räknesätt kan vi se vad som är väsentligt för en algoritm.

Jag skall bara nämna en sak: algoritmens regel kan uttryckas genom en *ändlig* uppsättning *instruktioner* som har en särskild karaktär. De är ”mekaniska” i följande mening: instruktionerna måste kunna följas endast genom att man slaviskt, ”som en idiot”, följer själva instruktionerna, utan tillämpning av egen insikt eller kreativitet. Sådana ”mekaniska” instruktioner lämpar sig alltså för maskiner. Algoritmernas mekaniska karaktär är deras utmärkande, och mest intressanta, egenskap.

Algoritmer behöver inte vara matematiska, de kan handla om vad som helst. Alla algoritmiska regler kan reduceras till beräkning av *funktioner*. I sådan ”får man ut” ett entydigt resultat när man ”stoppar in” något. De ovan nämnda räknesätten anger funktioner. Funktioner behöver inte handla om tal. Men alla algoritmiska funktioner måste vara ”diskreta”, dvs. de handlar om objekt som kan *numreras av heltal*.

Begreppet algoritm leder till djupa problem i en mängd områden: logik, matematik, fysik, människans intellekt, osv. För att kunna ta itu med dessa problem måste man emellertid precisera begreppet, och den stora svårigheten ligger i att precisera begreppet ”mekanisk”. Det gjorde Turing på 30-talet med sin mest berömda uppsats (Turing 1936) som gav en definition av den s.k. *Turing-maskinen*. Denna är en standardform som han antog att man kunde ge alla algoritmer. Med begreppet Turingmaskin bevisade han i denna artikel att alla diskreta funktioner inte kan beräknas med algoritmer – med Turingmaskiner. Jag skall säga att sådana funktioner är ”överalgoritmiska”.⁶

Begreppet ”överalgoritmisk” kan uppfattas som ett *matematiskt begrepp om kreativitet* – eftersom det handlar om överskridande av det mekaniska. Detta kan inte direkt identifieras med mänsklig kreativitet – dock kan man fundera på eventuella samband.

Begreppet ”överalgoritmisk funktion” har en viktig underavdelning, nämligen de överalgoritmiska funktioner som i någon mening kan ”beräknas” med specificerbara procedurer. Sådana funktioner har kommit att kallas ”hyperberäkningsbara” (eng. *hypercomputable*) och de har på senare tid bildat ett snabbt växande och kontroversiellt forskningsområde.⁷

⁶I analogi med mängdteoris ”överuppräknelig”.

⁷Till särskilt kända inslag i debatten på området hör Roger Penroses *The Emperor's New Mind* (1989) med efterföljare. Syropoulos 2008 ger en samlad framställning av området. En viktig kritiker av dessa teorier är den framstående logikern Martin Davis, se t.ex. hans ”The Myth of Hypercomputation” i Teuscher 2010.

5. HÅRDVARANS ABSTRAKTA EGENSKAP

Tillbaka till datorn: den kan alltså programmeras med en algoritm. Det betyder att instruktionerna lagras i datorn. Men nu kommer något viktigt: det har ingen betydelse vad de lagras *i*. Mediet, *hårdvaran*, är *oväsentlig*. I nutida teknik finns magnetisk lagring, men också optisk och elektronisk av olika slag. Det kommer säkert nya i framtiden. Det väsentliga i lagringen är att den består i att hårdvaran ges en *abstrakt strukturell egenskap*. Tänk på t.ex. hur ett program överförs, ”laddas ner”, från först en optisk disk ner i datorns hårddisk, som är magnetisk, och sedan i ett USB-minne som består av chips. *Samma* abstrakta egenskap har uppträtt i *olika* konkreta bärare.

Detta resonemang gäller inte bara datorns minnen utan kan också tillämpas på dess processor. Samma beräkningsegenskaper kan realiseras av olika hårdvara, olika konkret uppbyggnad, men med samma abstrakta egenskap. Allt det här kan inspirera till den ovan nämnda ”maximaltolkningen”.

6. DATORERS FÖRMÅGOR

Nu kommer jag äntligen fram till tesen ATT*. En dator styrs alltså av algoritmer. Kan sålunda ett mänskligt beteende vara framställt av en algoritm? Nu gäller det inte bara chatt-beteende utan rubbet. Man gör det alldeles för enkelt för sig om man omedelbart svarar ”Nej!”. Vi skall tänka på tre saker.

För det första: tesen ATT* handlar om vad som är möjligt *i princip*, inte om vad som är mänskligt möjligt. Det gäller alltså inte vad som en mänsklig programmerare rimligen kan åstadkomma med en dator. Vi får i stället tänka oss att det handlar om vad en gudomlig programmerare kan åstadkomma.

För det andra ska vi tänka på *uppdelningen i nivåer*. Datorn är strängt och mekaniskt lagbunden och kan beskrivas med en överblickbar uppsättning matematiska och naturvetenskapliga begrepp. Men så är det inte för en människa på ”mänsklig nivå”, dvs. så som hon uppfattas av andra människor. Vi beskriver människor, och uppfattar dem, med en uppsättning begrepp som inte är naturvetenskapliga. De förra har andra logiska egenskaper än de senare. Det förra slaget beskrivningar kan i allmänhet inte formaliseras. Och även om det finns några lagar för människor på mänsklig nivå är det inte rimligt att hävda att människor är *lagbundna* på denna nivå.⁸ Och vidare framstår människors egenskaper som otaliga och oöverblickbara. Det framstår ofta som orimligt att

⁸Detta är något som de flesta vanliga människor vet utan att formulera det, och som ibland uttryckligen har formulerats av filosofer, t.ex. av Donald Davidson.

försöka skapa fullständiga teorier om människor. Tänk bara på hur löjligt det skulle vara att försöka skapa sådana teorier om humor, konst, musik.

Men man kan inte av dessa uppräknade fakta *omedelbart* dra en slutsats att det hos människor finns något som överskrider alla algoritmer. Människan i vardagslivet befinner sig på en annan logisk nivå än datorn. Men detsamma kunde gälla *även en dator* med mycket avancerat program.

För det tredje: *om* människans beteende står i strid med ATT*, så kan det bara betyda att människan har *överalgoritmiskt* beteende. Det är en *extrem* egenskap, som implicerar att *naturlagarna* har överalgoritmisk karaktär. Det kan ju vara sant, och det diskuteras i nutida debatt av Roger Penrose och andra i hyperberäkningsvägen. Argumenten för det här bygger alltså på antagandet att vi människor i något avseende är kreativa i den *matematiska* meningen, och argumenten brukar utgå från några föregivna matematiska förmågor hos människor.⁹ Men det är ännu osäkert hur det faktiskt förhåller sig med denna sak.

Sammantaget finner jag att det ännu inte finns några starka argument mot ATT*. Denna tes framstår, åtminstone än så länge, som rimlig.

7. PARALLELLITET OCH IDENTITET

Låt oss nu i tanken genomföra det generaliserade Turing-testet. Vi antar alltså att dessa robotar lurar oss. Det intressanta är hur vi förhåller oss till dem *efter* att de avslöjats. Detta experiment är inte möjligt i verkligheten, med det är lätt att föreställa sig. Vi kan ju föreställa oss människor, och dessa robotar är som människor när de betraktas utifrån. Men sådana robotar finns också på film, t.ex. i "Alien"-serien. I den visas datorstyrda sociala robotar som först lurar människor och sedan avslöjas.

Vi kan i dessa filmer se hur en möjlig värld med sådana robotar kan vara, men också hur den *måste* vara – och det sistnämnda är viktigt: det blir klart att de som ingår i det sociala spelet med dessa robotar *måste* använda våra vanliga psykologiska och sociala ord, inte bara när det talar *om* robotarna, utan naturligtvis också när de samtalar *med* dem. Det gäller naturligtvis så länge de lurar oss. Men det viktiga är att det *också* gäller efter att de avslöjats.

Och dessa ord måste, i en viss bemärkelse, användas på samma *sätt* som vi gör i våra vanliga mänskliga situationer. Den "vissa bemärkelsen" är alla yttre, beteendemässiga och sociala, svar på "när säger vi vad?". Vi måste alltså använda samma ord, men det måste inte betyda att orden har samma innebörd i robotvärlden som i vår värld. Men det måste råda en

⁹Argumentationen har bl.a. av Penrose anknutit till Kurt Gödels logiska s.k. ofullständighetsteorem från 1931. De flesta logiker anser emellertid att sådana argument bygger på en missuppfattning av Gödels teorem.

parallellitet. Detta innefattar även psykologiska och sociala *förklaringar*. Vi måste ge dessa robotars uppförande psykologiska och sociala förklaringar som är parallella med våra förklaringar av människor, och med *samma ord* som vi använder om människor. Det jag säger nu gäller inte bara ord utan också den icke-verbala inställningen, inklusive den känslomässiga attityden, till robotarna.

Den *nödvändighet* med vilken allt det här måste gälla är *praktisk*. Det är inte praktiskt möjligt för dem som är medspelare i dessa sociala situationer att förhålla sig till robotarna som till elektroniska eller datalogiska system. Det senare kräver en distansering som förutsätter att man ställer sig utanför dessa situationer.

Allt det jag nu påpekat kan kallas "*parallellsatsen*". Den säger alltså något om dessa robotars "mänskliga nivå" som jag skrev om i föregående avsnitt. Det är inte mycket nytt i detta. Det är en generalisering av Daniel Dennetts teori om "intentional stance".¹⁰ Jag skall inte gå närmare in på detta, bara nämna en sak.

Dennett höll sig huvudsakligen till *intentionala* språkliga uttryck och förklaringar. Det gäller uttryck som "avse att", "önska att", "planera för", etc., alltså uttryck som anger ett subjekts inriktning mot ett objekt. Sådana uttryck följer andra logiska lagar än de som gäller för termerna på naturvetenskaplig nivå, varför intentionala uttryck inte kan översättas till naturvetenskapliga. Ändå är det, påpekade Dennett, mycket praktiskt att tala i intentionala termer även om en robot. Men det innebär inte att den har någon mystisk dubbelnatur. Vi har bara att välja att tala olika språk om en och samma verklighet. På samma sätt kan man tänka sig förhållandet mellan den "mänskliga" nivån i allmänhet och den naturvetenskapliga nivån.

Den viktiga frågan med det här tankeexperimentet är om man kan *förstärka parallellen till identitet*, alltså om man kan förstärka parallellsatsen till OM*. Alltså om de psykologiska orden om robotarna inte bara är paralleller till samma ord om människor utan har *samma innebörd* som de senare.

Wittgenstein hävdade en gång att det är alldeles meningslöst att tala om ett psyke hos en maskin – han menade att det vore lika meningslöst som att tala om färgen hos ett tal. Men det stora problemet med dessa robotar är att det verkar vara så svårt att *låta bli* att tala om dem som bärare av psyken. Det tycks som om Wittgenstein begick det misstag som han ofta anklagade andra för, nämligen det att hålla sig till alltför ensidiga exempel.

Vad är det egentligen för skillnad i psykologiskt avseende mellan dessa

¹⁰Dennett 1987.

robotar och oss. Vad är det egentligen som *fattas* hos dessa robotar? *Medvetandet* är naturligtvis problemet. Men vi måste använda samma vanliga psykologiska ord, och i samma sociala situationer, även om det *inre*, även om *upplevelserna* hos dessa robotar. Ja, vi skulle även bli tvungna att använda ordet ”smärta” om robotarna. Men problemet gäller inte användningen av dessa ord.

När kan det finnas en skillnad? Man kan räkna upp många fall som är relevanta. En viktig typ av fall är de *moraliska*. Kan man verkligen vara elak mot en sådan här robot? Det är elakt att tillfoga en människa smärta. Men skulle det vara elakt att göra det mot en robot när man måste använda ordet ”smärta” om den? Det tycks vara stor moralisk skillnad mellan att en robot *beter sig som om* den upplever smärta och att den *verkligen* upplever det. Visst kan man tänka sig en slags moral som gör det ”elakt” att bete sig på vissa sätt mot sådana här robotar i de fall som jag nu nämnt. I själva verket skulle det växa fram en sådan moral under de sociala sciencefiction-förhållanden som jag nu tankeexperimenterar om. Men det är inte säkert att en sådan morals problem är *allvarliga*. Skillnaden *konventionell moral/allvarlig moral* tycks bestämmas av frågan om dessa robotar verkligen har medvetande. Skulle en djupsinnig behavioristiskt lagd filosof kunna förneka det? Vad kunde Wittgenstein, som var behavioristiskt lagd och både djupsinnig och allvarlig, ha sagt? Jag vet inte. Antagligen skulle han ha sagt något om livsformer.

8. HUR DET ÄR ATT VARA EN ROBOT

Jag skall nu ta itu med problemet på ett annat sätt: hur skulle det vara om vi kunde *bli* sådana här robotar? *Medvetandet* i den filosofiskt relevanta mening som det nu gäller avser *hur det är att vara* ett subjekt, inte hur subjektet betar sig, och inte hur det talas om subjektet, varken av detta själv eller av andra. Medvetandet i denna bemärkelse är för det första ett *fenomenologiskt* begrepp. Det handlar om hur världen och subjektets inre *ter sig, finns till*, för subjektet. För det andra finns detta medvetande *i världen*, dvs. det är inte *endast* en ”transcendental förutsättning” för världen¹¹ – det är knutet till en kropp, närmast till en hjärna. Detta är det klassiska medvetandebegreppet som varit ett tema i västerländsk filosofi sedan början av 1600-talet. Det har under samma tid varit knutet till det s.k. psykofysiska problemet, dvs. *hur* detta medvetande är förbundet med materien, närmast materien i hjärnan.

Emellertid har det från 1900-talet varit en filosofisk stridsfråga om det alls finns något medvetande i denna bemärkelse. Denna strid är i

¹¹Tyvärr har jag inte här utrymme att förklara vad det sistnämnda betyder. Jag får hänvisa intresserade läsare till Husserl 2004.

och för sig viktig, men jag skall inte här gå in närmare på dess turer. Ett par av de djupaste angreppen mot begreppet kommer nog från Husserl i hans sena artiklar om vetenskapens kris och från Wittgensteins senare filosofi. Jag skriver ”nog” eftersom det inte är lätt att tolka dessa filosofer. En bra allmän begreppsutredning av och försvar för det klassiska medvetandebegreppet finns i början av Nagel 1993. Jag har tagit upp i stort sett samma idé och kort relaterat den till ”transcendental” filosofi av Husserls typ i min artikel i *FT* 2003.

Nå, nu skall vi återvända till Turings test och göra ännu en generalisering av det. Vi skall generalisera det till *hjärnprotes*-problemet: det pågår nu en utveckling av *neurala proteser* – konstgjorda anordningar som sätts in i nervsystemet och skall ta över dess funktioner. En *hjärnprotes* är en neural protes för någon del av hjärnan. Om man vill skapa en *hjärnprotes* för någon del av hjärnan som är förbunden med medvetandet måste man ta itu med det psykofysiska problemet och därmed med det klassiska medvetandebegreppet. Detta kommer helt säkert att inträffa om man vill skapa en protes för *hela* hjärnan. Låt oss alltså tänka oss in i ett sådant fall. Hur skulle man göra om man då drar ut konsekvenserna från Turings tankar i artikeln från 1950?

Vi antar att det gäller en man, Kalle. Han har av något skäl, rimligen sjukdom, anledning att byta ut sin hjärna mot en konstgjord protes. Kalle lägger på ett naturligt villkor på protesen: hans liv skall fortsätta efter operationen i mänsklig form och anknyta kontinuerligt till livet före. Han skall vara samma gamla Kalle som förut efter operationen, åtminstone till att börja med. Han skall vara densamma både för andra och för sig själv.

Då prövar vi Turings idéer på detta fall. En gudom som är hårdvaratekniker, programmerare och kirurg ger ett förslag på protes: den skall vara en *dator*, med *elektronisk* hårdvara, alltså mycket olik en hjärna i konkreta avseenden. Gudomen erbjuder sig vidare att skapa ett *simuleringsprogram* för Kalles hela beteende – som det var strax före operationen – och att ladda ned detta program i datorn, sedan ersätta Kalles hjärna med denna och slutligen göra alla ihopkopplingar till utgående och ingående nerver. I det väsentliga betyder det här att Kalle *blir* en sådan robot som jag skrivit om ovan. Att han har kvar sin biologiska kropp utanför hjärnan är oväsentligt. Skall Kalle gå med på det? Vi kan för enkelhets skull anta att det verkligen går att skapa simuleringsprogrammet. ATT* är alltså uppfyllt. Nu hänger allt på OM*.

Protesen klarar Turings allra mest generaliserade test. Den lurar alla, dvs. alla utom Kalle själv. Ingen som känner Kalle, men inte vet att han blivit opererad, kan misstänka att han blivit det. Och parallelltesen blir uppfyllt: alla (utom Kalle själv) förhåller sig till Kalle efter operationen

som till en vanlig människa, både språkligt och icke-språkligt. Och Kalle själv talar naturligtvis om sig själv som en vanlig medveten människa. Är inte allt frid och fröjd då? Nej, *naturligtvis* inte. Kalle bör förkasta denna protes!

Medvetandet kan identifieras med det totala, mest inklusiva, fenomenologiska begreppet: *livet i fenomenell mening*. För Kalle är det viktiga att hans *liv* skall fortsätta efter operationen, dvs. livet så som det är för honom när han lever det. Tycker du läsare att det finns något oklart med detta begrepp? Skärp dig då! Det ju är helt klart *vad* det är som Kalle kräver att det skall fortsätter efter operationen. Vi vet alla vad livet är i denna mening. Det är det där vanliga, vardagliga – ingenting annat. Det kan inte begreppsligt identifieras med livet som exakt bestämd biologisk process. Detta blir ju mycket förändrat av operationen, och det stora problemet är hur mycket man kan ändra det utan att livet ändras.

Enligt det här Turingska protesförslaget skall alltså livet i fenomenell mening identifieras med (ev. någon aspekt av) datorns mjukvara. Alltså med en *abstrakt strukturell egenskap* hos datorns hårdvara. Man kan då invända att livet *inte är abstrakt utan konkret* – och det går att precisera.

Men man kan påpeka något mycket enklare: livet är inte någon *egen-skap* alls (men det *har* naturligtvis egenskaper). Livet är ett *självständigt föremål*, eller kanske man kan säga ”*substans*” i någon av ordets filosofiska meningar. Det där låter kanske väldigt konstigt. Livet är inte någon sådan vanlig pryl som det finns så många av i vårt konsumtions-samhälle. Vad jag tänker på är ett självständigt föremål *i filosofisk mening*.

Ett sådant kan för det första inte prediceras om något och för det andra är det inte för sin existens väsensmässigt beroende av att något annat existerar. Egenskaper kan prediceras om något. Man kan predicera egenskapen ”röd” om en boll, bollen *är* röd, men bollen själv kan inte prediceras om något. I *denna* mening kan inte heller ett fenomenellt liv prediceras om något. Det här gäller inte användningen av ord i vanlig mening, utan *fenomenologi*. Vårt språk är ganska missvisande när det gäller denna sak. Vi skall dock inte hänga upp oss så mycket på detta utan se på det andra som utmärker ett självständigt föremål.

Här sitter jag nu framför min dator och skriver. Det är en liten bit av mitt liv. Jag ser datorn, känner tangenterna, är inne i skrivandet som aktivitet, skymtar mitt vardagsrum, känner min kropp inifrån, har vissa tankar och känslor, osv. Alla dessa ord måste nu ges fenomenell innebörd, de handlar alltså om *hur det är för mig*. Nu är klockan 20.30. Ungefär så här fortsätter mitt liv fram till klockan 21. Då är följande *filosofiskt* möjligt: När klockan blivit 21 dyker det upp en granne som hojtar: ”Å, så fantastisk att du är tillbaka Tom! Din kropp har varit obefintlig, har inte existerat, mellan 20.30 och 21! Detta har vi kunnat bevisa med vetenskapliga

instrument”. Då måste jag svara: ”Det var konstigt. Allt har varit som vanligt för mig under denna halvtimme. Jag har suttit och skrivit på min dator och ingenting särskilt har hänt!”

Jag skulle bli förvånad om något sådant skulle inträffa i verkligheten (men ibland blir man ju förvånad!). Det skulle strida mot mina föreställningar om naturen. Jag menar bara att det här är möjligt utifrån mitt fenomenella livs väsen, som visar sig för mig. Det betyder emellertid att mitt fenomenella liv är ett självständigt föremål, eller en substans. Likadant måste det vara för Kalle. Därför bör han förkasta den Turingska protesen som förutsätter att medvetandet är en (abstrakt) egenskap. Och därför är OM* falsk. Och följaktligen är Turings idéer i artikeln från 1950 falska.

Nu måste jag sluta denna artikel, men nu börjar de stora frågorna.

LITTERATUR

- Asplund, Johan. 2002. *Genom huvudet: Problemlösningens socialpsykologi*. Göteborg: Korpen.
- Davis, Martin. 2010. ”The Myth of Hypercomputation”. Ingår i Teuscher 2010.
- Dennett, Daniel. 1987. *The Intentional Stance*. Cambridge, Mass.: MIT Press.
- Eriksson, Tom. 2003. ”Om det psykofysiska problemet”. *Filosofisk tidskrift* nr 3.
- Husserl, Edmund. 2004. *Idéer*. Stockholm: Thales.
- Nagel, Thomas. 1993. *Utsikten från ingenstans*. Nora: Nya Doxa.
- Penrose, Roger. 1990. *The Emperor’s New Mind: Concerning Computers, Minds, and the Laws of Physics*. Oxford: Oxford University Press.
- Stenlund, Sören. 1990. *Language and Philosophical Problems*. London: Routledge.
- Syropoulos, Apostolos. 2008. *Hypercomputation*. New York: Springer.
- Teuscher, Christof, red. 2010. *Alan Turing: Life and Legacy of a Great Thinker*. Heidelberg: Springer.
- Turing, Alan M. 1936. ”On Computable Numbers, with an Application to the Entscheidungsproblem”. *Proceedings of the London Mathematical Society*. Ser. 2, Vol. 42, pt. 3–4 (Nov.–Dec. 1936). Finns på nätet.
- Turing, Alan M. 1950. ”Computing Machinery and Intelligence”. *Mind* 59, nr 236 (oktober), 433–60.

Medvetandefilosofins förbistring

1. INLEDNING

I sin bok *Medvetandets mekanik* (2004) skriver Tim Crane: ”Medvetenhetsdebatten är tydligen full av terminologisk förvirring” (s. 206). Jag skulle hellre kalla det ”förbistring” än ”förvirring”. Det kan mycket väl vara så att enskilda författare tycker sig ha reda i sina termer, men att olika författare inte förstår varandras terminologi fullt ut. Förbistringen gäller inte bara termerna, utan även begreppen. Redan i det här stycket ser vi ett exempel. Orden ”medvetande” och ”medvetenhet” används vanligen, men troligen inte alltid, i skilda betydelser. Dessa betydelser varierar från författare till författare och deras innebörd klagörs ofta inte.

I denna artikel vill jag ge ytterligare några exempel samt säga något om orsaker och konsekvenser. Det ser ut att finnas en huvudorsak, som alla de övriga orsakerna kan återföras på.

Jag ger dessutom några förslag till ändringar av termer och begrepp. Om dessa förslag inte vinner gillande, kanske de ändå stimulerar till diskussioner, som leder fram till bättre förslag.

Artikeln avslutas med några påpekanden om villospår som engelskspråkiga filosofer och psykologer lett in medvetandefilosofin på.

2. ORSAKER OCH KONSEKVENSER

Inom ett språkområde är orden och deras betydelser en gemensam egendom. Orden definieras vanligen inte utan lärs in i samband med att de används. Vuxna hjälper barnen att växa in i språket, dels genom att använda det, dels genom att kommentera olika ords betydelser, framförallt orden för de vanligaste objekten i omgivningen. Objekt i den näraliggande, yttre verkligheten är lätta att peka ut: stol, bord, boll, osv. Alla ord har en mer eller mindre vag innebörd – begreppen flyter in i varandra. En sorts boll kunde lika gärna kallas ”kula”. En sorts ask kunde alternativt benämnas ”låda” eller ”kartong”. Men vagheten i sådana konkreta begrepp medför sällan att människor får svårt att förstå varandra.

Vetenskapernas terminologi bygger vidare på allmänspråket. Fysiken talar om stötar, krafter, vågor osv., som ursprungligen härrör från allmänspråket, men som i fysiken har en mer precis innebörd. Motsvarande gäller för alla vetenskaper.

Många av de empiriska vetenskaperna är avknoppade från filosofin. De empiriska ämnesområden som idag är kvar i filosofin är sådana, där begreppen ännu inte hunnit bli så entydiga och skarpa att de kunnat bilda vetenskaper.

Medvetandefilosofin gränsar till ett område, kognitionsvetenskapen, som är en mycket ny vetenskap med beröringspunkter till bl.a. psykologi och artificiell intelligens (AI). Men inom den rena medvetandefilosofin är begreppen fortfarande inte färdigbildade. Allmänspråket är ju främst grundat på den yttre verkligheten, den som är gemensam för oss alla. Men ju längre vi tränger in i våra sjäsliv, desto mer privata blir våra erfarenheter, och desto svårare blir det att enas om relevanta begrepp och termer.

Denna svårighet avspeglas också i nationalspråkens utveckling. Till och med så närbesläktade språk som de inom den germanska språkgruppen, har utvecklats olika när det gäller termer för de mest privata företeelserna. Engelskans "mind" har ingen given motsvarighet i svenskan. Den engelska originaltiteln på Cranes bok är *The Mechanical Mind*. Översättaren hade här velat översätta "the mind" med "sinnet" (have in mind = ha i sinnet), men förläggaren gick inte med på det, eftersom "sinne" också har andra betydelser. Ordet "consciousness" blev på flera ställen översatt med "medvetenhet".

Engelskan har inget ord för upplevelse. Ordet "experience" översätts ibland med "erfarenhet", ibland med "upplevelse". En engelskspråkig person måste därför ha svårt att särskilja dessa två företeelser.

Dessa och andra liknande förhållanden gör att en bok på engelska om medvetandefilosofi inte går att ge en trogen översättning till svenska (och vice versa). Engelsk- och svenskspråkiga personer tänker med nödvändighet på olika sätt i detta ämne. En terminologi för ämnet, som är översättbar mellan de två språken, är än så länge utom räckhåll.

Även inom varje språkområde för sig är det långt till en enhetlig terminologi för medvetandefilosofin. Svårigheterna inom svenskan och engelskan är av olika slag, eftersom tänkandet i hög grad styrs av allmänspråket.

En debattör skrev nyligen att filosofer är mycket slarvigare med definitioner än naturvetarna. Jag tror att det är en feltolkning. Det är lätt att formulera definitioner inom ett område, där det redan finns skarpa begrepp att hänga upp definitionerna på. Helt andra är förutsättningarna inom ett område där man endast vagt anar innebörden i de till synes re-

levanta begreppen. Ska man komma närmare ett klargörande, både inför sig själv och andra, måste man ta till alla pedagogiska knep man kan hitta på: analogier, metaforer, suggestioner, . . ., allt för att leda in tankarna på fruktbara spår. När en disciplin fortfarande är på det stadiet, beror det inte på slarv om man använder termer som inte är strikt definierade. Det man måste göra i stället är att pröva sig fram, försöksvis artikulera och kommentera begrepp och deras relationer, i hopp om att successivt nå allt större klarhet.

Den nu rådande förbistringen hämmar ämnets utveckling, något som inte kan avhjälpas utan ett målmedvetet arbete. Den här artikeln är ett litet bidrag i det arbetet.

3. JAGFUNKTIONER

Ordet ”jag” är ursprungligen ett pronomen, men används i åtskillig filosofisk och psykologisk litteratur som substantiv.

Användningen av ”jag” som pronomen är oproblematiserad och föranleder sällan eller aldrig missförstånd. Pronomina ”jag”, ”du”, ”hon” etc. används för att markera individer och används som substitut för deras egennamn eller för demonstrativa pronomen eller andra utpekandeformer. Ett personligt pronomen avser *hela* individen.

När ”jag” används som substantiv (”mitt jag”, ”jaget”, osv.) avser det inte hela individen utan någon viss del av, aspekt på eller funktion hos individens psyke. Kant, Freud och många andra har introducerat sina egna jagbegrepp, vilka av allt att döma skiljer sig från varandra. För Freud tycks jaget främst vara en aktör inom individen, som står i vissa relationer till ”överjaget” och ”detet”. Hos somliga författare har jaget liknats vid en kapten, som står för de övergripande besluten inom en individ (men som delegerar många beslut till de underställda). Frågan, om en sådan central beslutsfattare inom individen verkligen existerar, är kontroversiell. Peter Gärdenfors (2000) skriver: ”Jag anser att det är ett misstag att anta ett centralt ’jag’ som en styrman för hjärnans processer och därmed för vad kroppen gör.” Åke Daun (2009) ansluter sig till den uppfattningen och menar att jaget är en illusion.

Ett annat jagbegrepp, som förekommer i litteraturen, kommer till följande uttryck hos Anders Jeffner i hans bidrag i boken *Om själen* (Ellegård 1994): ”Det femte drag som jag skall uppmärksamma i den spontana verklighetsupplevelsen är kanske mindre trivialt, eftersom de filosofiska tvivlen här är mer allmänt kända. Ändå är det lika uppenbart. Det är att vi har ett upplevelsecentrum som äger bestånd under en tidsperiod, dvs. ett jag.” Ordet ”upplevelsecentrum” är kanske inte helt välfunnet. Det för tankarna på hjärnans centra, vilket Jeffner knappast kan ha avsett.

Troligen menade han att det inom oss måste finnas neurala system eller strukturer som svarar för vår upplevelseförmåga.

Av tvivlarna tänkte Jeffner förmodligen särskilt på Hume. Denne noterade att han hade en rad enskilda förnimmelser, men att han inte kunde finna något ”jag”. Det är dock inte alldeles lätt att veta vilket jagbegrepp Hume laborerade med. Om det var samma som jag tolkar in hos Jeffner, blir Humes ståndpunkt så orimlig att man kan tro att själva språket ledde honom vilse. Hume kan i varje fall inte ha menat att vi saknar system som ger oss upplevelseförmåga. Det är tänkbart att Hume, som ju saknade ord för upplevelse, inte kunde formulera tankar om upplevelser. Hume har emellertid stöd hos Susan Blackmore (2010). Hon anser att upplevelser inte kräver någon som upplever något. Individer har emellertid upplevelseförmåga, och jag kallar tills vidare det eller de system, som ger oss den förmågan, ”det upplevande jaget”, UJ. Ingen människa, i varje fall ingen med normal upplevelseförmåga, kan betvivla att UJ i den meningen existerar. Vad jag tror att Daun egentligen menar är att vår föreställning, att UJ fattar beslut, är en illusion.

I varje fall finns det beslutsfattande funktioner inom oss. Till dess karaktären hos dessa funktioner är helt utredd är det inte orimligt att sammanfatta dem som ett agerande jag. Problemet, sådant det nu ter sig, är hur de två jagfunktionerna hänger ihop. Är de en och densamma? Om inte: hur förhåller de sig till varandra? Hur medverkar UJ i individens beslutsprocesser?

Det skulle behövas en teori som sätter in alla olika jagbegrepp i ett sammanhang. En monografi – kanske en doktorsavhandling – med en inventering av i litteraturen förekommande jagbegrepp vore ett steg på vägen. Jag anser dock att det är en olägenhet att ”jag” används som substantiv. Jag efterlyser ord som inte alls är sammansatta med ”jag”.

4. UPPLEVELSE, KVALIA

I medvetandedebatten används flitigt ord som ”uppleva” och ”upplevelse”. Ordet ”kvale” (plur. ”kvalia”) används (som jag tolkar det) för att beteckna komponenter i en upplevelse. I sin artikel i *Om själen* (Ellegård 1994) skriver Gunnar Svensson: ”Den latinska termen *qualia* är avsedd att beteckna *hur* det är att ha en viss upplevelse, dvs. just dess medvetna karaktär.” Här kopplas medvetande och upplevelse ihop, vilket kan synas naturligt. Men arten av denna koppling tar sig olika uttryck hos olika författare. I Wikipedia påstås rakt av: ”Medvetande och upplevelse är samma sak.” Detta är förmodligen en vanlig uppfattning.

Som Svensson uttrycker det, är det medvetna ett karakteristiskt drag hos en upplevelse. Hans Rosing (2009) visar en omvänd syn på saken:

”Medvetandet består av en mängd delupplevelser.” En ofta förekommande fras i medvetandefilosofin är ”medveten upplevelse”. Crane skriver så här (2004, s. 16): ”Men en fullständig beskrivning av det mekaniska medvetandet måste också säga något om problemet med medvetna upplevelser.” En svensk författare skulle knappast skriva så om han inte menade att det fanns omedvetna upplevelser också. Men frasen bör kanske inte övertolkas. På just det stället i boken var den en översättning av ”consciousness”. På andra ställen översätts som nämnts consciousness med ”medvetenhet”. Översättaren kunde inte översätta ”consciousness” med ”medvetande”, ty det ordet använder hon genomgående som översättning av ”mind”.

Min förmodan är att Crane, i brist på ett engelskt ord för upplevelse, använder ordet ”consciousness” i stället. Hade översättaren skrivit ”upplevelse” eller ”upplevande” i stället för ”medvetenhet” i följande två meningar (2004, s. 202 resp. 203), så hade de blivit lättare att förstå: ”Medvetenheten är samtidigt både den mest uppenbara egenskapen i det mentala livet och en av de svåraste att definiera eller karakterisera” och ”Vi är helt på det klara med skillnaden mellan medveten och icke-medveten i vardagslivet, men när vi frågar oss: *’Vad är medvetenhet?’* så står vi svarslösa”.

Vi som inte har engelska som modersmål har naturligtvis svårt att påverka engelskt språkbruk. Men det vore kanske möjligt att få någon engelskspråkig medvetandefilosof att inse behovet av engelska ord för uppleva och upplevelse. En sådan person kunde kanske föreslå engelska termer att användas åtminstone i texter som behandlar medvetandefilosofi.

5. MEDVETANDE

De flesta är säkert överens om att orden ”medvetande” och ”upplevande” har närliggande betydelser. I det här avsnittet ger jag ett förslag om deras förhållande till varandra.

Vad är medvetande? Någon regelrätt definition har jag inte funnit i någon av de texter jag läst. Men Dennett och Ellegård försöker sig på karakteriseringar. I *Om själen* skriver Ellegård så här: ”Medvetandet uppstår när djurets överlevnad främjas av att det kan mer i detalj jämföra aktuella erfarenheter med tidigare minnen, och det dessutom blir väsentligt att etablera erfarenheternas och minnenas tidsordning.” Den uppfattningen om medvetandets natur överensstämmer väl med vad jag kan utläsa från Dennett. I *Att förstå medvetandet* (1999) skriver han: ”Ett medvetande är i grunden något som förutser, som skapar förväntningar. Det gräver efter ledtrådar i nuet som det sedan bearbetar med hjälp av

sparat material från det förgångna. Det omvandlar ledtrådarna till förväntningar inför framtiden.”

Medvetandet är, enligt det synsättet, rätt och slätt ett informations-behandlingssystem i överlevandets tjänst. Men så som Dennett och Ellegård framlägger saken, har medvetandet inget nödvändigt samband med upplevelser. Sådana kontroll- och styrsystem som de avser, kan man mycket väl tänka sig hos en maskin eller varelse som saknar förmåga att uppleva. Vidare tyder all psykologisk erfarenhet på att människan dessutom har omedvetna funktioner som deltar i de processer som Dennett och Ellegård talar om. Deras beskrivningar passar alltså in på både medvetna och omedvetna (undermedvetna) kognitiva processer.

Hos Susan Blackmore möter man en helt annan syn på medvetandet. Hon påpekar (2010, s.14) att det inte finns någon allmänt accepterad definition av medvetande, men ger en karakterisering som bygger på fenomenella upplevelser – kvalia. Det tycks som om hon (i den svenska översättningen) i praktiken identifierar medvetande och upplevande, även om hon inte uttryckligen säger det. Om jag har rätt på den punkten, distanserar hon sig från Dennett, som på flera ställen i sina arbeten visar att han inte vill kännas vid tillvarons upplevandedimension. Han raljerar med läsare, som vänt sig till honom i förvåning över att upplevandedimensionen saknas i hans texter. I boken *Consciousness Explained* (1991) finns ett helt kapitel där han söker utmönstra kvalia ur medvetandedebatten.

Hos Dennett saknar jag alltså upplevandeaspekten på medvetandet, medan jag hos Blackmore saknar den funktionella aspekt som både Dennett och Ellegård lyfter fram. Jag önskar en definition som innehåller båda dessa aspekter.

Försöksvis föreslår jag att de kognitiva processer, som normala människor har upplevelse av, får vara dem vi kallar medvetna.

Jag understryker att detta inte är mer än ett försök i brist på en bättre definition av medvetande. Efter en utveckling av teorin, och medföljande precisering av begreppen, kanske det visar sig att gränsen mellan medvetna och omedvetna kognitiva processer inte exakt överensstämmer med gränsen mellan upplevda och oupplevda, i varje fall inte hos alla individer.

Det mesta som vi upplever är annat än kognitiva processer. Framförallt är det sinnesintryck och känslor. Av allt att döma har vi många upplevelser av den arten, som inte når det kognitiva systemet, i varje fall inte omedelbart. För att bli medvetna måste de nå det kognitiva systemet. Ser vi det så, kan vi alltså ha såväl medvetna som omedvetna upplevelser. Den som betvivlar detta kan tänka på hur vi beter oss när vi vill vara helt fokuserade på en upplevelse. För att riktigt njuta av ett vackert musik-

stycke, sitter vi gärna bekvämt avslappnade med slutna ögon för att inte störas av andra intryck, och vi försöker att hålla det kognitiva systemet så passivt som möjligt – ”leva i nuet”. Medvetande och upplevande är alltså två företeelser, som ofta hänger samman, men inte alltid gör det.

6. FILOSOFINS ZOMBIER

Existerar det varelser som har medvetande men saknar upplevelseförmåga? Med aldrig så litet fantasi kan man i varje fall föreställa sig det. Man kan till och med tänka sig en människa, som utåt sett fungerar normalt, och som tycks besitta normalt medvetande, men som likväl inte har någon förmåga att uppleva. För en sådan hypotetisk varelse har filosoferna (oegentligt) lånat ordet ”zombie”. Crane ger följande definition, som jag inte anser är lyckad: “. . . en zombie är en fysisk kopia av en medveten varelse, men den är inte medveten”. Kanske menar Crane att denna varelse saknar *medvetenhet* (tolkat som upplevelseförmåga), inte att den saknar *medvetande*. Men måste inte en i alla detaljer trogen kopia av en vanlig människa fungera precis som denna vanliga människa? I definitionen är dock ordet ”fysisk” väsentligt. Upplevelseförmågan kan ju tänkas ha ett icke-fysiskt ursprung. Ja, Crane drar faktiskt slutsatsen att ”om zombier är möjliga, så är fysikalismen falsk”. Den slutsatsen beror dock på Cranes val av definition. Hade en zombie definierats som en varelse med medvetande men utan upplevelseförmåga, så hade Cranes slutsats inte följt. Upplevelseförmågan kan ju ha sin grund i ett rent fysiologiskt system (det vi ovan betecknade UJ), som hos vissa individer kan vara så defekt att det inte alls ger några upplevelser, fastän det kognitiva systemet (i övrigt) är helt intakt.

Oavsett om zombier existerar eller inte, ja oavsett om zombier är möjliga eller inte, har idén om zombier ett värde som pedagogiskt hjälpmedel och i teoretiska diskussioner. En zombie tänks ha samma sorts sinnen som normala människor och samma system för bearbetning av sinnesdata. Därmed får zombien också samma sorts input från den yttre verkligheten. Zombien har smärtreceptorer och bearbetar smärtsignaler på samma sätt som andra, säger ”aj” och skriker och ger alltså samma uttryck för smärta som andra. Utåt sett ger han intryck av att uppleva smärta, men gör det ändå inte. Han upplever alls inga kvalia. Han har samma typer av känsloliv som andra: kan oro sig, känna ilska etc., ger adekvat uttryck för sina känslor, men upplever dem inte.

Idén om zombier tycks mig vara lätt att förstå. Ändå har jag läst texter vars författare haft svårt att ta till sig den idén. Frågan gäller då inte om zombier existerar, eller över huvudtaget kan existera, utan det är själva begreppet som ställs i tvivelsmål. Susan Blackmore tycker helt enkelt att

idén om zombier är en fånig tanke. De som delar hennes uppfattning tänker sig förmodligen att en zombie skulle sakna medvetande i Dennett-Ellegårds mening. Om jag lade den innebörden i "zombie" så skulle jag ansluta jag mig till Blackmore. Zombie-begreppet blir intressant bara om man, när man använder "medvetande" i definitionen, enbart avser fenomenellt medvetande, alltså förmåga att uppleva. Men Blackmore tycks mena att det funktionella och det fenomenella medvetandet är ouplösligt förknippade med varandra.

Idén om zombier kan emellertid föras vidare till följande allmänna frågeställning: *Vilka skillnader mellan olika människors upplevande-funktioner finns?* Mänsklig kommunikation bygger på gemensam erfarenhet. Orden i ett språk svarar mot företeelser som alla, eller i varje fall många, i språkområdet har gemensam erfarenhet av. Men blott till en viss gräns. Vi kan tala om t.ex. smaker och färger, därför att (nästan) alla har erfart smaker och färger. Men vi kan inte veta om våra upplevelser av dessa smaker och färger är lika. Det finns en ogenomtränglig barriär mellan två människor, som gör att de aldrig kan ha säkrare kunskap än gissningar om varandras upplevelser. De eventuella olikheterna mellan människors upplevelser är ibland svåra att upptäcka, ibland lättare. Exempel: innan vi fick testmetoder för färgsinnet, var det säkert många med röd-grön färgblindhet som inte visste om sin defekt.

Förmodligen finns det en rad olikheter mellan människors förmåga att uppleva, som ännu inte är upptäckta. Om det finns människor, som inte alls upplever något, men i övrigt fungerar helt normalt, så skulle det möjligen vara svårt att upptäcka. En sådan människa skulle ha normala reaktioner på stimuli utan att för den skull uppleva något. Kanske man samtalsvägen skulle upptäcka att en sådan människa är avvikande. Skulle en sådan människa kunna ta till sig begreppet kvalia? Inte så säkert.

7. HUR UPPSTÅR KVALIA?

I filosofernas spekulationer kring medvetandet har två huvudriktningar kristalliserat ut sig: *fysikalismen* och *dualismen*. Fysikalismen kan också kallas *materialistisk monism*. (En monistisk syn innebär att all verklighet ytterst vilar på en enda substans e.d.) Fysikalismen antar att all verklighet vilar på fysisk substans (materia och energi). Dualismen, som den traditionellt har uppfattats, är hypotesen att den yttersta verkligheten härrör från två olika sorts substanser, dels materiell, dels andlig.

En fysikalist förfäktar sålunda att även förmågan att uppleva har ett materiellt underlag, medan en traditionell dualist menar att man måste anta en andlig substans för att kunna förstå kvaliafenomenen. Av dessa två grundhypoteser är det få filosofer och vetenskapsmän som idag be-

känner sig till dualismen. Fysikalismen är den ståndpunkt som för närvarande är på modet. Det kan ta sig sådana uttryck som i följande uttalande av Dennett: ”Dualismen, tanken att medvetanden består av ett icke-fysiskt och ytterst gåtfullt material, och vitalismen, som anser att levande ting innehåller ett speciellt men lika gåtfullt material som kallas *élan vital*, har förvisats till historiens sophög – tillsammans med alkemin och astrologin.”

Det här är ett ganska typiskt exempel på Dennettsk polemik. I stället för att ange sina skäl mot dualismen jämför han den med ståndpunkter som inte längre har några seriösa förespråkare. Han jämför också med andra föråldrade uppfattningar som den att jorden är platt och att solen är en eldvagn, dragen av bevingade hästar. Alla dessa jämförelser saknar rim och reson, utom den med vitalismen. Den har följande poäng, som antagligen gått Dennett förbi:

På det stadium, där fysiken befann sig på 1800-talet, gav den inga möjligheter att förstå livsprocesserna. De insiktsfulla biologerna insåg att det behövdes något, ännu okänt, för en vetenskaplig förståelse av livsprocesserna på molkylärnivå. Vad detta okända kunde vara, spekulerade man naturligtvis över. Vitalismen är en samlingsbenämning på ståndpunkter som postulerar att naturen innehåller särskilda livskrafter som kan ge förklaringen. År 1900 kom första steget på lösningen i och med Max Plancks kvantteori. Den satte fart på utvecklingen av den nutida atomteorin och den därmed sammanhängande kvantmekaniken. Den kastar ljus över livsprocesserna på ett sätt som man givetvis inte kunde ana på 1800-talet. Vitalismen hade rätt i att det behövdes någon faktor som under den epoken var okänd, men hade fel i antagandet att denna faktor var en specifik livskraft. (Kvantmekanikens lagar är universella; de gäller levande materia såväl som icke levande.)

Problemet med att förstå förmågan att uppleva befinner sig nu (i början av 2000-talet) i en situation som i viss mån påminner om molekylärbiologins situation på 1800-talet. Det finns dock en väsentlig skillnad. På 1800-talet *insåg* man att den dåvarande fysiken inte räckte till för att förklara livsprocesserna. När det gäller upplevelseförmågan, finns det visserligen argument som tyder på att den nuvarande fysiken inte räcker till, men de argumenten är än så länge inte bindande. Frågan måste därför betraktas som helt öppen: Kan upplevelseförmågan förklaras med den nuvarande fysiken som grund, eller måste fysiken kompletteras med något, än så länge okänt, för att denna förmåga ska kunna förklaras/förstås?

Fysikalismen är ståndpunkten att den nuvarande fysiken räcker. Vad ska vi kalla ståndpunkten att något utöver den nuvarande fysiken behövs? Det är inte orimligt att kalla den ”dualismen”. Därmed får dualis-

men en vidare innebörd än den klassiska. Det vi har vunnit med detta är att fysikalismen och dualismen tillsammans uttömmar alla möjligheter, och att de motsäger varandra, varmed frågan om kvalias natur blivit mer renodlad. Man kan uppfatta de två ismerna som vetenskapliga hypoteser, mellan vilka vi söker ett avgörande. Något avgörande test kommer man kanske aldrig att finna. Men om man lyckas finna en trovärdig förklaring av kvalia utifrån den nuvarande fysiken, så har man i varje fall fått ett indicium för fysikalismen. Kan man i stället göra troligt att fysikalismen leder till orimliga konsekvenser, så skulle det tala för dualismen.

8. VILLOSPÅR I UTVECKLINGEN

Utöver det som nämdes i avsnitt 2 finns en rad historiskt betingade orsaker till den terminologiska och begreppsliga förbistringens inom medvetandefilosofin. En sådan är att psykologin är en ung vetenskap som ännu inte övervunnit de blindspår den halkat in på. När jag säger detta tänker många kanske på psykoanalysen. Freud ledde in psykologin på flera vägar som har fått lämnas, men hans negativa inflytande har alltmer övervunnits, medan hans många fruktbara idéer har överlevt.

Det största och mest genomgripande villospåret i psykologi och medvetandefilosofi synes mig vara *behaviorismen*. Då menar jag inte den experimentella behaviorismen (Pavlov, Skinner m.fl.) utan den paradigmatiska, för att inte säga dogmatiska, behaviorismen. Psykologer och filosofer, vars inflytande mer berodde på personlighet och polemisk förmåga än vetenskaplig skicklighet, kom att dominera tänkandet i årtionden. Jag tänker på personer som t.ex. Watson och Ryle. Deras inflytande blev en hämsko på medvetandefilosofi och kognitionsvetenskap genom att tona ner eller rentav förneka medvetandet och andra inre mekanismer. Intresset riktades mot stimulus och respons, mot det yttre beteendet, medan de processer som ledde från stimulus till respons uppfattades som ointressanta eller avfärdades som "spekulation". Ryle hörde till den s.k. Oxfordskolan, där man trodde att de filosofiska problemen kunde lösas upp genom språklig analys.

Min bedömning är att även Dennett är inne på villospår. Dennett är en arvtogare till behaviorismen, men han har gått sina egna vägar och verkligen intresserat sig för de inre processerna. Det han ärvt från behaviorismen är förnekandet, eller i varje fall nedtoningen av upplevelser, av kvalia. Han identifierar medvetandet med de processer han försöker rekonstruera. Han har gott om uppslag, han är verbalt skicklig och en god pedagog. Men hans argumentation i sak är ofta svag. När han svarar på frågor eller bemöter kritik, och inte kan ge sakskaäl, tar han gärna till polemiska grepp och drar sig inte för att håna de frågande. Han framträ-

der med stor självsäkerhet. Han verkar tro så fullt och fast på sina idéer att han framställer sina oprövade hypoteser som om de vore säkerställda fakta. Många låter sig duperas av sådant.

Förekomsten av upplevande har kallats *medvetandefilosofins svåra problem*. Man kan inte närma sig en lösning på det problemet genom att låtsas som om det inte finns. Det kan inte heller lösas genom polemiska grepp. Men det kanske är möjligt att närma sig en lösning via en in-trängande och stringent debatt. Ett nödvändigt steg på den vägen är att ämnets företrädare försöker enas om en gemensam begreppsapparat. För det behövs dels kreativa förslag, dels ett generöst, men för den skull inte okritiskt, debattklimat.

LITTERATUR

- Blackmore, Susan. 2010. *Kort om medvetandet*. Lidingö: Fri Tanke Förlag.
- Crane, Tim. 2004. *Medvetandets mekanik: En filosofisk introduktion till medvetanden, maskiner och mental representation*. Stockholm: Thales.
- Daun, Åke. 2009. *Arv och miljö*. Symposion.
- Dennett, Daniel. 1999. *Att förstå medvetandet hos människor och andra djur*. Stockholm: Natur och Kultur.
- Dennett, Daniel. 1991. *Consciousness Explained*. Amazon.
- Ellegård, Alvar, red. 1994. *Om själen*. Stockholm: Natur och Kultur.
- Gärdenfors, Peter. 2000. *Hur Homo blev sapiens: Om tänkandets evolution*. Nora: Nya Doxa.
- Rosing, Hans. 2009. *Vad är jag?* <http://hansrosing.blogspot.com/2009/09/vad-ar-jag-en-filosofisk-undersokning.html>

Recension

Filosoferna: Det västerländska tänkandet sedan år 1900

Svante Nordin

Bokförlaget Atlantis 2011. 851 sidor. ISBN 978-91-7353-480-2

Det här är en riktig tegelsten om 851 välmatade sidor. Författaren uppvisar en imponerande lärdom och han skriver – till skillnad från flera av de filosofer han behandlar – friskt och medryckande. Ibland riktigt flott.

Bokens titel är lite förbryllande. Den ger intrycket att filosoferna är de enda som har bidragit till det västerländska tänkandet sedan år 1900. Nog måste man väl ändå säga att t.ex. naturvetare, författare, ekonomer och andra humanister och samhällsvetare också har tänkt en del. Men det visar sig att många sådana också omnämns och kommenteras, även om de kanske inte får lika stor uppmärksamhet som filosoferna. Dostojevskij, Einstein, Freud, Joyce, T. S. Eliot, Thomas Mann, Eddington, Hayek, Heisenberg, Allan Bloom, Niels Bohr, J. M. Keynes, Joseph Ratzinger (sedermera påven Benedictus XVI), Francis Fukuyama, Lenin och Max Weber finns exempelvis med. Å andra sidan nämns inflytelserika filosofer som C. G. Hempel, Frank Jackson, Jürgen Mittelstrass, Peter Singer, Crispin Wright, Günther Patzig, J. J. C. Smart, Patrick Suppes, Timothy Williamson och Daniel Dennett (nästan) inte alls. Ekonomipristagarna Kenneth Arrow, Herbert Simon, John Harsanyi, Amartya Sen och Thomas Schelling saknas t.ex. också, trots att de delvis befinner sig rätt nära filosofiämnet.

Om det finns andra än filosofer som har bidragit till det västerländska tänkandet beror förstås på vad filosofi är för något. Som en enkel definition föreslår Nordin att filosofi är ”en reflexion över människan och världen i traditionen från Platon” (s. 796). Med denna definition kan det ju tyckas som om nästan alla vetenskaper faller inom filosofins område. Men de flesta som nämns hos Nordin är förstås ändå filosofer i mer akademisk mening.

Det kan ligga nära till hands att jämföra Nordins bok med den danska

och nästan lika långa *Vår tids filosofi* (redaktör Poul Lübcke) från 1982, som utkom i svensk översättning 1987 på förlaget Forum och som enligt sitt förord ”går igenom de viktigaste riktningarna inom amerikansk och europeisk filosofi under 1900-talet”. Den danska boken hade tolv olika författare och den kunde naturligtvis inte säga något om filosofin efter 1982. Den första halvan behandlar tysk och fransk filosofi, den andra ägnas anglosaxisk filosofi. Nordin skriver betydligt mer heltäckande, t.ex. även om spansk, italiensk, rysk, östeuropeisk och nordisk filosofi. Naturligtvis blir framställningen av enskilda filosofers tänkande då mer kortfattad. Husserl får t.ex. fyrtio sammanhängande sidor hos Lübcke mot endast lite över tio hos Nordin (men därutöver nämns han förstås i många andra sammanhang i bägge böckerna). Den tyska Erlangenskolan ägnas ett helt eget kapitel om tjugofem sidor hos Lübcke, men får bara sex rader hos Nordin. Quine får tio sidor hos Lübcke, fem hos Nordin. Dummett får sexton sidor hos Lübcke, tre hos Nordin. Å andra sidan: Nelson Goodman nämns inte alls hos Lübcke, men får flera sidor hos Nordin. Nordin har mer än dubbelt så många namn i sitt personregister som Lübcke, över tusen mot knappt femhundra.

Om man behandlar ett så oerhört omfattande material, så vore det väldigt konstigt om inte en del detaljer blev fel. Men jag har inte hittat så mycket att invända mot – vilket kanske säger mer om mig än om boken. (Willard Van Orman Quine heter dock ”Willard Van Orman Quine”, inte ”Willard van Orman Quine”, som Nordin skriver. Och mängdteoretikern A. A. Fraenkel heter ”Fraenkel”, inte ”Frankel”. Davidson kallas på något ställe ”Donaldson”.)

Mer otillfredsställande är t.ex. Nordins framställning av Goodmans idéer om kontrafaktiska villkorssatser. Där säger han sådant som att Goodman menar att dispositioner inte existerar och att ”ingenting är möjligt utom det som faktiskt inträffar”. Detta är åtminstone vilseledande. Och när han ska redogöra för den goodmanska egenskapen ”grue” (”blön”) säger han att den utmärker alla föremål som undersökts fram till en tidpunkt *t* och är gröna, men dessutom ”alla föremål efter tidpunkten *t* som är blå” (s. 474). Detta är direkt fel och dessutom knappast begripligt. Vad menas med ”alla föremål efter *t*”? Är det de som existerar efter *t*, eller de som endast existerar efter *t*, eller de som undersöks efter *t*, eller de som endast undersöks efter *t*? Med tanke på att sanningen hos vissa kontrafaktiska villkorssatser utgör ”en grundbult i hela vår världsbild”, som Nordin helt riktigt uttrycker det, så är detta olyckligt.

Det blir också lite underligt när Nordin ska redogöra för Quines berömda kritik av distinktionen mellan analytiska och syntetiska satser. Han noterar att distinktionen kan förklaras om vi kan avgöra när olika uttryck är synonyma. Men hur gör vi det? Nordin skriver: ”Skall vi hänvisa till

det allmänna språkbruket? Men förvandlar vi inte då de analytiska sanningarna till empiriska sanningar a posteriori?” (s. 470). Nej, de analytiska sanningarna blir inte empiriska bara för att man eventuellt med empiriska metoder skulle kunna avgöra vilka uttryck som är synonyma. Däremot skulle det då bli empiriskt avgörbart vilka satsers som *är* analytiska. Det skulle förstås empiristen Quine välkomna; det är bara det att just detta inte går. Det är det som är Quines poäng.

Nordins beskrivning av Quines tes om översättningens obestämdhet (eller ”underbestämdhet”, som Nordin säger) är också missvisande. Han säger att ”Quine utgick från en strikt behavioristisk syn på språk och ord” (s. 471). Men Quine *utgick* inte från en sådan syn, han argumenterade för att man *måste* göra det, eftersom man lär sig språk endast genom att observera andras verbala beteende och att det därför inte kan finnas något i språklig mening utöver vad som kan upptäckas på behavioristisk grund. När det gäller observationsatsers kan man förhållandevis lätt hitta lyckade översättningar; det är betydligt svårare med andra satsers. Här krävs mycket gissande. Kriteriet på lyckade översättningar är att de medger flytande konversation och en viss förutsägbarhet, dvs. att översättningen fungerar. Men obestämdhetstesen säger att flera olika översättningsmanualer kan vara lika lyckade, trots att de är oförenliga. Detta har inte så mycket med Quines kända exempel med ordet ”Gavagai” att göra, vilket Nordin hänvisar till; det exemplet gäller snarare *referensens* obestämdhet, vilket är något helt annat.

På tal om översättning säger Nordin också att satsen ”A är god” enligt moralfilosofen Charles Stevensons uppfattning skulle kunna översättas till ”Jag tycker om A, gör det också du!”. Värdesatser skulle alltså vara en sorts uppmaningar. Men detta är inte riktigt Stevensons åsikt. Det Nordin beskriver är bara en ”arbetsmodell”. Stevenson introducerar i stället ett pragmatiskt meningsbegrepp och han menar att värdesatser har ”emotiv mening”, ofta kombinerat med olika slag av ”deskriptiv mening”, som inte nödvändigtvis har med den talandes känsloläge att göra.

Nordin säger också på tal om Davidsons teori om handlingsförklaringar, att ”handlingens avsiktlighet och rationalitet hos Davidson inte spelar någon roll för förklaringen” och att en förklaring med hänvisning till den handlandes skäl eller avsikter ”betraktas som kausalt verkningslös” (s. 602–3). Detta stämmer inte. Tvärtom vill Davidson visa att det inte, som många trott, finns någon motsättning mellan rationella förklaringar och orsaksförklaringar. Den handlandes skäl kan både rationalisera och orsaka handlingen.

Ett påfallande drag i västerländsk filosofi under de senaste hundra åren är sprickan mellan analytisk (anglosaxisk) och kontinental filosofi, som uppstod någon gång kring andra världskriget. Nordin säger en hel del om

detta. Delvis har det väl att göra med att den analytiska filosofin först fick fäste i England (Moore, Russell, Broad, Wittgenstein) och att dess tidiga anhängare på kontinenten (de logiska empiristerna) tvingades flytta till England och USA när Hitler kom till makten. Man kan möjligen undra varför franska filosofer efter andra världskriget så starkt influerats av tysk filosofi (Marx, Nietzsche, Freud, Husserl, Heidegger) och nästan inte alls av anglosaxisk. I Tyskland finns numera många analytiska filosofer; Nordin rapporterar att de tyska filosofernas Gesellschaft für Analytische Philosophie år 2005 hade 800 medlemmar. Men franska filosofer har i allmänhet varit ointresserade av analytisk filosofi; ett undantag var Jacques Bouvresse, som därför stämplades som "ofransk". (Ett annat undantag var Jules Vuillemin, som inte nämns av Nordin.) Analytiska filosofer tyckte på motsvarande sätt illa om fransk filosofi och försökte bl.a. 1992 hindra att Derrida utsågs till hedersdoktor i Cambridge; i ett öppet brev till universitetet skrev de att i den mån Derridas påståenden alls var begripliga, så var de antingen falska eller triviala. Däremot har fransk filosofi haft visst genomslag i USA – men då huvudsakligen inom andra humanistiska ämnen, inte på filosofiinstitutionerna.

Nordins bok har en idéhistorisk snarare än filosofisk karaktär. Med det menar jag att den lägger förhållandevis mer vikt vid idéförändringar, skillnader och influenser än vid filosofisk kvalitet och förståelse. Det yttar sig bl.a. i att idéer som fått stort inflytande utanför filosofin ägnas mer uppmärksamhet – så att t.ex. Spengler, Jünger och Carl Schmitt ägnas tio sidor, ungefär lika mycket som Quine och Goodman. Beskrivningarna av de filosofiska idéerna blir också mer kortfattade. Ibland blir de därför också rätt svårbegripliga (om den italienske filosofen Giorgio Agamben sägs t.ex. utan vidare förklaringar: "I sprickan mellan rösten och språket vill Agamben finna möjligheter till en ny mänsklighet"; jag måste erkänna att jag inte alls begriper vad det betyder). Nordins disposition är också huvudsakligen kronologisk och geografisk. Lübckes bok får nog sägas ha en mer filosofisk karaktär.

Men även ur filosofisk synpunkt är Nordins framställning intressant och lärorik; den är också allmänbildande och förhållandevis lättläst. För egen del skulle jag dock gärna sett att den – liksom Lübckes bok – hade innehållit några bilder på de viktigaste filosoferna. Det bestående intrycket hos läsaren – åtminstone denna läsare – är dock en häpnad och beundran inför Nordins beläsenhet.

Notiser

PARADOXALT

Man brukar väl kunna säga att negationen av en falsk sats är sann. Men jag kom att tänka på en intressant paradox. Betrakta följande sats:

Den här satsen innehåller sju ord.

Satsen är ju falsk, i och med att den endast innehåller sex ord. Vi tar nu och negerar satsen enligt följande:

Den här satsen innehåller inte sju ord.

Den satsen är ju också falsk, i och med att den nu faktiskt innehåller sju ord. Jag tyckte detta var lite lustigt. . .

Hans Gabrielson

ETT AVSLÖJANDE SKÄMT

I sin senaste bok *On What Matters* (andra bandet, s. 612–20), diskuterar Derek Parfit pessimism eller optimism om framtiden. Han frågar: Kommer framtiden, i stort sett, vara bra? Den främste bland pessimister, Arthur Schopenhauer, har enligt Parfit en inkonsistent uppfattning om detta, då Schopenhauer anser att våra liv innehåller sådan enorm mängd lidande att det är värre än om man inte fanns alls och att vi snart kommer att vara döda, eftersom tiden går så fort. Att tro detta påminner Parfit om skämtet om två judar som äter på en restaurang de aldrig har varit tidigare på. Den ena säger efteråt: ”Maten var inte god.” Den andre svarar: ”Och portionerna var för små.”

Kanske inkohärensen blir tydligare via följande förtydligande: Om framtiden ser väldigt mörk ut – varför är det då inte bättre att ta självmord nu? David Benatars svar till Parfit, i hans bok med (den nedslående?) titeln *Better to Never Have Been* (s. 212), blir att vi inte har något skämt om samma ordväxling mellan judarna skulle inträffa i Auschwitz. Skämtet är bara roligt om vi förutsätter att de äter för nöjes skull eller om portionerna var tillräckligt stora.

Kan skämtet avslöja ett argument mot pessimisten? Pessimisten menar att vi redan är i en av de värsta av världar, så skämtet faller platt.

Finns det en neutral utgångspunkt som kan tillfogas skämtet? En utgångspunkt som är analog med att få tillräckliga portioner eller äta för nöjes skull? Finns det ändå inte något konstigt i att säga att framtiden ser väldigt mörk ut, men att jag fortfarande vill möta den? (Om vi antar, eventuellt för stora, förutsättningar; att självmord på rationella grunder inte är fel eller finns och att döden, som Epikuros tycks ha menat, inte innebär något dåligt tillstånd för oss eftersom vi inte finns då.)

När jag går på restaurang så har det hänt, inte ofta, att jag bett om lite extra mat. Självklart var jag hungrig då och maten var god. Var jag den naiva optimisten? Ska den koherenta pessimisten be om mer även fast det smakat illa och han är mätt? Den rationella optimisten, om han har fått en för liten portion av god mat, ska hoppas att nästa gång kyparen kommer förbi, om han nu kommer över huvud taget, så kommer optimisten få en påfyllning gratis? Som världen ser ut nu så verkar det högst rimligt att det inte finns några gratis luncher eller att påtåren eller påbackningen inte är inbakad i priset.

Parfit avslutar nämnda avsnittet med att säga att pessimisten ser på människans natur som fixerad. I en snar framtid så kommer vi mer och mer komma kunna förändra vår omvärld – kanske t.o.m. oss själva, så att vi blir något utöver människorna – han kallar det supra-människor. Kanske hoppas Parfit på att deras psyke inte innebär upplevelserna att tiden går fort eller att de inte längre upplever lidande. Är det genom att modifiera det som Schopenhauer såg som problematiskt – att vi alltid önskar oss något nytt om vi får det vi önskar oss eller att vi blir uttråkade efter att vi har uppnått ett mål?

Hur som helst kan inte supra-människan innebära att de blir modifierade till att kunna bli mätta även fast de har fått så lite som möjligt serverat till sig – oavsett om det är gott eller inte. En stor del av mänsklighetens historia har redan varit så. Vi får inte heller glömma att portionerna tycks bli mindre i framtiden eftersom de ska mätta ett allt större flertal (människor och eventuellt Parfits supra-människor). Kanske supra-människor blir euforiskt mätta, överlever, genom att bara suga på en tandpetare. Men finns det inte något illusoriskt över detta? Det verkar som om Parfit tror att supra-människorna ska sluta äta på restaurangen och i stället börja arbeta där. Men vilka ska då bli gästerna? Antingen får vi människor ta notan eller så lägger restaurangen ner eftersom det inte finns några gäster längre. Det verkar inte som om pessimisten kan avslöjas av ett skämt.

Rickard Vester Smetana

LITTERATUR

Benatar, David. 2006. *Better Never to Have Been: The Harm of Coming into Existence*. Oxford: Clarendon.

Parfit, Derek. 2011. *On What Matters*, Vol. 2. Oxford: Oxford University Press.

FILOSOFISK POLITIK

Torbjörn Tännsjö påstår (DN 10/6 2012) att den svenska lagstiftningen är inkonsekvent när den å ena sidan tillåter boxning men å andra sidan förbjuder sexköp och koppleri. I bägge fallen är det ju fråga om verksamheter där man riskerar att skada andra människor. Han anser att man bör eliminera inkonsekvensen och han lutar åt att man bör göra det genom att förbjuda boxningen.

Men det är ju inte alls klart att den svenska lagstiftningen *är* inkonsekvent när den tillåter boxning men förbjuder sexköp och koppleri. Det är ju inte fråga om samma sorts skadeverkningar i de bägge fallen. Man kan ju mycket väl vara emot vissa skador men acceptera andra.

Den svenska lagstiftningen tillåter ju inte bara boxning, utan också sådant som t.ex. ishockey, fotboll, biltrafik och bostadsbyggande. Även här är medverkan någorlunda frivillig och skador förekommer. Lutar Tännsjö åt att även vilja förbjuda dessa verksamheter?

Hur som helst, det man ur filosofisk synpunkt kan undra över är vad ”inkonsekvens” ska betyda i sammanhang som dessa. Det är ju inte fråga om någon verklig motsägelse. Snarare är man ”inkonsekvent” om man gör olika bedömning av företeelser, som inte är olika i något *relevant* avseende. Men vilka olikheter är ”relevanta”? Det är frågan. De som vill förbjuda sexköp och tillåta boxning kan säkert peka på någon olikhet mellan dessa verksamheter som i deras ögon är relevant. Till exempel att sexköp, men inte boxning, är förnedrande för kvinnor.

Men Tännsjö anser tydligen inte att det finns någon sådan relevant skillnad. Annars kunde man ju tro att han, som är hedonistisk utilitarist, borde lägga vikt vid mängden av lust och olust som blir resultatet av dessa verksamheter. Men om sexköp ger mer olust än lust och boxning ger mer lust än olust – vilket kanske är ganska troligt, om än svårt att bevisa; man kan i alla händelser luta åt att tro det – så är detta en relevant skillnad enligt Tännsjö. Och det är då ingen inkonsekvens, ens ur *hans* perspektiv, att förbjuda det förra men inte det senare.

I själva verket borde eventuella inkonsekvenser inte spela någon som helst roll för utilitarister som Tännsjö. För att säga det lite slagordsmässigt: det är ju *konsekvenser*, inte *inkonsekvenser*, som är relevanta för utilitarister.

När det gäller lagstiftning ska ju beslut tas i demokratisk ordning, det menar även Tännsjö. Men från vilket perspektiv ska man då avgöra vad som är ”relevant”? Så länge den svenska författningen inte anger att hedonistisk utilitarism eller någon annan moralfilosofisk doktrin ska vara avgörande, så kan man tycka att var och en bör utgå från sin egen uppfattning eller från det man kan komma fram till i en allmän diskussion.

Men här finns förstas en komplikation: när man vill bedöma vad en person *bör* göra i en demokratisk beslutsprocess, så kan olika moral-

filosofiska doktriner ge olika svar. Utilitaristen säger väl att var och en bör fatta beslut om vad som är relevant på det sätt som ger de bästa konsekvenserna (mest lust över olust), men andra kan säga att man bör utgå från sin egen uppfattning eller sitt eget samvete eller vad man råkar luta åt. Eller från någon rättighetskatalog. Eller från något kategoriskt imperativ. Och så vidare.

SVAR

Min diskussion rörde frågan om man ska vara tillåten att skada andra (allvarligt). Som jag förstått saken förbjuder vi köp av sexuella tjänster, då sådana köp antas skada säljaren. Men då borde vi förbjuda boxning, då boxning innebär att utövarna skadar varandra. Det är ju t.o.m. möjligt att vinna en match genom att slå motståndaren medvetlös (ja ihjäl). En helt annan sak är verksamheter, där människor riskerar att komma till skada. Där väger vi nyttan av verksamheten mot onyttan hos de skador som faktiskt uppkommer. Vi försöker också reducera dessa skador på olika vis. Det är sant att det förekommer inom fotboll att en spelare miss-handlar en annan, ungefär som i boxning. Det är emellertid ett regelbrott. Man kan inte vinna en fotbollsmatch genom att slå en motståndare medvetlös. För sådant utmäts straff. Är förseelsen grov nog leder den till straffrättsliga påföljder. Men den som på ett regelrätt sätt slår ihjäl sin motståndare i en boxningsring döms inte till någon påföljd över huvud taget – han eller hon tilldöms i stället segern i matchen.

Torbjörn Tännsjö

LOGIK I TEOLOGINS TJÄNST

Är det möjligt att vara både allsmäktig och allvetande? Kan en allsmäktig varelse låta bli att göra det hen vet att hen kommer att göra? Det kan förefalla omöjligt. Men egentligen är det väl inte så svårt. Vad som är omöjligt är att *göra* det man inte gör. Men man behöver nog inte ens vara allsmäktig för att *kunna* göra det man inte gör.

A PRIORI ENLIGT WIKIPEDIA

Enligt Wikipedia är satsen ”Alla greker är människor” sann a priori, vilket sägs innebära att ”Satsens sanningsvärde är enbart beroende av de ingående begreppens betydelser, och har ingenting att göra med om det existerar människor eller greker”. Men är inte även de grekiska gudarna greker? De är ju grekiska. Och de är väl inte människor. Om de existerar eller inte ska ju inte ha med saken att göra. Här är det något som inte stämmer!

LYCKLIGA LÄNDER

De tio lyckligaste länderna i världen – med avseende på sådant som tillfredsställelse, inkomst, hälsa, god miljö, fritid och liknande – är enligt en rätt färsk undersökning följande:

1. Danmark,
2. Kanada,
3. Norge,
4. Australien,
5. Nederländerna,
6. Sverige,
7. Schweiz,
8. Finland,
9. Israel,
10. Österrike.

Ska man verkligen tro på detta? Man skulle vilja veta hur mätningarna har gått till. Kanske har man bara frågat folk hur lyckliga de är.

IDÉHISTORISK LYCKOFORSKNING

Under rubriken ”En depressiv samtid” skriver idéhistorikern Karin Johannisson bland annat följande: ”Uppenbart finns ingen relation mellan välfärd och upplevt välbefinnande” (*Axess* nr 4, 2012). Kan det verkligen stämma? Om man funderar över detta måste man först skilja mellan två olika tolkningar av ”upplevt välbefinnande”. Dels kan det syfta på vad man *säger* om hur man mår, dels kan det syfta på hur man *mår*. Det första är ju relativt lätt att undersöka, det andra är betydligt svårare. Johannisson erkänner att det finns ett ”hermeneutiskt problem att påstå att individen mår sämre i en viss tid än i en annan”, men hon anser ändå att ”förskjutningar är möjliga att studera via diagnosstatistik, patientjournaler, offentlig debatt och medial exponering av olika känslotillstånd”. Det är dock svårt att tro att sådana studier skulle kunna fastställa läget hos (ett representativt urval av) en befolkning, vare sig beträffande hur den mår eller hur den skulle säga sig må.

Att folk inte blir lyckligare av att välfärden ökar vore kanske inte så underligt – något i den stilen menade för övrigt Tage Erlander i sitt tal om ”de stigande förväntningarnas missnöje” – men detta är i alla händelser fullt förenligt med att folk skulle bli *olyckligare*, och också *säga* att de mår sämre, om välfärden skulle minska (t.ex. genom att ersättningsnivåerna i olika försäkringssystem sänks).

Kanske är det också så att välbefinnandet successivt skulle *minska* om välfärden inte hela tiden *ökade*. I så fall skulle det finnas ett skäl att fortsätta att förbättra välfärden, trots att välbefinnandet inte därmed ökar.

Och om detta vore fallet, så skulle det ju finnas en viktig ”relation mellan välfärd och upplevt välbefinnande”, vilket just är det som Johannisson förnekar. Omvänt skulle vi kanske inte ha något skäl att förbättra välfärden, om hon har rätt. Eller finns det något annat skäl, som inte har med välfärd att göra?

SVENSK EXPORT

The American Philosophical Association har delat ut sitt pris för bästa filosofiska artikel under de senaste två åren till den svenska filosofen Anna-Sara Malmgren, som numera är anställd vid Stanford University. Priset delas ut vartannat år till en yngre forskare och är på 2000 dollar. Uppsatsen är ”Rationalism and the Content of Intuitive Judgements”, *Mind* 120 (2011), s. 263–327.

OM FASTIGHETSSKÖTARE

Jag börjar med att läsa *FT* bakifrån, som jag som barn alltid slog upp tidningarnas sista sidor med serier. Nu är det notiser och recensioner som får mig till det.

I nummer 3/12 finner jag ”Ofta bra att lita på sinnen”, skrivet av Tom Eriksson under just *Notiser*. Det handlar om en fastighetsskötare som uppenbarligen inte vill det.

Kan det vara så att just fastighetsskötare generellt inte gör det, tänker jag? De kanske har ett metafysiskt ideal där allt fungerar som de vill att det skall göra? Jag har en liknande erfarenhet som Toms.

Jag ringer upp fastighetsskötaren och säger att min porttelefon inte fungerar. Han svarar: Du har ingen porttelefon. Jag blir stum, vet inte vad jag ska svara på det, men hämtar mig och insisterar; det har jag visst! Han lägger från sig luren och återkommer snart och säger då att han frågat dom andra (fastighetsskötarna) och dom säger att det inte finns någon porttelefon.

Kan du inte gå och titta, försöker jag (vädjande till ett hopp om sinnesnärvaro)? Det tycker han inte behövs.

John Falkenberg

Medarbetare i detta nummer av *Filosofisk tidskrift*: *Anders Berglund* är administratör vid institutionen för idé- och samhällsstudier, Umeå universitet, *Lars Bergström* är professor emeritus i praktisk filosofi, *Tom Eriksson* är doktorand i teoretisk filosofi i Stockholm, *Anders Lundberg* är professor emeritus i matematik vid Försvarshögskolan och *Torbjörn Tännsjö* är professor i praktisk filosofi i Stockholm.

Instruktioner till skribenter

Filosofisk tidskrift har som syfte att bidra till en allsidig och fruktbar diskussion av filosofiska problem, samt att på ett lättfattligt sätt informera om aktuell filosofisk forskning. Den vänder sig inte enbart till fackfilosofer, utan vill framför allt nå en bredare läsekrets av filosofiskt intresserade personer

Tidskriften står öppen för olika filosofiska inriktningar, men den vill undvika bidrag som man inte kan tillgodogöra sig utan speciella förkunskaper eller tekniska färdigheter. Utöver längre artiklar på omkring 10–15 sidor, tar tidskriften gärna emot även kortare bidrag och inlägg av notiskaraktär.

Manuskript till *Filosofisk tidskrift*:

- skickas normalt med datorpost till lars.bergstrom@philosophy.su.se
- kan också skickas till
Lars Bergström, Telegrafgränd 1 B, 111 30 Stockholm
– men i så fall skall det dessutom bifogas på diskett
- skall utformas i överensstämmelse med den typografi som är normal i tidskriften, utan onödig formatering
- skall vara försedda med namn och adress
- skall som regel vara skrivna på svenska, och citat från andra språk bör översättas till svenska
- särskild litteraturförteckning upprättas vid behov i alfabetisk ordning och placeras sist i manus
- för icke beställt material ansvaras ej
- korrektur läses i regel endast av redaktören
- ändringar mot manuskript bekostas icke av förlaget; författaren debiteras kostnad för sådana ändringar
- införda bidrag honoreras inte för närvarande
- i stället för särtryck erhåller författaren, gratis, 10 exemplar – för recensioner 5 exemplar – av det nummer av tidskriften i vilket bidraget varit infört