

Recension

Derek Parfit: *Reasons and Persons*, Clarendon Press, Oxford, 1984.

Oxfordfilosofen Derek Parfits sedan länge förebådade bok *Reasons and Persons* utgör ett otvetydigt belägg för att Parfit är en av de mest idérika filosoferna i samtiden. Boken innehåller ett överdåd av snillrika tankar – ja, så många att Parfit har måst ge avkall på kompositionens krav för att få med dem alla. Det är ingen stor överdrift att karakterisera verket mer som en redovisning av vad Parfit har tänkt än som en tematiskt enhetlig konstruktion.

Bokens huvudsyfte är att kritisera omhuldade och grundläggande föreställningar rörande praktisk rationalitet, moral och personlig identitet. Det är bara ifråga om personlig identitet denna kritik växer ut till en positiv teori. Vad moralen beträffar lämnar Parfit medvetet läsaren i sticket, men även för rationalitetens vidkommande gäller att Parfits positiva förslag framstår som högst fragmentariska. Icke desto mindre är boken fängslande läsning, dels därför att den drivs fram av en aldrig sinande ström briljanta infall, dels därför att ett otal fantasieggande och science-fictionartade illustrationer förlämnar framställningen charm och åskådlighet.

I den första av verkets fyra delar undersöks om vissa allmänna antaganden rörande rationalitet och moral är självupphävande i en begränsad mening. För moralens del består detta bl a i en granskning av den konsekvensetiska tesen att man bör handla så, att resultatet blir så gott som möjligt. Parfits slutsats är att detta syfte motverkas om man låter sitt handlande direkt styras av denna princip. Bättre konsekvenser uppnås om ens beteende i stället regleras av ett detaljerat regelsystem i stil med det sunda förnuftets moral. Parfit kommer här nära R M Hares dubbeldäckarteori i *Moral Thinking* (Oxford, 1981).

Den rationalitetsapproach som skärskådas är egenintresseteorin, S, vilken går ut på att det för var och en finns ett rationellt mål framför andra: att det egna livet förlöper så väl som möjligt. Uttrycket ”så väl som möjligt” kan specificeras på olika vis, t ex i termer av maximal tillfredsställelse av ens viljeattityder livet igenom. Så tolkad blir S ingen egoistisk teori, eftersom det inte utesluts att ens viljeattityder gäller andras välfärd. Parfits argument är här att om en agent ständigt har sitt eget bästa för ögonen, så tenderar resultatet att bli sämre för honom än det annars kunde ha blivit. Personligen är jag inte övertygad av detta resonemang, men även om det hade ägt riktighet, hade denna ”självupphävande” effekt inte varit förödande för S, ty – som Parfit också påpekar – S fordrar inte att ens beteende i livets alla skeenden direkt kontrolleras av egenintressets kalkyler, lika lite som en konsekvensetisk grundsyn hindrar att man för det mesta följer färdiggjorda tumregelsnormer i stället för att hänge sig åt komplicerade konsekvensberäkningar.

S är således inte vederlagd, och det är därför befogat att fortsätta den kritiska granskningen av den i del II, vilken ägnas åt hur tidsaspekten kommer in i rationalitet. Parfits strategi är här att försöka visa hur svårt det är för en anhängare av S att förklara varför man inte skall vara partisk gentemot vissa *tidpunkter* – t ex det innevarande ögonblicket – om man skall vara partisk gentemot en viss *person*, sig själv. Termerna ”jag” och ”nu” är semantiskt likartade, varför då privilegiera den ena? Som en rival till S ställer Parfit upp en teorityp, P, vilken gör gällande att en rationell person bör söka den maximala tillfredsställelsen av sina *nuvarande* viljeattityder. I kontrast till P favoriserar S inte nuvarande attityder på bekostnad av attityder i det förflutna eller framtiden.

Denna temporala neutralitet hos S kan emellertid betyda två skilda saker som Parfit inte håller isär med erforderlig klarhet. Enligt en mer beskedlig uttolkning innebär den (1) att om man hyser två viljeattityder, vilka är identiska i alla avseenden förutom att den första är nuvarande och den andra framtida (eller förfluten), har man inget skäl att föredra stillandet av den första attityden (eller tvärtom). Den andra, mer långtgående utläsningen är (2) att om man hyser två attityder som är av samma styrka, men som har olika objekt, och en är nuvarande och den andra framtida (eller förfluten), har man inget skäl att föredra stillandet av den förra framför den andra (eller tvärtom). Uppenbarligen går (2) utöver (1) i det att den diskvalificerar hänsyn rörande en attityds *objekt* som skäl för att föredra tillgodoseendet av en nuvarande viljeattityd. Det står också utom tvivel att Parfit förstår S så, att den omfattar även (2) (se t ex s 154).

Parfit avlossar (s 130–2, 149–58) förödande invändningar mot tesen (2). Helt klart är det möjligt att en rationell person, som förutser att han kommer att utvecklas i en viss riktning om han inte vidtar mått och steg, kan fördöma denna utveckling som degeneration, göra sitt bästa för att förhindra den och se till, att om han misslyckas med detta, så skall hans förfallna jags vilja inte kunna tillfredsställas. Däremot är de av Parfits argument som berör (1) – d v s hans diskussion huruvida våra spontana böjelser, att bry oss mer om den nära framtiden än den mer avlägsna och mer om framtida välbehag och smärta än om de förflutna, är rationella eller ej (s 158–86) – inkonklusiva, trots all sin genialitet. Detta är också anledningen, tror jag, till att Parfit till slut medger (s 188–9) att ett krav på temporal neutralitet kan infogas i hans version av P, den kritiska versionen, CP. Förmodligen avses här bara neutralitet i betydelsen (1).

Resultatet av del II blir att S är definitivt vederlagd endast om den inkorporerar (2). Men i så fall överdriver Parfit vikten av sin slutsats när han gör anspråk på att ha kullkastat en doktrin som ”de flesta människor har trott på i mer än två årtusenden” (s 194). Det tycks mig som om en så speciell tes som (2) med säkerhet bara kan tillskrivas ett fåtal nutida filosofer. Kanhända har det varit ett stående antagande under ett par årtusenden att människor är i grunden själviska och att det är oförnuftigt att offra långsiktiga större fördelar

till förmån för kortsiktiga mindre. Men detta rör (1) snarare än (2), förutom att det är ett misstag att förväxla egenintresse i S's form med själviskhet. Och även om S viker för P så garanteras inte att altruistiskt beteende är rationellt, såvida inte Parfit lyckas bevisa sin version av CP, enligt vilken altruistisk motivation är *intrinsiskt* rationell – en tanke som han leker med (s 121, 133, 194), men aldrig försöker föra i hamn.

I viss mån fortsätter prövningen av S i del III där ämnet är personlig identitet, för detta har relevans för en bedömning av S-teorins betonande av jaget. Parfits mästerliga behandling av föreställningen om en persons identitet genom tiden innehåller tre huvudteser. (a) Personer är ingenting utöver sina "hjärnor och kroppar och skilda interrelaterade fysiska och mentala händelser" (s 216). Personlig identitet måste följaktligen bestå i psykiska och fysiska processers kontinuitet i tid och rum – Parfit framhäver speciellt den psykiska kontinuiteten. (b) Det finns inte alltid ett icke-godtyckligt svar på frågan om personlig identitet föreligger eller ej. Det vill säga, den psykiska kontinuitet som föreligger mellan två personer, vilka existerar vid olika tidpunkter, kan vara mer eller mindre omfattande – de kan ha fler eller färre gemensamma minnen och karaktärsdrag – och det finns en gränsszon där kontinuiteten varken är tillräckligt stark för att man skall kunna hävda att vederbörande är samma person eller tillräckligt svag för att detta skall kunna förnekas med bestämdhet. (c) "Personlig identitet är inte vad som spelar roll. Det som spelar roll är i grunden . . . /psykisk kontinuitet/ . . . , med vilken orsak som helst" (s 217); fysisk kontinuitet, att samma (delar av en) kropp fortlever, är nästintill betydelselöst för ens känslomässiga inställning till en person. Själva kan jag skriva under på mycket i teserna (a)–(c), men på en kärnpunkt vill jag ifrågasätta Parfits position: är inte psykisk kontinuitet, bevarandet av minnen och karaktärsdrag, ur emotionell synvinkel lika oväsentlig som fysisk kontinuitet? Låt oss emellertid först följa hur Parfit når fram till sin ståndpunkt(c), att det är psykisk kontinuitet snarare än personlig identitet som spelar roll för en rationell persons känslor.

Gör detta antagande: om *corpus callosum*, förbindelselänken mellan hjärnans två hemisfärer, skärs av, är var och en av de två isolerade hjärnhalvorna nog för att vidmakthålla full psykisk kontinuitet. Antag vidare att jag är en individ i en grupp identiska trillingar. I en olycka skadas vitala organ i min bål, men min hjärna förblir oskadd. För mina två bröder är det precis tvärtom: deras hjärnor totalförstörs, men för övrigt är deras kroppar intakta. Neurokirurgin har utvecklats till den grad att man är i stånd till att klyva hjärnor och transplantera delarna. Min hjärna tudelas, och mina bröders kroppar får varsin halva. Föreställ er nu två fortsättningar på den här historien: (1) bara en av hjärnhalvstransplantationerna lyckas eller (2) båda transplantationerna lyckas. Om jag före operationen förutser utfallet (1), skulle jag då inte behandla den resulterande personen som om han var identisk med mig själv (se till att han får tillgång till min egendom etc)? Kvalitativt sett är han ju precis lik mig. Men kanske är den resulterande personen verkligen identisk med mig (vilket

är Parfits åsikt), och i så fall indikerar inte förekomsten av mina omsorger om honom att vad som i realiteten spelar roll är något annat än personlig identitet. Detta uppenbaras först när man skärskådar versionen (2). Här kan jag, menar Parfit (s 260), inte vara identisk med någon resulterande person, för jag kan omöjlig vara identisk med båda, och det vore godtyckligt att identifiera mig med den ena. Icke desto mindre, kan jag annat än att bry mig om dessa två dubbelgångare lika mycket som den ende i (1)? Ingenting skiljer ju dem åt. Därför måste den relation som bär upp nämnda komplex av känslor vara en annan relation än personlig identitet. Förslagsvis psykisk kontinuitet, ty de två personerna i (2) är fullt kontinuerliga med mig i detta hänseende, men fysiskt kontinuerliga med mig bara i begränsad utsträckning (i form av en hjärnhalva).

Parfits uppfattning – att psykisk kontinuitet, men inte fysisk, är väsentlig – ter sig emellertid egendomlig när det går upp för en att ifrågavarande psykiska kontinuitet förutsätter en fysisk. Observera att vi här inte talar om mental kontinuitet i betydelsen en ström av upplevelser utan avbrott, ty i så fall skulle den kärvånliga inställningen till det egna jaget inte kunna överleva perioder av medvetlöshet och drömlös sömn. Nej, vad som här åsyftas är en kontinuitet som utgörs exempelvis av att man minns vad en person vid en tidigare tidpunkt tänkte och varselev. Men att minnas en tidigare upplevelse är inte bara att ha en ”bild” som troget återger denna upplevelse; det innebär också att denna ”bild” står i en viss orsaksrelation till den föregående upplevelsen. Normalt består den kausala bakgrunden av att den första upplevelsen lämnar ett ”spår” i hjärnan, ett spår som existerar kontinuerligt så länge man kan erinra sig upplevelsen. Parfit föreställer sig att psykisk kontinuitet kan bibehållas även om de bakomliggande kausala mekanismerna är andra än de gängse – som jag redan har noterat talar han om psykisk kontinuitet ”med vilken orsak som helst”. T ex tänker han sig en maskin som exakt registrerar tillståndet hos en levande kropps alla celler och neuroner, förstör kroppen och därefter återskapar en perfekt kopia av ny materia. Även detta kausala förlopp, resonerar Parfit, upprätthåller psykisk kontinuitet.

Det sista påståendet är inte självklart sant, men jag är beredd att acceptera det för resonemangets skull. Tesen, att psykisk kontinuitet vilar på fysisk, står ändå oemotsagd, för i detta, liksom i Parfits andra science-fiction exempel, hänvisar orsaksförklaringen till kontinuerliga materiella processer. Om man eliminerar denna referens till fysisk kontinuitet, reducerar man inte då psykisk kontinuitet till enbart en kvalitativ likhet mellan tidigare och senare upplevelser? Jag kan inte annat än se att detta blir följden. Men om nu appellen till fysisk kontinuitet utgör den enda skillnaden mellan psykisk kontinuitet och psykisk likhet, och fysisk kontinuitet är känslomässigt oväsentlig, följer det att på sin höjd psykisk likhet kan ha emotionell signifikans.

Ett tankeexperiment kan belysa kontinuitetens periferia ställning. Antag att en slumpmässig händelse av detta slag äger rum då och då: de elementarpartiklar som konstituerar en persons kropp förskingras, men en bråkdel av en

sekund senare bildar andra elementarpartiklar en konstellation som är en fulländad kopia av denna kropp. Här finns ingen kontinuitet mellan de två kropparna, men ändå kan jag inte förstå hur en rationell person skulle kunna frukta denna händelse eller hysa andra känslor för kopian än för det framtida jag den kopierar. I förlängningen av detta resonemang finns en konsekvens som kan hända är ännu mer omskakande. Föreställ er att resultatet av den slumpmässiga händelsen inte är en kopia av den person man faktisk *är* utan en materialisering av den person man *skulle vilja vara*, av ens ideal- eller dröm-jag – skulle en rationell person då inte hoppas och önska att den slumpmässiga händelsen inträffar och att han ”efterträds” av denna idealperson snarare än att han fortsätter att existera i sin medelmåttighet? Detta vore ett grundskott mot S.

I fjärde och sista delen tar Parfit upp några moraliska problem som uppstår i vårt förhållande till kommande generationer. Man kan tycka att det är självklart att en handling är moraliskt fel endast om den förvärrar situationen för någon. Parfit vill kasta tvivel över denna princip genom följande typ av exempel. Antag att vi har att besluta hur en viss naturtillgång skall exploateras. Om vi hårdexploaterar den, kommer mänskligheten att åtnjuta en behaglig välfärd 300 år framöver, men därefter kommer resursen att vara nästan uttömd, och de därpå följande generationerna kommer att få genomlida en drastisk sänkning av levandstandard – dock inte så drastisk att deras liv inte blir bättre än icke-existens. Med alternativet en mer begränsad exploatering under de närmsta 300 åren åsamkas de då levande generationerna en viss förlust i välbefinnande jämfört med vad hårdexploateringen skulle medföra, men denna policy innebär betydligt bättre levnadsvillkor för de följande generationerna. Det är naturligt att tycka, att om vi väljer hårdexploatering, handlar vi moraliskt förkastligt: för att tillskansa oss en mindre fördel förvärrar vi situationen högst betydligt för dem som kommer att leva om 300 år. Men det finns här ett förbiseende som Parfit sätter fingret på. Det är sannolikt att vårt val av policy får genomgripande effekter på samhället i stort. Tex kommer andra män och kvinnor att träffas om vi verkställer den ena policyn snarare än den andra. Resultatet blir att de personer, som existerar om 300 år om vi väljer hårdexploatering, inte är identiska med dem som skulle existera om vi fastnade för det andra alternativet. Valet av hårdexploatering förvärrar därför inte situationen för någon person, eftersom livet om 300 år fortfarande är bättre än icke-existens.

Jag är inte beredd att sträcka vapen inför detta argument. Om vårt val i dag av hårdexploatering är moraliskt fel, kan det inte vara för dess konsekvenser om 300 år, för dessa konsekvenser inställer sig bara om de mellanliggande generationerna ansluter sig till vår policy, och vårt val förblir fel även om de bryter med den och introducerar en mer restriktiv policy. Anledningen till att vårt beslut är förkastligt tycks snarare vara att det i ett visst avseende förvärrar situationen för *nästa* generation: det krävs större uppoffringar av dem än av oss för att de skall kunna iscensätta en hushållspolicy som är lika effektiv

som den vi hade kunnat initiera en generation tidigare. Om våra barn i sin tur skyggar för detta offer och fortsätter på den inslagna vägen, försämrar de situationen på ett analogt sätt för sina barn o s v. Är detta synsätt – som förklarar policyvalets felaktighet med hänvisning till dåliga konsekvenser för nästa generation – korrekt, glider man undan Parfits problem. Det är nämligen inte rimligt att hävda att valet av policy påverkar identiteten av nästa generation, ty denna har redan fötts vid valögonblicket.

Därmed har jag eliminerat ett av de problem Parfit anser en adekvat moralteori måste komma tillrätta med. Ett annat krav han ställer på en dylik teori är att den skall undvika ”den motbudande konklusion”, vilken går ut på att en ökning av en populations storlek kan kompensera en sänkning av välfärdsgenomsnittet. Pondera att vi har två populationer, A och B, där B är dubbelt så stor som A. Alla varelser i en och samma population åtnjuter samma grad av välfärd: i A är den +100 och i B +60. Det är inte orimligt att säga att B är bättre – eller åtminstone inte sämre – än A. En välkänd form av utilitarism säger just precis detta. Men i så fall bör också en population C, vilken är tre gånger A's storlek och vilken har en välfärdsnivå på +45, vara bättre än A. På så vis erhåller man den motbudande konklusionen att ett enormt stort antal individer med liv som är precis värda att levas bildar en helhet som är bättre än en mindre population med avsevärt bättre levnadsvillkor.

Naturligtvis har det framlagts många teorier som inte leder till denna motbudande slutsats, t ex den variant av utilitarism som vill höja välfärdsgenomsnittet så mycket som möjligt. Men Parfit argumenterar bestickande för att de kan belastas med andra kontraintuitiva konsekvenser. Så kanske man i stället bör fråga sig om den motbudande konklusionen verkligen är motbudande i den bemärkelse som är relevant vid testningen av moralteorier, nämligen *moraliskt* motbudande. Att en sänkning av vars och ens välfärd är motbudande för de drabbade ur ett *egoistiskt* perspektiv är oomtvistligt, men också ointressant. En omfördelning av jordens tillgångar, vilken består i att en stor del av I-ländernas överflöd överflyttas till U-länderna, ter sig givetvis motbudande för I-länderna ur deras själviska synvinkel. Likväl kan åtgärden vara moraliskt motiverad. I ljuset av denna tankegång förefaller det inte alltför långsökt att tänka sig att moralen också kräver, att därest alla existerande varelser lever liv som är mer än väl värda att levas, bör de, om så är möjligt, reproducera sig och fördela sitt värdeöverskott över sin avkomma tills genomsnittslivet blir endast minimalt bättre än icke-existens. Observera att gränsen för moralens krav på uppoffringar ligger vid ett liv som är precis *värt att levas* och inte vid ett liv på existensminimum eller dylikt. Moralens måste tillerkänna en rätten att leva ett liv värt att levas, annars undermineras fundamentala förpliktelser som den att rädda eller bevara liv. Genom att ställa stora fordringar på ett liv värt att levas kan man således hålla moralens pockande på uppoffringar i schack.

Parfit söker en moralteori som dels löser det tidigare nämnda problemet om kommande generationers identitet, dels undviker denna motbjudande konklusion, men tvingas erkänna sitt sökande som förgäves. Om min diskussion är riktig, framstår detta misslyckande som lättbegripligt: Parfit har ställt felaktiga adekvanskrav. Men med tanke på de många intrikata problem och insiktsfulla observationer som hopas i Parfits bok verkar ändå en tillfredsställande moralteori mer avlägsen än någonsin.

Parfit har, sammanfattningsvis, skrivit en utomordentligt stimulerande och uppslagsrik, men något ostrukturerad, bok – en bok som knappast någon med ett seriöst intresse för filosofi kan unna sig att ignorera, ty den kommer säkerligen att för lång tid framåt prägla den anglosaxiska diskussionen av moralteori, rationalitet och personlig identitet.

Ingmar Persson

Notiser

Vid filosofiska institutionen i Lund arrangerades 29 oktober 1985 en Temadag, som syftade till att tala om för omvärlden vilka intressanta saker som filosofer idag håller på med, även ur samhällets synvinkel. Det övergripande temat "Filosofi—Samhälle" ville understryka att filosofin behandlar angelägna problem. Åtta forskare av olika generation presenterade forskning om vitt skilda saker som medicinsk etik, samhällsfilosofi, vetenskapsteori, människokännedom, argumentationsanalys, kvalitetskriterier m m. Inbjudna från HSFR, BFR, näringsliv och kommunala förvaltningsorgan samt allmänheten lyssnade i Carolinasalen från 9—17.00 med en avslutande allmän diskussion om filosofin i samhället. Med över sextio inbjudna deltagare blev Temadagen en framgång, som befäste intrycket att Humaniora idag möter en allt bredare respons. Temadagen har stötts av Crafoordska Stiftelsen, för vilket arrangörerna är mycket tacksamma. Även Rektorsämbetet stödde initiativet på ett mycket aktivt sätt. Syftet är att följa upp med en Proceedings-skrift av mindre format.

Bertil Mårtensson

Stellan Welin har redigerat en antologi med uppsatser av svenska vetenskapsteoretiker: *Att förstå världen*, Doxa 1984.

Tore Strömberg har utgivit *Rättsfilosofins historia*, andra upplagan, Studentlitteratur 1985.