

LARS BERGSTRÖM

Utilitarismens fel och förtjänster

Man kan säga att utilitarismen fått ett icke obetydligt genomslag bland svenska filosofer. Delvis har jag kanske själv bidragit till det. Jag har åtminstone länge tyckt, och tycker fortfarande, att utilitarismen har stora förtjänster. Annars är det väl i Sverige främst Ingemar Hedenius som har gjort propaganda för utilitarismen – och på senare tid Torbjörn Tännsjö.

För drygt tjugo år sedan höll jag ett offentligt föredrag i vilket jag visserligen hävdade att utilitarismen ”måste utarbetas närmare” för att undvika allvarlig kritik av olika slag, men att den ändå är det ”mest lovande” och ”mest acceptabla” av alla de moralsystem som föreslagits under historiens lopp. Idag är jag mer tveksam.

Utilitarismen kan utformas på flera olika sätt. I mitt föredrag hävdade jag att den borde uppfattas som *riktighetskriterium* (inte som beslutsmetod) och som *handlingsutilitarism* (snarare än som regelutilitarism). För övrigt menade jag att den borde innebära att en handling moraliska status bestäms av dess *totala, faktiska* konsekvenser (snarare än av t ex dess förutsägbara eller sannolika konsekvenser). Så uppfattas den också av t ex G E Moore och i Sverige av Torbjörn Tännsjö. Det innebär alltså att en enskild handling är moraliskt riktig om och endast om den leder till minst lika bra konsekvenser som varje annan handling som agenten kunde ha utfört i stället. Det är denna egenskap som gör riktiga handlingar riktiga. Hur man sedan i praktiken skall bära sig åt för att bestämma sig för vad som är riktigt är en annan sak. Det är inte alls säkert att man enligt utilitarismen skall försöka beräkna de olika alternativens konsekvenser och deras relativa värden. Som redan Henry Sidgwick påpekade är det kanske även ur utilitarismens synpunkt bättre att folk i allmänhet följer vissa enkla moralregler om att hålla löften och tala sanning och så vidare. Sådana

konventionella moralregler kan betraktas som en sorts ”tumregler” som man bör lära ut till sina barn och själv efterleva. Därmed blir troligen konsekvenserna bättre än om var och en försöker handla på grundval av noggranna utilitaristiska konsekvensanalyser av varje alternativ.

Tanken är alltså att utilitarismen blir *rimligast* om den tolkas på det sätt jag här har skisserat. Inom denna ram finns sedan en rad andra tolkningsvarianter – se exempelvis Erik Carlsons utmärkta avhandling *Consequentialism Reconsidered* (Kluwer 1995) – som jag dock inte skall trassla in mig i här.

1. Utilitarismens förtjänster

Det enda som ytterst spelar roll är kanske om utilitarismen är *sann*. Detta är dock inte lätt att avgöra. Det syns ju inte direkt på den – lika lite som man direkt kan se på vetenskapliga teorier om de är sanna. Hur skall man då gå tillväga för att värdera utilitarismen som moral-lära? Vilka krav kan vi ställa på den? Vilka förtjänster har den? Vilka nackdelar? Mycket har sagts om detta. Jag skall här sammanfatta det hela, på mitt eget sätt, i fem fördelar och fem nackdelar. Låt oss börja med fördelarna.

Den är moraliskt tilltalande. I framför allt två avseenden är den en moraliskt attraktiv lära. För det första tar den fasta på hur det faktiskt går för dem som berörs av våra handlingar, hur deras välbefinnande påverkas eller hur deras intressen tillgodoses. Därigenom generaliserar den åsikten att man bör behandla andra som man själv skulle vilja bli behandlad. För det andra är den opartisk. Den avvisar alla former av egoism, inklusive gruppegoism, rasegoism och nationalism, och den insisterar på att alla som berörs skall tillmätas samma betydelse (*each to count for one*), oavsett om de är djur eller människor och oavsett om de lever nu eller tillhör kommande generationer.

Den är hittills optimal. Möjligen kan man säga att den än så länge är bättre än de alternativa moralteorier som har föreslagits. Med detta kan man i så fall mena, dels att den är normativt mer tilltalande (även om smaken här är olika), och dels att den är mer teoretiskt avancerad. En vanlig tanke i vetenskapsteori är ju att man inte bör överge en teori, ens om den har allvarliga problem, innan man har hittat en bättre. Något liknande kunde kanske anföras som stöd för utilitarismen. Framför allt är den konfliktlösande, djärv, och teoretiskt fruktbar. Detta räknar jag som de återstående tre fördelarna.

Den är konfliktlösande. Den säger något om hur man skall lösa moraliska konflikter, dvs situationer där olika moraliska överväganden pekar åt olika håll. Detta är kanske det viktigaste kravet på en moralteori, ty en moralteori skall ju lösa moraliska problem, och moraliska problem uppstår ju framför allt i de fall där olika moraliska överväganden pekar åt olika håll. Andra moralteorier, som t ex Kantianism, rättighetsteori, klassisk dygdeetik och kristen moral, lämnar en här i sticket.

Den är djärv. Man kan nog säga att utilitarismen är en "djärv gissning" (*bold conjecture*) i ungefär den mening som Karl Popper avsåg när han hävdade att bra vetenskapliga teorier skall vara djärva gissningar. Andra moralteorier (som t ex Kantianism, kristen etik, existentialism, etc) är mer intetsägande och oprecisa. Därför är det också svårare att säga vad de påbjuder i enskilda fall, åtminstone om dessa enskilda fall är tänkta exempel av det slag som moralfilosofer brukar diskutera. Utilitarismen tycks ge klarare utslag, och därför är den också lättare att kritisera – vilket Popper, i andra sammanhang, ser som en stor förtjänst.

Den är teoretiskt fruktbar. Utilitarismen är teoretiskt fruktbar i den meningen att den ger upphov till en mängd intressanta problem och till mycket diskussion. Det beror nog bl a på att den bättre än andra teorier lämpar sig för applikation på tänkta exempel, samt på att den – åtminstone vid första anblicken – tycks ge upphov till vissa moraliskt obehagliga konsekvenser som det då för moralteoretikern gäller att bortförklara.

2. Utilitarismens nackdelar

Emellertid har utilitarismen även en hel del nackdelar. Dessa påpekas ofta av dess motståndare, men i viss utsträckning erkänns de även av dess anhängare. Jag skall som sagt nämna fem nackdelar.

Den är moraliskt stötande. Den vanligaste typen av invändning mot utilitarismen består av något tankeexperiment som skall visa att den har moraliskt orimliga konsekvenser. Exempelvis att man bör straffa en oskyldig och onyttig person hellre än en som är skyldig men samhällsnyttig, om det går att övertyga allmänheten om att den oskyldige är skyldig. Eller att man med våld bör beröva en frisk person vissa vitala organ, som sedan kan transplanteras in på andra och därmed rädda flera liv, om detta kan göras utan att folks förtroende för sjuk-

vården rubbas. Eller att man bör öka jordens befolkning så länge alla får en dräglig tillvaro hellre än att ha en mindre befolkning där alla har det mycket bättre, men där den totala välfärden är något lägre. Dessutom tycks utilitarismen ofta kunna tillåta sådant som skattefusk, svartjobb, etc. Det är dock något oklart vad dessa exempel visar och hur realistiska de är.

Den är orimligt krävande. Eftersom utilitarismen kräver att jag alltid gör det som har de absolut bästa konsekvenserna, så verkar det rimligt att tro att vi alla alltid handlar moraliskt orätt enligt utilitarismen. Förmodligen kommer vi sällan ens i närheten av det som är rätt. Hellre än att lägga ut sextio kronor på en biobiljett till mig själv bör jag nog enligt utilitarismen ge pengarna till något nödhjälpsprojekt i tredje världen. Det låter i och för sig inte helt fel, men konsekvensen tycks bli att vi aldrig får gå på bio. Det är mindre rimligt. Och samma resonemang tycks kunna visa att vi inte heller får äta oss riktigt mätta eller ge våra barn nya kläder. Detta verkar ganska orimligt.

Den är ofullständig. Utilitarismen förutsätter att det finns objektivt riktiga svar på frågor om hur förbättringar för den ene förhåller sig till försämringar för den andre. Sådana frågor måste ha sanna svar om man på ett meningsfullt sätt skall kunna tala om *sammanlagda* välfärdsvinster eller välfärdsförluster med olika handlingsalternativ. Utilitarismen förutsätter med andra ord *interpersonella välfärdsjämförelser*, och jag tror man kan säga att ingen ännu lyckats visa hur sådana är möjliga. Att vi i praktiken ändå måste göra sådana jämförelser är inget försvar för teorin. Den är ofullständig så länge frågan om interpersonella jämförelser inte är löst på ett teoretiskt tillfredsställande sätt.

Den är grundlös. Det finns såvitt jag kan se *ingen evidens* för utilitarismen, ingenting som talar för att den är sann. Ofta tänker man sig ju att man har evidens för en teori om den utgör den bästa förklaringen av föreliggande fakta, men utilitarismen förklarar egentligen ingenting alls. I en uttunnad mening kan man kanske säga att utilitarismen ”förklarar” sina egna konsekvenser – alltså det som kan härledas från utilitarismen i kombination med empiriska tilläggspremissor om enskilda situationer – men dessa konsekvenser ger knappast något stöd för utilitarismen eftersom deras enda grund i sin tur är just utilitarismen. (Att det är rationellt att godta konsekvenserna av vad man godtar är en

annan sak. Detta har ingenting med evidens eller induktiva slutledningar att göra.)

I ett föredrag i Stockholm någon gång i mitten av 80-talet kritiserade Dagfinn Føllesdal mig för denna uppfattning om utilitarismens grundlöshet (som jag tidigare framfört i *Filosofisk tidskrift* nr 1, 1980). Han hänvisade då till John Rawls' idé om ett "reflektivt ekvilibrium" mellan principer och enskilda omdömen, och han hävdade att de senare inte behöver något oberoende stöd för att de skall kunna utgöra ett stöd för de förra. Till en del håller jag med om detta. De enskilda omdömena behöver inget oberoende stöd – men det viktiga är att de inte själva enbart är motiverade av det de skall utgöra evidens för. Poängen är alltså att de inte kan utgöra något stöd för vissa principer, som t ex utilitarismen, om de i sin tur är baserade just på dessa principer (och om vi alltså inte hade accepterat omdömena om vi inte hade härlett dem från principerna).

För övrigt är det ju, som Folke Tersman visat (i *Southern Journal of Philosophy* vol 29, 1991), omöjligt att underbygga utilitarismen i ett reflektivt ekvilibrium, eftersom man inte kan ha tillgång till välgrundade empiriska premisser av det slag som skulle behövas (utom kanske i vissa tankeexperiment – men det faktum att en utilitarist kan tänka ut abstrakta situationer i vilka han skulle hålla fast vid utilitarismen ger knappast något stöd för utilitarismen). Vi kan alltså inte veta vilka konsekvenser utilitarismen har i enskilda fall. Detta leder över till den femte och sista nackdelen med utilitarismen.

Den är oanvändbar. Den är med andra ord *svår eller omöjlig att tillämpa* i praktiken, eftersom vi aldrig kan veta tillräckligt mycket om de totala konsekvenserna av våra handlingar. Även om vi kanske ibland kan ha skäl att tro att en handling har bättre konsekvenser än en annan, så har vi troligen aldrig goda skäl att tro att en handling har bättre, eller minst lika bra, konsekvenser som varje möjligt alternativ.

Utilitarister är naturligtvis i allmänhet medvetna om detta problem. Men de brukar slingra sig undan det genom att hävda att ett riktighetskriterium inte behöver kunna tillämpas; det avgörande är om det är sant. Hur man sedan skall bete sig i praktiken för att avgöra hur man bör handla är en annan sak. Det är här man behöver en "beslutsmetod", och en sådan är något annat än ett riktighetskriterium. Denna distinktion går som sagt tillbaka till Sidgwick, och den motsvaras hos exempelvis R M Hare av distinktionen mellan en "intuitiv" och en

”kritisk” nivå i det moraliska tänkandet. Hos Thomas Wetterström (*Basic Ethics*, Oxford 1986) görs på motsvarande sätt en distinktion mellan ”basic” och ”pragmatic” etik. För egen del är jag dock rätt skeptisk mot dessa distinktioner. (Detta förklarar kanske att jag ibland kritiserats – t ex av Sören Halldén och Torbjörn Tännsjö – för att inte upprätthålla dem.) Låt oss se lite närmare på detta.

3. Utilitarismen i praktiken

Jag har inget speciellt för mig på eftermiddagen och funderar på att gå på bio. Men så slår det mig att ett bättre alternativ är att jag köper den vinnande lotten på penninglotteriet i stället, ty konsekvensen blir då att jag senare *både* kan gå på bio och skicka en miljon till välgörande ändamål. Alltså bör jag göra det.

Mot detta kan man invända att jag *inte kan* köpa den vinnande lotten, och att ”bör” implicerar ”kan”. Alltså kräver inte moralen av mig att jag köper den vinnande lotten.

Men *varför* kan jag inte köpa den vinnande lotten? Den är ju ännu osåld (kan vi tänka oss). Ja, men jag *vet inte hur jag skall bära mig åt* för att köpa just den. Och detta medför att det följande inte gäller: om jag skulle försöka, så skulle jag lyckas. Och därför kan jag inte göra det (i denna hypotetiska mening av ”kan”, som möjligen är, åtminstone en del av, den relevanta).

Men varför spelar det någon roll? Eftersom vinnarlotten ännu inte är såld, så är det ju ändå *möjligt* att jag köper den. Varför skulle ”bör” implicera ”kan” i någon starkare mening? (Faktum är att jag i levande livet har stött på två utilitarister som anser att ”bör” inte implicerar ”kan” i någon starkare mening, och att jag i exemplet alltså bör köpa vinnarlotten. Men jag är övertygad om att de flesta utilitarister, liksom Moore i sin klassiska framställning av detta i *Ethics*, anser att ”bör” implicerar mer än bara ”möjlig”).

Svaret måste väl vara att om ”bör” bara implicerar ”möjlig”, så skulle moralens normer i många fall sakna *praktisk betydelse*, på så vis att de inte skulle ge någon som helst *vägledning* till agenten. Detta torde vara skälet till att man inte kan hävda att jag bör köpa den vinnande lotten, trots att denna handling har de bästa konsekvenserna (av alla möjliga handlingar). Och detta visar i sin tur att även riktighetskriterier – och inte bara beslutsmetoder – måste ha någon sorts vägledande funktion.

En utilitarist skulle emellertid kunna invända mot detta resonemang

att man med hjälp av sin moraliska intuition helt enkelt kan *inse* att "bör" implicerar "kan" i en starkare betydelse än "möjlig", och att detta inte alls beror på att riktighetskriterier (eller moraliska principer) måste ha en vägledande funktion.

Men detta hjälper inte. Ty vi kan då anta att lotteriet i mitt exempel är sådant att jag kan välja lottnummer när jag köper en lott, och att den lott som kommer att vinna har nummer 4711. Utan tvekan *kan* jag nu köpa den vinnande lotten (dvs lott nr 4711), i den meningen att om jag försöker så lyckas jag. Utilitaristen måste väl då anse att jag *bör* göra det (givet att denna handling har de bästa totala konsekvenserna). Men i praktiken är jag ju precis lika hjälplös som om jag skulle plocka en lott på måfå ur en tombola. Ty jag *vet* ju inte, och *kan* inte veta, vilken lott som kommer att vinna. Detta gäller i bägge fallen. Det är därför, ur moralisk synpunkt, *lika orimligt* att hävda att jag (i stället för att gå på bio) bör köpa lott 4711, som att jag bör köpa den vinnande lotten. Den princip som påbjuder detta har ingen vägledande funktion, på grund av att agenten inte kan veta vilken handling den påbjuder i det enskilda fallet.

I detta läge skulle man kanske vilja hävda att "bör" inte bara implicerar "kan", utan dessutom "kan veta att man bör". Men det vore att gå för långt (bland annat kanske för att det skulle förutsätta att moraliska normer kan vara sanna). Däremot skulle jag vilja föreslå att "bör" implicerar "det finns någon egenskap hos handlingen, som gör att den bör utföras, och man kan veta, eller åtminstone ha förnuftiga skäl att tro, att handlingen har denna egenskap". (Sven Danielsson har frågat mig vem denna "man" i min princip är. Den jag främst hade i tankarna var den som kan utföra handlingen, men kanske borde man försvaga principen genom att ersätta "man" med "någon".)

Men om min princip är riktig, så är utilitarismen felaktig! Ty i själva verket vet man väl *aldrig* vilket alternativ som har de bästa totala konsekvenserna, och trots det menar ju Moore och andra utilitarister att man ändå bör utföra just det.

Utilitarismen är alltså inte ett acceptabelt riktighetskriterium. Det är inte heller sant att om man *vet* att en viss handling har bättre konsekvenser än varje alternativ, så bör man göra den – ty dess optimalitet kan bero på att man i framtiden kommer att göra något som man inte borde ha gjort (se min uppsats "Utilitarianism and future mistakes", *Theoria* vol 43, 1977).

4. En mer modest utilitaristism

En utilitarist behöver dock inte hänga läpp över detta. Jag föreslår att han eller hon i stället retirerar till att betrakta utilitarismen – eller någon mer tillämpar variant av den – som en beslutsmetod. Eller som en del av en beslutsmetod. Eller som en tumregel. Eller som en moralregel bland andra.

Men hur skall då utilitarismen tolkas för att alltså kunna tillämpas i praktiken? Jag skall inte här gå närmare in på detta. Jag föreställer mig att man till en del helt enkelt kan överta de beslutsregler som brukar figurera inom beslutsteori – ”maximin”, ”minimax risk”, ”maximering av förväntat värde”, osv. Det ur praktisk synpunkt kanske största problemet med dessa regler är att de för sin tillämpning förutsätter att man redan har tillgång till en strukturering av beslutsituationen i form av en matris, där alternativ, möjliga utfall och dessas värden på någon lämplig skala är givna. Hur man fastställer en relevant matris säger däremot beslutsteoretiker inte mycket om. Ur utilitaristisk synpunkt bör naturligtvis värderingen av utfall göras på ett *opartiskt* sätt, utfallen får inte begränsas till något speciellt perspektiv – de bör i princip inkludera *alla relevanta* konsekvenser – och i den mån sannolikheter skall vägas in, så bör de vara *objektiva* (i någon mening) snarare än subjektiva. Detta är en mycket stor beställning. I praktiken får man väl nöja sig med att konstruera någon matris som åtminstone inte är *uppenbart* vilseledande eller partisk. Sedan kan man t ex säga att en handling är moraliskt rätt om den maximerar förväntat värde relativt till en sådan matris – såvida det inte finns andra, starkare skäl för att den är moraliskt orätt.

Kanske kunde man också tänka sig en regel av följande slag: Om de *förutsägbara* konsekvenserna av en handling är bättre än de förutsägbara konsekvenserna av varje *rimligt* alternativ, så bör handlingen utföras – såvida det inte finns andra, starkare skäl som talar mot detta. Detta kan ju sägas vara en utilitaristisk regel, och den kan antagligen ge en viss vägledning i många situationer. Dessutom är det en regel som troligen skulle kunna accepteras av de allra flesta.

Skall man betrakta en sådan regel som ett ”riktighetskriterium” eller som en ”beslutsmetod” eller som något som varken är det ena eller det andra? Vilket *ord* man här väljer kan knappast spela någon roll. Det viktiga är att man inser att utilitaristiska handlingsregler inte identifierar någon precis, empiriskt avgörbar egenskap som är gemensam

och specifik för alla handlingar som bör utföras. De kräver för sin tillämpning subjektiva bedömningar, t ex vad gäller interpersonella välfärdsjämförelser, relevansen hos givna beslutsmatriser och den relativa resistensen mot icke-utilitaristiska synpunkter. Helst bör de nog betraktas som moraliska tumregler, som kan komma i konflikt med andra moraliska tumregler. För egen del anser jag att de som sådana bör väga tungt.

5. Måste tumregler baseras på ett riktighetskriterium?

Om utilitarismen i sin normativt rimligaste utformning består av en samling tumregler – kanske bland andra – måste det då inte finnas något moraliskt riktighetskriterium i bakgrunden, något som utgör grunden eller motiveringen för dessa tumregler? Hänger de inte annars i luften?

Jag tror inte det. Jag tror snarare att detta är en utilitaristisk fördom – som delvis sammanhänger med svårigheterna att tillämpa det utilitaristiska riktighetskriteriet. Den utilitaristiska uppfattningen om förhållandet mellan riktighetskriteriet och våra mer konventionella moralregler (tumregler) är problematisk. Vilken är egentligen denna uppfattning?

(a) Vi har ju ingen som helst anledning att tro att tumreglerna alltid eller oftast eller ens någonsin pekar ut just de handlingar som det utilitaristiska riktighetskriteriet pekar ut. Tvärtom, att döma av de vanliga normativa invändningarna mot utilitarismen.

(b) Vi har inte heller någon anledning att tro att de mer konventionella tumreglerna utgör ett regelsystem som har högre acceptansvärde än varje alternativt regelsystem. (Ett regelsystems acceptansvärde är värdet av konsekvenserna av att alla accepterar systemet och försöker tillämpa det i sitt handlande.) Om det är svårt att veta att en given handling har bättre konsekvenser än varje alternativ, så är det säkerligen ännu mycket svårare att veta att ett regelsystem har maximalt acceptansvärde. Man kan inte heller förvänta sig att evolutionen har sett till att våra vanliga regler har maximalt acceptansvärde. Evolutionen har nog inte haft tillräckligt lång tid på sig i mänskliga samhällen, och dessa har oftast varit konservativa och föga experimentella.

(c) Inte heller har vi någon anledning att tro att ett system av mer uttalat utilitaristiska tumregler – av det slag som antytts ovan – skulle ha maximalt acceptansvärde. Därtill är de bland annat alltför obestämda.

(d) Även om ett visst regelsystem skulle ha ett maximalt acceptansvärde, så är det ingalunda uppenbart att man därför bör tillämpa det i enskilda fall. Ty konsekvenserna av att göra det kan bero på om även andra tillämpar systemet. Gör de inte det, så blir kanske konsekvenserna bättre om man själv inte heller gör det.

(e) Torbjörn Tännsjö tycks anse att om ett visst regelsystem har maximalt acceptansvärde *för en agent* (dvs om konsekvenserna av att agenten tillämpar dem konsekvent i alla situationer han hamnar blir bättre eller minst lika bra som de konsekvenser som skulle uppstå om han konsekvent tillämpade något annat regelsystem), så är det *rationellt* för denna agent att tillämpa detta regelsystem i enskilda fall, *om* han accepterar det utilitaristiska riktighetskriteriet. Detta är inte särskilt övertygande. Vad som är rationellt för agenten i en enskild situation tycks bero på om han tillämpar regelsystemet i andra fall. Och även om han skulle göra det, så kan det ju mycket väl vara så att han enligt det utilitaristiska riktighetskriteriet inte bör tillämpa systemet i ett enskilt fall. Hur skall han då bete sig? Hur skall man bete sig om ett alternativ är rationellt och ett annat är moraliskt riktigt? Följande är inte något bra svar: man bör bete sig rationellt, om man vill bete sig rationellt, och man bör bete sig moraliskt riktigt, om man vill bete sig moraliskt riktigt. Ett rimligt moralsystem borde ställa större krav än så.

(f) Man kan möjligen ha skäl att tro att många konventionella moraliska regler har ett *visst* positivt acceptansvärde. Ty det förefaller som om de skulle kunna underlätta social samvaro. Och det är väl delvis därför vi lär ut dem till våra barn. Detta är en sorts utilitaristiskt skäl, men det är inte av maximeringstyp.

(g) På liknande sätt tror vi väl att mer direkt utilitaristiska regler också har ett visst positivt acceptansvärde. Det är svårt att säga om det är högre eller lägre än för konventionella regler. Sidgwick trodde väl att det är lägre, medan den mer sentida utilitaristen J J C Smart tycks mena att det är högre. Men här kan vi knappast ha mer än fromma förhoppningar.

(h) Hur som helst tycks det vara så – om man bortser från fall (a) ovan – att det utilitaristiska riktighetskriteriet överhuvud taget inte har någon funktion att fylla när det gäller att motivera en beslutsmetod eller moraliska tumregler. Vad som spelar roll är snarare acceptansvärde eller något liknande. Men en motivering av tumregler i termer

av acceptansvärde blir inte mer vägande om man därutöver lägger till det utilitaristiska riktighetskriteriet. Snarare tvärt om.

Jag anser med andra ord att moraliska beslutsmetoder inte behöver stöd av något riktighetskriterium – åtminstone inte på det sätt som utilitarister tycks tänka sig. Att utilitarister gör ett så stort nummer av distinktionen mellan riktighetskriterium och beslutsmetod beror väl dels på att just deras riktighetskriterium saknar praktisk användning, och dels på att distinktionen tycks kunna användas för att bemöta de vanligaste formerna av moralisk kritik mot utilitarismen. Denna kritik brukar nämligen bemötas genom påpekandet att det utilitaristiska riktighetskriteriet inte skall tillämpas i praktiken, utan att man i stället bör följa mer konventionella moralregler (eller liknande), vilka antas ha högt acceptansvärde. (Här hänvisar man till empiriska hypoteser som är ytterst svårkontrollerade.) Men ett sådant försvar för utilitarismen innebär ju att man i själva verket gör den immun mot moralisk kritik. Detta är knappast mer än ett retoriskt knep.

Jag vill dessutom tillägga, att tron på att det finns ett moraliskt riktighetskriterium – en bestämd egenskap, som är gemensam och specifik för alla handlingar som ur moralisk synpunkt bör utföras – i sig själv mycket dubiös. Sådillvida kan den jämföras med den lika tveksamma (men kanske också lika vanliga) åsikten att det finns ett *sanningskriterium*, en bestämd egenskap som är gemensam och specifik för alla sanna satsar eller för alla sanna påståenden. Varför skulle man tro det?

6. *Moralens grund*

Kontentan av ovanstående är att vi knappast har något behov av ett moraliskt riktighetskriterium som inte med viss lätthet kan tillämpas i praktiken. Vi klarar oss lika bra med enbart tumregler, eller med andra ord med moraliska regler av mer konventionellt slag.

Men på vilken grund vilar då dessa regler? Såvitt jag kan se behöver de ingen grund. Det finns säkerligen *orsaker* till att vi tror på de regler vi tror på. Men det behöver inte finnas några mer bakomliggande *skäl* till att vi tror på dem.

Det finns inga *andra* regler som implicerar att vi bör acceptera just våra regler. Det finns inte heller några *mer generella* regler från vilka våra regler kan härledas. Det kanske inte heller finns några generella regler om hur regelkonflikter bör lösas – bortsett från den självklara

regeln att viktigare regler skall väga tyngre än mindre viktiga.

Och även om det funnes mer fundamentala regler, som utgjorde en grund för våra vanliga moralregler, så skulle det inte göra någon större principiell skillnad. Ty i så fall kunde man på samma sätt fråga sig om *de* kan accepteras, utan att ha en grund i ännu mer fundamentala regler, osv.

Därför kan vi lika gärna godta de *regler som vi faktiskt tror på* (och som vi är benägna att hålla fast vid även om de utsätts för kritik) som moralens grund. Liksom i vetenskap har vi inget annat val än att utgå från det vi faktiskt tror på. Otto Neurath beskrev situationen i en berömd liknelse: vi är som sjömän i en båt på öppna havet; det gäller att hålla sig flytande och om vi behöver reparera båten (dvs modifiera våra åsikter), så måste det göras gradvis, medan det mesta behålls intakt – annars går vi under helt och hållet. Att försöka att hitta en oberoende grund för moralen, från vilken man kan bedöma våra faktiska uppfattningar, är lika orimligt som att konstruera en ”första filosofi” som grund för vetenskapen.

Till dessa våra vanliga moralregler hör sådant som: man bör hålla löften och avtal, tala sanning, hjälpa folk i nöd, följa demokratiskt fattade lagar, glädja sina medmänniskor, skydda och uppfostra sina barn, inte skattefuska, inte ta emot socialbidrag som man inte behöver eller som man inte har rätt till, inte sparka folk som ligger försvarslösa på marken, behandla andra som man själv vill bli behandlad, respektera andras autonomi, inte mobba arbetskamrater, inte bära falsk vittnesbörd mot sin nästa, behandla djur humant, osv. Vi uppfattar också olika regler som mer eller mindre viktiga – att hjälpa folk i svår nöd är t ex viktigare än att hålla löften i mer triviala sammanhang. Reglerna kan naturligtvis komma i konflikt, och då får man bedöma deras relativa viktighet.

Man kunde kanske vilja hävda att när olika regler kommer i konflikt, så skall man *enbart* lägga vikt vid utilitaristiska överväganden. Men det vore nog ingen bra regel. Ty utilitaristiska överväganden kan behöva kompletteras med andra regler. Flera av de vanliga motexemplen mot utilitarism visar just detta.

Att ”moralens grund” alltså inte är något annat än den moral vi faktiskt godtar innebär dock inte att vi nu har bundit oss för ett nytt, *konventionalistiskt* riktighetskriterium, som innebär att en handling bör utföras om och endast om den är i överensstämmelse med de moral-

regler som accepteras av agenten (eller i agentens samhälle, eller i vårt samhälle). Det inser man om man betänker att denna princip *inte* tillhör det som vi faktiskt tror på.

7. Svåra fall

I vissa fall pekar våra regler som sagt åt olika håll. De kommer i konflikt. Sådana konflikter kan ibland lösas genom att vi tar hänsyn till de olika reglernas relativa viktighet. I vissa fall löser dock olika bedömare konflikten på olika sätt, och ibland kan vi kanske inte hitta några lösningar alls.

Att en uppsättning regler, eller värden, inte är tillräcklig för att garantera ett bestämt resultat i varje enskilt fall är emellertid inte något speciellt för moral. Det gäller även i *estetik* (ett konstverk är bättre i vissa avseenden men sämre i andra), *juridik* (vissa skäl talar för friande dom, andra för fällande, etc), *vetenskap* (en teori är bättre i vissa avseenden men sämre i andra), *texttolkning* (evidens av olika slag kan peka åt olika håll), *socialpolitik* (ett lagförslag är bättre i vissa avseenden men sämre i andra) och för övrigt i en mängd olika fall av *klassifikation* (det kan finnas olöslig oenighet om huruvida något är komiskt, dramatiskt, demokratiskt, rättvist, ett etablerat faktum, sentimentalt, riskabelt, godkänt, välgrundat, kunskap, stort, kallt, tungt, meningslöst, säker slutförvaring, tillförlitligt, ett lyckligt liv, trafikfarligt, etc).

Man kan fråga sig om det i dessa svåra fall (*hard cases*) finns en entydig *sanning*, eller om det snarare är fråga om sanningsvärdesluckor där våra påståenden varken är sanna eller falska. Detta är en filosofiskt spännande fråga, men inget speciellt för moralfilosofi, och ingenting som bör få bli avgörande för frågan om utilitarismen är sann eller falsk. Även om det skulle vara så att det utilitaristiska riktighetskriteriet ger en lösning av alla svåra fall – vilket bland annat förutsätter att problemet om interpersonella jämförelser får en lösning – så kan detta knappast användas som ett argument för att utilitarismen är sann.