

1. RIKTIGHETSKRITERIER OCH BESLUTSMETODER

”Om utilitarismen är riktig är det ofta omöjligt att ta reda på hur man bör handla. Detta eftersom det nästan alltid verkar omöjligt att veta vilket handlingsalternativ som faktiskt skulle ha bäst konsekvenser.” Ätminstone sedan Sidgwick (1907, s. 413) har utilitarister försökt parera invändningar av detta slag genom att skilja mellan utilitarismen tolkad som ett riktighetskriterium respektive som en beslutsmetod. I den förra tolkningen, som påstås vara den mest rimliga, uttalar sig utilitarismen endast om en handlings egenskap att vara moraliskt riktig. Hur man sedan i praktiken ska gå till väga för att välja rätt, dvs. vilken beslutsmetod man bör tillämpa, är en annan fråga.

Låt oss titta lite närmare på distinktionen mellan riktighetskriterier och beslutsmetoder, och undersöka vad som gör en (aktualistisk handlings-) utilitaristisk beslutsmetod B_1 *bättre* än en annan B_2 . Ur utilitaristisk synvinkel torde detta rimligen ha något att göra med vilka konsekvenser tillämpandet av B_1 respektive B_2 faktiskt skulle ha. Antag att vi tillämpar beslutsmetoderna B_1 och B_2 på tusen beslut. Om B_1 pekar ut den handling som är riktig enligt riktighetskriteriet i niohundra av de tusen fallen, och B_2 den handling som är riktig enligt riktighetskriteriet i åttahundra av de tusen fallen, kunde det ligga nära till hands att hävda att B_1 är en bättre utilitaristisk beslutsmetod än B_2 eftersom kvoten mellan det faktiska antalet korrekta råd och det möjliga antalet korrekta råd är större för beslutsmetod B_1 än för B_2 . En nackdel med detta kriterium är emellertid att det bara tar hänsyn till antalet riktiga rekommendationer, inte till hur allvarliga fel de olika rekommendationerna ger. Ett annat problem är att kriteriet förutsätter att själva valet av beslutsmetod inte har några konsekvenser eller påverkar vilka framtida beslutssituationer agenten hamnar i. Låt oss därför modifiera kriteriet och säga att B_1 är bättre än B_2 om och endast om produkten av antalet felaktiga råd

¹Jag vill tacka Jonas Gren, Per Sandin och deltagarna i Lars Bergströms seminarium för värdefulla synpunkter.

och deras genomsnittliga grad av felaktighet är mindre för B_1 än motsvarande produkt för B_2 , samt begränsa dess räckvidd till fall där själva valet av beslutsmetod inte påverkar vilka beslut beslutsmetoden appliceras på. Detta kan förhoppningsvis fungera som en fruktbar konkretisering av den grundläggande idén att man bör välja en beslutsmetod som har bäst totala konsekvenser.

Nåväl, givet det föreslagna kriteriet tror jag vi kan vara tämligen säkra på att det inte går att hitta någon *perfekt* beslutsmetod, dvs. en metod som alltid genererar korrekta råd och är sådan att vi kan veta vilket handlingsalternativ vi bör välja. För om en beslutsmetod ska vara perfekt måste den i alla lägen peka ut exakt samma mängd handlingsalternativ som riktighetskriteriet. Men att tro att alla handlingsalternativ med bäst konsekvenser skulle ha någon ytterligare egenskap gemensamt, som är lättare att upptäcka (t.ex. egenskapen att utföras med en viss avsikt), verkar minst sagt överoptimistiskt. Noga taget kan det väl ofta vara så att de handlingsalternativ som har bäst konsekvenser – och därför är riktiga – skiljer sig åt i *alla* andra relevanta empiriska avseenden.

Ur praktisk synvinkel förefaller dock inte avsaknaden av perfekta beslutsmetoder så viktig. Rimligen kunde man känna sig rätt nöjd ifall man fann en beslutsmetod som man visste pekade ut rätt handling för det mesta, eller åtminstone ibland. Men inte heller det tycks möjligt. Ja, det tycks faktiskt inte ens möjligt att ange någon beslutsmetod för vilken man kan veta att den åtminstone någon gång har pekat, eller kommer att peka ut, rätt handling. För att se detta kan vi resonera på följande vis. Huruvida en beslutsmetod B är bra, dvs. träffar rätt ofta, beror enligt resonemanget ovan på ifall de rekommendationer B genererar sammanfaller med de handlingsalternativ riktighetskriteriet pekar ut. För att ta reda på detta måste vi således jämföra de handlingsalternativ som genereras av B med de handlingsalternativ som pekas ut av riktighetskriteriet. Men hur skulle vi ens någonsin kunna ta reda på ifall de handlingsalternativ som B genererar sammanfaller med de handlingsalternativ som riktighetskriteriet pekar ut? Vi *vet* ju inte vilka handlingsalternativ riktighetskriteriet rekommenderar; det var ju därför vi införde idén om en beslutsmetod!

Följande exempel kan eventuellt göra argumentet tydligare. Antag att jag vill ta reda på ifall Bibeln är en bättre beslutsmetod än principen om förväntad nyttoximering när det gäller beslut i kärlekslivet. Med lite tur, och kunskaper i teologi, kan jag kanske ta reda på hur Bibeln ställer sig till att jag lägger in en stöt mot P. Vi är inte släkt, och av olika kön, osv. Säg att Bibeln rekommenderar mig att lägga in en stöt mot P. Då måste jag, för att kunna jämföra denna rekommendation med det kor-

rekta svar som ges av det utilitaristiska riktighetskriteriet *känna till* det senare svaret. Men det gör jag ju inte, och kan inte heller någonsin göra enligt resonemanget ovan. Därför kan jag aldrig, för någon beslutsmetod – vare sig Bibeln, principen om förväntad nyttoximering, eller någon hittills okänd beslutsmetod – veta att den åtminstone en gång har pekat ut eller kommer att peka ut rätt handlingsalternativ. Och således kan det kriterium vi vill använda för att jämföra olika beslutsmetoder aldrig tillämpas.²

2. SKÅL FÖR HANDLING

Nu kunde utilitaristen givetvis bita ihop och hävda att hans teori är rimlig och intressant även om vi aldrig kan rättfärdiga någon beslutsmetod, och därigenom få råd om hur vi bör handla (jfr Bales 1971, s. 261). Detta tror jag dock vore en alltför macho inställning. Ty en rimlig moralisk teori måste ha åtminstone något att säga om hur vi bör gå tillväga för att handla rätt. Att *bara* på ett teoretiskt plan förklara vari moralisk riktighet består är inte tillräckligt. Mitt argument för denna ståndpunkt lyder i korthet som följer: (1) Att en handling bör utföras innebär att man har mest skäl att utföra denna handling. (2) Skäl för handling är interna snarare än externa, och därför vetbara. (3) Alltså kan det inte vara så att man bör utföra en handling utan att man också kan (i den nedan angivna betydelsen av "kan") veta det. (4) Alltså är utilitarismen falsk.

För att kunna bedöma detta argument behöver vi självfallet utveckla det lite mer i detalj. Låt oss först undersöka påståendet att en handling bör utföras endast ifall man har mest skäl att utföra handlingen ifråga. Själv är jag nog beredd att hävda att det i betydelsen hos "bör" ingår att man bör göra det man har mest skäl att göra. Som bekant har dock en del utilitarister i stället ansett att det i betydelsen hos "bör" ingår att man bör göra det som skulle realisera mest värde. (Nuförtiden har denna idé rätt få anhängare; moderna utilitarister brukar anse att de förespråkar en normativ och inte en begreppslig tes.) Men varför är mitt språkbruk mer rimligt än de gamla utilitaristernas? Ett lite lummigt argument, som man bör handskas försiktigt med, är att tillämpa Moores öppna fråga. Jämför följande frågor: (i) Ifall man utan att veta om det de facto realiserar mest värde genom att göra X, bör man då också göra X? (ii) Ifall man allting sammanvägt har mest skäl att göra Y, bör man då också göra Y? Om vi kunde visa att (i) men inte (ii) är en öppen fråga skulle vi

² Jag har uppmärksammat på att Kihlbom (2000) gör en liknande poäng. (Se avsnittet om "Decisiveness".)

ha ett argument för att acceptera tesen att man bör göra det man har mest skäl att göra, snarare än det som de facto skulle realisera mest värde. Men kan vi då visa att (i) men inte (ii) är en öppen fråga? Tja, det tycks i alla fall klart att (i) är en öppen fråga. Det visar om inte annat de senaste tvåhundra årens omfattande diskussion om handlingsutilitarismens riktighet: bör man verkligen göra det som de facto realiserar mest värde? Konsekvensialisterna svarar ja, medan t.ex. dygdetikerna menar att det är fel fråga som ställs. Men visst, givetvis kan en slipad filosof få även (ii) till en öppen fråga. Min poäng är bara att detta verkar svårare. Huruvida man bör göra det man har mest skäl att göra är betydligt mindre omstritt, och kräver mer filosofisk träning för att ifrågasättas. Alltså är (i) en *mer* öppen fråga än (ii), vilket talar för att det i betydelsen hos "bör" ingår att man bör göra det man har mest skäl att göra.

Hur är det nu med den andra premissen, är skäl för handling interna (och därmed vetbara) eller externa? Antag med Prichard att du närmar dig en korsning med din bil och funderar på om du har några skäl att stanna, eller åtminstone sakta ner.³ Om du är internalist är det klart att du, givet att dina önsknings- och trosföreställningar är någorlunda sunda, har skäl att stanna: Du har en önskan att inte krocka, samt tror att det finns en risk att det kommer korsande trafik. Men är du externalist måste du nog gå med på att du bara har skäl att stanna ifall det faktiskt är så att det kommer korsande trafik. Att säga att ditt externa skäl att stanna är grundat i en objektiv sannolikhet för korsande trafik duger inte. För vi kan lätt bygga ut exemplet med antagandet att den korsande vägen har varit blockerad i bägge riktningarna de senaste tio åren, utan din vetskap. Då är den objektiva sannolikheten för korsande trafik noll, samtidigt som du fortfarande tycks ha skäl att stanna.

Eventuellt kunde man hoppa över dessa inledande steg, och i stället argumentera direkt för (3). Lars Bergström hävdar i uppsatsen "Reflections on Consequentialism" att (3) kan rättfärdigas med hjälp av tesen "bör implicerar kan": En rimlig tolkning av ordet "kan" måste här vara starkare än "kan, men vet inte hur"; mer precist ska man tolka "kan" som "kan, och vet hur". Bergström använder följande exempel för att illustrera denna poäng. I den svaga meningen av "kan" kan du köpa den vinnande lotten i penninglotteriet, även fast du inte vet vilken den är, och troligen skulle misslyckas om du försökte. Men i den starka meningen av "kan" kan du det inte, eftersom du inte vet hur du ska gå tillväga. Ska du ta den nere till höger, eller kanske den som ligger högst upp till vänster? Eftersom du inte kan köpa den vinnande lotten, eller utföra

³ Jag har fått exemplet från Dancy (2000, s 56).

den handling som har bäst konsekvenser, i den relevanta meningen av ”kan”, kan vi genom att tillämpa modus tollens konstatera att du inte är skyldig att köpa den vinnande lotten eller utföra den handling som har bäst konsekvenser. Dancy sammanfattar denna tanke i påståendet att skäl för handling måste passera genom ett *epistemiskt filter*, och att det utilitaristiska skälet inte klarar detta test.⁴

3. BESLUTSTEORI OCH MORAL

Kärnpunkten i resonemanget så här långt är att distinktionen mellan riktighetskriterier och beslutsmetoder inte förmår förse utilitaristen med någon handlingsvägledning, tvärtom vad många förespråkare tycks tro (se t. ex. Tännsjö 1998, s. 36–37), och att teorin därför är falsk enligt (1)–(3) i föregående avsnitt. Den uppenbara utmaningen är nu att föreslå en bättre teori.

På sätt och vis har vi redan en sådan teori, nämligen premiss (1) i föregående avsnitt: *Man bör göra det man har mest skäl att göra*. Vår uppgift är att fylla detta stängda påstående med en smula innehåll. För detta ändamål tror jag vi bör ta hjälp av beslutsteorin, vilket eventuellt kan tyckas en smula märkligt då man väl ofta har uppfattat beslutsteorins uppgift som att hitta en lämplig beslutsmetod att komplettera ett exogent givet riktighetskriterium med.⁵ Men den tanken tror jag alltså är förfelad. Som antytts ovan tror jag i stället att beslutsteorin bör uppfattas som en direkt konkurrent till andra moralfilosofiska teorier. Ifall principen om maximering av förväntat värde (vilken vi snart ska se är den bästa bestämningen av tesen att man bör göra det man har mest skäl att göra) kan ges ett tillfredsställande stöd gäller det således att den pekar ut den mängd handlingsalternativ som agenten *bör* utföra; någon distinktion mellan riktighetskriterier och beslutsmetoder behövs inte.

När man uppfattar principen om maximering av förväntat värde som en moralisk teori bör man naturligtvis väga in det förväntade värdet för *alla* berörda personer, inte bara det förväntade värdet för en själv. Till skillnad från utilitarismen erbjuder principen om maximering av förväntat värde, hävdar jag, en hel del handlingsvägledning – vilket också är ett av de viktigaste skälen för att acceptera den. (Detta är inte någon originell tanke. Många andra har varit inne på liknade spår, men ofta utan att ge något påtagligt argument för sin ståndpunkt.)

Att en beslutsteoretisk princip erbjuder mer handlingsvägledning än

⁴ Dancy (2000, s. 57).

⁵ Denna tanke utvecklas och kritiserar i Hansson (2001).

utilitarismen är nu inget okontroversiellt påstående. För på sätt och vis har vi ju bara ersatt ett problem med flera nya. Förr var problemet att vi inte kunde veta vilka konsekvenser det faktiskt skulle ha att utföra ett visst handlingsalternativ, t.ex. att delta i ett lotteri. Nu är problemet att vi nog inte kan känna till alla de relevanta tillstånd som världen kan befinna sig i, samt den sannolikhet med vilket varje tillstånd inträffar, osv. För att råda bot på detta problem tror jag vi bör acceptera en tes vi kan kalla för *global subjektivism*. Den globala subjektivismen hävdar, i likhet med vanlig beslutsteoretisk subjektivism, att de sannolikheter beslutsfattare bör bry sig om när de fattar beslut är subjektiva snarare än objektiva sannolikheter. (Exakt hur vi ska förstå begreppet subjektiv sannolikhet kan vi bortse från här.) Men dessutom hävdar globala subjektivisterna att begreppen alternativ, naturtillstånd och nytta ska förstås subjektivt. Det är inte de faktiska handlingsalternativen agenten bör försöka välja mellan, utan de hon *tror* sig kunna välja mellan (i betydelsen ”om jag försökte skulle jag troligen lyckas”). Hur världen råkade se ut, dvs. om en viss handling faktiskt var ett alternativ för agenten spelar mindre roll. Likaså för naturtillstånd och nytta; det är inte den nytta som folk faktiskt skulle uppleva om ett visst naturtillstånd var fallet och jag utförde en viss handling som spelar roll, utan den som jag (på tillräckligt goda grunder) tror att de skulle uppleva, osv. Således kan vi, givet att vi har tillgång till våra egna trosföreställningar och är duktiga på att räkna, alltid ta reda på vilket handlingsalternativ som principen om maximering av förväntat värde pekar ut.

4. PRINCIPEN OM MAXIMERING AV FÖRVÄNTAD NYTTA

Låt oss avslutningsvis ventilera frågan varför man överhuvudtaget bör maximera förväntat värde, snarare än vara riskaversiv eller risksökande. Så vitt jag känner till finns det i huvudsak två typer av argument som är intressanta här. Den första bygger på de stora talens lag (först formulerad och bevisad av Jacob Bernoulli på 1600-talet) och påpekandet att det går bäst i långa loppet ifall vi maximerar förväntat värde. Problemet med argument av detta slag är givetvis den bristande relevansen. För i praktiken är ju de flesta beslutssituationer unika, dvs. de uppträder bara en eller ett fåtal gånger. Och varför ska man i en unik beslutssituation bry sig om vad som skulle hända i långa loppet?

Moderna argument för principen om maximering av förväntat värde, t.ex. von Neumann & Morgenstern (1947), Savage (1954, 1972), Jeffrey (1965, 1983), Oddie & Milne (1991), Broome (1999) och Peterson (2002), har ingen koppling till de stora talens lag. I stället utgår de från ett litet

antal postulat som påstås vara rimliga vart och ett för sig. Tanken är att ifall man accepterar postulaten måste man också acceptera principen om maximering av förväntat värde, eftersom den senare följer ur de förra genom formell logisk deduktion. Naturligtvis behöver man inte tro att postulaten är självvedenta, eller något slags *fundamentala* sanningar. Idén är bara att vi måste välja: Antingen ger vi upp något av postulaten, eller så accepterar vi den slutsats som följer ur dem (principen om maximering av förväntat värde).⁶ Givet att postulaten är tillräckligt plausibla blir det ”billigare” att acceptera principen om maximering av förväntat värde än att ge upp något av postulaten.

Nedan följer en förenklad variant av ett formellt argument av detta slag, som presenteras mer utförligt i Peterson (2002). Till skillnad från tidigare argument tar mitt argument inte omvägen via ett representationsteorem, utan går ”direkt på” det som ska visas. Vidare vilar inte argumentet på något s.k. oberoende-postulat eller ”sure thing”-princip. För att kunna formulera argumentet utan formler har jag varit tvungen att göra postulaten något oprecisa, samt utelämna beviset.

Låt oss säga att två beslutsproblem X_1 och X_2 är lösningsekvivalenta om och endast om varje handlingsalternativ som är normativt rimligt att välja i X_1 också är normativt rimligt att välja i X_2 och vice versa.

- P 1: Antag att X_1 och X_2 är två lösningsekvivalenta beslutsproblem. Då spelar det ingen roll om man väljer att lösa X_1 eller X_2 .
- P 2: Om två beslutsregler r_1 och r_2 ger precis samma resultat när de appliceras på en mängd M av beslutsproblem, och det är normativt rimligt att applicera r_1 på alla problem i M , då är det också normativt rimligt att applicera r_2 på alla problem i M .
- P 3: Antag att två beslutsproblem X_1 och X_2 skiljer sig åt endast i ett avseende, nämligen att en av kolumnerna i X_1 är ersatt i X_2 av två identiska kolumner som tillsammans har samma sannolikhet som den ursprungliga kolumnen. Då är X_1 och X_2 lösningsekvivalenta.
- P 4: Låt a vara ett handlingsalternativ som har minst två möjliga utfall. Då gäller det att ifall vi gör det ena utfallet *lite* sämre (vi kan säga att försämringen är högst ϵ), så finns för varje sådan

⁶ Vi kan givetvis också välja att ge upp de logiska lagar som ligger till grund för deduktionen. Priset för detta är dock mycket högt.

försämring, en möjlig förbättring av det andra utfallet, som får lov att vara hur stor som helst, så att det modifierade handlingsalternativet a' är precis lika bra som det ursprungliga a .

- P 5: En normativt rimlig beslutsregel ska, där så är möjligt, välja ett handlingsalternativ som dominerar. (Ex: a_1 dominerar a_2 [svagt] ifall det, oavsett hur världen råkar se ut, går minst lika bra ifall jag gör a_1 som om jag gör a_2 .)

Alltså: Det är normativt rimligt att applicera principen om maximering av förväntat nytta på varje beslutsproblem under risk.

REFERENSER

- Bales, R.E. 1971. "Act-Utilitarianism: Account of Right-Making Characteristics or Decision-Making Procedure?", *American Philosophical Quarterly* 8, ss. 257-65.
- Bergström, L. 1996. "Reflections on consequentialism" *Theoria* 62, ss. 74-94.
- Broome, J. 1999. "Bolker-Jeffrey expected utility theory and axiomatic utilitarianism", in *Ethics out of economics*, Cambridge University Press.
- Hansson, S. O. 2001. "The Modes of Value", *Philosophical Studies*, 104, ss. 33-46.
- Jeffrey, R. C. 1965. *The Logic of Decision*, 2:a uppl. 1983, University of Chicago Press.
- Kihlbom, U. 2000. "Guidance and Justification in Particularistic Ethics", *Bioethics*, 14(4).
- von Neumann, J. & O. Morgenstern. 1947. *Theory of Games and Economic Behavior*, 2:a uppl, Princeton University Press.
- Oddie, G. & P. Milne. 1991. "Act and Value: Expectation and the Representability of Moral Theories", *Theoria* 57 (1-2), ss. 42-76.
- Peterson, M. 2002. "An Argument for the Principle of Maximizing Expected Utility", *Theoria* 68 (2), ss. 122-138.
- Savage, L. J. 1954. *The Foundations of Statistics*, Wiley and Sons. 2:a uppl. 1972, Dover.
- Sidgwick, H. 1907. *The Methods of Ethics*. Macmillan.
- Tännsjö, T. 1998. *Hedonistic Utilitarianism*. Edinburgh University Press.