

Om att gå runt på en gräsmatta

En populär typ av argument mot utilitarismen representeras av gräsmattefallet:

Den närmaste vägen mellan två platser, mellan vilka en stor mängd, säg 100, personer skall förflytta sig, går över en gräsmatta, som är mycket vacker. Var och en av de hundra kan göra en liten tidsvinst, av litet, men dock av värde, genom att gå över gräsmattan i stället för att gå den längre vägen på sidan om. Men om tillräckligt många går över gräsmattan blir den förstörd, och den värdeförlusten är större än det sammanlagda värdet av alla de enskilda tidsvinsterna.

Här säger nu utilitarismen, menas det, *var och en för sig* bör gå över gräsmattan, oavsett vad de andra gör, för genom att gå gör han en liten tidsvinst och genom att gå orsakar han ingen förlust under den estetiska aspekten, eftersom just *hans* gående inte innebär att gräsmattan förstörs.

Nu är ju utilitarismen ett sätt att bestämma plikter utifrån en värdeordning av konsekvenserna. Vad som utilitarismen säger i ett enskilt fall beror på vilken värdeordning man matar in. Det är lätt att se att den värdeordning man i det här argumentet stoppar in, är en *cyklisk* värdeordning, dvs sådan att man genom att hela tiden gå från sämre till bättre kan komma tillbaka till samma nivå.

De möjliga konsekvenserna tänks här som helt bestämda av antalet personer som korsar gräsmattan. Det betyder att de alternativa utfall som skall värderas helt enkelt kan identifieras genom alternativen "100 går", "99 går", . . . "2 går", "1 går", "0 går". Och man antar att det är bättre att 0 går än att 100 går, för värdet av en bevarad gräsmatta är större än sammanlagda värdet av tidsvinsterna. Vidare antar man att det är bättre att 1 går än att 0 går, för däremellan finns ingen skillnad i gräsmattans kvalitet, men väl en tidsvinst. Generellt antar man att det är bättre att n personer går än att $n-1$ personer går. Följaktligen är det bättre att 100 går än att 99 går. Men det var ju bättre att ingen går än att hundra går. Vi har alltså, genom att gå från sämre till bättre kommit tillbaka till samma nivå.

Man bör icke förvånas över att utilitaristen säger konstiga saker om man matar honom med cykliska värdeordningar. Genom att välja lämpliga

cykliska ordningar torde man ur utilitarismen kunna dra slutsatser mot vilka gräsmattefallet framstår som höjden av moral och förnuft.

Inte heller är det så, som sagts, att gräsmattefallet skulle peka på utilitaristiska problem som är speciella för situationer där ett kollektivs beteende är väsentligt.

Antag till exempel att jag ägde tusen tallar i Burselberget. Jag planerar en avverkning. Jag får betalt per kubikmeter och marginalkostnaderna faller med antalet träd, så ur ekonomisk synpunkt är det alltid bättre att fälla ännu ett träd. Men om jag kalhugger Burselberget, då slår fågelskådarna ihjäl mig, så ur den, dominanta, aspekten är det bättre att låta skogen stå orörd än att kalhugga. Å andra sidan är begreppet kalhygge ett vagt begrepp, det finns inget bestämt antal träd per arealenhet som utgör gränsen för kalhuggenhet. Därför är det alltså bättre att fälla 0 träd än 1 000 bättre fälla 1 än 0, bättre 2 än 1 etc, upp till bättre 1 000 än 999.

(Det här exemplet skiljer sig från gräsmattefallet bara däri att handlingsmöjligheterna här inte är distribuerade över ett antal olika personer – ersätter man här den ende agenten med tusen agenter, som kan fälla högst ett träd, har vi gräsmattefallet.)

Om vi accepterar att utilitarismen i gräsmattefallet säger att envar bör gå över gräsmattan, då måste vi i det här fallet acceptera att vad jag än gör, så handlar jag enligt utilitarismen orätt, för det finns alltid ett bättre alternativ, som jag kunde ha valt i stället.

Men naturligtvis säger utilitarismen inte i gräsmattefallet att envar bör gå över gräsmattan. För "bättre än" är nu en gång inte någon cyklisk relation, om än en del behaviouristiskt lagda beslutsteoretiker i denna fråga låtit sig duperas av sin egen finurlighet när det gäller att hitta på serier av valsituationer där försökspersonen med bestämdhet väljer sig tillbaka till utgångsläget.

Vad som eventuellt kan behöva förklaras, det är hur man ur till synes rimliga värdeförutsättningar kan hamna i antagandet att man genom att gå från sämre till bättre kan komma tillbaka till samma nivå.

Problemet är uppenbarligen ett amalgameringsproblem. Vi har två olika värdeordningar som vi skall väga samman till en. Vi har en estetisk ordning, och en tidsvinstordning. Låt oss, för korthets skull, tala om "estetiskt bättre" och "tidsvinstbättre". I amalgameringsproblem brukar vi ju oftast ha att göra med ordningar av samma logiska typ, men här är en av ordningarna, den estetiska, vag. Att x personer går är estetiskt bättre än att y personer går om och bara om det förra ger en oförstörd och det senare en förstörd gräsmatta. Men det finns ingen bestämd gräns mellan oförstörd och förstörd. Om vi lånar terminologin från Bertil Rolfs avhandling *Topics on Vagueness*, Lund 1981, om vaghet kan vi säga att det finns par av alternativ (x, y) om vilka det är *fundamentalt ovisst* huruvida x är estetiskt

bättre än y eller inte, dvs våra kriterier för förstördhet och oförstördhet räcker inte till för att avgöra frågan.

Man skulle kanske vilja hävda att inga sådana problematiska par är av typen (" x går", " $x+1$ går"). Alla sådana par hör, skulle man kanske vilja hävda, klart *inte* till extensionen för relationen "estetiskt bättre än". Men det är en svårförklarlig ståndpunkt. För om man accepterar att det finns fundamentalt oklara individuella fall, dvs alternativ " x går" som våra kriterier inte förmår klassa vare sig som fall av oförstördhet eller fall av förstördhet, då verkar det inte rimligt att man menar att kriterierna ändå förmår avgöra att det *inte* är så att " $x-1$ går" är ett fall av oförstördhet och " x går" ett fall av förstördhet.

Och om vi antar att alla par av typen $(x, x+1)$ är lika i fråga om oförstördhet, då är vi redan i utgångsläget invecklade i en version av "den skallige", då kan vi ur "gräsmattan är oförstörd om ingen går" härleda "gräsmattan är oförstörd om 100 går", då är problemet med gräsmattefallet inte något särskilt moralfilosofiskt.

Om vi nu accepterar att det finns fundamentalt ovissa fall av typen $(x, x+1)$ i den estetiska ordningen, vad har det för konsekvenser för sammanvägningen? Ja vi kunde kanske vara frestade att tillämpa Paretoprincipen, som här skulle säga att om x är bättre än y i en aspekt och inte sämre än y i den andra, då är x totalt bättre än y . Då skulle $x+1$ vara bättre än x även i de problematiska fallen och den totala ordningen cyklisk. Det är väl skäl nog att avvisa Paretoprincipen här, acykliciteten hos "bättre än" förefaller ju så ofantligt mycket säkrare än Paretoprincipen. Det förefaller också rimligare att säga att vagheten dominerar i det här fallet, därför att vagheten finns i den dominanta aspekten. Då skulle vi alltså säga att det också totalt är fundamentalt oviss om $x+1$ är bättre än x i de problematiska fallen.

Man kan kanske finna talet om "fundamental ovisshet" en smula suspekt. Då har man ju alltid den möjligheten att anta att för somliga par $(x, x+1)$ är x estetiskt bättre än $x+1$ fast inte *märkbart* estetiskt bättre. Det skulle innebära att vi antar att övergången från en oförstörd till en förstörd gräsmatta sker vid en viss punkt, men att det inte går att se var. Då skulle vi ha den situationen att $x+1$ är tidsvinstbättre medan x är estetiskt bättre (fast inte märkbart estetiskt bättre), och då följer det förstås inte säkert att $x+1$ är totalt bättre – i några fall skulle det faktiskt vara x som blev totalt bättre. I förstone kan det kanske förefalla konstigt att en skillnad som är mindre än märkbar i den estetiska ordningen kan vara större än en som är märkbar i tidsvinstordningen. Men det är inte märkligare än att vi tex uppfattar finare differenser i fråga om höjd och bredd än i fråga djup, när vi bedömer fysiska föremåls storlek. Vi kan ordna en serie lådor så att varje låda har en märkbart större höjd än föregående, men för övrigt inte är

direkt märkbart olik, samtidigt som vi genom att gå ett tillräckligt antal steg kan komma fram till en låda som är direkt märkbart mindre till sin volym.

Den varianten kan tyckas strida mot en mycket rimlig princip som säger att *endast märkbara skillnader är värder relevanta*. Om x och y inte är märkbart olika, så har de inte olika värde. Låt oss kalla denna princip för principen om det omärkligas irrelevans.

Den principen kan emellertid förstås på åtminstone två olika sätt. Principen om det *direkt* omärkligas irrelevans säger att x har ett annat värde än y endast om det finns en skillnad mellan x och y som kan uppfattas oberoende av observationer av något tredje objekt z . Men principen om det totalt omärkligas irrelevans säger bara att x har ett annat värde än y bara om det finns någon observerbar skillnad mellan x och y (en skillnad som även kan bestå i att x har en annan relation till z än y har).

Principen om det totalt omärkligas irrelevans är åtminstone tänkvärd, men principen om det direkt omärkligas irrelevans förefaller inte alls rimlig utom möjligen vad gäller *intrinsiskt* värde, som ju brukar sägas bero på egenskaper som inte är relationsegenskaper.

I gräsmattefallet finns enligt hypotesen inga direkt observerbara skillnader mellan x och $x+1$ i fråga om förstördhet. Men det finns kanske indirekt observerbara skillnader. I det enklaste fallet är det så att det mellan x och $x+2$ finns en märkbar förslitning, fast det inte finns någon mellan x och $x+1$ och inte heller mellan $x+1$ och $x+2$. Skall vi inte då säga att det finns en observerbar relevant skillnad mellan x och $x+1$? Vi har ju medelst observationer konstaterat en sådan skillnad. Den är visserligen inte *direkt* observerbar i den meningen att vi kan se den om vi bara betraktar x och $x+1$, men nog är den väl observerbar?

I ett riktigt realistiskt exempel är det kanske inte bara fråga om så enkla fall; beroende på relationens typ kan vi behöva tillgripa olika många jämförelseobjekt för att konstatera skillnaden, men det är av sekundärt intresse här, det viktiga är att det här inte tycks finnas några bra skäl för principen om det *direkt* omärkbaras irrelevans.

Vad utilitarismen säger i gräsmattefallet beror på vilket av dessa två alternativ vi väljer som förklaring. Om vi antar att den totala ordningen på några punkter är fundamentalt oviss, så säger utilitarismen om några, men ovisst vilka, varken att de bör gå eller att det inte är fallet att de bör gå. Väljer vi den andra varianten så säger utilitarismen om några, men fortfarande ovisst vilka, att de inte bör gå.

De problematiska fallen ligger, så mycket vet vi, någonstans i det område där det börjar bli oklart om gräsmattan är förstörd eller inte. Närmare besked får vi inte. Stoppas vi in vaga förutsättningar får vi vaga besked.