

# TORBJÖRN TÄNNSJÖ

## *Om rationalitet*

---

Jag tänker gå ut. Ska jag ta med mig paraplyet? Det kanske kommer att börja regna. Men det kanske inte börjar regna. Det är obekvämt att bära på paraplyet. Men det är obehagligt att bli blöt, om det skulle börja regna. Hur ska jag göra?

Jag frågar kanske en beslutsteoretiker. Han ber mig att tänka över de olika möjliga tillstånd världen kan vara i och han ber mig försöka värdera de motsvarande utfallen, som blir resultatet av respektive handlings sätt i tillståndet. Antag att han är framgångsrik med mig. Jag kan inte bara säga att av de olika möjliga utfallen så är det värsta att sakna paraply när det börjar regna och det bästa att inte ha paraply om det inte börjar regna utan jag kan också ta ställning till olika lotterier, som placerar in hur det är att ha paraply om det regnar mellan dessa alternativ. Detsamma kan han göra med hur det är att ha paraply om det inte börjar regna. Beslutsteoretikern ber mig nu representera min situation på följande sätt:

	Det börjar regna	Det börjar ej regna
Ta med paraply	0,7	0,5
Ta ej med paraply	0	1

Beslutsteoretikern förklarar för mig att siffrorna i matrisen representerar styrkan i mina preferenser, som jag uttryckt dem genom att ta ställning till de olika lotterierna. "Måtten är unika upp till positiva linjära transformationer", tillägger han.

Beslutsteoretikern ber mig nu tänka över ännu en aspekt av min situation. Hur sannolikt menar jag det är att det ska börja regna resp inte börja regna. Jag tänker noga över frågan, konsulterar väderleksrapporten i tidningen och kommer fram till att sannolikheten är fifty-

fifty. Men nu börjar jag bli otålig. Jag vill veta hur jag bör göra.

"Vänta ett ögonblick", säger beslutsteoretikern, "så ska jag räkna ut vad som är rationellt för dig att göra". Han multiplicerar värden och sannolikheter och summerar för varje rad i matrisen och konstaterar:

"Du ska ta med paraplyet".

"Varför ska jag göra det?", undrar jag.

"Jo, väntevärdet av att ta med det (0,55) är högre än väntevärdet av att inte ta med det (0,5)", svarar beslutsteoretikern.

Vad slags råd är det jag fått av beslutsteoretikern? Vad kan han rimligen mena med att det är "rationellt" för mig att ta med paraplyet? Vad är hans råd värt?

Betyder det att jag handlar rätt om jag tar med paraplyet? Det vore djärvt att påstå något sådant. Jag kanske borde låna ut mitt paraply åt någon som behöver det bättre än vad jag gör.

Kan man säga att, av de alternativ som föreligger, och om vi tänker bort andra som kan beröras av beslutet, så är att ta med paraplyet vad jag bör göra? Det är också djärvt att påstå något sådant. Jag kanske skulle bli lyckligare om jag blev blöt. Detta är kanske inte något jag vill bli, men något som vore bra för mig om jag blev. Jag kanske förtjänar att bli blöt!

Kanske ska svaret inte alls förstås som normativt. Kanske uttrycker det snarare ett estetiskt ideal. Om jag tar med paraplyet, så uppnår jag en viss sorts konsistens mellan min önsknings, mina trosföreställningar och mina handlingar.

Är det ett bra svar? Jag tycker inte det. Man kunde väl ha frågat mig en enklare fråga: *vill* du ta med paraplyet eller ej? Antag att jag hade svarat att jag inte vill ta med det. Om jag då hade avstått från att ta med det, så hade mina önsknings och mina handlingar "hängt ihop".

Beslutsteoretikern tycker kanske att detta slags sammanhang är för ytligt. Jag bör också ta hänsyn till vad jag tycker om de olika möjliga utfallen och om hur sannolika de framstår för mig.

Men *varför* ta hänsyn till just detta? Mera bestämt: varför ta hänsyn till vad jag tycker om de olika utfallen (till mina preferenser beträffande utfallen) om man inte ska ta hänsyn till vad jag tycker om själva handlingsalternativen?

Men mina preferenser för handlingsalternativen är inte ultimata, svarar kanske beslutsteoretikern. Sådana (instrumentella) preferenser kan just kritiseras ur rationalitetssynpunkt.

Det svaret är gott och väl då det gäller att visa att mina preferenser för handlingsalternativen är irrelevanta ur rationalitetssynpunkt. Men hur är det då med mina preferenser för utfallen? Är inte de också instrumentella?

Om jag inte vill bli blöt så beror det väl på att jag tänker att det är obehagligt att bli blöt, eller att jag kommer att ha besvär med att torka mina kläder, när jag väl har blivit blöt, osv. Jag kanske tror att jag kommer att göra en slät figur då jag anländer till en viss person, dryppande blöt. Men vad vet jag vid närmare eftertanke om allt detta? Vad *vet* jag om i vilken grad de olika "utfallen" representerade i matrisen tillfredsställer mina ultimata önskningar? Är jag bara en smula skeptiskt lagd måste jag medge att jag inte vet ett dugg om den saken. Här handlar det ju bl.a. om effekter på lång sikt. Jag kanske kommer över en del sociala hämningar, om jag blir blöt. Personen jag vill göra intryck på kanske får ett bättre intryck av mig om jag kommer blöt än torr.

Beslutsteoretikerns argument tidigare, att mina preferenser för handlingsalternativen är irrelevanta för frågan om rationaliteten av mitt beslut tycks pressa honom ännu ett steg tillbaka. Mina preferenser för de olika utfallen är lika irrelevanta, liksom mina ställningstaganden till de lotterier beslutsteoretikern presenterade mig med.

Vad beslutsteoretikern bör fråga mig är i stället vilka värden jag ser som fundamentala. För vart och ett av utfallen i matrisen ovan måste jag gå med på att det kan gå på många olika sätt med tillfredsställelsen av mina fundamentala preferenser. På så många olika sätt, i själva verket, att det är svårt att alls föreställa sig att vi kan rada upp dem och åsätta dem sannolikheter.

Kanske kan man för vart och ett av utfallen urskilja ett värsta och ett bästa möjliga utfall beträffande mina fundamentala preferenser? Kanske har de olika utfallen olika värsta möjliga utfall beträffande tillfredsställelsen av mina fundamentala preferenser. I så fall kan vi omvandla det som såg ut som ett enkelt beslut under risk till ett komplicerat beslut under osäkerhet. Men det finns inga okontroversiella sätt inom beslutsteorin att handskas med denna typ av *värderingsmässig* osäkerhet.

Det är också möjligt att vi tvingas konstatera att, för vart och ett av utfallen gäller, att vad som helst är möjligt, då det gäller tillfredsställelsen av våra fundamentala intressen. Värsta och bästa möjliga

utfall beträffande tillfredsställensen av våra fundamentala preferenser är desamma för de olika utfallen i matrisen ovan. Nu är det ännu mera uppenbart att beslutsteorikern är oförmögen att säga oss något om vad det är rationellt för oss att göra i situationen.

Har vi gått för långt då vi kräver att fundamentala preferenser, preferenser som vi har för olika tillstånd för dessa tillstånds egen skull, inte för några konsekvenser de kan tänkas föra med sig, ska beaktas av ett rationellt beslut? Borde vi ha tagit våra preferenser för utfallen i matrisen för givna?

Tja, det låter sig sägas. Men i så fall låter det sig också sägas att vi kunde ha tagit våra preferenser för *alternativen* för givna. Det är svårt att se något bra argument för att förkasta det senare men acceptera det förra. Går vi på den senare linjen uppnår vi också en avsevärd förenkling av beslutsteorin.