

P har lovat att posta ett brev. Det finns dessutom ingenting i den situation S som P befinner sig i som medför att han inte behöver uppfylla sitt löfte. Då gäller att:

(1) P bör (i S) posta brevet.

En i den gängse deontiska logiken accepterad princip, där a, b är godtyckliga handlingar, är:

(P₁) Om en agent X bör utföra a (i S) och den som utför a (i S) därmed (nödvändigtvis) också utför b, så bör X utföra b (i S).

Det klart att den som postar brevet därmed (nödvändigtvis) postar eller bränner det. Enligt (P₁), följer alltså av (1) att

(2) P bör (i S) posta eller bränna brevet.

Redan på denna punkt höjs kanske ett och annat ögonbryn: följer (2) verkligen av (1)? Detta är den ursprungliga (svagare) versionen av Ross' paradox.

En annan allmän princip som, åtminstone vid första påseende, kan verka övertygande är följande. Om a, b är handlingar, låter vi $\neg a$ vara (den negativa) handlingen att inte utföra a och $a \vee b$ (den disjunktiva) handlingen att utföra a eller utföra b. Vidare är $(a \rightarrow b) = (\neg a \vee b)$.

(P₂) Om en agent X bör utföra $a \vee b$ (i S) och inte utför a (i S), så får (kanske till och med bör) X utföra b (i S).

Av detta tillsammans med (2) följer att:

(3) Om P inte postar brevet, får han bränna det.

Vi har alltså med hjälp av (P₁) och (P₂) härlett (3) från (1). Men det är förståeligt att (3) inte följer av (1). (Detta är en annan, kanske mer övertygande, version av Ross' paradox.) En av (P₁) och (P₂) är alltså icke-giltig. Men vilken, om inte båda? Och varför?

Låt oss skriva O(a/S) för "agenten X bör utföra a i S", eller med andra

ord ”det är obligatoriskt för X att utföra a i S”. (Eftersom det här aldrig är fråga om mer än en agent kan vi utelämna alla referenser till X.)

Låt $S \Rightarrow a$ betyda att i S är a oundviklig. $S \Rightarrow \neg a$ betyder alltså att a är omöjlig i S och $S \Rightarrow (a \rightarrow b)$ betyder att den som i S utför a utför därmed (nödvändigtvis) b. (P₁) kan nu skrivas:

(P₁) Om $O(a/S)$ och $S \Rightarrow (a \rightarrow b)$, så $O(b/S)$.

Det är klart att $S \Rightarrow (a \rightarrow (a \vee b))$ för alla S. Av (P₁) följer alltså att:

(R₁) Om $O(a/S)$, $O(a \vee b/S)$,

dvs. den svagare formen av Ross' paradox.

Låt S-a vara en (den) situation som (i alla relevanta hänseenden) sammanfaller med S bortsett från det faktum att det i S-a är omöjligt (t.ex. för sent) att utföra a (t.ex. posta brevet samma dag).

Låt $T(a/S)$ vara en förkortning av $\neg O(\neg a/S)$. $T(a/S)$ betyder alltså att det är tillåtet för X att utföra a i S. En tänkbar formalisering av (P₂) är då:

(P₂) Om $O(a \vee b/S)$, så $T(b/S-a)$.

Detta tillsammans med (P₁) ger:

(R₂) Om $O(a/S)$, så $T(b/S-a)$,

(en version av) den starkare formen av Ross' paradox.

Det råder delade meningar om huruvida den svagare Ross-paradoxen verkligen är en paradox, dvs. om (R₁) är giltig eller inte. När det gäller (R₂), som naturligtvis är oproblematiskt paradoxal, har man föreslagit att detta visar att vi i själva verket rör oss med (två) olika begrepp *bör* (eller *obligatorisk*). För ett av dessa gäller (P₁) men inte (P₂) och för det andra gäller (P₂) men inte (P₁). Om detta vore sant skulle vi alltså ha kommit tillräta med den starkare Ross-paradoxen. Men det finns, såvitt jag kan se, inget (oberoende) skäl att tro att det förhåller sig så.

En annan utväg skulle kunna vara att ersätta (P₂) med en svagare princip, som möjligen bättre svarar mot våra intuitioner och som tillsammans med (P₁) inte implicerar (R₂) eller någon annan (liknande) paradoxal slutsats. Men detta har visat sig vara lättare sagt än gjort. Jag skall därför föreslå en annan lösning av problemet.

Ett *normativt system* för handlingar av en viss typ uttalar sig om vilka handlingar av denna typ som är påbudna, tillåtna etc. i vilka situationer.

Vi kan tänka oss detta i abstrakta termer på följande sätt. Låt M vara en godtycklig (ändlig) mängd. Elementen i M kallar vi *atomära handlingar*, dvs. handlingar som är, för det aktuella syftet, maximalt specificerade. I Ross' exempel är t. ex. att posta brevet (samma dag) en atomär handling;

det spelar t.ex. ingen roll vilken brevlåda P väljer (inom rimliga gränser). Att bränna brevet är likaledes i Ross' exempel en atomär handling, etc. Atomära handlingar är alltså parvis oförenliga. Med en *handling* (i mer teknisk mening) skall vi förstå en icke-tom mängd a av atomära handlingar. (När det gäller atomära handlingar a skiljer vi alltså inte mellan a och {a}.) Att *utföra* a är då detsamma som att utföra en av de i a ingående atomära handlingarna. En *situation* S är (också den) en icke-tom delmängd till M. Elementen i S är precis de atomära handlingar som är tillgängliga (möjliga) för X i situationen i fråga. (S är alltså också den handling som med nödvändighet blir utförd i S.) Jag hoppar över sådana komplikationer som hänger samman med att agenten ifråga inte alltid vet vilken situation han befinner sig i.

Om a och b är handlingar låter vi nu $\neg a$ vara komplementet av a med avseende på M och $a \vee b$ vara unionen av a och b. $S-a$ är komplementet av a med avseende på S.

Låt $t(S)$ vara mängden av de element i S som är *tillåtna* i S (enligt det normativa system vi har i tankarna). Med hjälp av t kan vi nu definiera ett (naturligt) begrepp *obligatorisk* (för X) i S på följande sätt:

(O) $O(a/S) =_{\text{def}} t(S)$ är inkluderad i a.

Det är då klart att med denna definition är (P₁) en giltig princip. (P₂), å andra sidan, är inte giltig.

För att förklara varför (P₂) trots detta kan förefalla plausibel skall vi titta på ett exempel. Vi kan betrakta t.ex. schack som ett normativt system. En atomär handling består då i att göra ett drag, dvs. att flytta en pjäs (av rätt färg) från en ruta till en annan. En situation är nu vad som vanligen kallas en ställning. I varje situation är vissa atomära handlingar tillåtna och andra inte.

Antag att A håller på med att lära B spela schack. En regel som B har vissa besvär med är att han inte får ställa sin kung i schack. B's kung står i schack och inte kan garderas. Man kan då tänka sig följande tre replikskiften.

- (I) A: Nu måste du flytta din kung till e1 eller e2.
 B: Då flyttar jag den till e1.
 A: Det får du inte eftersom den på e1 står i schack.
 B: Men du sa ju ...
- (II) A: Du kan flytta din kung till e1 eller e2.
 B: Då flyttar jag den till e1.
 A: Det får du inte eftersom den på e1 står i schack.
 B: Men du sa ju ...

(III) A: Du kan flytta din kung till e_1 eller e_2 .

(B flyttar sin kung till e_2 .)

B: Kunde jag inte också ha flyttat kungen till e_3 ?

A: Jovisst.

B: Varför sa du inte det?

I dessa exempel är B's respons på vad A säger fullt rationell. Det kan alltså tyckas som om B i t.ex. (I) förutsätter att vad A säger inte bara är att det är obligatoriskt, i den i (O) definierade meningen, att flytta kungen till e_1 eller e_2 , utan att detta är obligatoriskt i någon annan (starkare) mening.

Men en annan möjlighet är att A använder det i (O) definierade begreppet men att B väntar sig och tar för givet att A förser honom med så mycket information som B vet att A vet att B är intresserad av, som alltså är relevant för B, i den givna situationen. B vill veta exakt vilka drag som är tillåtna och detta vet A. Han vet också hur B kommer att uppfatta vad han säger: B kommer att ta för givet att det inte är förbjudet att flytta kungen till e_1 , eftersom detta skulle innebära att B måste flytta kungen till e_2 , något som B tar för givet att A skulle upplysa honom om. Därför är B's slutsats fullt rationell och vad A säger vilseledande även om det är bokstavligen sant. Och detta är i själva verket bara en instans av en allmän princip, som vi kan kalla principen om Maximal Relevant Information (MRI), som bestämmer hur vi uppfattar vad andra upplyser oss om och därmed vilka språkliga utsagor (påståenden) som är vilseledande (utan att kanske vara direkt falska). Och det är inte svårt att hitta på hur många andra exempel som helst på denna princip. (Jag hoppar över den komplikationen att avsändaren inte har tillgång till all relevant information.) T.ex.:

- När brukar du gå och lägga dig?
- Det är olika, men nästan alltid före två.
- Så sent!

Det är nu också klart vad som händer i replikskiftena (II) och (III). Om A säger att det är tillåtet att flytta kungen till e_1 eller e_2 i den givna situationen drar B, liksom i (I), i överensstämmelse med MRI, slutsatsen att exakt de drag som A nämner är de tillåtna dragen.

Att B's slutsats i (II), och liknande slutsatser i liknande sammanhang, är naturliga har många lagt märke till men, såvitt jag vet, ingen förklarat. Man har till och med påstått att det inte finns något att förklara: Det är helt enkelt så att "Det är tillåtet att a eller b" i vissa, men naturligvis inte alla, sammanhang, av någon anledning, betyder detsamma som "Det är tillåtet att a och det är tillåtet att b"; det är ingenting som behöver (eller kan) förklaras. Men varför, och i vilka sammanhang, och varför just dessa?

Det är nu också lätt att förstå varför redan den svagare versionen av Ross' paradox kan förefalla problematisk. I en situation i vilken det är korrekt att hävda att P bör posta brevet är det, enligt MRI, inte (alltid) korrekt att hävda att P bör posta brevet eller bränna det.