

Den Gyllene Regeln och Intra- och Interpersonella Viljekonflikter

Daniel Rönnedal

Abstrakt

Enligt den s.k. gyllene regeln bör vi behandla andra så som vi själva vill bli behandlade. Denna historiskt mycket inflytelserika princip är fortfarande en av de mest spridda och allmänt accepterade normer som någonsin formulerats. Alla världsreligioner tycks innehålla någon variant av denna maxim och mängder av filosofer har också accepterat den i en eller annan form. Men regeln har också kritiserats och det tycks finnas många möjliga tolkningar som har problematiska konsekvenser. I den här uppsatsen diskuterar jag en mängd argument emot den gyllene regeln som alla i någon mening går ut på att principen tillsammans med vissa andra rimliga antaganden är inkonsistent. Därefter visar jag hur dessa argument kan besvaras och hur den gyllene regeln kan försvaras. I ljuset av vår diskussion tycks vi kunna dra följande slutsatser. (1) Det finns tolkningar av den gyllene regeln som har problematiska konsekvenser och som vi därför troligtvis bör undvika. (2) Inte alla tolkningar av denna princip drabbas av de argument som diskuteras i den här uppsatsen. (3) Då man försöker avgöra värdet hos denna norm bör man försöka fokusera på de bästa möjliga tolkningarna.

1. Introduktion

Enligt den s.k. gyllene regeln bör vi behandla andra så som vi själva vill bli behandlade. Denna historiskt inflytelserika princip är en av världens mest spridda och accepterade normer. Alla världsreligioner, och många andra religioner, tycks innehålla någon variant av denna regel¹, och åtskilliga filosofer med olika moralfilosofiska teorier har också accepterat den i en eller annan form². Men regeln har också kritiserats och det tycks finnas många möjliga tolkningar som har problematiska konsekvenser.³ I den här uppsatsen diskuterar jag en mängd argument emot den gyllene regeln som alla i någon

¹ Neusner och Chilton (red.) (2008).

² Se t.ex. Hobbes (1985), Kapitel XIV, s. 190, Pufendorf (1964), Bok 2, 3:13, Mill (1987), Kapitel 2, s. 28, och Gensler (1996), särskilt Kapitel 5.

³ Se t.ex. Gensler (2013) för en sammanfattning av några potentiella invändningar. Gensler försvarar en särskild tolkning av den gyllene regeln mot dessa invändningar.

mening går ut på att principen tillsammans med vissa andra rimliga antaganden är inkonsistent. Därefter visar jag hur dessa argument kan besvaras och hur den gyllene regeln kan försvaras. I ljuset av vår diskussion tycks vi kunna dra följande slutsatser. (1) Det finns tolkningar av den gyllene regeln som har problematiska konsekvenser och som vi därför troligtvis bör undvika. (2) Inte alla tolkningar av denna princip drabbas av de argument som diskuteras i den här uppsatsen. Med andra ord, det finns tolkningar av den gyllene regeln som undviker de problem vi tar upp i den här artikeln. (3) Då man försöker avgöra värdet hos denna maxim bör man försöka fokusera på de bästa möjliga tolkningarna.

Den här uppsatsen är indelad i fem avsnitt. Avsnitt 2 innehåller en kort introduktion till den gyllene regeln. I Avsnitt 3 och 4 diskuterar jag sammanlagt sju olika argument emot olika versioner av denna princip. Avsnitt 3 innehåller tre argument som utgår ifrån *intrapersonella* viljekonflikter och Avsnitt 4 innehåller fyra argument som utgår ifrån *interpersonella* viljekonflikter. Alla argument bygger i någon mening på att man kan härleda en kontradiktion från någon version av den gyllene regeln om vi antar att det *finns* eller *kan finnas* intra- eller interpersonella viljekonflikter. Även om de olika argumenten liknar varandra, så skiljer de sig åt på väsentliga punkter, vilket motiverar att vi diskuterar dem alla. Problemet med interpersonella viljekonflikter omnämns bl.a. av Richard Whately (1856), och Don Locke (1981). Harry Gensler tar upp en variant av problemet i Gensler (2013), ss. 211–212. Men de få diskussioner som återfinns i litteraturen är relativt skissartade och jag känner inte till någon annan som har diskuterat dem närmare. Flera varianter av de argument jag behandlar tycks vara helt nya. Diskussionen i den här uppsatsen fyller därför ett gap i litteraturen. Avsnitt 5 innehåller en sammanfattning av uppsatsen och några slutsatser.⁴

2. Den gyllene regeln

Enligt den s.k. gyllene regeln bör vi behandla andra så som vi själva vill bli behandlade. Men hur skall man förstå detta och vad betyder de olika uttrycken i denna princip?⁵ Den gyllene regeln kan preciseras på en mängd

⁴ För mer information om den gyllene regeln, se t.ex. Blackstone (1965), Bruton (2004), Cadoux (1912), Carson (2010), Kapitel 6, Carson (2013), Duxbury (2009), Gensler (1986), Gensler (1996), särskilt Kapitel 5, (2013), Gewirth (1978), Gould (1980), Hare (1963), Hertzler (1934), Hirst (1934), Hoche (1978), Huang (2005), Neusner och Chilton (red.) (2008), Reinikainen (2005), Rönnedal (2015), Singer (1963), Wattles (1996) och Weiss (1941).

⁵ Rönnedal (2015) tar upp ett antal frågor som är relevanta då man försöker förstå den gyllene regeln. Se också Rönnedal (2016).

olika sätt och vi kommer i den här uppsatsen att nämna flera olika alternativ. Vi skall börja med att utgå från följande bokstavliga formulering:

(\square BGR). Det är nödvändigt att: Det gäller för alla individer x och y och alla handlingar H att: Om x vill att y utför handling H mot x , så bör x utföra H mot y .

(\square BGR) medför t.ex. följande satser: om du vill att din partner är ärlig mot dig, så bör du vara ärlig mot din partner; och om din partner vill att du behandlar henne rättvist, så bör din partner behandla dig rättvist. Vi antar i den här uppsatsen att den gyllene regeln medför den s.k. silverregeln, som säger att vi bör undvika att behandla andra på sätt som vi själva inte vill bli behandlade. Mer precist,

(\square BSR). Det är nödvändigt att: Det gäller för alla individer x och y och alla handlingar H att: Om x vill att y inte utför handling H mot x , så bör x inte utföra H mot y .

(\square BSR) medför t.ex. följande satser: om du vill att din granne inte stjäla från dig, så bör du inte stjäla från din granne; om du vill att din skolkamrat inte slår dig, så bör du inte slå din skolkamrat; och om Magdalena vill att Yvonne inte sprider lögn om henne, så bör Magdalena inte sprida lögn om Yvonne. Dessa konsekvenser av (\square BGR) och (\square BSR) tycks vara rimliga. Men vi skall nedan undersöka några problematiska implikationer av den gyllene regeln tolkad på detta sätt.

Notera att både \square BGR och \square BSR innehåller uttrycket ”Det är nödvändigt att”. Det finns flera olika typer av nödvändighet: logisk, metafysisk, naturlig, historisk m.m. Men för argumentens giltighet spelar det inte så stor roll hur vi väljer att precisera detta uttryck. Vi antar att nödvändigheten är en s.k. ”S5-nödvändighet”.

3. Argument från *intrapersonella* viljekonflikter

De tre första argumenten vi skall undersöka i den här uppsatsen utgår ifrån att det *finns* eller åtminstone tycks vara *möjligt* att det finns *intrapersonella* viljekonflikter. En *intrapersonell* viljekonflikt är en konflikt som rör en enskild persons vilja, till skillnad från en *interpersonell* viljekonflikt som involverar flera olika personer.

3.1. Argument 1: argumentet från en motsägelsefull vilja

Vårt första argument går ut på att det tycks vara möjligt att det finns personer som har en motsägelsefull vilja. Antag att det är så. Då kan vi – tillsammans med vissa andra rimliga antaganden – härleda en motsägelse från (\square BGR). Låt oss betrakta ett konkret exempel.

Antag att vi har två personer, a och b, som lever i ett stormigt förhållande med varandra. Båda har varit otrogna och båda misstänker att den andre har varit otrogen. b frågar nu a om a har varit otrogen. Bör a berätta sanningen för b eller inte? Vad säger den gyllene regeln? Antag att a har en motsägelsefull vilja; a vill att b berättar sanningen för a och a vill att b inte berättar sanningen för a. a vill att b berättar sanningen, eftersom a plågas av sina tvivel och eftersom a tror att det är bättre att sanningen kommer fram än att leva i en lögn. Samtidigt vill a att b inte berättar sanningen, eftersom det kommer att vara så oerhört smärtsamt om det visar sig att b har varit otrogen och eftersom det troligen kommer att innebära det definitiva slutet för förhållandet. Vi skulle också kunna säga att en del av a vill att b berättar sanningen och att en annan del av a vill att b inte berättar sanningen. Detta tycks vara möjligt. Människor tycks ibland vara ambivalenta, de tycks ibland vilja att något är fallet samtidigt som de vill att det inte är fallet. Men om detta är möjligt, kan vi härleda en motsägelse från (\square BGR) på följande sätt.

Argument 1

1. $\diamond(VaTba \wedge Va\neg Tba)$ i w_1 [Antagande]
2. $VaTba \wedge Va\neg Tba$ i w_2 [1, ML]
3. $VaTba$ i w_2 [2, SL]
4. $Va\neg Tba$ i w_2 [2, SL]
5. $VaTba \rightarrow OTab$ i w_2 [\square BGR, ML, PL]
6. $Va\neg Tba \rightarrow O\neg Tab$ i w_2 [\square BGR, ML, PL]
7. $OTab$ i w_2 [3, 5, SL]
8. $O\neg Tab$ i w_2 [4, 6, SL]
9. $(OTab \wedge O\neg Tab) \rightarrow O(Tab \wedge \neg Tab)$ i w_2 [AG, ML]
10. $O(Tab \wedge \neg Tab)$ i w_2 [7, 8, 9, SL]
11. $O(Tab \wedge \neg Tab) \rightarrow \diamond(Tab \wedge \neg Tab)$ i w_2 [BK, ML]
12. $\diamond(Tab \wedge \neg Tab)$ i w_2 [10, 11, SL]
13. $\neg \diamond(Tab \wedge \neg Tab)$ i w_2 [ML]
14. Falsum [12, 13, SL]

T_{xy} i argumentet ovan står för "x berättar sanningen för y". Vx , O och \diamond är satsoperatorer som tar satser som argument och ger satser som värde. VxA läses "x vill att A", OA läses "Det bör vara fallet att A" ("Det är obligatoriskt

att A”), och $\diamond A$ läses ”Det är möjligt att A”. ” \rightarrow ” är vanlig materiell implikation, ” \wedge ” vanlig konjunktion, och ” \neg ” vanlig negation. w_1 och w_2 står för möjliga världar. ”SL” är en förkortning av ”satslogik” och innebär att steget är satslogiskt giltigt. ”PL” är en förkortning av ”(högre ordningens) predikatlogik” och innebär att steget är predikatlogiskt giltigt. ”ML” är en förkortning av ”modallogik” och innebär att steget är giltigt i alla normala modallogiska S5-system. ”BK” står för ”Bör kan principen”, dvs. följande maxim:

(BK) Det är nödvändigt att: Det bör vara fallet (Det är obligatoriskt) att A endast om det är möjligt att A. $OA \rightarrow \diamond A$, för alla A.

”AG” är en förkortning av ”Agglomeration”, dvs. följande princip:

(AG) Det är nödvändigt att: Om det bör vara fallet att A och det bör vara fallet att B, så bör det vara fallet att A och B. $(OA \wedge OB) \rightarrow O(A \wedge B)$, för alla A och B.

Argumentet är uppenbart giltigt och vi har endast använt oss av mycket väletablerade ”slutledningsregler”. Om vi vill förkasta slutsatsen, måste vi alltså ge upp minst en sats i härledningen. Följaktligen måste vi överge minst en av satserna 1, 5, 6, 9, 11. Steg 9 följer ur (AG) och steg 11 följer ur (BK). Om vi överger (AG) kan vi således ge upp steg 9, och om vi överger (BK) kan vi ge upp steg 11. Både (AG) och (BK) är emellertid mycket rimliga principer. (AG) kan bevisas i alla s.k. normala deontiska system och (BK) är en intuitivt mycket rimlig tes. Jag har i andra arbeten försökt försvara uppfattningen att moralen är ”konsistent” (se Rönndal (2012), ss. 73–96 för mer information). Om (AG) och (BK) är sanna, måste vi förkasta 5, 6 och/eller 1. 1 är helt enkelt en formalisering av vårt ursprungliga scenario, och detta tycks vara möjligt. Men kanske är denna intuition inte riktig, kanske är det inte möjligt att det finns personer med en motsägelsefull vilja. I så fall kan vi ge upp steg 1 och hålla fast vid $(\square BGR)$. Personligen är jag dock benägen att tro att det är möjligt att ha en motsägelsefull vilja. Notera att det är tillräckligt för att vårt argument skall gå igenom att det är *möjligt* med en motsägelsefull vilja; argumentet kräver inte att någon *faktisk* har en sådan vilja. Men om steg 1 är sant, måste vi antingen förkasta steg 5 eller 6. Och steg 5 och 6 är logiska implikationer av $(\square BGR)$. Det följer att vi i så fall måste förkasta $(\square BGR)$. Vi kan emellertid inte sluta oss till att vi måste överge den gyllene regeln. För det är inte uppenbart att $(\square BGR)$ är den enda eller bästa tolkningen av denna maxim. Vi skall nu undersöka några andra

möjliga tolkningar av denna princip som undviker argument 1. Jag skall nämna fem möjliga interpretationer.

(i) Vi kan försvaga den gyllene regeln och hävda att den inte är nödvändig. Istället för (\square BGR) tolkas den gyllene regeln på följande sätt:

(BGR). Det gäller för alla individer x och y och alla handlingar H att:
Om x vill att y utför handling H mot x , så bör x utföra H mot y .

(\square BGR) är starkare än (BGR); (\square BGR) medför (BGR), men (BGR) medför inte (\square BGR). Ett argument som är problematiskt för (\square BGR) innebär därför inte nödvändigtvis några problem för (BGR). (\square BGR) medför 5 och 6. Men 5 och 6 kan inte härledas ur (BGR). Tolkar vi den gyllene regeln på detta sätt, undviker vi alltså den kontradiktoriska slutsatsen.

Låt mig nämna två möjliga kontraargument. För det första: allt annat lika, bör vi alltid formulera och tro på så starka teorier som möjligt. (\square BGR) är starkare än (BGR). Alltså bör vi föredra (\square BGR) framför (BGR). Detta är knappast något konklusivt argument för (\square BGR), eftersom vi just har sett hur vi kan härleda en motsägelse ur denna princip (tillsammans med andra rimliga premisser). Men om det finns andra möjliga lösningar på argument 1, kan det vara rimligt att överväga dessa istället.

För det andra, det tycks som om vi kan formulera ett giltigt argument som påminner om argument 1 där vi utgår från premissen att det *faktiskt* är så att a vill att b utför handling H mot a och att a vill att b *inte* utför handling H mot a . Och i så fall kan vi på nytt härleda en motsägelse med hjälp av (BGR). Detta kontraargument bygger på antagandet att det inte bara är *möjligt* att det finns någon med en motsägelsefull vilja, utan att det *faktiskt* finns någon med en sådan vilja. Personligen är jag benägen att tro att även detta är fallet (åtminstone vid något tillfälle). I så fall finns det problem även med (BGR).

Det är dock inte uppenbart att detta kontraargument är korrekt. Om det *faktiskt* inte finns någon med en motsägelsefull vilja, så drabbas inte (BGR) av de problem som tycks drabba (\square BGR). Det räcker inte med att visa att det är *tänkbart* att det finns någon med en motsägelsefull vilja eller att det är *möjligt* att det finns någon sådan individ för att vederlägga (BGR). Vi måste hitta ett verkligt exempel. Finns det faktiskt någon med en motsägelsefull vilja? Låt mig nämna hur en anhängare av (BGR) skulle kunna förneka detta och på så sätt undvika det aktuella problemet.

En anhängare av (BGR) skulle kunna göra en distinktion mellan en *allt taget i beaktande* vilja och vad vi skulle kunna kalla en *prima facie* vilja. Hon

kan vidare hävda att (BGR) handlar om en allt taget i beaktande vilja och inte en prima facie vilja. (BGR) skall med andra ord tolkas på följande sätt.

(BGR'). Det gäller för alla individer x och y och alla handlingar H att: Om x (allt taget i beaktande) vill att y utför handling H mot x , så bör x utföra H mot y .

Vi kan acceptera att det finns någon som prima facie vill att A och prima facie vill att inte- A , men förneka att det finns någon som allt taget i beaktande vill att A och allt taget i beaktande vill att inte- A . Att vilja något allt taget i beaktande är någonting annat än att drömma om att A eller att önska att A eller att anse att det finns vissa faktorer som talar för att A . Om man allt taget i beaktande vill att A , så är man fast besluten att A och har en intention att, om möjligt, se till att A . Och det är inte alls uppenbart att det finns någon som har en motsägelsefull vilja i denna mening. I vårt exempel ovan kan vi t.ex. hävda att a (prima facie) vill att b berättar sanningen och att a (prima facie) vill att b inte berättar sanningen. Men a har inte bestämt sig för vad a vill *allt taget* i beaktande. Det är inte så att a allt taget i beaktande vill att b berättar sanningen och att a allt taget i beaktande vill att b inte berättar sanningen; a velar fram och tillbaka. För att vederlägga (BGR') räcker det inte med att visa att det finns någon, a , som *prima facie* vill att någon, b , utför handling H mot a och *prima facie* vill att det inte är fallet att b utför H mot a . Vi måste hitta någon, a , som *allt taget i beaktande* vill att någon, b , utför handling H mot a , och *allt taget i beaktande* vill att det inte är fallet att b utför H mot a . Det är därför inte säkert att (BGR) (tolkad som (BGR')) drabbas av det aktuella problemet.

(ii) Vi kan tolka den gyllene regeln som en tumregel. Vi går med på att (\square BGR) inte är bokstavligt talat sann. Men den kan ändå vara användbar om de flesta instanser av den är sann. Denna position kan tyckas vara märklig. Om en sats inte är sann, hur kan den då vara användbar? En analogi kanske kan förklara tankegången. Antag att det finns en miljon svanar i världen. Alla dessa svanar är vita, utom en som är svart. Då är satsen: "Alla svanar är vita" falsk, men den kan ändå vara användbar för att göra förutsägelser om färgen hos enskilda svanar. Från påståendet att alla svanar är vita följer det att den svan *Kassandra* äger är vit. Det är förstås möjligt att den svan *Kassandra* äger är svart, men detta är mycket osannolikt. Om vi därför använder tumregeln att alla svanar är vita för att sluta oss till att *Kassandras* svan är vit, så är det mycket troligt att vår trosföreställning att *Kassandras* svan är vit är sann. På

liknande sätt kan det förhålla sig med normer som inte är sanna i en strikt mening.

Vi vill förstås helst formulera teorier som är strikt sanna. Så, om det är möjligt att formulera en tolkning av den gyllene regeln som är sann i en strikt mening är en sådan tolkning av den gyllene regeln att föredra framför en tolkning som endast är approximativt sann, allt annat lika. Tumregler och approximativt sanna principer bör emellertid inte föraktas. Filosofer har ibland varit alltför upptagna av att hitta absolut nödvändiga, universella och undantagslösa normer. Men faktum är att de flesta normer som människan har formulerat tycks vara ett slags tumregler som endast är approximativt sanna. Och sådana principer kan likväl vara mycket användbara.

(iii) Vi kan hävda att den gyllene regeln handlar om prima facie plikter och inte allt taget i beaktande plikter. Vi tolkar alltså den gyllene regeln på följande sätt.

(□PFGR). Det är nödvändigt att: Det gäller för alla individer x och y och alla handlingar H att: Om x vill att y utför handling H mot x , så är det en prima facie plikt att x utför H mot y .

Hur hjälper detta oss att undvika en motsägelse? (AG) är en rimlig princip om vi talar om allt taget i beaktande plikter, men inte om vi refererar till prima facie plikter. Det är möjligt att det är prima facie obligatoriskt att A och att det är prima facie obligatoriskt att B , samtidigt som det är falskt att det är prima facie obligatoriskt att A och B . Steg 10 i argument 1 är då inte berättigat. Därmed går deduktionen inte igenom.

(iv) Vi kan begränsa den gyllene regeln så att den endast handlar om den rationella viljan.

(□RGR). Det är nödvändigt att: Det gäller för alla individer x och y och alla handlingar H att: Om x vill att y utför handling H mot x och x 's vilja är rationell, så bör x utföra H mot y .

Med hjälp av denna princip kan vi inte längre härleda 5 och 6. Det är rimligt att anta att a 's vilja inte är rationell om a vill att A och a vill att inte- A . Vi kan kräva att alla fullständigt rationella personer, x , uppfyller följande villkor: $(\forall xA \wedge \forall xB) \rightarrow \forall x(A \wedge B)$ och $\forall xA \rightarrow \Diamond A$, dvs. om x vill att A och x vill att B , så vill x att A och B ; och x vill att A endast om det är möjligt att A . Antagandet att a är fullständigt rationell, att a vill att b berättar sanningen och att a vill att b inte berättar sanningen leder då till en kontradiktion, varför

det inte kan vara sant. Argument 1 går alltså inte igenom om vi använder denna tolkning av den gyllene regeln.

(v) Vi kan tolka den gyllene regeln som en s.k. *vid* villkorlig norm och inte en s.k. *snäv* villkorlig norm, dvs. vi antar att O-operatorn i den gyllene regeln har *vid* och inte *snäv* räckvidd. Vi kan med andra ord förstå den gyllene regeln på följande sätt:

(\square VGR). Det är nödvändigt att: Det gäller för alla individer x och y och alla handlingar H att det bör vara fallet att: Om x vill att y utför handling H mot x, så utför x H mot y.

Tolkad på detta sätt är den gyllene regeln ett slags konsistens-princip som talar om för oss hur vi skall vara konsistenta i vilja och handling, men den ger inte någon direkt vägledning om hur vi bör handla. Principen hävdar att följande kombination är otillåten: att vi vill att x utför H mot oss samtidigt som vi inte utför H med x. Om vi tolkar den gyllene regeln på detta sätt, kan vi inte längre härleda steg 5 och 6 från denna maxim. Alltså går argumentet inte igenom givet denna tolkning.

Med hjälp av alla dessa strategier kan vi alltså undvika argument 1. Man kan också tänka sig kombinationer av dessa, t.ex. av (i) och (iv). Vi kan t.ex. försvaga den gyllene regeln så att den inte är en nödvändig sanning och samtidigt begränsa den så att den endast varierar över personer med en rationell vilja. Vilken strategi som är bäst beror förstås också på vilka övriga argument för och emot den gyllene regeln det finns, och är en fråga som vi inte skall gå in på närmare här.

3.2. Argument 2: argumentet från en *direkt* motsägelsefull vilja

Vårt andra argument liknar det första och vår utvärdering av detta argument kan därför vara mer kortfattad. Det tycks vara möjligt att det finns personer som har en *direkt* motsägelsefull vilja (och inte endast en motsägelsefull vilja). Och om det är möjligt, så kan vi – tillsammans med vissa övriga rimliga antaganden – härleda en motsägelse från (\square BGR). Låt oss betrakta följande variant av det konkreta exemplet i Avsnitt 3.1.

Antag att vi har två personer, a och b, som befinner sig i exakt samma situation som personerna i Avsnitt 3.1. Den enda skillnaden är att vi antar att a har en *direkt* motsägelsefull vilja, dvs. a vill att b berättar sanningen för a och att b inte berättar sanningen för a. Detta tycks vara möjligt. Men om det är så, kan vi härleda en motsägelse från (\square BGR) på följande sätt.

Argument 2

1. $\diamond \text{Va}(\text{Tba} \wedge \neg \text{Tba})$ i w_1 [Antagande]
2. $\text{Va}(\text{Tba} \wedge \neg \text{Tba})$ i w_2 [1, ML]
3. $\text{Va}(\text{Tba} \wedge \neg \text{Tba}) \rightarrow \text{O}(\text{Tab} \wedge \neg \text{Tab})$ i w_2 [\square BGR, PL, ML]
4. $\text{O}(\text{Tab} \wedge \neg \text{Tab})$ i w_2 [2, 3, SL]
5. $\text{O}(\text{Tab} \wedge \neg \text{Tab}) \rightarrow \diamond(\text{Tab} \wedge \neg \text{Tab})$ i w_2 [BK, ML]
6. $\diamond(\text{Tab} \wedge \neg \text{Tab})$ i w_2 [4, 5, SL]
7. $\neg \diamond(\text{Tab} \wedge \neg \text{Tab})$ i w_2 [ML]
8. Falsum [6, 7, SL]

Argument 2 liknar argument 1, men det är ändå värt att ta upp argument 2 separat, eftersom inte exakt samma lösningsförslag tycks kunna användas i de båda fallen. Lösningsförslag (iii) i Avsnitt 3.1 tycks inte kunna användas i detta fall. Argument 2 innehåller ingen motsvarighet till steg 7–10 i argument 1. Så även om det skulle vara sant att (AG) inte gäller för prima facie plikter och den gyllene regeln antas handla om prima facie plikter, så kan vi inte undvika argument 2.

Man skulle kunna hävda att (BK) endast gäller för allt taget i beaktande plikter och inte för prima facie plikter. Om detta är riktigt, så är det möjligt att det är en prima facie plikt att A trots att det inte är möjligt att A. Antag att den gyllene regeln gäller för prima facie plikter. Då följer inte längre steg 6 i argument 2 från 4 och 5. Och på så sätt skulle vi kunna undvika den problematiska slutsatsen i argument 2. Personligen är jag emellertid benägen att tro att (BK) gäller även för prima facie plikter. Det är därför tveksamt om detta är en tillfredsställande invändning mot argument 2.

Kanske är det något mer rimligt att ifrågasätta premiss 1 i argument 2 än premiss 1 i argument 1. Kanske finns det *faktisk* ingen som har en direkt motsägelsefull vilja av det slag som premiss 1 i argument 2 antar. Men det tycks ändå vara *möjligt*. Och det räcker för att argumentet skall gå igenom.

Alla andra lösningsförslag som diskuterades i Avsnitt 3.1 kan emellertid även tillämpas för att undvika slutsatsen i argument 2. Argument 2 utgör därför inte ett problem för *alla* möjliga tolkningar av den gyllene regeln.

3.3. Argument 3: argumentet från en splittrad vilja

Vårt tredje argument är en version av ett argument som har diskuterats av Gensler. Så här skriver Gensler i en fotnot på s. 211 i Gensler (2013).

Suppose you're a broccoli-hating waiter and Becky orders broccoli. By Pyrite 1 [If you want X to do A to you then do A to X], you shouldn't serve her broccoli (since you don't want others to serve you

broccoli) and you should serve her broccoli (since you want others to bring you what you order).

Låt oss uttrycka en mer precis variant av detta argument. Vi antar att följande scenario är möjligt. Du är en servitris på en restaurang. Becky beställer broccoli av dig. Du gillar inte broccoli. Du vill att det inte är fallet att Becky serverar dig broccoli (eftersom du inte gillar broccoli). Du vill att om det vore fallet att du skulle beställa broccoli av Becky, så skulle Becky servera dig broccoli (du vill att andra serverar dig vad du beställer). Becky beställer broccoli av dig. Om detta är sant, så tycks vi kunna härleda en motsägelse ur (\square BGR) givet vissa andra rimliga antaganden. Låt oss uttrycka detta med hjälp av formella symboler. "Sxy" läses "x serverar y broccoli", och "Bxy" läses "x beställer broccoli av y", "d" refererar till dig och "b" refererar till Becky. "A $\square \rightarrow$ B" läses "Om A vore fallet, så skulle B vara fallet".

Argument 3

1. $Vd \rightarrow Sdb \rightarrow O \rightarrow Sdb$ [(\square BGR), PL, ML]
2. $Vd \rightarrow Sdb$ [Antagande]
3. $O \rightarrow Sdb$ [1, 2, SL]
4. $Vd(Bdb \square \rightarrow Sdb) \rightarrow O(Bdb \square \rightarrow Sdb)$ [(\square BGR), PL, ML]
5. $Vd(Bdb \square \rightarrow Sdb)$ [Antagande]
6. $O(Bdb \square \rightarrow Sdb)$ [4, 5, SL]
7. $O(Bdb \square \rightarrow Sdb) \rightarrow O(Bdb \rightarrow Sdb)$ [KD]
8. $O(Bdb \rightarrow Sdb)$ [6, 7, SL]
9. Bdb [Antagande]
10. $Bdb \rightarrow \square Bdb$ [HN]
11. $\square Bdb$ [9, 10, SL]
12. $(O(Bdb \rightarrow Sdb) \wedge \square Bdb) \rightarrow OSdb$ [AD]
13. $OSdb$ [8, 11, 12, SL]
14. $OSdb \wedge O \rightarrow Sdb$ [3, 13, SL]
15. $(OSdb \wedge O \rightarrow Sdb) \rightarrow O(Sdb \wedge \neg Sdb)$ [AG]
16. $O(Sdb \wedge \neg Sdb)$ [14, 15, SL]
17. $O(Sdb \wedge \neg Sdb) \rightarrow \diamond(Sdb \wedge \neg Sdb)$ [BK]
18. $\diamond(Sdb \wedge \neg Sdb)$ [16, 17, SL]
19. $\neg \diamond(Sdb \wedge \neg Sdb)$ [ML]
20. Falsum [18, 19]

Den här deduktionen liknar tidigare härledningar. Men argument 3 innehåller vissa steg som inte har några motsvarigheter i argument 1 och 2. De viktigaste nya stegen är 7, 10 och 12. Övriga nya satser i härledningen följer

med hjälp av väletablerade härledningsregler. 7, 10 och 12 är logiska konsekvenser av vissa bakomliggande allmänna antaganden, nämligen (KD) ("Kontrafaktisk deontisk logik"), (HN) ("Historisk nödvändighet") och (AD) ("Aletisk deontisk logik").

(KD) Om det är obligatoriskt att om A vore fallet så skulle B vara fallet, så är det obligatoriskt att om A så B. $O(A \square \rightarrow B) \rightarrow O(A \rightarrow B)$, för alla A och B.

(HN) Om A är en atomär sats, så är det historiskt nödvändigt att A. Om A är atomär, så $\square A$, för alla A.

(AD) Om det är obligatoriskt att A implicerar B och det är (historiskt) nödvändigt att A, så är det obligatoriskt att B. $(O(A \rightarrow B) \wedge \square A) \rightarrow OB$, för alla A och B.

Eftersom argument 3 förutsätter fler satser än argument 1 och 2, finns det ett antal nya strategier att kritisera denna deduktion. Låt oss se närmare på dessa.

(i) Vi kan överge (KD). Utan denna princip kan vi inte härleda steg 7, $O(Bbd \square \rightarrow Sdb) \rightarrow O(Bbd \rightarrow Sdb)$, i argument 3. Och utan detta steg är argumentet inte giltigt.

(KD) tycks emellertid vara en rimlig princip. Enligt de mest populära tolkningarna av kontrafaktiska villkorsatser är det logiskt sant att $(A \square \rightarrow B)$ implicerar $(A \rightarrow B)$.⁶ Och i alla s.k. normala deontiska system gäller det att om "A \rightarrow B" är ett teorem, så är "OA \rightarrow OB" ett teorem. Kombinerar vi "kontrafaktisk logik" med "deontisk logik", kan vi alltså härleda (KD). Vi tycks således ha goda skäl att hålla (KD) för sann. Strategi (i) kan därför ifrågasättas.

(ii) Det är möjligt att förkasta (AD) och om vi gör det kan vi inte härleda steg 12, $(O(Bbd \rightarrow Sdb) \wedge \square Bbd) \rightarrow OSdb$, från denna princip. Är det verkligen sant att om det är obligatoriskt att A implicerar B och det är historiskt nödvändigt att A, så är det obligatoriskt att B? Om steg 12 inte är sant, så går argument 3 inte igenom.

(AD) kan emellertid bevisas i alla s.k. normala aletisk-deontiska system som innehåller den s.k. mål medel principen⁷, och vi tycks därmed ha mycket

⁶ Se t.ex. Stalnaker (1968) och Lewis (1973).

⁷ Enligt mål medel principen gäller det att om det bör vara fallet att A och det är nödvändigt att A implicerar B, så bör det vara fallet att B.

goda skäl att acceptera (AD). Personligen tycker jag att (AD) är en intuitivt tilltalande maxim. Denna strategi kan alltså ifrågasättas.

(iii) Vi kan överge (HN). (HN) är onekligen en mer kontroversiell princip än många andra som används i argumentet och den är inte uppenbart korrekt. Om vi förkastar (HN), kan vi inte längre härleda steg 10, $Bbd \rightarrow \Box Bbd$, från denna princip. Och utan steg 10 faller argumentet. Är det sant att allt som faktiskt är fallet är historiskt bestämt? Om det är det, så misslyckas denna invändning mot argument 3. Det är rimligt att acceptera (HN) om det är rimligt att anta att nuet är historiskt bestämt. Nuet kan vara historiskt bestämt även om framtiden är öppen.

(iv) Ett annat sätt att undvika den kontradiktoriska slutsatsen i argument 3 är att begränsa den gyllene regeln så att den endast gäller för enkla handlingar och en kategorisk vilja, och inte för komplexa handlingar och en villkorlig vilja. Vi kan, med andra ord, formulera den gyllene regeln på följande sätt:

(\Box KGR). Det är nödvändigt att: Det gäller för alla individer x och y och alla *enkla* handlingar H att: Om x vill att y utför handling H mot x , så bör x utföra H mot y .

Tanken här är att vi *inte* kan instansera H med vilka handlingspredikat som helst. Om vi använder denna tolkning av den gyllene regeln, kan vi inte längre härleda steg 4 ifrån denna princip. Och utan steg 4 så faller argument 3.

Vi behöver alltså inte överge den gyllene regeln. Men detta svar är likväl problematiskt, även om vi inte längre kan härleda en motsägelse från (\Box KGR) på samma sätt som ovan. För om vi använder denna tolkning av den gyllene regeln tycks den medföra vissa kontraintuitiva påståenden i vårt exempel. Vi kan inte bevisa steg 4 från (\Box KGR), men vi kan fortfarande sluta oss till steg 1. Den gyllene regeln medför då att du *inte* bör servera Becky broccoli i vårt exempel. Men rent intuitivt så är det sant att du *bör* servera Becky broccoli. Om du arbetar som servitris, är det väsentliga inte vad *du* gillar och inte gillar utan vad *kunden* beställer.

(v) Ytterligare ett lösningsförslag är att kräva en ”samma situation-klausul” i den gyllene regeln. Från det faktum att du vill att Becky *inte* skall servera dig broccoli följer det då *inte* att du *inte* bör servera Becky broccoli. Det enda som är väsentligt är vad du skulle vilja om du befann dig i Beckys situation (och Becky befann sig i din situation) (eller vad du nu vill skulle

vara fallet om du befann dig i Beckys situation (och Becky befann sig i din situation)). Denna idé kan preciseras på minst två olika sätt:

(\square PGR). Det är nödvändigt att: Det gäller för alla individer x och y och alla handlingar H att: Om det är sant att om x befann sig i y 's situation (och y i x 's situation) så skulle x vilja att y utför handling H mot x , så bör x utföra H mot y .

(\square SGR). Det är nödvändigt att: Det gäller för alla individer x och y och alla handlingar H att: Om x vill att om x skulle befinna sig i y 's situation (och y i x 's situation) så skulle y utföra handling H mot x , så bör x utföra H mot y .

Även om alla nödvändiga sanningar är ekvivalenta, så är innehållen i dessa preciseringar inte ekvivalenta. Skillnaden mellan principerna är subtil, men betydelsefull. I (\square SGR) har vilje-operatoren vid räckvidd, i (\square PGR) har den snäv räckvidd. Jag kallar (\square PGR) för " \square PGR" eftersom innehållet i denna maxim tycks vara ekvivalent med den s.k. platinaregeln, som säger att om y vill att x utför H mot y , så bör x utföra H mot y . Om vi använder någon av dessa tolkningar av den gyllene regeln, kan vi inte längre härleda steg 1 i argument 3 från denna princip. Och utan steg 1 faller argumentet. Det kan finnas andra oberoende skäl att anta att den gyllene regeln bör begränsas på detta sätt. Lösningförslaget förefaller därför vara mycket intressant.

För att undvika argument 3 kan man alltså använda någon av strategierna (i)–(v) ovan. Förslag (i)–(iv) förefaller vara problematiska av olika anledningar, men (v) är ett förslag som är prima facie tilltalande. Dessutom tycks det som om man kan använda åtminstone några av strategierna från Avsnitt 3.1. Om vi t.ex. antar att den gyllene regeln handlar om prima facie plikter, kan vi förkasta steg 15 i argument 3. På så sätt kan vi undvika den kontradiktoriska slutsatsen. Olika kombinationer av de enskilda lösningarna är också möjliga. Däremot är det inte säkert att lösningförslag (iv) i Avsnitt 3.1 kan användas för att undvika slutsatsen i argument 3, åtminstone inte om inte detta förslag modifieras. Låt mig säga lite mer om detta.

Det är rimligt att anta att någon som är fullständigt förnuftig inte både vill att A och vill att inte- A . Men det aktuella exemplet tycks inte innehålla någon sådan "uppenbar" motsägelse. Vi har endast antagit att du vill att Becky inte serverar dig broccoli och att du vill att om du skulle beställa broccoli så skulle Becky servera dig broccoli, i symboler: $Vd \rightarrow Sbd$ och $Vd \rightarrow (Bdb \rightarrow \square)$

Sbd). Men detta tycks inte vara irrationellt, åtminstone inte på ett uppenbart sätt; $\{Vd \rightarrow Sbd, Vd(Bdb \square \rightarrow Sbd)\}$ tycks vara konsistent. Är det inte möjligt att din vilja är rationell trots att du vill att det inte är fallet att Becky serverar dig broccoli samtidigt som du vill att om du hade beställt broccoli från Becky, så hade Becky serverat broccoli till dig? Denna vilja tycks inte vara oförnuftig. I så fall kan argument 3 inte undvikas genom att begränsa viljan på det sätt som vi gjorde i lösningsförslag (iv) i Avsnitt 3.1.

Eftersom $\{Vd \rightarrow Sbd, Vd(Bdb \square \rightarrow Sbd)\}$ åtminstone inte på ett uppenbart sätt är inkonsistent, kallar jag det här argumentet för ”argumentet från en *splittrad* vilja” och inte ”argumentet från en *motsägelsefull* vilja”. Viljan är splittrad i den meningen att din kategoriska vilja pekar i en annan riktning än din hypotetiska vilja. Du vill att Becky *inte* serverar dig broccoli, samtidigt som du vill att *om* du hade beställt broccoli så skulle Becky ha serverat dig broccoli. Dessutom vill du *inte* bli behandlad på samma sätt av Becky som Becky vill bli behandlad av dig. Du vill att Becky *inte* serverar dig broccoli, men Becky vill att du serverar henne broccoli. Det tycks som om vi kan härleda en motsägelse ur $\{Vd \rightarrow Sbd, Vd(Bdb \square \rightarrow Sbd)\}$ endast tillsammans med den gyllene regeln. Det är inte viljan i sig som är problematisk. I så fall tycks det som om det inte hjälper att anta att den gyllene regeln handlar om en rationell vilja för att undvika argument 3.⁸

Detta problem påverkar emellertid inte lösningsförslag (v). Utgår vi ifrån (\square SGR) eller (\square PGR) kan vi förkasta steg 1 i argument 3. Och i så fall går argument 3 inte igenom.

4. Argument från *interpersonella* viljekonflikter

Våra fyra sista argument utgår ifrån att det tycks finnas *interpersonella* viljekonflikter, eller att sådana konflikter åtminstone tycks vara *möjliga*. Om det är riktigt, så kan vi – tillsammans med vissa andra rimliga antaganden – härleda en motsägelse ur vissa former av den gyllene regeln. Det grundläggande problemet ifråga finns omnämnt i litteraturen. Vi skall börja med att titta på Richard Whatelys, Don Lockes, och Harry Genslers formuleringar. Sedan skall vi undersöka fyra olika möjliga preciseringar av resonemanget.

⁸ Det finns sätt att modifiera denna typ av lösning. Om vi antar att en förnuftig vilja också är opartisk i en viss mening, kan vi undvika argumentet. Ett alternativ är att vi explicit anger att den gyllene regeln endast kan tillämpas på en opartisk vilja. Jag skall inte i detalj gå in på hur denna lösning kan utarbetas eftersom det skulle ta oss allt för långt bort från den här uppsatsens huvudtema.

Så här uttrycker Whately problemet:

... if you had to decide between two parties who were pleading their cause before you, you might consider that *each* of them wished for a decision in his *own* favor. And how, then, you might ask, would it be possible to apply the rule ? since in deciding *for* the one party you could not but decide *against* the other. A literal compliance with the rule, therefore, would be, in such a case, *impossible*. Whately (1856, s. 26)

Locke nämner ett par problem med den gyllene regeln i Locke (1981). Det är Lockes andra argument som intresserar oss i den här uppsatsen. Så här formulerar Locke detta problem:

The second difficulty is that adhering to the Golden Rule in respect of one person may require us to break the Golden Rule in respect of another: doing unto one person as we would have him do unto us may involve our doing unto some other person as we would not have that other person do unto us. The Golden Rule inevitably breaks down in those situations, and there will be many, where the preferences of the different parties conflict with each other. In fact, it will break down in just those situations where moral problems most commonly arise. Locke (1981, s. 536)

Gensler ger oss lite mer kött på benen eftersom han diskuterar ett konkret exempel på en interpersonell viljekonflikt. Den grundläggande poängen är emellertid densamma.

GR can issue contradictory commands when your action affects X and Y – since you may want A done when you put yourself in X's place but want A not done when you put yourself in Y's place.

Suppose I own a store and need to hire just one worker. Alice and Betty apply, and I must choose between them. Both are qualified, but Alice more so. When I consider myself in Alice's place, I desire that I be hired instead of Betty; so GR tells me to hire Alice and not Betty. But when I consider myself in Betty's place, I desire that I be hired instead of Alice; so GR tells me to hire Betty and not Alice. So GR gives contradictory commands... When our actions affect several

people, who gain or lose depending on our choice, GR may give contradictory commands. Gensler (2013, s. 211–212)

4.1. Argument 4:

argumentet (emot (\square SGR)) från en *indirekt* interpersonell viljekonflikt

Vi skall nu undersöka fyra möjliga preciseringar av detta grundläggande problem. Enligt de två första argumenten finns det en *indirekt* konflikt mellan olika personers viljor, och enligt de två andra argumenten finns det en *direkt* konflikt mellan olika personers viljor.

Argument 4

1. Om jag vill att om jag befann mig i Alices situation och Alice befann sig i min situation så skulle Alice anställa mig, så bör jag anställa Alice. [\square SGR]
2. Om jag vill att om jag befann mig i Bettys situation och Betty befann sig i min situation så skulle Betty anställa mig, så bör jag anställa Betty. [\square SGR]
3. Jag vill att om jag befann mig i Alices situation och Alice befann sig i min situation så skulle Alice anställa mig. [Antagande]
4. Jag bör anställa Alice. [1, 3, SL]
5. Jag vill att om jag befann mig i Bettys situation och Betty befann sig i min situation så skulle Betty anställa mig. [Antagande]
6. Jag bör anställa Betty. [2, 5, SL]
7. Om jag bör anställa Alice och jag bör anställa Betty, så bör jag anställa Alice och Betty. [AG]
8. Jag bör anställa Alice och Betty. [4, 6, 7, SL]
9. Jag bör anställa Alice och Betty endast om det är möjligt att jag anställer Alice och Betty. [BK]
10. Det är möjligt att jag anställer Alice och Betty. [8, 9, SL]
11. Det är inte möjligt att jag anställer Alice och Betty. [Antagande]
12. Falsum. [10, 11]

Detta argument innebär problem för (\square SGR). Steg 1 och steg 2 följer ur (\square SGR). Steg 3 och steg 5 är ”konsekvenser” av beskrivningen av vårt exempel. Steg 11 har ingen direkt motsvarighet i tidigare argument. I Genslers formulering av argumentet förekommer påståendet att jag måste välja mellan Alice och Betty. Detta påstående kan tolkas på lite olika sätt. Enligt en möjlig interpretation är det ekvivalent med steg 11. Övriga satser i argumentet bevisas i princip på samma sätt som tidigare. Det är relativt enkelt att formulera ett liknande argument emot (\square PGR) (se argument 5 nedan). Eftersom argument 4 innehåller flera nya satser, som inte har någon

motsvarighet i tidigare argument, kan detta argument också potentiellt kritiseras på ett antal nya punkter.

Flera av de lösningsförslag som tidigare har presenterats kan även användas för att undvika argument 4, t.ex. det förslag som går ut på att den gyllene regeln handlar om prima facie plikter. Tolkar vi den gyllene regeln på detta sätt, kan vi endast dra slutsatsen att det är prima facie obligatoriskt att jag anställer Alice och prima facie obligatoriskt att jag anställer Betty. Från detta följer det inte att det är prima facie obligatoriskt att jag anställer Alice och Betty. Inte heller följer det någonting om vad jag har för plikt allt taget i beaktande.

I det här skedet kan det vara värt att ta upp ett potentiellt problem med denna lösning. Tolkas den gyllene regeln på detta sätt, kan den inte ge någon vägledning i många fall. Bör jag t.ex. anställa Alice eller Betty? Om den gyllene regeln endast handlar om prima facie plikter ger den ingen entydig vägledning. Och många situationer tycks påminna om denna. Finns det inte nästan alltid interpersonella viljekonflikter som man bör ta hänsyn till?

Den här invändningen är emellertid inte konklusiv. I många fall kan det finnas andra faktorer i omgivningen som tillsammans med den gyllene regeln ger vägledning. Man kan hävda att man måste väga in alla faktorer som är relevanta i en viss situation för att komma fram till vad som är allt taget i beaktande obligatoriskt. I Genslers exempel är t.ex. Alice mer kvalificerad än Betty. Detta talar för att jag allt taget i beaktande bör anställa Alice och inte Betty. Liknande faktorer kan finnas i andra situationer.

(i) Låt oss nu undersöka en möjlig lösning som är ”unik” för argument 4. Vi har antagit att det inte är möjligt att ge jobbet både till Alice och till Betty. Men det är inte uppenbart att detta är sant. Om vi förkastar denna premiss, faller argumentet. Kanske *bör* jag inte anställa båda, men att hävda detta är inte detsamma som att hävda att det är *omöjligt*. Kanske måste jag inte välja mellan Alice och Betty. Kanske kan Alice och Betty dela på tjänsten; kanske kan de arbeta halva tiden var.

Vi kan emellertid kringgå detta kontraargument om vi antar att det *inte bör* vara fallet att jag anställer både Alice och Betty. Då kan vi med hjälp av steg 8 i argument 4 direkt härleda en kontradiktion. Även om det inte är *omöjligt* att jag anställer både Alice och Betty, så *bör* jag kanske inte göra det, eftersom jag inte har råd (låt oss stipulera att det förhåller sig så). Det skulle driva min firma i konkurs, det skulle ruinera mig och göra både Alice och Betty arbetslösa.

Vi kan också formulera en variant av argument 4 som innehåller en direkt interpersonell konflikt. Detta för oss över till argument 6 (se nedan). Först skall vi emellertid se hur vi kan formulera ett liknande argument emot (\square PGR) snarare än emot (\square SGR).

4.2. Argument 5: argumentet (emot (\square PGR)) från en *indirekt* interpersonell viljekonflikt

Argument 5

Vi byter ut steg 1, 2, 3, och 5 i argument 4 ovan mot 1', 2', 3', och 5' nedan. Övriga steg är desamma som i argument 4. Resultatet är argument 5.

1'. Om det är sant att om jag befann mig i Alices situation och Alice befann sig i min situation så skulle jag vilja att Alice anställer mig, så bör jag anställa Alice. [\square PGR]

2'. Om det är sant att om jag befann mig i Bettys situation och Betty befann sig i min situation så skulle jag vilja att Betty anställer mig, så bör jag anställa Betty. [\square PGR]

3'. Om jag befann mig i Alices situation och Alice befann sig i min situation så skulle jag vilja att Alice anställer mig. [Antagande]

5'. Om jag befann mig i Bettys situation och Betty befann sig i min situation så skulle jag vilja att Betty anställer mig. [Antagande]

Detta argument påminner om argument 4, och i stort sett samma invändningar som kan riktas mot argument 4 kan också riktas mot argument 5. Vi behöver därför inte säga mer om argument 5 i den här uppsatsen.

4.3. Argument 6: argumentet (emot (\square SGR)) från en *direkt* interpersonell viljekonflikt

Jag kallade argument 4 för ett argument från en *indirekt interpersonell* viljekonflikt. Anledningen till detta är att vi tänker oss att Alice vill att jag anställer Alice och Betty vill att jag anställer Betty samtidigt som det inte är möjligt att anställa båda. Att detta argument handlar om en *interpersonell* konflikt är lätt att se, eftersom det inte nödvändigtvis tycks vara fallet att Alices vilja *i sig* är motsägelsefull eller Bettys vilja *i sig* är motsägelsefull. Att det är en *indirekt* konflikt beror på att vi har antagit att det är *omöjligt* att jag anställer både Alice och Betty. Om det inte vore omöjligt för mig att anställa båda, skulle det inte finnas någon konflikt mellan Alices vilja och Bettys vilja. Att det rör sig om en *konflikt* beror på att det inte är möjligt att både Alices vilja och Bettys vilja tillfredsställs.

Argument 6 och 7 påminner om argument 4 och 5. Men i argument 6 och 7 antar vi en *direkt* viljekonflikt. Vi antar att Alice inte bara vill att jag skall anställa Alice utan även att jag *inte* skall anställa Betty, och att Betty inte bara vill att jag skall anställa Betty utan även att jag *inte* skall anställa Alice. Då kan vi formulera argument 6 på följande sätt.

Argument 6

1. Om jag vill att om jag befann mig i Alices situation och Alice befann sig i min situation så skulle Alice anställa mig och inte Betty, så bör jag anställa Alice och inte Betty. [(□SGR)]
2. Om jag vill att om jag befann mig i Bettys situation och Betty befann sig i min situation så skulle Betty anställa mig och inte Alice, så bör jag anställa Betty och inte Alice. [(□SGR)]
3. Jag vill att om jag befann mig i Alices situation och Alice befann sig i min situation så skulle Alice anställa mig och inte Betty. [Antagande]
4. Jag bör anställa Alice och inte Betty. [1, 3, SL]
5. Jag vill att om jag befann mig i Bettys situation och Betty befann sig i min situation så skulle Betty anställa mig och inte Alice. [Antagande]
6. Jag bör anställa Betty och inte Alice. [2, 5, SL]
7. Om jag bör anställa Alice och inte Betty, så bör jag inte anställa Betty. [OD]
8. Jag bör inte anställa Betty. [4, 7, SL]
9. Om jag bör anställa Betty och inte Alice, så bör jag anställa Betty. [OD]
10. Jag bör anställa Betty. [6, 9, SL]
11. Om jag bör anställa Betty och det bör vara fallet att jag inte anställer Betty, så bör det vara fallet att jag anställer Betty och att jag inte anställer Betty. [AG]
12. Det bör vara fallet att jag anställer Betty och att jag inte anställer Betty. [8, 10, 11, SL]
13. Det bör vara fallet att jag anställer Betty och att jag inte anställer Betty endast om det är möjligt att jag anställer Betty och att jag inte anställer Betty. [BK]
14. Det är möjligt att jag anställer Betty och att jag inte anställer Betty. [12, 13, SL]
15. Det är inte möjligt att jag anställer Betty och att jag inte anställer Betty. [ML]
16. Falsum [14, 15]

(i) Detta argument innehåller ett antal nya steg som kan ifrågasättas. Argumentets hållbarhet beror framför allt på två nya steg, steg 7 och 9. Steg 7 och 9 följer ur principen (OD) (O-Distribution), dvs. från följande sats:

(OD) Det är nödvändigt att: Om det bör vara fallet att A och B, så bör det vara fallet att A och det bör vara fallet att B. $O(A \wedge B) \rightarrow (OA \wedge OB)$, för alla A och B.

Så, om vi förkastar (OD) kan vi inte längre härleda steg 7 och steg 9 i argument 6 och utan dessa steg går argumentet inte igenom. (OD) är emellertid en intuitivt mycket rimlig princip och den kan också bevisas i alla s.k. normala deontiska system. Detta lösningsförslag kan därför ifrågasättas.

Argument 6 innehåller en *direkt interpersonell* viljekonflikt eftersom följande mängd satser är inkonsistent: {Jag anställer Alice och inte Betty, Jag anställer Betty och inte Alice}. Innehållet i Alices vilja är därför inkonsistent med innehållet i Bettys vilja. Det är inte logiskt möjligt att båda deras viljor tillfredsställs. Steg 15 kan bevisas i alla s.k. normala modallogiska system.

4.4. Argument 7: argumentet (emot (\square PGR)) från en *direkt interpersonell viljekonflikt*

Argument 7

Vi byter ut steg 1, 2, 3, och 5 i argument 6 ovan mot 1', 2', 3', och 5' nedan. Övriga steg är desamma som i argument 6. Resultatet är argument 7.

1'. Om det är sant att om jag befann mig i Alices situation och Alice befann sig i min situation så skulle jag vilja att Alice anställer mig och inte Betty, så bör jag anställa Alice och inte Betty. [\square PGR]

2'. Om det är sant att om jag befann mig i Bettys situation och Betty befann sig i min situation så skulle jag vilja att Betty anställer mig och inte Alice, så bör jag anställa Betty och inte Alice. [\square PGR]

3'. Om jag befann mig i Alices situation och Alice befann sig i min situation så skulle jag vilja att Alice anställer mig och inte Betty. [Antagande]

5'. Om jag befann mig i Bettys situation och Betty befann sig i min situation så skulle jag vilja att Betty anställer mig och inte Alice. [Antagande]

Detta argument påminner om argument 6. I princip kan samma invändningar som kan riktas mot argument 6 också riktas mot argument 7. Ett antal lösningsförslag som redan har nämnts kan användas för att undvika slutsatsen

i argument 6 och 7. Vi kan t.ex. anta att den gyllene regeln handlar om prima facie och inte allt taget i beaktande plikter.

Ett annat möjligt svar på argumenten i Avsnitt 4 utgår ifrån förslag (v) i Avsnitt 3.1. Enligt detta förslag är den gyllene regeln en s.k. *vid* plikt, inte en *snäv* plikt. Det här förslaget kan kombineras med andra förslag; vi kan t.ex. även kräva att den gyllene regeln innehåller en samma situation-klausul. Då kan denna princip formuleras på följande sätt:

(\square SVGR). Det är nödvändigt att: Det gäller för alla individer x och y och alla handlingar H att det bör vara fallet att: Om x vill att om x befann sig i y 's situation (och y i x 's situation) så skulle y utföra handling H mot x , så utför x H mot y .

Om vi vidare antar att (HN) inte gäller för satser som uttalar sig om vad någon vill, så kan vi använda denna version av den gyllene regeln istället för (\square PGR) och (\square SGR) för att undvika alla argument i Avsnitt 4. Utan (\square PGR) och (\square SGR) kan vi inte härleda vissa väsentliga premisser i de olika argumenten. Från (\square SVGR) följer det att det bör vara fallet att om jag vill att om jag befann mig i Alices situation (och Alice befann sig i min situation) så skulle Alice anställa mig, så anställer jag Alice. Det följer också att det bör vara fallet att om jag vill att om jag befann mig i Bettys situation (och Betty befann sig i min situation) så skulle Betty anställa mig, så anställer jag Betty. Antag att det bör vara fallet att jag inte anställer både Alice och Betty (eller att det är omöjligt att jag anställer båda och att den s.k. mål medel principen är sann). Då följer det att det bör vara fallet att jag inte både vill att om jag befann mig i Alices situation (och Alice befann sig i min situation) så skulle Alice anställa mig och vill att om jag befann mig i Bettys situation (och Betty befann sig i min situation) så skulle Betty anställa mig. Om jag vill att om jag befann mig i Alices situation (och Alice befann sig i min situation) så skulle Alice anställa mig samtidigt som jag vill att om jag befann mig i Bettys situation (och Betty befann sig i min situation) så skulle Betty anställa mig, så kommer jag då att bryta mot en norm. Men detta är förstås inte detsamma som att vi kan härleda en motsägelse ur våra antaganden. Det är inte logiskt omöjligt att bryta mot normer. För att undvika att bryta mot denna härledda norm måste den gyllene regeln s.a.s. satisfieras i relation till alla individer.

Ett potentiellt problem med denna lösning är att (\square SVGR) inte omedelbart är handlingsvägledande. (\square SVGR) är endast ett slags konsistens-princip som handlar om hur vi måste uppnå konsistens mellan vår vilja och

våra handlingar. Detta problem är emellertid inte konklusivt. Det är inte uppenbart att den gyllene regeln måste vara direkt handlingsguidande. Den gyllene regeln är intressant även som ett slags konsistens-princip.

(i) Låt mig slutligen nämna ett möjligt ”unik” lösningsförslag på argumenten i Avsnitt 4, utan att i detalj gå in på hur detta skulle kunna utvecklas. Man skulle kunna hävda att den gyllene regeln kan tillämpas endast på personer som *inte* är inbegripna i en interpersonell viljekonflikt. Om denna idé kan göras begriplig och försvaras, så tycks det som om argumenten i Avsnitt 4 skulle kunna bemötas. Uppenbarligen skulle mycket mer behöva sägas för att försvara en sådan lösning. Men jag skall nöja mig med att omnämna detta förslag i den här uppsatsen.

5. Slutsats

Enligt den s.k. gyllene regeln bör vi behandla andra så som vi själva vill bli behandlade. Detta är en mycket populär princip, men regeln har också kritiserats. I den här uppsatsen har jag diskuterat en rad argument emot den gyllene regeln som alla i någon mening går ut på att principen tillsammans med vissa andra rimliga antaganden är inkonsistent. Sammanlagt undersökte jag sju olika argument. Därefter försökte jag visa hur dessa argument kan besvaras och hur den gyllene regeln kan försvaras. I ljuset av vår diskussion tycks vi kunna dra följande slutsatser. (1) Det finns tolkningar av den gyllene regeln som har problematiska konsekvenser och som vi därför troligtvis bör undvika. (2) Inte alla tolkningar av denna princip drabbas av de argument som diskuteras i den här uppsatsen. (3) Då man försöker avgöra värdet hos denna norm bör man försöka fokusera på de bästa möjliga tolkningarna.

Låt mig slutligen säga någonting mera generellt om de olika lösningsförslagen i den här uppsatsen.

Det försvar som går ut på att tolka den gyllene regeln som en tumregel kan användas för att ”undvika” alla argument i den här uppsatsen. Denna ”lösning” innebär dock att vi antar att den gyllene regeln inte är bokstavligt talat sann. Allt annat lika är det förstås intressantare om vi kan hitta någon tolkning av den gyllene regeln som är bokstavligt talat sann.

Det förslag som hävdar att den gyllene regeln handlar om prima facie plikter och inte allt taget i beaktande plikter ”löser” alla argument förutom argumentet från en direkt motsägelsefull vilja (argument 2). Den tolkning av den gyllene regeln som förslaget bygger på är därför helt klart intressant.

Det försvar som går ut på att ”begränsa” den gyllene regeln till att handla om en rationell vilja kan användas för att ”undvika” argument 1 och 2. Om

man antar att rationalitetsbegreppet innefattar opartiskhet, kan man eventuellt även lösa vissa andra argument. Detsamma gäller om man ”utökar” rationalitetsbegreppet så att det innefattar interpersonell konsistens. Jag har dock inte i detalj försökt utarbeta dessa förslag i den här uppsatsen.

Det förslag som bygger på att vi inför en samma situation-klausul förefaller i sig endast lösa argument 3. Det kan därför tyckas som om detta förslag är relativt ointressant. Det kan emellertid finnas andra goda skäl att införa en sådan klausul. Tillsammans med andra modifieringar av den gyllene regeln är förslaget mycket intressant.

Om vi kombinerar det förslag som går ut på att den gyllene regeln handlar om en rationell vilja med det förslag som hävdar att den gyllene regeln handlar om prima facie plikter och inte allt taget i beaktande plikter, kan vi formulera en version av den gyllene regeln som undviker alla argument i den här uppsatsen. Denna form av den gyllene regeln ser ut på följande sätt:

(\square RPFGR). Det är nödvändigt att: Det gäller för alla individer x och y och alla handlingar H att: Om x vill att y utför handling H mot x och x 's vilja är rationell, så är det en prima facie plikt att x utför H mot y .

Den lösning som går ut på att tolka den gyllene regeln som en *vid* och inte en *snäv* plikt tycks kunna användas för att undvika alla argument i den här uppsatsen. (\square SVGR) är en mycket intressant tolkning av den gyllene regeln som också innehåller en samma situation-klausul.

Andra möjliga kombinationer kan också vara intressanta.

Vilken tolkning av den gyllene regeln är den bästa? Jag kommer inte att försöka besvara denna fråga i den här uppsatsen. Det finns en mängd andra argument för och emot olika versioner av den gyllene regeln som vi inte har diskuterat i den här uppsatsen. Personligen är jag benägen att tro att det finns flera olika tolkningar av den gyllene regeln som är rimliga.

Referenser

- Blackstone, W. T. (1965). The Golden Rule: A Defense. *Southern Journal of Philosophy*, ss. 172–177.
- Bruton, S. V. (2004). Teaching the Golden Rule. *Journal of Business Ethics*, Vol. 49, No. 2, ss. 179–187.
- Cadoux, A. T. (1912). The Implications of the Golden Rule. *International Journal of Ethics*, Vol. 22, Nr. 3, ss. 272–287.

- Carson, T. L. (2010). *Lying and Deception: Theory and Practice*. Oxford: Oxford University Press.
- Carson, T. L. (2013). Golden Rule. i Hugh LaFollette (red.) *The International Encyclopedia of Ethics*, ss. 2186–2192.
- Duxbury, N. (2009). Golden Rule Reasoning, Moral Judgement and Law. *Notre Dame Law Review* 84, ss. 1529–1605.
- Gensler, H. J. (1986). Ethics is Based on Rationality. *The Journal of Value Inquiry* 20, ss. 251–264.
- Gensler, H. J. (1996). *Formal Ethics*. London and New York: Routledge.
- Gensler, H. J. (2013). *Ethics and the Golden Rule*. New York and London: Routledge.
- Gewirth, A. (1978). The golden rule rationalized. *Midwest Studies in Philosophy*, 111, ss. 133–147.
- Gould, J. A. (1980). Blackstone's Meta-Not-So-Golden-Rule. *The Southern Journal of Philosophy*, Vol. 18, Issue 4, ss. 509–513.
- Hare, R. M. (1963). *Freedom and Reason*. Oxford: Oxford University Press.
- Hertzler, J. O. (1934). On Golden Rules. *International Journal of Ethics*, Vol. 44, Nr. 4, ss. 418–436.
- Hirst, E. W. (1934). The Categorical Imperative and the Golden Rule. *Philosophy*, Vol. 9, Nr. 35, ss. 328–335.
- Hobbes, T. (1985). *Leviathan*. Penguin Books. (red. C. B. Macpherson). (Ursprungligen publicerad 1651.)
- Hoche, H.-U. (1978). Die Goldene Regel. Neue Aspekte eines alten Moralprinzips. *Zeitschrift für philosophische Forschung*, Bd. 32, H. 3, ss. 355–375.
- Huang, Y. (2005). A Copper Rule versus the Golden Rule: A Daoist-Confucian Proposal for Global Ethics. *Philosophy East and West*, Vol. 55, Nr. 3, ss. 394–425.
- Lewis, D. (1973). *Counterfactuals*. Oxford: Basil Blackwell.
- Locke, D. (1981). The Principle of Equal Interests. *The Philosophical Review*, Vol. 90, Nr. 4, ss. 531–559.
- Mill, J. S. (1987). *Utilitarianism*. Buffalo. New York: Prometheus Books. (Ursprungligen publicerad 1863).
- Neusner, J. och Chilton, B. (red.) (2008). *The Golden Rule: The Ethics of Reciprocity in the World Religions*. Continuum.
- Pufendorf, S. (1964). *On the Law of Nature and Nations*. New York: Wildy and Sons. (Ursprungligen publicerad 1672).

- Reinikainen, J. (2005). The Golden Rule and the Requirement of Universalizability. *The Journal of Value Inquiry* 39, ss. 155–168.
- Rönnedal, D. (2012). *Extensions of Deontic Logic: An Investigation into some Multi-Modal Systems*. Department of Philosophy, Stockholm University.
- Rönnedal, D. (2015). The Golden Rule and The Platinum Rule. *The Journal of Value Inquiry*, Volume 49, Issue 1, ss. 221–236.
- Rönnedal, D. (2016). Den Gyllene Regeln och Substitutionsfunktioner. *Filosofiska Notiser*, Årgång 3, Nr 2, Augusti, ss. 53–80.
- Singer, M. G. (1963). The Golden Rule. *Philosophy*, Vol. 38, Nr. 146, ss. 293–314.
- Stalnaker, R. C. (1968). A Theory of Conditionals. I N. Rescher (red.), *Studies in Logical Theory*. Oxford: Blackwell, 1968.
- Wattles, J. (1996). *The Golden Rule*. New York, Oxford: Oxford University Press.
- Weiss, P. (1941). The Golden Rule. *The Journal of Philosophy*, Vol. 38, Nr. 16, ss. 421–430.
- Whately, R. (1856). *Introductory Lessons on Morals*. Cambridge: John Bartlett.

Daniel Rönnedal
Filosofiska institutionen
Stockholms universitet
daniel.ronnedal@philosophy.su.se