ANαliZA

časopis za kritično misel

Letnik 13, številka 3, november 2009 · ISSN 1408-2969

SVOBODA IN DETERMINIZEM

Peter van Inwagen

Modalni argument za inkompatibilizem

prevedel Danilo Šuster

Citiranje: Peter van Inwagen, »Modalni argument za inkompatibilizem«, *Analiza: časopis za kritično misel* 13, št. 3 (2009): 45–57.

Prevod. Licenca Creative Commons, ki velja za arhiv revije, za prevode ne velja; pravice izvirnika pripadajo njegovim imetnikom.

Izdajatelj: Društvo za analitično filozofijo in filozofijo znanosti, Ljubljana.

Elektronska različica je izvirna tiskarska datoteka številke, ki ji je bil leta 2026 dodan ta naslovni list. Strani in prelomi ustrezajo tiskani izdaji.

Peter van Inwagen
Prevedel Danilo Šuster

Modalni argument za inkompatibilizem*

3.1

Glavno vprašanje, okoli katerega se bije boj v sodobnih razpravah o svobodni volji, ni, kot bi pričakovali, ali *imamo* svobodno voljo. Vprašanje je, ali je svobodna volja združljiva z determinizmom. Zdi se mi, da sta svobodna volja in determinizem nezdružljiva, kar bom skušal dokazati v tem poglavju.

Razprave o tem vprašanju običajno niso na zelo visoki ravni. V večini primerov gre za dela kompatibilistov, ki inkompatibilistom pripisujejo kakšno prav otročjo napako (zamenjava »deskriptivnih« in »preskriptivnih« zakonov; nerazlikovanje med vzročno nujnostjo in prisilo; izenačevanje svobode in gole naključnosti). Donald Davidson sodi v to tradicijo, ko piše:

> Ne bom se neposredno ukvarjal z [argumenti za nezdružljivost svobode in vzročne determiniranosti], saj ne poznam nobenega, ki bi bil kaj več kot zgolj navidez prepričljiv. Hobbes, Locke, Hume, Moore, Schlick, Ayer, Stevenson in mnoštvo drugih so opravili vse, kar je bilo mogoče ali naj bi bilo sploh kadarkoli potrebno za odstranitev zmede, zaradi katere bi se lahko zdelo, da je determinizem v nasprotju s svobodo.[1]

V tej knjigi ni moj namen, da bi zagovarjal predhodne avtorje pred obtožbami za napake v argumentaciji. Moji lastni argumenti bodo jasni in vsaka zmota, ki jo zagrešijo, bi zato morala biti takoj vidna. (Dvomim pa, da so kogarkoli res zavedle zmote, za katere so ponavadi obtoženi inkompatibilisti; če je kdo zares dosegel takšno raven filozofske nesposobnosti, vsaj zase lahko rečem, da mi tega ni mogoče očitati). Ločnica med zagovarjanjem določenega stališča in pripisovanjem napak v sklepanju nasprotnemu stališču je morda v teoriji jasna, toda v praksi je pogosto zamegljena. Pozitivni argumenti za združljivost svobode in determinizma so zapredeni v obtoževanje nasprotnikov za napake in zmedo. Kolikor lahko te argumente razmotamo iz diagnoziranj bolezni, ki jih meni

* Prevedeno po Peter van Inwagen (1983). »Three Arguments for Incompatibilism.« V van Inwagen, P., *An Essay on Free Will*. Oxford: Clarendon Press, str. 55-56, in str. 93-105.

[1] »Freedom to Act«, v Ted Honderich, ur., *Essays on Freedom of Action*, (London: 1973), str. 139.

ni mogoče pripisati, če ta zla sploh obstajajo, se bom z njimi ukvarjal v četrtem poglavju.[2]

3.2

V prvem poglavju sem navedel naslednji enostaven argument za inkompatibilizem, Argument iz posledic:

> Če je determinizem resničen, tedaj so naša dejanja posledice zakonov narave in dogodkov v daljni preteklosti. Toda to, kar se je pripetilo, preden smo se rodili, ni odvisno od nas [*it is not up to us*] in zakoni narave prav tako niso odvisni od nas. Torej so tudi posledice teh stvari (vključno z našimi sedanjimi dejanji) od nas neodvisne.

To je, vsaj tako mislim, *dober* argument. Priznam pa, da je začrtan le v glavnih potezah. V tem poglavju ga bom skušal dopolniti na tri različne načine. (Naslov poglavja je »Trije argumenti za inkompatibilizem«, toda načelo individuiranja argumentov je precej nejasno, poglavje bi lahko imelo tudi naslov »Argument za inkompatibilizem – na tri načine«). Ti trije argumenti ali različice enega samega argumenta ali karkoli že so, naj bi se med seboj podpirali. Čeprav imajo v bistvu vsi isto poanto, se vendarle po strukturi in uporabljenih pojmih med seboj razlikujejo. Razlike med njimi bom uporabil na naslednji način. Če ima kdorkoli kakšne nejasne zadržke zaradi enega od argumentov – recimo zato, ker ni izražen v ničemer, kar bi bilo podobno običajnemu jeziku ali ker je njegova logična struktura precej zapletena in se lahko zazdi, da prikriva neko logično zmešnjavo, potem bo, vsaj tako *upam*, odkril, da se zadržki ne bodo pojavili, ko presoja enega od obeh drugih ali oba druga argumenta. Mislim, da se bo s tem pokazalo, da je vzrok za njegov občutek nelagodja neka *posebnost v načinu predstavitve značilnosti*, ki so skupne vsem trem argumentom in ne temelji na teh skupnih značilnostih samih. Zadnje možnosti seveda ne moremo logično izključiti: vzrok za občutek nelagodja, ki ga sproža eden od argumentov, je lahko še vedno prisoten, vendar pa ne deluje, ko razsojamo o drugem argumen tu, kajti ta vzrok, ki bi lahko bil ali pa tudi ne prava napaka v argumentih, je zdaj zakrit z znaãilnostmi drugega argumenta.

…

Op. Prev. Prvi argument v izvorni razliãici nastopa v van Inwagen, P. »Nezdružljivost svobodne volje in determinizma« v Šuster, D. ur. (2007). O svobodni volji. Od Leibniza do Libeta, Maribor: Aristej.

[2] Poznam samo eno zmoto, za katero obtožujejo inkompatibiliste, ki ni otročja. V delu *Will, Freedom and Power* (Oxford: 1976), str. 155-156, Anthony Kenny obtožuje inkompatibiliste, da uporabljajo naslednje pravilo sklepanja, ki je zanj neveljavno: »Ne morem storiti tega in tega. Storiti tako in tako je, v tem primeru, storiti to in to. *Torej*, ne morem storiti tako in tako«. (Inkompatibilist po Kennyju postavi »kršiti zakon narave« namesto »storiti to in to« in izraz, ki označuje neko neizvršeno dejanje, namesto »storiti tako in tako«). Naj bo ta oblika sklepanja veljavna ali ne – mislim, da Kenny ne predstavi nobenega jasnega protiprimera –, ni videti, da bi bil katerikoli od argumentov v tem poglavju od nje odvisen, čeprav so moji argumenti odvisni od naše nesposobnosti, da bi naredili za neresnično katerokoli propozicijo, ki je zakon narave.

Van Inwagen v Prvem argumentu takole definira determinizem:

»Izraz bo označeval konjunkcijo naslednjih dveh tez:

> Za vsak trenutek časa obstaja propozicija, ki izraža stanje sveta v tem trenutku.
> Če sta *p* in *q* propoziciji, ki izražata stanji sveta v določenih trenutkih, tedaj iz konjunkcije *p* in zakonov narave sledi *q*.

Zdi se mi, da ta definicija zajame vsaj eno od tez, ki jo imenujemo 'determinizem'. Determinizem je, intuitivno, teza, da glede na preteklost in zakone narave obstaja samo ena prihodnost. Iz te definicije to zanesljivo sledi. Posledica te definicije je tudi, da prihodnost določa edinstveno preteklost. To pa me ne skrbi. Edine *fizikalne* teorije, za katere vemo, da so deterministične 'iz preteklosti v prihodnost' (klasična mehanika dveh delcev in določene, tej matematično podobne teorije) so takšne, da so deterministične tudi 'iz prihodnosti v preteklost'. Obstajajo 'teorije' v določenem širokem smislu te besede, ki so deterministične samo 'v eni smeri' – recimo 'teorije', ki opisujejo obnašanje določenih Turingovih strojev. Ampak bil bi vsaj blago presenečen, če bi katerakoli plavzibina *fizikalna* teorija imela to značilnost. Če koga to zares moti, lahko po potrebi doda v definicijo determinizma besedici 'kasneje kot'. Ta dodatek ne bo vplival na mojo uporabo te definicije« (Van Inwagen, 1983: 65).

...

3.10

Tretji argument je modalni; v njem nastopa modalni operator, operator, ki ga damo pred stavke, ki imajo (ali izražajo propozicije, ki imajo) resničnostne vrednosti in z operatorjem tvorijo stavke, ki imajo resničnostne vrednosti. Pri tem pa resničnostna vrednost stavka, ki smo ga tvorili na tak način, ni vedno funkcija resničnostne vrednosti izhodiščnega stavka, ki smo mu dodali operator. Ta operator je '*N*'. Za vsak stavek *p* lahko tisto, kar dobimo, ko stavku *p* damo predpono '*N*', obravnavamo, kot da bi ob bok stavku *p* postavili »in nihče nima ali nikdar ni imel nobene izbire glede tega, da«. Tako je

> *N* Vsi ljudje so smrtni.

okrajšava za

> Vsi ljudje so smrtni in nihče nima ali nikdar ni imel nobene izbire glede tega, da so vsi ljudje smrtni.

'*N*' je zelo zanimiv operator in naloga izgradnje plavzibilne in popolne logike za '*N*' bi bila zelo zanimiva. Še bolj zanimivo bi bilo, če bi slovar te logike vključeval običajne »aletične« modalne operatorje. A za naše namene ni potrebno, da dejansko zgradimo takšno logiko.[3] Menim, da bi naslednji sklepanji morali biti veljavni pravili te logike:

(α) $\Box p \vdash Np$

(ß) $Np, N(p \rightarrow q) \vdash Nq$

Noben pravilno oblikovani izraz, ki vsebuje '*N*', ni logično resničen, razen tistih, ki so izpeljivi po teh dveh pravilih iz prazne množice premis. Operator '$\Box$' v pravilu (α) predstavlja »logično nujnost v širšem smislu«.

Tretji argument bo imel naslednjo logično zgradbo: iz logične posledice determinizma (ta je opredeljen tako kot v Prvem argumentu) bomo izpeljali, da nihče nikdar ni imel o ničemer nobene izbire.

V Tretjem argumentu bomo spet uporabljali simbola 'P_0' in '*L*', ki sta nastopala v predstavitvi Prvega argumenta. Ta simbola pa bosta imela rahlo drugačen pomen. V Prvem argumentu sta bila 'P_0' in '*L*' imeni določenih propozicij ['P_0' označuje propozicijo, ki izraža stanje sveta v trenutku T_0, in '*L*' označuje propozicijo, v kateri so konjunktivno združeni vsi zakoni narave (prim. van Inwagen, 1983: 70) – op. prev.]. V Tretjem argumentu ju bomo uporabljali kot okrajšavi za stavka, ki izražata ti propoziciji. V tej rabi imata 'P_0' in '*L*' slovnične značilnosti stavkov in ne terminov. V rabi Prvega argumenta je bil stavek 'P_0 je resnično' slovnično pravilno oblikovan, '$P_0 \rightarrow L$' pa ne (iz enakega razloga, kot je gramatikalno nedopusten izraz 'Quine → Davidson'). V sedanji rabi pa je '$P_0 \rightarrow L$' slovnično pravilno, izraz 'P_0 je resnično' pa ne (iz enakega razloga, kot je gramatikalno nedopusten izraz 'Vsi ljudje so smrtni je resnično').

Še enkrat bomo uporabili črko '*P*'. Gre za shematsko črko, namesto katere lahko postavimo poljuben stavek, če le izraža resnično propozicijo, recimo »Quina bega kvantificirana modalna logika« ali »Pariz je glavno mesto Francije«.

Če je determinizem resničen, potem sledi, da je resnično:

(1) $\Box [(P_0 \wedge L) \rightarrow P)]$

Iz (1) lahko po elementarni modalni in stavčni logiki izpeljemo

(2) $\Box [P_0 \rightarrow (L \rightarrow P)]$

[3] Predlagal bom sicer semantiko za '*N*', a le za ponazoritev poante, da bo vsaka zanimiva semantika za ta operator sporna.

Ko uporabimo pravilo (α) na (2), dobimo

(3) $N\,[P_0 \rightarrow (L \rightarrow P)]$

Zdaj uvedemo premiso:

(4) $N\,P_0$

Iz (3) in (4) po pravilu (ß) sledi:

(5) $N\,(L \rightarrow P)$

Uvedemo še drugo premiso:

(6) $N\,L$

Potem iz (5) in (6) po pravilu (ß) sledi:

(7) $N\,P$

Ta izpeljava kaže, da če je determinizem resničen, tedaj nihče nikdar ni imel nobene izbire glede ničesar, saj lahko vsak stavek, ki izraža poljubno resnico, nadomesti '*P*'. Vzemimo, denimo, vprašanje, ali je imel kdorkoli kakšno izbiro glede tega, ali bodo Richardu Nixonu oprostili kakršnokoli žalitev, ki jo je morda zagrešil kot predsednik. Po tem, če je determinizem resničen,

> $\Box$ [$(P_0 \wedge L) \rightarrow$ (Richardu Nixonu oprostijo poljubno žalitev, ki jo je morda zagrešil kot predsednik)]

> Če sta potemtakem pravili (α) in (ß) veljavni in če velja $N\,L$ ter $N\,P_0$, potem imamo:

> Nihče ni nikdar imel nobene izbire glede tega, ali Richardu Nixonu oprostijo poljubno žalitev, ki jo je morda zagrešil kot predsednik.

Večina bi nas zavrnila tak sklep kot neresničen, vsaj tako domnevam. Večina nas misli, da je Gerald Ford imel izbiro glede tega, ali Nixonu oprostijo. (Celo če kdo meni, da je Fordova dejanja povsem nadzirala neka tajna združba, še vedno misli, da je *nekdo* imel izbiro glede tega, ali Nixonu oprostijo). Ampak zgornja izpeljava (Tretji argument) kaže, da moramo, če želimo zavrniti sklep, zavrniti eno od naslednjih petih propozicij:

Determinizem je resničen;
$N\ P_0$;
$N\ L$;
Pravilo (α) je veljavno;
Pravilo (ß) je veljavno.

Moja izbira je seveda zavrnitev determinizma. A oglejmo si alternative.

Ne uvidim, kako bi lahko kdorkoli zavrnil *'N* P_0' ali *'N L*'. Moji razlogi so v bistvu tisti, s katerimi sem podpiral premisi *MAA* in *MAB* v Drugem argumentu ter premisi (5) in (6) v Prvem argumentu. Propozicija, da P_0, je propozicija o daljni preteklosti. Če želimo, lahko postuliramo, da gre za propozicijo o razporeditvi in gibalnih količinah atomov in drugih delcev v začetnih, predzvezdnih meglicah. Zato zanesljivo nihče nima nobene izbire glede P_0. Propozicija, da *L*, je propozicija, ki »zabeleži« zakone narave. Če je zakon narave, da se kotna gibalna količina ohranja, potem nihče nima nobene izbire glede tega, ali se kotna gibalna količina ohranja. Bolj splošno, ker je zakon narave, da *L*, nihče nima nobene izbire glede *L*.

Ne uvidim, kako bi lahko kdorkoli zavrnil pravilo (α). Če je (α) neveljavno, potem bi lahko kdo imel izbiro glede tega, kar je nujno resnično. Komaj kdo, izvzemši Descartesa, je bil pripravljen priznati kaj takega celo Bogu. Nihče, kolikor je meni znano, ni nikdar predlagal, da bi človeška bitja lahko imela izbiro glede tega, kar je nujno resnično. (Mislim, da »konvencionalistične« teorije modalnosti v resnici ne *implicirajo* teze, da imajo človeška bitja izbiro glede tega, kar je nujno resnično; vsaj ne v smislu, ki je relevanten za naše trenutne namene).

Obravnavati moramo le še pravilo (ß). Veljavnost (ß) je, mislim, izmed premis Tretjega argumenta premisa, ki jo je najtežje obraniti. Kako bi jo lahko branili? Kako, na sploh, pokažemo, da je pravilo sklepanja veljavno? Zdi se, da obstajata dva načina.

Najprej bi lahko uporabili metode formalne semantike. V našem primeru, ko je '*N*' modalni operator, bi se zdele obetajoče metode semantike *možnih svetov*. Tu je skica, kako bi lahko uporabili te metode na (ß).[4] Najprej določimo neko množico možnih svetov *W* in rečemo, da je *Np* resnično samo v primeru, ko je *p* resničen v vseh teh svetovih. To bi bila semantična definicija '*N*'. Bila bi definicija takšne vrste, kot je tista, ki jo podamo, ko rečemo, da je propozicija fizikalno nujna, če je resnična v vseh svetovih, v katerih veljajo zakoni narave, ali ko rečemo, da je propozicija moralno obvezujoča, če je resnična v vseh moralno dopustnih svetovih. Na primer, rekli bi lahko, da je *Np* resnično, če je *p* resnično v dejanskem svetu in v vseh nedejanskih svetovih, o katerih dejanskosti človeška bitja imajo ali so nekoč imela izbiro. Zanimivo je, da vprašanje, kateri svetovi sodijo v *W*, nima nobenega *formalnega* pomena. Če sprejmemo katerokoli definicijo *Np* oblike »*Np* je resnično, če in samo če je *p* resnično v vseh svetovih, takih, da ...«, kjer v manjkajočem pogoju (...) ni omenjen *p*,

[4] V preostanku tega odstavka bom malo manj pazljiv glede rabe in omembe, kot sem običajno.

»izide« (ß) kot veljavno pravilo. (Očitno, če je *p* resničen v vsakem elementu *W* in če je $p \rightarrow q$ resnično v vsakem elementu *W*, potem je tudi *q* resničen v vsakem elementu *W*). Kljub svoji popolni trivialnosti ta rezultat ni brez svoje prepričevalne moči. Vendar pa je daleč od tega, da bi bil odločilen. Odvisen je od domneve, da obstaja *neka* množica svetov, takih, da lahko o njih plavzibilno mislimo, da *Np* zatrjuje, da je *p* resničen v vsakem elementu te množice. Ta domneva se mi zdi pravilna, nimam pa zanjo nobenega argumenta in oseba, ki se je odločila, da bo zavrnila (ß), bi lahko to domnevo zavrnila.

Drugič, lahko bi poskusili pokazati, da je pravilo sklepanja veljavno, tako da ga »zvedemo« na določena splošno sprejeta pravila sklepanja. Torej tako, da njegovo veljavnost izpeljemo iz domnev, da so ta splošno sprejeta pravila veljavna. A zdi se malo verjetno, da lahko s to metodo pokažemo, da je pravilo (ß) veljavno: ne morem si misliti nobenega splošno sprejetega sklepanja, ki bi se zdelo relevantno za (ß), še manj pa takega, iz katerega bi ga lahko izpeljali.

Obeti, da bi *pokazali*, da je (ß) veljavno, se tako zdijo slabi, čeprav morda nič slabši od obetov, da bi *pokazali* karkoli filozofsko zanimivega. Priznati moram, da ima moje prepričanje v (ß) samo dva vira, eden je nesporočljiv, drugi pa ni odločilen. Prvi je tisto, kar filozofi radi imenujejo »intuicija«: ko pazljivo motrim (ß), se mi zdi veljavno. A ne morem pričakovati, da bo kdorkoli impresioniran nad tem dejstvom. Intuicije so navsezadnje ljudi vodile to sprejetja raznih vrst norih propozicij in mnogih normalnih, a neresničnih (načelo neomejene združitve reči v eno celoto v teoriji množic in Galilejev zakon seštevanja hitrosti v fiziki sta dobra primera propozicij, ki sodijo v drugo kategorijo). Drugi vir pa je, da si ne morem zamisliti nobenih uprimerjanj (instanc) pravila (ß), ki imajo ali bi lahko imela resnične premise in neresničen sklep. Vzemimo, denimo, naslednji uprimerjanji (ß):

> Alenka ima astmo in nihče nima ali je kdajkoli imel kakršnokoli izbiro glede tega, ali ima Alenka astmo.
>
> Če ima Alenka astmo, potem ima včasih težave z dihanjem in nihče nima ali je kdajkoli imel kakršnokoli izbiro glede tega, ali, če ima Alenka astmo, tedaj ima včasih težave z dihanjem.
>
> *Torej,* Alenka ima včasih težave z dihanjem in nihče nima ali je kdajkoli imel kakršnokoli izbiro glede tega, ali ima Alenka včasih težave z dihanjem.
>
> Sonce bo eksplodiralo leta 3000 in nihče nima ali je kdajkoli imel kakršnokoli izbiro glede tega, ali bo Sonce eksplodiralo leta 3000.
>
> Če Sonce eksplodira leta 3000, tedaj se bo vse življenje na Zemlji končalo leta 3000 in nihče nima ali je kdajkoli imel kakršnokoli izbiro glede tega, ali se bo, če Sonce eksplodira leta 3000, tedaj vse življenje na Zemlji končalo leta 3000.

Torej, Vse življenje na Zemlji se bo končalo leta 3000 in nihče nima ali je kdajkoli imel kakršnokoli izbiro glede tega, ali se bo vse življenje na Zemlji končalo leta 3000.

Ta dva argumenta sta očitno veljavna. Enostavno ni načina, kako bi lahko stvari uredili na način, ki bi bil zadosten za resnico premis in neresnico sklepa. Vzemimo drugi argument. Zamislimo si lahko, da bi lahko nekaj storili, da bi preprečili eksplozijo Sonca. Potem bi morda sklep argumenta bil neresničen. Toda potem bi prva premisa bila prav tako neresnična! Morda bi lahko zgradili ogromen ščit, ki bi varoval Zemljo pred eksplozijo Sonca. Če bi to lahko storili, bi bil sklep neresničen. Ampak potem bi bila druga premisa prav tako neresnična. Morda bi lahko razpršili nek strup, ki bi uničil življenje na Zemlji že pred letom 3000, a v tem primeru bi bila neresnična tudi druga premisa. Ne morem se znebiti občutka, da bo bralec, ki resno poskuša zgraditi protiprimer sklepanju (ß), začel ceniti intuicijo, ki sem jo izrazil z besedami, »ko pazljivo motrim (ß), se mi zdi veljavno«, čeprav je morda ne bo delil z mano.[5]

Zanimivo je opozoriti, da se zdi, da pravilo (ß) nastopa v nedavnih razpravah o filozofskih in socialnih implikacijah sociobiologije in vprašanju, ali so določene zelo razširjene značilnosti človeškega socialnega obnašanja genetsko determinirane. Vzemimo, denimo, propozicijo (*P*), da so določena dela (ki sta jih oba spola fizično sposobna opravljati), ki v vsaki družbi pripadajo skoraj izključno ženskam. Denimo, da sociobiolog trdi, da obstaja določeno dejstvo ali množica dejstev (*F*) o evolucijski zgodovini naše vrste, ki pojasni, zakaj je *P* resnično. Kdorkoli to trdi, bo verjetno tarča kakšne kritike naslednje vrste:

> To, kar praviš, je, da je »vloga žensk« genetsko determinirana in da so zato vsi poskusi, da bi spremenili družbeno vlogo žensk, obsojeni na neuspeh zaradi biologije. Tak nazor pa je poguben. Ti sploh nisi znanstvenik, ampak ideolog in ideologija, s katero krošnjariš, iz tebe dela uporabno oporo obstoječemu sistemu.

Kdorkoli bi bil žrtev takšne kritike, bi se verjetno branil na način podoben temule:

[5] Thomas Mckay je predlagal zanimiv možni protiprimer sklepanju (ß). Običajno bi rekli, da če kdo vrže pošteno kocko v pošteni igri, potem nihče nima izbire glede izida meta. Večina pa nas misli, da večina kockarjev ima izbiro glede tega, ali se vdajajo svojemu grehu. Toda potem bi naslednje uprimerjanje (ß) lahko enostavno imelo resnične premise in neresničen sklep:

N Alfred vrže šestico.

N (Alfred vrže šestico → Alfred meče kocke).

Torej, *N* Alfred meče kocke.

(Druga premisa je resnična, ker je pogojnik, ki nastopa znotraj '*N*', nujno resničen). *Respondeo*: v strogem smislu Alfred ima izbiro glede tega, ali vrže šestico, vsaj če ima izbiro glede tega, ali sploh meče kocko. Lahko se izogne temu, da bi vrgel šestico tako, da se sploh izogne igranju. Tisto, glede česar nima izbire, je, ali, *glede na to, da meče kocko*, vrže šestico. Se pravi, da velja, če privzamemo, da tudi nihče drug nima glede tega nobene izbire in da Alfred vrže šestico: *N* (Alfred meče kocko → Alfred vrže šestico).

> Ni res. Čeprav sem prepričan, da *F* pojasni, zakaj je *P* resnično, ne trdim, da je *P* zaradi *F neogibno* resnično. Glede na *F* obstaja tendenca, da je *P* resničen, a tendencam se lahko upremo. Ne trdim, da je »biologija usoda.« Čisto mogoče je, da imamo izbiro glede tega, ali se bomo obnašali v skladu s tendenco, s katero nas je obdarila naša biološka dediščina.

Ne zanima me toliko, kdo ima prav v razpravah te oblike – navsezadnje je to lahko odvisno od tega, kaj postavimo za '*F*' in '*P*' –, kot me zanima logika, ki je v ozadju razprave. Prepričan sem, da je logični skelet razprave nekaj takega kot:

KRITIK

> Iz tvojega stališča sledi, da so premise naslednjega veljavnega argumenta resnične:
>
> *N F* nastopi
>
> *N* (*F* nastopi → *P* je resnično)
>
> *Torej*, *N P* je resnično.

SOCIOBIOLOG

> Prva premisa je zanesljivo resnična, ampak druga *ne* sledi iz stališča, ki ga zastopam in je zelo verjetno lahko neresnična.

Sociobiologu verjetno nikdar ne bi padlo na pamet, da bi zanikal, da sklep argumenta, ki se mu pripisuje, dejansko sledi iz premis navedenega argumenta. In če bi to zanikal, potem bi ga kritik z vso pravico obtožil sofistke, saj če privzamemo, da nihče nima izbire glede tega, da če je naša zgodovina takšna in takšna, tedaj ravnamo tako in tako, in če prav tako privzamemo, da naša zgodovina *je* takšna in takšna in mi nimamo izbire glede tega, tedaj očitno sledi, da nihče nima izbire glede tega, ali ravnamo tako in tako. Iz tega seveda še ne sledi, da je (ß) veljavno, saj bi se lahko primerilo, da je to uprimerjanje (ß) veljavno, druga uprimerjanja pa ne. Ampak mislim, da kdorkoli bi rekel, da je argument, ki ga je formuliral kritik, veljaven, pravilo (ß) pa ni veljavno, bi s tem izrekel nekaj, kar se očitno ne zdi plavzibilno. Veljavnost (ß) se v vsej svoji splošnosti zanesljivo zdi del splošnega ozadja sociobiološke razprave, ki sem si jo zamislil. (Čeprav gre za zamišljeno razpravo, je njena logična zgradba značilna za razprave o biološkem determinizmu). Se pravi, zdi se, da načelo (ß) sprejemata, in to z vso pravico, obe strani v razpravi. Razprava se zdi enostavno o tem, glede *česa*, glede na sociobiologijo, sploh imamo kakšno izbiro in ne o vprašanjih glede veljavnosti sklepanj, v katerih nastopa »imeti izbiro glede«. Tisti, ki sprejemajo, ali so vsaj obtoženi, da sprejemajo »specialne determinizme«, se pravi teorije, ki pravijo ali jih včasih vsaj tolmačijo, kot da pravijo, da je nek pomemben aspekt človeškega obnašanja določen s tem ali onim

dejavnikom izven našega nadzora, se pogosto zapletejo v razprave o svobodi volje. Sociobiolog je moj primer takšne osebe, pač glede na trenutek mojega pisanja, ampak razprava o »biologija je usoda« ni edina te vrste. Če bi pisal pred nekaj leti, bi moral zgraditi primer, v katerem nastopa kak freudovec ali marksist, ampak poanta mojega primera bi bila enaka. Kdorkoli zanika veljavnost (ß), se mora, vsaj tako se mi zdi, odzvati na takšne razprave na enega od dveh načinov. Imeti jih mora za psevdorazprave, ki nastanejo zgolj zato, ker obe strani v razpravi napačno privzemata (ß), ali pa mora priznati, da razprava navsezadnje sploh ni o (ß). V prvem primeru bi moral sociobiologu priporočiti, da kritiku odgovori na naslednji način:

> Priznam, da je *F* dejstvo o zgodovini naše vrste in da gre za nekaj, glede česar nimamo nobene izbire. Priznam, da *če* je *F* dejstvo o zgodovini dane vrste, tedaj v družbah, zgrajenih iz pripadnikov te vrste, določena dela pripadajo skoraj izključno ženskam in da nihče nima glede *tega* nobene izbire. Ampak še vedno *imamo* izbiro glede tega, ali v naši družbi te zaposlitve pripadajo skoraj izključno ženskam.

Ne morem si zamisliti, da bi kdorkoli to lahko resno trdil. V drugem primeru pa nam je sociobiolog dolžan pojasnilo teh razprav, ki kaže, kako je lahko navidezno obojestransko sprejetje (ß) zgolj navidezno.

Poanto razprave lahko povzamem v vprašanju: zakaj ni nihče od udeležencev v razpravah o biološkem determinizmu kompatibilist? Morda je odgovor v tem, da je za udeležence v teh razpravah ideja biološkega determinizma mnogo bolj resna, kot pa je običajno za filozofe ideja »univerzalnega« ali »laplacovskega« determinizma in da je kompatibilizem glede na dano vrsto determinizma možen zgolj za tiste, ki tega determinizma ne jemljejo zelo resno.[6]

---

[6] Obstaja argument za fatalizem, ki je analogen predstavljenemu argumentu za genetski determinizem. Argument za fatalizem, ki smo ga obravnavali v prejšnjem poglavju te knjige, lahko predstavimo na način, v katerem eksplicitno nastopa načelo (ß). Naj bo *S* (resnična) propozicija, da se bom jutri obril in naj bo *t* nek trenutek v daljni preteklosti. S pomočjo pravila (ß) fatalist dokazuje:

N S je bilo resnično v *t*.
N (*S* je bilo resnično v *t* → jutri se bom obril)
Torej, *N* jutri se bom obril.

V prejšnjem poglavju (2.5) sem dokazoval, da prva premisa tega argumenta nima nobenega pomena ali pa da je, glede na določene predpostavke o tem, kaj bi lahko pomenila, dvomljiva. (Očitno nobena od teh možnosti ni na voljo v zvezi z argumenti tega poglavja. Propozicija, da P_0, dejansko *je* o preteklosti in torej o tem, kar je zunaj našega nadzora. Za propozicijo, da je *S* bilo resnično v *t*, pa se samo po nekem jezikovnem triku zdi, kot da bi bila o preteklosti. Zato je problem fatalizma problem, ki ga lahko razrešimo s pazljivo obravnavo jezika o času in resnici, problem svobodne volje in determinizma pa je globok, težko rešljiv in tak, da te spravlja ob pamet). To je teza, o kateri bi lahko imel s fatalistom resno filozofsko razpravo. A denimo, da bi, namesto da bi to rekel, jaz enostavno zanikal, da je zgornje uprimerjanje načela (ß) veljavno. Potem, mislim, bi me fatalist z vso pravico lahko obtožil, da odklanjam resno razpravo in s takšno obtožbo bi se strinjal tudi vsak antifatalist, ki bi zares želel priti do dna uganki fatalizma.

Zgoraj sem omenil, da si ne morem zamisliti nobenega uprimejanja (ß), ki bi imelo ali bi lahko imelo resnične premise in neresničen sklep. S tem sem seveda menil, da si ne morem zamisliti uprimerjanja (ß), ki bi imelo resnične premise in neresničen sklep neodvisno od vprašanja, ali je svobodna volja združljiva z determinizmom. Če je svobodna volja združljiva z determinizmom, in če je determinizem resničen, tedaj ima, domnevno, vsaj eno od naslednjih dveh uprimerjanj (ß) resnične premise in neresničen sklep.

N (P_0 → (L → Nixonu oprostijo))
N P_0
Torej, N (*L* → Nixonu oprostijo).

N (*L* → Nixonu oprostijo)
N L
Torej, N Nixonu oprostijo.

A bilo bi lepo videti protiprimer (ß), ki ne bi predpostavljal združljivosti svobodne volje in determinizma. Navsezadnje, primeri, s katerimi sem podprl veljavnost (ß), niso predpostavljali *nezdružljivosti* svobodne volje in determinizma. Mislil bi si, da če so *kakršnikoli* protiprimeri (ß), potem bi vsaj nekateri od njih bili takšni, da bi šteli kot protiprimeri neodvisno od vprašanja, ali sta svobodna volja in determinizem združljiva.

Morda je težko verjeti, a skoraj zagotovo obstajajo filozofi, ki bi rekli, da dejstvo, da iz determinizma in kompatibilizma sledi, da obstajajo protiprimeri k (ß), kaže, da moja raba (ß) v argumentu za inkompatibilizem predpostvalja, kar bi šele moral dokazati. Ampak če bi bila ta obtožba upravičena, potem je težko videti, kako se lahko katerikoli argument izogne temu, da ne bi že predpostavljal, kar šele mora dokazati. Če kdo predstavi argument za neko propozicijo in če je ta propozicija neresnična, potem mora biti *nek* korak v argumentu napačen (uporabo neresnične premise štejem kot »napačen korak«). In o nekem določenem koraku v argumentu smo lahko prepričani, da če je katerikoli korak napačen, potem je to *ta* korak. A iz tega še težko sledi, da s tem, ko uporabimo ta korak, že predpostavljamo, kar bi šele morali dokazati. Lahko da predpostavljamo, kar bi šele moral dokazati, karkoli že to natanko pomeni, toda to, da predpostavljamo, kar bi šele morali dokazati, ni avtomatsko posledica golega obstoja »najšibkejšega člena« v verigi sklepanja.

Vprašanja o tem, kaj pomeni predpostavljati to, kar bi šele morali dokazati, in kdo nosi breme dokaza, so zelo težavna. Oglejmo si jih iz drugačne perspektive. Denimo, da bi kompatibilist rekel nekaj takega: »Tretji argument je odvisen od veljavnosti (ß). A to pravilo je neveljavno. To dokazujem na naslednji način:

Kompatibilizem je resničen,
torej, (ß) je neveljavno.

Sam priznavaš, da sklep tega argumenta sledi iz njegove premise«. (To zares priznam). »Morda ne sprejemaš premise tega argumenta, a to je *tvoj* problem, kajti ta premisa je *resnična*. Še več. Težko ugovarjaš temu mojemu kratkemu argumentu rekoč, da že predpostavlja, kar bi šele morali dokazati. V tem pogledu ni nič slabši od tvojega argumenta, ki je v bistvu naslednji:

(ß) je veljavno,
torej, kompatibilizem je neresničen.«

Kaj naj rečem na to? Mislim, da mi ne preostane drugega, kot da se skličem na intuicije mojih bralcev. Takole so, po moje, videti te stvari (in ali niso takšne tudi za vas?): pravilo (ß) se zdi očitno pravilno, kompatibilizem pa se ne zdi očitno pravilen. Če sta si dve načeli v nasprotju in se eno zdi očitno pravilno, drugo pa se ne zdi očitno pravilno, potem, če moramo izbirati, naj bi sprejeli tistega, ki se zdi očitno pravilno.

No, denimo, da kompatibilist reče, da se mu zdi kompatibilizem očitno pravilen. Domnevno ta trditev temelji na takojšnji in intuitivni zvezi s kompatibilizmom – on trdi, da *vidi*, da je resničen, prav tako kot jaz trdim, da vidim, da je (ß) veljavno – ali pa temelji na nekem argumentu za kompatibilizem. Oglejmo si najprej primer filozofa, ki trdi, da intuitivno uvideva resnico kompatibilizma. Argumenti, tako kot pojasnila, se morajo nekje končati. Če morda obstaja kak tak filozof, potem on in jaz tvoriva avtentičen primer konflikta temeljnih intuicij. A moram reči, da je zame trditev, ki sem si jo zamislil, neverjetna. Kompatibilizem se mi zdi tista vrsta stvari, v katero bi lahko bili prepričani zgolj zaradi argumentov v njegovo podporo. Enostavno ne uvidim, kaj bi se lahko dogajalo v glavi tistega, ki bi trdil, da intuitivno pozna resnico kompatibilizma. Ne vem, kako bi se *občutilo* kaj takega.

Filozof, ki sprejema kompatibilizem zaradi *argumenta*, zame ni tako skrivnosten. A želel bom vedeti, kaj so premise njegovega argumenta. In o premisah bom postavil naslednje vprašanje: Ali so zares intuitivno bolj plavzibilne od (ß)? Zelo težko verjamem, da so kakšne propozicije, iz katerih sledi kompatibilizem, ki so bolj plavzibilne kot (ß). A kdo ve? Morda pa so. Kaj drugega lahko storimo, kot da preučimo primer za primerom? V razdelku 4.3 bomo primerjali (ß) s premisami priljubljenega argumenta za kompatibilizem, ki sem ga v prvem poglavju imenoval »Argument iz pogojne analize«.[7]

S tem je zaključena moja predstavitev Tretjega argumenta in moja obramba njegovih premis. Zdaj sem predstavil tri argumente za nezdružljivost svobodne volje in determinizma. Če kompatibilist želi zavrniti te argumente – in, seveda,

[7] V poglavju IV bomo obravnavali še dva druga argumenta za kompatibilizem, »Argument iz paradigmatskega primera« in »Argument iz revije *Mind*«. Premisi teh dveh argumentov ne bom primerjal z (ß). Izkazalo se bo, da so Argument iz paradigmatskega primera in dve različici Argumenta iz revije *Mind* pomanjkljivi na način, ki je veliko bolj resen, kot pa je sklicevanje na premise, ki so manj plavzibilne od (ß). V preostali različici Argumenta iz revije *Mind* – obstajajo tri različice – pa kot premisa nastopa veljavnost (ß).

nič ga ne sili, da to stori –, potem mora narediti tole: navesti bo moral neko množico propozicij, ki so intuitivno bolj plavzibilne od premis teh argumentov in pokazati, da iz teh propozicij sledi kompatibilizem. Ali pa bo moral iznajti argumente za neresnico nekaterih premis, ki nastopajo v tem poglavju, te argumente pa mora biti mogoče presojati in oceniti kot zdrave neodvisno od vprašanja, ali sta svobodna volja in determinizem združljiva.

Nekateri kompatibilisti, ko so soočeni z argumenti za nezdružljivost svobodne volje in determinizma, porečejo nekaj takega kot: »Tvoji argumenti enostavno kažejo, da ko ti uporabljaš fraze kot 'bi lahko ravnal drugače' ali 'ima izbiro glede', jim daješ pomen, ki se razlikuje od pomena, ki ga imajo v naših dejanskih razpravah o moralni odgovornosti«. Ta kritika je, *mutatis mutandis*, aplikabilna na vse tri argumente za inkompatibilizem. In moj odgovor je v bistvu enak v vsakem primeru. Ampak ta odgovor lahko predstavim na zelo zgoščen in učinkovit način v terminologiji Tretjega argumenta. Zatorej bom odgovoril samo na to obtožbo: »Ko uporabljaš frazo 'ima izbiro glede', ji daješ pomen, ki je precej različen od pomena v dejanskih razpravah o moralni odgovornosti«. Bralcu bom prepustil mehanično delo uporabe tega odgovora zahtevam Prvega in Drugega argumenta. Moj odgovor je enostavna reinterpretacija '*N*':

> $Np =_{df} p$ in v pomenu imeti izbiro, ki je relevanten za razprave o moralni odgovornosti, nihče nima ali ni nikdar imel nobene izbire glede p.

Če je sploh kaj na ugovoru, ki ga zdaj obravnavamo, tedaj je vsaj ena od naslednjih štirih propozicij,

$N P_0$,
$N L$,
Pravilo (α) je veljavno,
Pravilo (ß) je veljavno.

neresnična glede na to, da '*N*', ki nastopa v (α) in (ß), interpretiramo na zgoraj navedeni način. Ampak zdi se mi, da to ni res. Če pazljivo premotrimo vsak korak Tretjega argumenta, bomo, mislim, ugotovili, da noben korak ne postane dvomljiv v tem novem tolmačenju operatorja '*N*'. Sklep tega poglavja je torej, da če je determinizem resničen, tedaj nihče nima nobene izbire o ničemer, in to natanko v smislu *imeti izbiro*, ki je relevanten v razpravah o moralni odgovornosti.