ANαliZA

časopis za kritično misel

Letnik 8, številka 4, december 2004 · ISSN 1408-2969

SVOBODA IN VOLJA

Danilo Šuster

Hierarhija volje in svoboda

Pedagoška fakulteta Univerze v Mariboru

Citiranje: Danilo Šuster, »Hierarhija volje in svoboda«, *Analiza: časopis za kritično misel* 8, št. 4 (2004): 46–58.

To delo je objavljeno pod licenco Creative Commons Priznanje avtorstva 4.0 Mednarodna (CC BY 4.0), https://creativecommons.org/licenses/by/4.0/deed.sl.

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

Elektronska različica je bila leta 2026 rekonstruirana iz izvirnih stavnih datotek in preverjena s tiskano izdajo. Številke strani ustrezajo tiskani izdaji; prelomi strani znotraj prispevka se lahko nekoliko razlikujejo.

Danilo Šuster
Pedagoška fakulteta Univerze v Mariboru

HIERARHIJA VOLJE IN SVOBODA

Za zagovornike združljivosti svobode in determinizma smo svobodni takrat, kadar delamo to, kar hočemo. Toda, ali ni prava svoboda *volje* šele v tem, da imamo voljo, ki jo hočemo? Po Frankfurtovi *hierarhični* teoriji volje svobodno dejanje izvira iz želje, ki jo oseba hoče imeti in je usklajena z višjimi elementi dejavnikovega avtentičnega sebstva. V članku posebej opozorim na večkrat spregledano razlikovanje med tem, ali uživamo svobodo volje, tj. ravnamo po volji, s katero se *poistovetimo* in ki zadošča za moralno odgovornost, ter svobodno voljo, ki zahteva obstoj alternativ. Hierarhična teorija volje ima dve glavni težavi. Problem *regresa"^*, zakaj sploh sprejemamo hotenja drugega reda. Iskanje odgovora v hotenjih vedno višjih redov gre lahko v neskončnost. Problem manipulacije pa je, da so hotenja drugega in višjih redov lahko proizvod sil, nad katerimi nimamo nadzora. Problem regresa Frankfurt rešuje podobno kot fundacionalisti zaustavijo regres upravičevanja vedenja. Svobodno oseba se iskreno poistoveti s svojo voljo na tak način, da odobri hotenje višjega reda, s katerim je povsem *zadovoljna* in s tem zaustavi regres. Na problem možne manipulacije in vzročne determiniranosti hotenj višjega reda pa Frankfurt kot dosledni kompatibilist odgovarja, da je svoboda volje povsem interna strukturna lastnost, pri kateri zgodovina njenega nastanka ni pomembna.

Ključne besede: Frankfurt, oseba, svoboda volje in svobodna volja, hierarhija volje, problem regresa, problem manipulacije.

1.

> Za kompatibiliste, zagovornike združljivosti svobode in determinizma, svoboda tipično pomeni moč delovanja v skladu z določitvami lastne volje, kot pravi že Hume. Dejanja osebe so svobodna, kadar so odvisna od njenih motivov in ne od zunanjih sil (grožnje, hipnoze ...). Toda včasih zares ravnamo tako, kot želimo, pa vendar se nam zdi, da naše ravnanje ni bilo svobodno, saj nas vodi volja, ki si je niti najmanj ne želimo. Pomislimo na zasvojenosti, fobije in obsedenosti, ki ne

zajemajo le odvisnosti od raznih substanc, pač pa, denimo, tudi pretirano redoljubnost, zaradi katere ne prenesemo prahu in zato dneve preživljamo v obsedenem čiščenju vseh mogočih površin. Včasih nas prav naše želje ovirajo v tem, da bi storili tisto, kar *zares* hočemo (recimo, da bi bili malo manj redoljubni). Ali ni prava svoboda *volje* šele v tem, da imamo voljo, ki jo hočemo, in ne le, da delamo to, kar hočemo?

Harry Frankfurt (prevedeni članek, objavljen leta 1971) dokazuje, da sodobni kompatibilizem lahko vključi tudi refleksijo o željah in zgradbi volje. Po *hierarhični* teoriji volje (mimogrede – na Freuda tu pozabite!) svoboda pomeni, da so želje med sabo usklajene in "vpete" druga v drugo kot sistem zobatih kolesc pri uri, zato to teorijo včasih imenujejo tudi *hierarhična teorija zobatih koles* [*hierarchical mesh theory*]. Svobodno dejanje izvira iz želje, ki je usklajena z višjimi in bolj obsežnimi elementi dejavnikovega sebstva [*self*]. V tem članku bom predstavil Frankfurtovo teorijo v širšem okviru sodobnega kompatibilizma in komentiral nekatera težja mesta prevedenega članka. Posebej bom opozoril na razlikovanje med tem, ali uživamo svobodo volje, tj. ravnamo po volji, s katero se *poistovetimo*, in svobodno voljo, ki zahteva obstoj alternativ. V drugem delu pa bom analiziral dve težavi hierarhične teorije volje, problem regresa in problem manipulacije.

2.

Volja je po Frankfurtovi tehnični opredelitvi tista želja, ki dejavnika motivira za dejanje, dejansko učinkovita želja, zaradi katere stori, kar pač stori. Morda si že leta želim svojemu šefu končno po pravici povedati, kar mu gre, pa se to nikdar ne zgodi. Ta želja ne postane moja volja. Kadar pa želim prečkati cesto, to v veliki večini primerov tudi naredim in s tem uspešno udejanim svojo voljo. Predmet želje je običajno neko dejanje (prečkanje ceste), lahko pa je predmet želje tudi želja sama – želim si, da bi imel željo po jutranji telovadbi ali da ne bi imel želje po čokoladi. V tem primeru gre za željo *drugega reda. Hotenje* drugega reda pa je za Frankfurta želja drugega reda, da bi neka želja prvega reda postala moja volja. Od višje razvitih živali oseb ne loči dejstvo, da imajo želje in motive, pa tudi ne sposobnost izbiranja in odločanja. Bistvena značilnost *oseb* je, da imajo hotenja drugega reda, da so sposobne oblikovanja takšnih želja, s katerimi se poistovetijo. Samo oseba si lahko želi, da bi imela drugačno voljo. Volja zato ni samo neko psihološko stanje, ampak izraz značaja, trajnih dispozicij, kaj in kakšna oseba kdo je.

Frankfurtovo teorijo o zgradbi in svobodi volje lahko ponazorimo s tabelo, v kateri nastopa pet likov. V tabeli so naslednji simboli: 'A' in 'B' označujeta dejanji, ki sta predmet želje; '↓' označuje željo, ki si jo S želi; '⇓' označuje željo, za katero S hoče, da deluje; '||' označuje voljo (delujočo željo); '|' označuje naravnanost neke želje, ki ne postane volja; '×' pa označuje nasprotje med željami.

Odisej	*Hotnež*	*Oseba, ki uživa svobodo volje*	*Upirajoči se zasvojenec*	*Prostovoljni zasvojenec*
i) želi ($Ž_2$) željo $ž_1$ po A; ii) noče, da $ž_1$ deluje; iii) $ž_1$ ne deluje.	i) želi $ž_1$ po A in želi $ž_1$' po B; ii) zmaga močnejša želja.	i) hoče (H_2) željo $ž_1$ po A; ii) hoče, da $ž_1$ deluje; iii) $ž_1$ deluje.	i) hoče (H_2) željo $ž_1$ po A; ii) hoče, da $ž_1$ deluje; iii) deluje $ž_1$' po B.	i) hoče (H_2) željo $ž_1$ po A; ii) hoče, da $ž_1$ deluje; iii) $ž_1$ deluje, a ne (le) zaradi H_2.
$Ž_2$ ↓ ⇓ $ž_1$ × $ž_1$' | || A B	 $ž_1$ × $ž_1$' || | A B	H_2 ⇓↓ $ž_1$ × $ž_1$' || | A B	H_2 ⇓↓ $ž_1$ × $ž_1$' | || A B	H_2 ⇓↓ $ž_1$ × $ž_1$' || fiziološka A ← nujnost

V znani epizodi epa imajo vsi mornarji razen *Odiseja* vosek v ušesih, junak pa je privezan na jambor, ker želi slišati pesem siren in občutiti hrepenenje, s katerim te njihova pesem navda. Odisej je privezan, ker ne želi, da bi hrepenenje delovalo in ga izročilo sirenam v roke (saj bi ga te požrle), zato nima *hotenja* drugega reda, ampak samo željo drugega reda, da bi imel ustrezno željo prvega reda. Odisej si ne želi, da bi bila želja, ki jo želi imeti, tudi njegova volja. Oseba, ki je na dieti in ponoči hodi po prigrizke v hladilnik, pa morda zares hoče, da bi bila njena želja po hujšanju učinkovita, da bi to bila njena volja. Vendar ji to ne uspe. Teoretskih omejitev za hierarhijo želja ni. Oseba, ki je na dieti in hoče, da bi bila njena želja po hujšanju učinkovita, lahko po razmisleku ugotovi, da je njeno hotenje drugega reda plod propagande in manipulacij. Mediji in okolje ji vsiljujejo podobo vitkosti, ta ideal pa vodi do raznih anomalij in zdravstvenih težav (anoreksija ...). Saj je še kaj drugega v življenju bolj pomembno in zato si lahko želi, da bi ne imela hotenja drugega reda, da bi bila želja prvega reda po stradanju učinkovita.

Hotnež je bitje brez "drugega in višjih nadstropij", brez želja ali hotenj drugega reda. Ni mu mar za njegovo voljo, za "želenost" njegove želje. Takšno bitje je lahko zelo racionalno. Če ga, recimo, vodi želja po zbiranju znamk, si lahko zamislimo, da je tako inteligenten, da izumi stroj, ki vsako stvar, ki jo v stroj vtakneš, pretvori v vžigalice. Toda vsej silni inteligenci navkljub takšno bitje ni oseba in nikdar ne ravna po svoji svobodni volji, saj mu manjka osnovni predpogoj: hotenja drugega reda.

Osebo, ki uživa svobodo volje, v delovanju vodi želja, za katero hoče, da deluje. Svoboda dejanj je zadostitev želja prvega reda: dejanje je svobodno, kadar delam to, kar hočem. Svoboda volje je za Frankfurta v izpolnitvi hotenj drugega reda: dejanje je izvedeno *po* moji lastni svobodni volji, kadar gre za voljo, ki jo hočem imeti. Svoboda volje je v usklajenosti med različnimi ravnmi želja – hotenja drugega reda so usklajena z voljo, torej je želja, ki učinkuje, tista, za katero hočem, da me vodi v delovanju. Željo, ki me motivira, imam zaradi hotenja drugega reda, ker tako hočem hoteti. Za svobodo volje ni pomembno, ali obstaja

nek konflikt med nasprotujočimi si željami prvega reda, ki se na koncu razreši v prid hotenja drugega reda, ali pa, kot pravi Frankfurt, me motivira prijaznost, takoj ko si zaželim biti prijazen.

Svobodne so lahko samo osebe, hotnež pa zaradi odsotnosti hotenj drugega reda sploh ni oseba. Nesvobodna oseba je tista, ki je odtujena od svoje volje ali pa je popolnoma pasivna in nemočna opazovalka sil, ki jo ženejo. To je primer *upirajočega se zasvojenca*, ki vzame mamilo proti svoji volji. Nesvoboda je v razkoraku med voljo, ki jo hočemo imeti, in željo, ki zares deluje. Takrat prvi red deluje neodvisno od drugega in ne *zaradi* drugega reda. Kadar ropar zagrozi "Denar ali življenje!", takrat večina ljudi izroči denar. A ne po svoji volji, saj si ne želi, da bi želja po izročitvi denarja delovala ali da bi sploh imeli to željo (hotenje drugega reda bi denar raje zadržalo). Situacije nesvobode so pri osebah situacije notranjega konflikta. Tudi konflikt med hotenji drugega reda nas lahko uniči kot osebe, če nam prepreči, da bi se poistovetili s katerokoli od želja prvega reda in s tem paralizira našo voljo in nam odvzame moč delovanja.

Zanimiva in teoretsko zelo pomembna figura med Frankfurtovimi "liki" je *prostovoljni zasvojenec*. Ta oseba si želi vzeti mamilo, toda s to željo se tudi poistoveti. Morda sprejema načelo "Carpe diem!" in si po razmisleku izbere kratko, a užitkov polno življenje. Želja po mamilu tako postane del njenega značaja, del tega, kar je in želi biti. Volja upirajočega se zasvojenca ni svobodna, saj je v sporu z njegovimi hotenji višjega reda. Prostovoljni zasvojenec pa vzame mamilo po lastni svobodni volji, saj mu ta volja ni bil vsiljena, pri njenem oblikovanju ni bil zgolj pasivni opazovalec. Res pa je, da bi želja po mamilu zaradi fiziološke zasvojenosti *nujno* delovala, pa naj on to hoče ali ne! Kako gre potem lahko za svobodno voljo?

Frankfurt razlikuje med primerom, ko ravnamo z voljo, ki je svobodna [*acting with a will that is free*], in primerom, ko ravnamo po svoji lastni svobodni volji [*acting of one's own free will*]. V prvem primeru smo si voljo sami izbrali med različnimi alternativami. Kakršnakoli že je bila naša volja, lahko bi bila drugačna. Toda Frankfurta v večini besedila zanima samo drugi primer – kdaj oseba *poseduje* svobodo volje, kdaj *uživa* svobodo volje, kdaj ravna *po* svoji svobodni volji, kdaj *izpriča* svobodo volje. Svobodo volje uživam, če ravnam po volji, ki je meni *lastna*, s katero se *poistovetim* in jo *hočem*. To pa še ne pomeni, da je volja tistega, ki *uživa svobodo* volje, zares *svobodna*. Vprašanje, ali je volja dane osebe svobodna, je povezano z vprašanjem *alternativ*. Volja osebe je svobodna, če je glede na katerokoli željo prvega reda ta oseba svobodna, da naredi to željo za svojo voljo ali pa naredi kako drugo željo prvega reda za svojo voljo. Gre za pomembno razliko, ki jo zlahka spregledamo, in tudi Frankfurt sam nas zavede, ker pravi, da ga v njegovi teoriji *o* svobodi volje zanima vprašanje, ali je volja dane osebe svobodna in to vprašanje prevede v vprašanje, ali oseba ima voljo, ki jo hoče imeti zaradi hotenja drugega reda. Izkaže pa se, da usklajenost s hotenji drugega reda določa samo to, ali ravnamo *po* svoji svobodni volji, ne pa tudi tega, ali je bila ta volja svobodna v smislu, da je oseba izbirala med različnimi odprtimi alternativami, ki so ji bile na voljo tako v prvem kot v kakem višjem redu.

Ko Frankfurt pravi, da je svoboda volje v tem, da svobodno hočem, kar hočem hoteti, to pomeni, da želim (prvi red), kar hočem želeti (drugi red) – upirajoči se zasvojenec nima te svobode. Tradicionalno pa libertarci zagovarjajo nekoliko drugačno pojmovanje svobode: moja volja je svobodna, kadar je v moji moči tako, da hočem storiti X, kot da nočem storiti X. Karkoli sem izbral, vedno bi lahko izbral drugače. Libertarci tudi zahtevajo, da je svobodni dejavnik sam zadnji vzrok svojih dejanj in svoje volje (prim. Šuster, 1997 in 2002).

Frankfurt enkrat govori o tem, da je volja osebe svobodna [*a person's will is free*], drugič pa, da oseba uživa svobodo volje [*freedom of will*]. Mislim, da je vprašanje, ali je volja osebe svobodna, najboljše tolmačiti kot vprašanje, ali je volja svobodna v tradicionalnem, libertarskem smislu obstoja alternativ in vzročnega porekla te volje. Odgovor na vprašanje o svobodi ali nesvobodi volje pa odčitamo iz trenutnega posnetka zgradbe volje. Svoboda volje je tako povsem interna strukturna lastnost, pri kateri zgodovina nastanka te lastnosti ni pomembna, je t.i. "snapshot property". Definira jo sklad med željami različnih nivojev in ustrezna vzročnost: željo prvega reda imam *zaradi* hotenja drugega reda. Hotenje drugega reda mora nastopiti v vzročni verigi, ki vodi do svobodnega dejanja. Nesvoboda volje je tako ali odsotnost hierarhije ("hotnež") ali neintegriranost, zaradi katere volja, ki jo hočem, ne deluje.

3.

Razlikovanje med svobodno voljo in svobodo volje je pomembno še za eno veliko temo v razpravah o svobodi, za moralno odgovornost. Zgodovinsko dejstvo je, da je bil problem moralne odgovornosti glavni motor, pravzaprav *ratio essendi*, razprav o svobodi volje. Tradicionalno velja, da je svobodno dejanje tisto, za katerega smo moralno odgovorni. Dejanje osebe S pa šteje kot svobodno, če je avtor in zadnji izvor tega dejanja oseba S, ki bi lahko ravnala drugače, kot je.

Pomislimo na tipična opravičila, s katerimi se skušamo oprati moralne odgovornosti za dano dejanje. "Grozili so mi s smrtjo, zato sem jim izročil denar," se opravičuje bančni uslužbenec. Ker ni mogel ravnati drugače, ni moralno odgovoren. "Porinili so me, zato sem vas pohodil," se vam opravičuje potnik v natrpanem avtobusu. Dejanje ni bilo pod njegovim nadzorom, zato ni kriv. Če pa bi v na pol praznem avtobusu isti potnik rekel: "Skozi okno sem si ogledoval obleke v izložbi, zato nisem bil pozoren na vašo nogo," mu tega najbrž ne bi oprostili. Lahko bi vendar ravnal drugače, malo bolj pozorno premikanju po avtobusu je bilo zanj izvedljivo. Zdaj lahko predstavimo standardna pogoja za pripisovanje moralne odgovornosti:

> Oseba je moralno odgovorna za svoje dejanje, samo če ima nadzor nad tem dejanjem (dejanje je odvisno od njene volje).
>
> Oseba je moralno odgovorna za svoje dejanje, samo če bi lahko ravnala drugače, kot je.

Tradicionalno je veljalo, da sta pogoja ekvivalentna in izražata isto idejo:

svoboda volje je nujni pogoj za moralno odgovornost. Drugi pogoj ima tudi posebno ime: *Načelo alternativnih možnosti* (NAM). Podobno, kot je Gettier pretresel filozofsko sceno s tezo, da vedenje ni resnično upravičeno prepričanje, je Frankfurt pretresel filozofsko sceno s tezo, da načelo alternativnih možnosti ni pogoj za pripisovanje moralne odgovornosti (Frankfurt, 1969). Prostovoljni zasvojenec nima alternativ glede svoje volje, saj bi ga zasvojenost privedla do mamila, pa naj on to hoče ali ne. In vendar je moralno dogovoren za svoje dejanje. Kako je to mogoče? Začnimo z znamenitim primerom iz razprav o svobodi:

> "Denimo, da nekega moža, ki čvrsto spi, odnesejo v sobo, kjer je oseba, ki jo že dolgo želi videti in s katero želi govoriti. Sobo trdno zaklenejo, tako, da je mož ne more zapustiti. Mož se zbudi in je vesel, da je v želeni družbi, kjer ostane po svoji volji. Raje ostane v sobi, kot pa bi jo zapustil. Vprašam vas, ali ni ostal v sobi po lastni volji [*voluntary*]? Sodim, da nihče v to ne bo dvomil. Pa vendar, glede na to, da je zaklenjen v sobi, je očitno, da nima svobode [*liberty*] ne ostati, saj nima svobode [*freedom*], da bi sobo zapustil. Zato svoboda ni ideja, ki sodi k hotenju [*volition*] ali temu, kaj raje delamo, ampak k temu, ali oseba ima moč delovanja ali nedelovanja, tako kot se njen duh odloči ali usmeri. Naša ideja svobode seže samo tako daleč kot ta moč in nič dalje." (John Locke, *An Essay Concerning Human Understanding*, Knjiga 21, Razdelek 10)

Za Locka je pogoj za svobodo sposobnost drugačnega ravnanja. Če oseba ne more ravnati drugače, kadar bi to želela, potem je njeno dejanje sicer lahko hoteno [*voluntary*], vendar ni svobodno. Mož v sobi nima moči, da bi delal, kot hoče, zato za Locka njegova volja *ni* svobodna. Toda Locke je vseeno tipični kompatibilist, zanj je svobodno tisto dejanje, ki ga povzroča dejavnikova želja po njegovi lastni presoji. Vzročnost je zanj združljiva s sposobnostjo drugačnega ravnanja: dejanje osebe je lahko vzročno povzročeno z okoliščinami in hotenji, pa vendar bi ta oseba lahko ravnala drugače, če bi hotela drugače.

Za Frankfurta pa sposobnost drugačnega ravnanja ni pogoj ne za svobodo volje ne za moralno odgovornost. Ker mož ostane v sobi po lastni volji, je moralno dogovoren za svoje dejanje. Če pa mož tudi hoče imeti željo po tem, da ostane, potem po Frankfurtu ostane v sobi celo po svoji svobodni volji. Vzemimo enostaven primer iz vsakdanjega življenja. V trgovini kupujem kruh in me vprašajo, ali hočem belega ali črnega. Odločim se za črnega, potem pa mi rečejo: "Oh, še dobro! Belega je tako in tako zmanjkalo." V tem primeru sem dobil, kar sem hotel, nakup je bil po moji volji. Morda gre celo za reflektirano voljo: odločil sem se, da bom živel bolj zdravo, vem pa, da je črni kruh bolj zdrav. Beli se mi je sicer vedno zdel bolj okusen, toda hočem imeti željo po črnem. V tem primeru sem kupil kruh po svoji *svobodni* volji, čeprav alternativ sploh nisem imel!

Tezo, da načelo alternativnih možnosti ni pogoj za pripisovanje moralne odgovornosti, Frankfurt vpeljuje z naslednjim primerom (Frankfurt, 1969):

> "Denimo, da nekdo, naj bo to Black, želi, da Jones stori neko dejanje.

> Black se je pripravljen zelo potruditi, da bi dosegel svoj namen, vendar se ne želi razkriti, če to ni nujno potrebno. Black zato čaka, dokler se Jones ne odloči, kaj bo storil, in ne stori ničesar, razen če se bo Jones odločil, da bo ravnal drugače, kot to hoče Black (ki je izvrsten v presoji o tem, kako se bo Jones odločil). Če pa Black ugotovi, da se bo Jones odločil za nekaj drugega, ukrepa tako, da zagotovi, da se bo Jones odločil in storil tisto, kar hoče Black. Kakršnekoli so že izhodiščne preference, ki jih ima Jones, na koncu se bo zgodilo tako, kot hoče Black ... Predpostavimo pa, da se Blacku nikdar ni treba razkriti, saj se Jones zaradi lastnih razlogov odloči, da bo izvedel in tudi zares izvede tisto dejanje, ki ga želi Black. V tem primeru se mi zdi povsem jasno, da bo Jones moralno odgovoren za to, kar stori, v natanko tisti meri, kot bi bil, če ne bi bilo Blacka, ki ukrepa tako, da zagotovi, da bo Jones ravnal tako, kot hoče Black."

Ta primer je vir za številne bolj ali manj fantastične scenarije. Običajno Jones sovraži svojo ženo, recimo Barbaro, in zato načrtuje njen umor. Black, skrivni ljubimec, ki se je Barbare naveličal, prav tako sovraži Barbaro in hoče njeno smrt, vendar želi ostati skrit v ozadju. Black je zelo spreten znanstvenik nevrofiziolog, ki v glavo Jonesa, ne da bi ta to vedel, vgradi nek (biokemični?) mehanizem, ki bi ga Black sprožil takoj, ko bi, če bi, Jonesa minila mržnja in želja po umoru. Mehanizem deluje tako, da bi v tem primeru Jonesa zopet navedel k mržnji in umoru Barbare. Dejansko pa Blackova intervencija ni potrebna, Jones sledi svoji morilski želji in zagreši zločin sledeč svoji volji. Ali je moralno odgovoren za svoja dejanja? Seveda, bi rekla večina. Jones ni mogel ravnati drugače, kot je, to zagotavlja vgrajeni mehanizem, pa vendar bi ga vsako sodobno sodišče obsodilo. Splošna shema Frankfurtovih primerov je naslednja:

> Zaradi okoliščin C ("belega kruha ni več") je neogibno za S, da stori X ("kupi črnega"). Toda C ni razlog, zaradi katerega S stori X, S ravna po lastnem nagibu.

Ti primeri so, tako kot Gettierevi, bizarni in nenavadni. Vendar kažejo na nekaj pomembnega: moralna odgovornost je odvisna od *dejanske* zgodovine dejanja in ne od obstoja alternativnih scenarijev. Za moralno odgovornost je pomembno, da ima dejavnik ustrezno vlogo v tem, kar je dejansko vodilo do dejanja. Odgovornost pripisujemo na osnovi verige razmišljanj in notranjih motivov, ki so pripeljali do dejanja. Spomnimo se prostovoljnega zasvojenca. Z voljo, ki ga vodi do dejanja, se povsem poistoveti. Četudi bi lahko imel drugačno voljo, ne bi želel, da je njegova volja drugačna, kot je. Zato je moralno odgovoren za svoje dejanje, čeprav sta v tem primeru dejansko delovala dva neodvisna vzroka: njegova volja in stanje zasvojenosti, zaradi katerega sploh ni mogel ravnati drugače.

Frankfurt tako sprejema:

> Oseba je moralno odgovorna za svoje dejanje, samo če ima nadzor nad tem dejanjem (dejanje je odvisno od njene volje).

In ta pogoj razume kot:

> Oseba je moralno odgovorna za svoje dejanje, samo če je dejanje izvršila po svoji lastni volji [*freely, of one's own will)*.

Vendar to načelo loči od NAM:

> Moralna odgovornost za dejanje X *ne* zahteva, da je oseba svobodna storiti X v smislu, da lahko stori X ali pa lahko ne stori X, in ne zahteva, da je oseba imela možnost imeti voljo, kakršnokoli je hotela.

J. M. Fischer (1994 in 1999) pravi, da ima oseba, ki ravna po svoji lastni volji, *nadzor upravljanja* z dejanji [*guidance control*] in tak nadzor loči od *nadzora urejanja* dejanj [*regulative control*]. Nujni pogoj za nadzor urejanja je zares načelo NAM, toda za pripisovanje moralne odgovornosti zadošča nadzor upravljanja, kot kažejo Frankfurtovi primeri. Inkompatibilisti trdijo, da je NAM pogoj za pripisovanje moralne odgovornosti. Toda oseba bi po njihovem dokazovanju lahko ravnala drugače samo, če je teza o determinizmu neresnična. Če pa je determinizem resničen, ni še nihče nikdar bil odgovoren za ničesar, kar je kadarkoli storil, in ne bo nikdar imel moralne dolžnosti, da karkoli stori. Frankfurt je popolni kompatibilist: zanj sta tako svoboda volje kot moralna odgovornost združljivi z determinizmom – prva zahteva le ustrezno zgradbo volje in vzročno učinkovitost višjih redov, druga pa ne zahteva obstoja alternativ. Toda razlika med nadzorom upravljanja z dejanji in nadzorom urejanja dejanj pušča odprt prostor za novo možnost, *polkompatibilizem* [*semicompatibilism*]. Determinizem zares ne dopušča odprtih možnosti in zato izključuje obstoj prave svobode, zato popolni kompatibilizem ni resničen. Toda determinizem je vseeno združljiv z moralno odgovornostjo, saj ta zahteva samo nadzor upravljanja z dejanji. Polkompatibilizem, ki ga zagovarja Fischer, je ena od najbolj vročih tem v sodobnih razpravah o svobodi in moralni odgovornosti.

4.

Frankfurtovo teorijo volje so označili tudi kot teorijo *pravega sebstva* [*Real Self Theory*] (Wolf, 1990: 29). V hotenjih drugega reda (ali višjih redov) se manifestira pravo sebstvo, ki se po refleksiji poistoveti z dano željo prvega reda. Motivi in vrednote višjega reda zrcalijo avtentično sebstvo, ki je sposobno refleksivnega samoovrednotenja in samooblikovanja. S tem, kako gledam na svoje motive, s katerimi se poistovetim in katerimi ne, se v delovanju razkrije moj "globoki jaz" [*deep self*]. Tako so povezana vprašanja sebstva, osebe in svobode.

Frankfurtovi teoriji volje in moralne odgovornosti vsi priznavajo, da gre za pomembno izboljšavo kompatibilizma, ki pa ni brez težav. V tem sestavku bom obravnaval samo kritiko teorije volje, ne pa obširnih razprav o moralni odgovornosti in obstoju alternativ, ki so jih sprožili Frankfurtovi primeri. Hierarhična teorije volje ima dve glavni težavi. Problem *regresa* je, zakaj sploh sprejemamo hotenja drugega reda. Iskanje odgovora v hotenjih vedno višjih redov gre lahko v neskončnost. Problem *manipulacije* pa je, da samo hierarhija volje še ni dovolj za svobodo. Hotenja drugega in višjih redov so lahko proizvod

sil, nad katerimi nimamo nadzora, lahko so vsajena v nas ali pa smo tako "sprogramirani". Denimo, da se poistovetim z željo prvega reda: "Žrtvuj življenje, če tvoj guru to zahteva," ker sem tako indoktriniran od otroštva. Če je poistovetenje popolnoma slepo, ali sem kaj bolj svoboden od navadnega "hotneža", se sprašuje Double (1991: 35)?

Za osebo nastopi problem svobodne voljo takrat, ko so njene želje prvega reda v medsebojnem nasprotju. Hotenja drugega reda naj bi ta konflikt razrešila. Toda, če nastopi nasprotje in s tem problem svobode v prvem redu, ali ni možen konflikt tudi v drugem redu in še višje? Ali svoboda volje ne zahteva, da je tudi drugi red reflektiran in skladen z željami višjih redov? Toda možnosti refleksije so omejene in množenje vedno višjih redov ne more iti v nedogled. Frankfurt v izvornem članku pravi, da je takšno zaporedje mogoče prekiniti, ne da bi ga samovoljno presekali. Ko se oseba odločilno poistoveti z eno od želja prvega reda, ta zavezanost "odzvanja" skozi potencialno neskončno vrsto višjih redov. Odločilnost pomeni, da se je oseba odločila, da ni več nobenih vprašanj o hotenjih drugega reda, ki bi si jih oseba postavljala s kakšnega višjega reda. Ampak ta odgovor ni zadovoljil kritikov. "Želeli bi vedeti, kaj preprečuje hotnost glede hotenj višjega reda. Zakaj so ravno ta v posebni zvezi s sebstvom? Odgovor, da gre za odločilno zavezanost, ne pomaga, če to pomeni le to, da ne bomo dovolili neskončnega vzpona po hierarhiji hotenj. To je arbitrarno." (Watson, 1975: 17).

S problemom regresa se je Frankfurt ukvarjal v več sestavkih, njegovo končno stališče pa bomo morda najlaže razumeli, če problem regresa hotenj primerjamo s starim epistemološkim problemom regresa upravičevanja. Antični skeptik je začel svoj skeptični argument s tezo, da oseba S ve, da *p*, samo če ima ustrezen razlog za *p*. Toda, ali je ta razlog sam utemeljen ali pa je nekaj, kar S sprejema brez razloga? Če prvo, potem grozi neskončni regres utemeljevanja. Če drugo, ali ni to povsem arbitrarno, čisti dogmatizem – brez razloga bi lahko sprejeli že začetno propozicijo (prim. Šuster, 1997: 69)! Frankfurtova dilema pa je: dejanje *A* svobodno storim zaradi želje prvega reda, ki jo imam zaradi hotenja drugega reda. Toda kaj utemeljuje hotenje drugega reda? Spet se zdi, da sta alternativi le neskončni regres ali pa arbitrarnost.

Frankfurt je za prekinitev regresa najprej zahteval *iskrenost* [*wholeheartedness*] odločitve in *aktivno* poistovetenje z voljo, oseba in njena volja morata biti popolnoma *integrirani* (Frankfurt, 1987). Oseba se aktivno poistoveti z željo nižjega reda, kadar ima hotenje višjega reda, da bi delovala v skladu z željo nižjega reda, ki ji nič ne nasprotuje, in oceni, da nadaljnja presoja ne bi vplivala na njeno odločitev. Vzemimo spet prostovoljnega zasvojenca. Denimo, da je njegovo hotenje drugega reda učinek jemanja mamil, ki je vplivalo na njegova vrednotenja v drugem redu. Po Frankfurtovi prvotni teoriji deluje po lastni svobodni volji in je moralno odgovoren. Ampak to se ne zdi preveč verjetno. Prostovoljni zasvojenec, ki se odpove boju in *pasivno*, kot nekaj zunanjega, kar se mu je enostavno pripetilo, sprejema hotenje drugega reda, da bi imel željo po mamilu, po izboljšani teoriji ni svoboden. Iskrenost pomeni enotnost hotenj, nobenega spora ni med hotenji drugega in višjih redov, volja ni razdeljena ali

ambivalentna. Tudi iskrene osebe lahko naletijo na ovire svoji volji, toda kljub naporu in boju niso ambivalentne, ne želijo imeti drugačne volje.

Odločilen za zaustavitev regresa pa je kasnejši dodatek, pojem *zadovoljstva* [*satisfaction*].

Svobodna oseba se poistoveti s svojo voljo na tak način, da odobri [*endorse*] hotenje višjega reda, s katerim je povsem zadovoljna (Frankfurt, 1992: 13). Zadovoljstvo s samim seboj ne zahteva nekega novega *kognitivnega* odnosa, ki bi to zadovoljstvo utemeljeval in zato vodil v regres vedno višjih redov. Mislim, da lahko pogoj "biti popolnoma zadovoljen s stanjem svoje volje" primerjamo z osnovnimi prepričanji fundacionalistov, ki zaustavijo regres upravičevanja. Takšno osnovno prepričanje ni iracionalno, vendar ne rabi nadaljnjega upravičevanja. Zakaj ne? Ker je samoevidentno, pravijo nekateri; ker smo se odločili, da ga bomo sprejeli kot nedvomnega, pravijo drugi; ker dobi evidenco od nepojmovne vsebine določenih psiholoških stanj, ki sama niso prepričanja (zaznava, spomin), pravijo tretji. Frankfurtova rešitev je še najbolj podobna zadnjemu tipu epistemoloških fundacionalistov. Na vprašanje, zakaj imam hotenje reda *n*, lahko odgovorim: hotenje reda *n* imam *zaradi* hotenja reda *n* +2. V tem primeru pristanem v regresu. Lahko pa odgovorim: hotenje reda *n* imam, ker sem z njim popolnoma zadovoljen in tako končam razpravo. Regres zaustavi psihološko stanje zadovoljstva z danim hotenjem, vendar po Frankfurtu nisem niti iracionalen niti neupravičen v odobravanju takega hotenja.

Frankfurtovi končni rešitvi problema regresa tudi inkompatibilisti priznavajo vsaj delno uspešnost (Kane, 1996: 64). Menijo pa, da Frankfurt ne uspe rešiti problema manipulacije, ki zadeva izvor želja drugega reda. Po Frankfurtovi teoriji lahko uživam svobodo dejanj (delam, kar hočem), vendar nisem moralno odgovoren za svoja dejanja, ker so želje, ki tvorijo mojo voljo, "odrezane" od pravega sebstva – če sem hipnotiziran, če sem bil podvržen "pranju" možganov ... Toda tudi želje drugega reda so lahko manipulirane, pridobljene s pogojevanjem v otroštvu, določene z dejavniki, nad katerimi nihče nima vpliva. Z manipulacijo volje lahko dosežemo, da, recimo, sužnji po svoji "hoteni" volji delajo to, kar od njih hočejo gospodarji. Wolfova (1990) navaja naslednji primer. Jojo je sin Joprvega, zlobnega sadista in diktatorja v majhni in nerazviti državi. Deležen je posebne vzgoje, oče je njegov idol, ki ga povsod spremlja. Jojo sprejema očetove vrednote in kot njegov naslednik počne vse tisto, kar je delal oče: pošilja ljudi v ječo, odreja mučenja in terorizira svoje podanike, kakor se mu zahoče. Jojo ravna po svojih željah, ki pa so želje, ki si jih želi. Zaradi vzgoje in dednosti, ki nista pod njegovim vplivom, se zdi Wolfovi zelo vprašljivo, ali je JoJo moralno odgovoren za svoja dejanja. V njegovem položaju bi vsakdo postal perverzni sadist. Toda Jojo ima želje, ki jih hoče imeti, v tem je popolnoma iskren in zadovoljen s svojim stanjem. Torej, po Frankfurtovi teoriji, ravna po svoji svobodni volji. Ampak zdi se nam, da Jojo ni svoboden, da ne nadzira svojega lastnega življenja.

Problem nastanka svobodne volje je klasično orožje inkompatibilistov v boju s kompatibilisti. Že Hobbes je začrtal enega od možnih kompatibilističnih

odgovorov, ko se je spraševal, zakaj bi nas skrbela božja vnaprejšnja vednost in vnaprejšnja določenost naših dejanj. Ker po Hobbesu ne vemo, kaj nam je Bog vnaprej določil, moramo razmišljati o najboljšem načinu, kako živeti, kot da vnaprejšnje določenosti ni. Edina svoboda, ki jo lahko imamo, je svoboda, da storimo to, kar hočemo, to pa lahko imamo, četudi je vse določeno vnaprej (Kane, 1996: 67). Podobni meni Skinner: če smo zadovoljni s stanjem svoje volje, če nismo podvrženi zunanji prisili, zakaj je potem še pomembno, kako je do tega stanja prišlo? Ko bi imeli učinkovito znanost o človeškem obnašanju, bi bilo nesmiselno, da se ji odpovemo samo zato, ker bi morda načrtovanje srečne družbe vključevalo manipuliranje s človeško voljo.

To stališče Kane imenuje "trdi kompatibilizem", pripisuje pa ga tudi Frankfurtu (Kane, 1996: 67). Ali je neko dejanje svobodno in takšno, da smo zanj moralno odgovorni, je po Frankfurtu odvisno samo od usklajenosti med voljo in hotenji drugega reda, zato ga ne vznemirja vzročna zgodovina nastanka svobode volje. Celo možni "manipulator", ki dani osebi oblikuje povsem nov značaj, po njegovem mnenju ne ogrozi svobode in moralne odgovornosti:

> "Saj tudi nas oblikujejo in vzdržujejo okoliščine, nad katerimi nimamo nadzora. Vzroki, ki na nas delujejo, nas lahko radikalno spremenijo, pa vendar ostanemo moralno odgovorni dejavniki. Povsem irelevantno je, ali ti vzroki delujejo prek naravnih sil, ki oblikujejo naše okolje, ali pa delujejo zaradi premišljenega manipulatorskega načrta drugih človeških dejavnikov." (Frankfurt, 2002: 28)

V nekaterih primerih se determinirani dejavniki ne razlikujejo od tistih, ki so na kak drug način manipulirani. Pa vendar nekateri delujejo po lastni svobodni volji, drugi pa ne. Frankfurt mora sprejeti takšno stališče, če želi, da je njegovo pojmovanje svobode volje nevtralno glede problema determinizma. Morda je bilo to, da je oseba svobodna, da hoče, kar hoče, povzročeno s silami, nad katerimi nihče nima nadzora. Morda pa je svoboda volje nastala po naključju ali pa na kak tretji način. Dejstvo, da oseba uživa svobodo volje, je za Frankfurta združljivo tako z determinizmom kot z indeterminizmom. Veliko vprašanje pa je, ali je svobodna volja, ki zahteva obstoj alternativ in je takšna, da je dejavnik sam njen zadnji vzrok, sploh združljiva s čimerkoli, pa naj bo to determinizem ali indeterminizem (prim. Šuster, 2002). Morda gre v hlepenju po takšni svobodi zares samo za atavistično hrepenenje po neki misteriozni globoki svobodi in popolnem nadzoru nad svojim življenjem, ki ga tako in tako nikdar ne moremo imeti, kot meni Skinner (navaja Kane, 1996: 67). Oseba kot zadnji vzrok in izvor svojih hotenj je po prepričanju inkompatibilistov temelj osebne avtonomije in dostojanstva, za kompatibilista pa je to zahteva po metafizično nemogoči samokreaciji. Ali, kot pravi Frankfurt (2002: 28): "Smo pač takšne osebe, kot smo, ampak šteje to, kar smo, ne pa zgodovina našega nastanka."

FREEDOM AND THE HIERARCHY OF WILL

According to compatibilists freedom of action is typically the freedom to do what one wants to do. But does not real freedom of will consists of being free to have the will one wants to have? Harry Frankfurt in his seminal 1971 article develops a hierarchical theory of will, sometimes called hierarchical mesh theory. Free action is motivated by the desire one wants to have and a person enjoys freedom of the will when he is free to will what he wants to will. A distinction, often overlooked and not clearly expressed by Frankfurt himself, between acting of one's own free will and acting with a will that is free is explained. The second notion requires alternative possibilities, but the first, which is sufficient for moral responsibility, does not. Frankfurt's theory faces two main problems. The problem of *regress* arises when we ask ourselves about the reasons for accepting second-order volitions. Based on an analogy with the skeptical regress of justification I claim that Frankfurt in his later work presents a successful "foundationalist" response to the problem. The problem of *manipulation* concerns the history of higher-order volitions, they could well be manipulated or programmed. But according to Frankfurt a freely willed action for which an agent is morally responsible is purely a function of the relation between an agent's will and her second-order volitions, so it does not matter in any way how an agent came to have that particular mesh.

Keywords: Frankfurt, person, freedom of will and free will, hierarchy of will, regress problem, manipulation problem.

Literatura

Buss, S. in Overton, L. (ur.) (2002). *Contours of Agency: Essays on Themes from Harry Frankfurt*. Cambridge, Mass: MIT Press.

Double, R. (1991). *The Non-Reality of Free Will*. New York: Oxford University Press.

Fischer, J. M. (1994). *The Metaphysics of Free Will*. Oxford: Blackwell

Publishers.

Fischer, J. M. (1999). “Recent Work on Moral Responsibility”. *Ethics*, 110, str. 93-139.

Frankfurt, H. (2002). “Reply to John Martin Fischer”. V Buss in Overton (ur.), str. 27-31.

Frankfurt, H. (1988). *The Importance of What We Care About*. Cambridge: Cambridge University Press.

Frankfurt, H. (1987). “Identification and Wholeheartedness”. Ponatis v Frankfurt, H. (1999), *Necessity, Volition, and Love*, Cambridge: Cambridge University Press.

Frankfurt, H. (1971). “Freedom of the Will and the Concept of a Person”. *Journal of Philosophy*, 68, str. 5-20.

Frankfurt, H. (1969). “Alternate Possibilities and Moral Responsibility”. *Journal of Philosophy*, 66, str. 829-839.

Kane, R. (1996). *The Significance of Free Will*. Oxford: Oxford University Press.

McKenna, M. (2004). “Compatibilism”. *The Stanford Encyclopedia of Philosophy (Summer 2004 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/sum2004/entries/compatibilism/>.

Šuster, D. (1997). “Svoboda kot sposobnost”. *Analiza*, 1/1, str. 7-23.

Šuster, D. (1997). “Logika in skepticizem”. *Analiza*, 2/1, str. 71-86.

Šuster, D. (2001). “Sestavljanka o svobodi”. *Analiza*, 3/V, str. 62-72.

Watson, G. (1975). “Free Agency”. *Journal of Philosophy*, 72, 205-220.

Wolf, S. (1990). *Freedom within Reason*. Oxford: Oxford University Press.