ANαliZA

časopis za kritično misel

Letnik 5, številka 3, marec 2002 · ISSN 1408-2969

SVOBODA IN DETERMINIZEM

J. J. C. Smart

Ali ima robot lahko svobodno voljo?

prevedel Danilo Šuster

Citiranje: J. J. C. Smart, »Ali ima robot lahko svobodno voljo?«, *Analiza: časopis za kritično misel* 5, št. 3 (2001): 57–61.

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 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.

J. J. C. Smart
prevedel Danilo Šuster

ALI IMA ROBOT LAHKO SVOBODNO VOLJO?*

Ali iz gole naključnosti sledi svoboda volje in ali iz odsotnosti gole naključnosti sledi odsotnost svobodne volje? Torej, ali sta svoboda volje in determinizem nezdružljiva? Dokazoval bom, da je vprašanje gole naključnosti ali determinizma nepomembno za vprašanje svobodne volje. Še več, ne le, da svobodna volja in determinizem nista nezdružljiva, nek dober približek determinizmu na makronivoju je celo potreben za obstoj svobodne volje.

Ruleta in Geigerjev števec: deterministični in indeterministični izbori

Vzemimo dva računalnika, *A* in *B*, ki sprejemata odločitve. Recimo, da sta narejena z namenom, da med osebjem organizacije izbirata kandidate za izobraževalni tečaj za nov vodstveni kader. (Ne zagovarjam stališča, da naj bi izbor kandidatov bil avtomatiziran na tak način: ta zgled je zgolj ponazorilo metafizične poante.) Vhodna enota računalnika sprejema informacijo prek preluknjanih kartic. Informacijo sestavljajo natančna poročila o rezultatih preizkusov, ocene značaja osebe (npr. "pošten", "zmerno pošten" ipd., "vzkipljiv", "blagega značaja" ipd.) in vse druge informacije, ki bi bile zanimive za človeškega ocenjevalca. Računalnika sta programirana tako, da bi se z njunim izborom strinjali (npr. nihče od izbrancev ni preveč neumen, preveč nepošten, preveč vroče krvi).

Denimo, da je 20 prostih mest, računalnik pa najprej izbere 19 kandidatov. Za zadnje, dvajseto mesto sta dva enakovredna kandidata. Zadnji je zato izbran s pomočjo naključnostnega mehanizma. Denimo, da sta Kovač in Novak kandidata, med katerima se po vseh merilih, ki so na voljo, stroj ne more odločiti. Naključnostni mehanizem je tak, da je verjetnost izbora za vsakega kandidata ½. Pri stroju *A* je mehanizem ruleta, na kateri je električni stik. Če se kolo rulete nasproti fiksnemu električnemu stiku zaustavi na rdečem polju, to povzroči, da stroj izbere Kovača, če pa se zaustavi na črnem polju, je izbran Novak. Pri stroju *B* je naključnostni mehanizem Geigerjev števec v bližini majhne količine radioaktivne snovi, tako da je verjetnost ½, da bo števec reagiral v *n*-ti sekundi po

* Prevedeno po Smart, J. J. C., *Between Science and Philosophy: An Introduction to the Philosophy of Science*, Random House, 1968. Mednaslovi so prevajalčevi.

tem, ko začne stroj delovati, kjer je *n* liho število in verjetnost ½, da bo števec reagiral v *m*-ti sekundi po tem, ko začne stroj delovati, kjer je *m* sodo število. Če je Geigerjev števec aktiviran po sodem številu sekund, stroj izbere Novaka. Privzemimo, da sta dejansko oba stroja *A* in *B* izbrala Kovača.

Stroj *A*, kot smo ga opisali, lahko opazujemo kot deterministični stroj, saj njegovo obnašanje lahko opišemo z dovolj dobrim približkom s sredstvi klasične fizike. Rekli bi lahko, da je bil *A*-jev izbor Kovača determinističen. Če bi poznali vse vhodne podatke ter začetno stanje in delovanje stroja, bi lahko napovedali, da bo stroj izbral Kovača in ne Novaka. Stroj *B* pa je indeterminističen, saj je njegov izbor Kovača odvisen od indeterminističnega kvantnomehanskega efekta.

Ali je bil izbor Kovača s strani stroja *A* kaj bolj ali manj "racionalen", kot je bil izbor Kovača s strani stroja *B*? Seveda ne. Ali je bil izbor Kovača s strani stroja *B* "bolj racionalen", kot je bil izbor prvih 19 kandidatov, pri katerih uporaba naključnostnega mehanizma ni bila potrebna? Rečemo lahko, da ni indeterministični stroj *B* nič bolj in nič manj "svoboden" kot deterministični stroj "*A*". Zato menim, da načela kvantno mehanske nedoločenosti nimajo nič s problemom svobodne volje. Naj bodo naši možgani (približno) deterministični stroji ali pa naj vsebujejo mehanizme "čistih naključij", ne v enem ne v drugem primeru niso zato nič manj in nič bolj posestniki svobodne volje.

Ugovor – svobodno ravnanje je *ne*programirano delovanje zaradi razlogov

Nekateri filozofi bi ugovarjali implicitni domnevi, da je izbor nekaj, kar zadeva nek v glavnem vzročno komputacijski proces. Rekli bi, da v svobodnem izboru ravnamo zaradi razlogov, ne pa zaradi vzrokov, ravnanje po razlogih pa ni niti povzročeno niti ni zadeva čistega naključja. To se mi zdi neinteligibilno. Stroj, ki je izbral dvajsetega kandidata s pomočjo kvantno-mehanskega naključnostnega mehanizma, tega ni napravil zaradi nekega vzroka, ampak po golem naključju. Nobenega razloga ni imel, da bi dal prednost enemu kandidatu pred drugim. Stroj pa se je v izboru prvih 19 kandidatov obnašal na determinističen način, rekli bi lahko, da so bili njegovi izbori povzročeni. Programiran je bil tako, da so njegovi vzroki ustrezali razlogom, recimo, programiran je bil tako, da izbere najbolj inteligentnega, občutljivega ipd. kandidata. Torej je bil programiran tako, da deluje v skladu s tem, kar bi mi rekli "dobri razlogi". Ravnal je po razlogih natanko zato, ker je bil povzročen.

Tu bi lahko ugovarjali in opozorili na neko težavo. Strojevo ravnanje je bilo *programirano*, ampak mi ravnamo svobodno. V odgovor lahko navedem dve dodatni pojasnili. Najprej, ali nismo mi sami "programirani"? Na osnovi genetskih lastnosti, skupaj z vplivi okolja – že ko smo bili še v stanju embria in od takrat dalje – je vzročno določeno, da imajo naši možgani določeno zgradbo. Torej računajo na določene načine. Ljudje z določeno zgradbo možganov bodo sprejemali take in take odločitve, ljudje z drugačno zgradbo možganov pa drugačne. In drugič, če iz nazora, da smo v glavnem deterministični stroji, sledi, da nimamo svobodne volje, potem lahko priznamo, da je temu zares tako, da zares nimamo svobodne volje v smislu, na katerega meri ugovor. Ni pa jasno, za

kakšen smisel tu sploh gre: svobodni izbor naj bi ne bil determinističen in ne stvar golega naključja, tako kot je kvantni skok golo naključje. Svobodni izbor naj bi bil golo naključje v smislu, da "ni determiniran", ampak po tem predlogu je hkrati res tudi to, da ni zgolj slučajen, ter da v svobodnem izboru "ravnamo zaradi razlogov". Predhodni odstavek pa nakazuje, da ravnanje po razlogih ni golo naključje ravno zato, ker gre za povzročeno ravnanje.

Deterministova razlikovanja

Prav mogoče je, da običajni pojem svobode vsebuje nezdružljive značilnosti. Od svobodne volje zahteva, da je naključna in ni naključna obenem. V tem primeru lahko brez vsakega občutka obžalovanja zanikamo obstoj svobodne volje v tem običajnem pomenu. To pa še ne pomeni, da nimamo na voljo sredstev, s katerimi lahko izrazimo vsa razlikovanja, ki jih navadno delamo, ko pripisujemo ali zanikamo svobodno voljo. Determinist še vedno lahko razlikuje med položajem človeka, ki gre na sprehod, ker si to želi (določen je s svojo željo, ki jo lahko opišemo kot stanje nevronskih povezav v možganih) in človekom v jetniški celici, ki ne gre na sprehod, čeprav si to srčno želi. (Pri slednjem bi stanje njegovih možganov povzročilo, da bi šel na sprehod, a mu to preprečujejo zapahi na vratih njegove celice.) Biti povzročen, na prvi način, z notranjim stanjem želje, je, *ceteris paribus*, prijetno, biti onemogočen z zunanjimi ovirami pa je, *ceteris paribus*, neprijetno. Na podoben način ne zanikamo, da obstaja pomembna razlika med tem, da naredimo X, ker si to želimo in nimamo med seboj nasprotnih si želja, in primerom, ko naredimo X zaradi neke grožnje. Mož, ki da prispevek v dobrodelne namene, ker želi pomagati, deluje na način, ki je relativno prijeten. Mož, ki izroči denar na sodišču, ker mora plačati prekršek, v določenem smislu prav tako naredi tisto, kar želi, toda denar želi izročiti samo zato, ker mu ne ugaja alternativa, zaporna kazen. Tudi v tem primeru lahko determinist napravi ustrezna razlikovanja, pri tem pa ni preveč pomembno, ali zato uporablja besede "svoboden" in "nesvoboden".

Kaj pa razlikovanje med psihopatskim storilcem (denimo kleptomanom) in običajnim prestopnikom (recimo vlomilcem)? V obeh primerih je želja povzročila ravnanje. Ljudje z ulice in kazenski zakonik ločujeta med kleptomanom in navadnim vlomilcem tako, da zanikata, da bi bil kleptoman zares popolnoma svoboden. Razlikujemo (kar je vse, kar potrebujemo v kazenskem zakoniku) lahko tako, da rečemo, da na kleptomana ne moremo zlahka vzročno vplivati z grožnjo kazni, pri navadnem vlomilcu pa takšne grožnje delujejo. Zato lahko racionalno upravičimo, zakaj obravnavamo kleptomane drugače kot navadne vlomilce, ne da bi pri tem uporabili problematični pojem svobodne volje. Še enkrat lahko ugotovimo, če pojem svobodne volje, ki ga imajo navadni ljudje, zanika tako determinizem kot naključnost, potem ta navidezna nekonsistenca ni tako zelo pomembna, saj tudi brez tega pojma lahko izrazimo vsa razlikovanja, ki jih potrebujemo in navadno delamo.

Vzemimo zdaj znanstveno fantastično zgodbico. Namesto da bi bili naši možgani "programirani" z našimi geni in okoljem, privzemimo, da jih lahko sprogramira

ali preprogramira nek stroj. Denimo, da bi bilo mogoče spremeniti celotni značaj osebe tako, da na njegovo glavo postavimo nekaj elektrod, tako da so na nek (zaenkrat tehnološko nepojmljivi) način povezave med njegovimi nevroni (in po možnosti tudi tisto, kar tvori njegov spomin) radikalno predrugačene. Morda bi lahko takšen stroj uporabili na humanitarnem pesniku in ga spremenili v diaboličnega tehnologista, ki je sposoben iznajditelj smrtonosnih orožij. Nekateri bi rekli, da bi takšen reprogramirani mož, čeprav bi še tako malo deloval v nasprotju s svojimi željami, ne bil svoboden in si ne bi zaslužil ne hvale ne graje za svoja dejanja. V odgovor na to bi lahko rekli, da bi zaslužil grajo ali hvalo v natanko isti meri, kot bi si ju zaslužila običajna oseba (oseba, katere možgani niso spremenjeni). Če bi bilo koristno (zaradi zastraševanja in preprečevanja) kaznovati običajno osebo, potem bi bilo koristno (zaradi zastraševanja in preprečevanja) kaznovati reprogramirano osebo. (Tu sem zanemaril možnost, da bi v izmišljenem primeru dosegli boljše rezultate, če bi kaznovali programerja in ne programiranca.) Edini dober razlog za kaznovanje ali nagrado, hvalo ali grajo, se mi zdijo socialni učinki. Filozof, ki misli, da ni prav, da grajamo ali kaznujemo običajno osebo, zanesljivo pozablja, da je tudi običajna oseba programirana z dednostjo in okoljem. Ko bi razmišljal o izprijenem morilcu, bi tak filozof lahko imel nenavadno idejo, da če bi *on* imel morilčeve slabe možgane in slabo okolje, potem bi on vendarle ravnal drugače. Mislim, da je takšno stališče nerazumljivo.

Moj sklep je torej tak, da četudi so možgani deterministični mehanizem, še vedno lahko opravimo pomembna razlikovanja, ki jih normalno naznačimo, ko neko obnašanje razglasimo za “svobodno” ali pa “nesvobodno”. Odločitev, ali bomo rekli, da pojem svobode, ki ga imajo navadni ljudje, nasprotuje determinizmu, ter da po nazoru, da je duševnost deterministični (ali v glavnem deterministični) stroj sledi, da nimamo svobodne volje, ali pa tega ne bomo rekli, je nekoliko arbitrarna. Precej je odvisna od tega, katerega navadnega človeka bomo ujeli in ali bo le-ta metafizično razpoložen ali ne. Tudi če bomo rekli, da je pojem svobodne volje, ki ga imajo navadni ljudje, nezdružljiv z determinizmom, lahko še vedno rečemo, da naše zanikanje svobodne volje nima strašljivih posledic, ki bi jih imelo, če bi iz njega sledila zanikanja pomembnih razlikovanj, ki smo jih obravnavali zgoraj. Lahko pa se odločimo drugače in uporabljamo izraza “svoboden” in “nesvoboden” za to, da z njima izrazimo zgornja razlikovanja; v tem primeru pa je determinizem povsem združljiv s svobodno voljo.

Svoboda in znanost o možganih

Kaj nam znanost lahko pove o vprašanju, ali so možgani deterministični stroj? Precej filozofov je menilo, da indeterminističnost kvantne mehanike pušča prostor za svobodno voljo. Naša razprava o svobodni volji bi morala pokazati, kaj je s tem narobe. Abstraktno je mogoče, da bi možgani vsebovali indeterministični sprožilni mehanizem, kot je Geigerjev števec v stroju *B*. Ampak to bi pomenilo samo to, da možgani vsebujejo indeterministični generator naključij. Če bi vsebovali deterministični generator naključij kot v stroju *A*, kako bi to lahko vplivalo na vprašanje njihove svobode? Možnosti zelo majhnih

dogodkov, morda pod ravnijo kvantno mehanske nedoločenosti, ki nastopijo v sinaptičnih izrastkih posameznih nevronov, potem pa jih nevronski mehanizmi ojačijo tako, da vodijo do učinkov v obnašanju, ne moremo povsem izključiti. Nič manjša avtoriteta, kot je nevrofizolog sir John Eccles, je bila tista, ki je zares predlagala tak mehanizem. Eccles misli, da na tak način dogodki iz nesnovne duševnosti lahko vplivajo na možgane. Mnogo znanstvenikov bi morda menilo, da zamisel o nesnovni duševnosti proizvaja več težav, kot pa jih rešuje. Najprej se zdi, da postuliranje nesnovne duševnosti povzroča težave genetski teoriji evolucije po naravnem izboru. Ni težko uvideti splošnega načina, kako mutacije genov vodijo, prek biokemije embriologije, do rahlo drugače zgubanih možganov in nazadnje do zmožnosti obnašanja, zaradi katere pravimo, da ima nek organizem duševnost. Težko pa je uvideti, kako bi nastanek nesnovne duševnosti uskladili s kemičnim pristopom h genetiki. Za Ecclesa je obstoj nesnovne duševnosti nekaj, za kar nam jamči introspekcija, ampak tu se bomo izognili razpravi, ki sproža velika in sporna vprašanja v filozofiji duha. Naj le omenimo, da postuliranje nesnovne duševnosti ne pomaga razsvetliti problema svobodne volje. Še enkrat se lahko vprašamo, ali so dogodki v nesnovni duševnosti povzročeni ali naključni. Če pa niso niti povzročeni niti naključni, potem lahko vprašamo, v čem je ta tretja možnost.