

Danilo Šuster

Univerza v Mariboru

Prediktor, fatalizem in miselni eksperimenti¹

V članku analiziram kratko zgodbo *Kaj se pričakuje od nas* (Chiang, 2005) kot primer filozofskega miselnega eksperimenta, ki domnevno dokazuje, da ni svobodne volje. Pri tem uporabim členitev po stopnjah, ki jo je za miselne eksperimente predlagal Miščević (2012: 2017): najprej scenarij in hipoteza, potem model in razumevanje, vživetje in izpeljava posledic, spontani intuitivni odgovor, uvedba variacij. Miselni eksperiment nas vodi do tega, da v domišljiji 'vidimo' in razložimo pojmovne komponente tam, kjer jih običajno ne. *Prediktor* uvede ločitev smeri časa in smeri vzročnosti. Ta obrat smeri vodi do kolizije našega doživljanja časa in sebe kot vzročnega dejavnika v času. Dejanski red našega delovanja je vzročni, izkustveni red pa je časovni. Pa vendar miselni eksperiment ne uspe dokazati, da naše izbire niso svobodne, saj temelji na zmotni logiki argumentov za fatalizem.

Ključne besede: svobodna volja, fatalizem, vzročnost, čas, miselni eksperiment.

¹ Besedilo je nastalo s podporo raziskovalnega program št. P6-0144, ki ga je sofinancirala Javna agencija za raziskovalno dejavnost Republike Slovenije.

1.

Sodobni severno-ameriški pisatelj kitajskih korenin Ted Chiang ne objavlja veliko, a je prejemnik številnih nagrad na področju znanstvene fantastike. Velja za perfekcionista, ki lahko leta 'medi' eno samo kratko zgodbo. *Zgodba tvojega življenja* (*Story of Your Life*), po kateri je bil posnet znani film *Prihod* (*Arrival*, 2016), je tako 'mukoma' nastajala štiri leta (po preučevanju lingvistike in piljenju stavka za stavkom). Njena tema je sicer jezik, a glavno filozofsko vprašanje, značilno za vrsto Chiangovih novel, je starodavna dilema svobodne volje in časa: če že veš, kaj se bo zgodilo, kako boš potem doživljal svoje bodoče odločitve in izbire? Chiang je eden redkih sodobnih avtorjev, ki poudarja, da znanstvena fantastika vsebuje besedo 'znanost', saj raziskuje meje spoznanja: loteva se miselnih eksperimentov in raziskuje filozofska vprašanja². Prav to je tudi značilnost prevedene kratke zgodbe, spet miselni eksperiment o času in svobodni volji. Objavljena je bila v reviji *Nature* in citirana celo v resni znanstveni literaturi kot primer miselnega eksperimenta o »mašini za napovedovanje odločitev«, ki bi jo bilo morda mogoče realizirati na kak drug način (Haynes, 2011: 172).

Zgodba je v svojem minimalizmu lepa ponazoritev uporabe miselnih eksperimentov, pa tudi zmotne 'logike' fatalizma in naše naravne uročenosti z 'nespremenljivo' preteklostjo, ki tako kot mitološka Meduza razteguje svoje kačaste lovke naprej v prihodnost. Kdor ji pogleda v oči, konča v stanju 'akinetičnega mutizma'. Morda pa to sploh ni končna poanta zgodbe in gre za nekakšnega vsevednega ironičnega Pripovedovalca, ki se pri sebi hahlja in ironizira človeško 'zaslepljenost'? Ampak tu me zanima filozofija, ne pa literarna kritika in navadna proza, ne pa več-značnost literarnega besedila. Naj takoj razkrijem svoje karte: glavni miselni eksperiment je neuspešen, saj sklep, da zamišljena naprava dokazuje, da nimamo svobode volje, ne sledi.³ Toda v ozadju enostavnega miselnega eksperimenta se skriva še eno vprašanje: ali naše *izkustvo* vnaprejšnje vednosti nujno vsebuje elemente fatalizma kot neke življenjske drže? Tu je Chiang bolj prepričljiv.

2.

Miselni eksperimenti sodijo v tisti del filozofske orodjarne, ki je najbliže umetniškemu delu. Če je metafora sanjsko delo jezika, kot pravi Davidson, potem so miselni eksperimenti laboratorijske vaje domišljije. Gre za konkretne in nazorne scenarije, ki so hipotetični ('protidejstveni') in zelo premišljeno zgrajeni. Ko se 'odvijejo' v naši domišljiji ali ko v mislih preigramo korake, h katerim

² Prim. Clark, 2015.

³ Podobno zgrešene se mi zdijo Haynesove izpeljave, da rezultati Libetovih eksperimentov ovržejo 'naivno' prepričanje v svobodno voljo, a to je druga zgodba – sicer povezana z idejo Prediktorja, kot je opazil Haynes. O Libetu več na koncu članka.

smo napoteni, nas končni rezultat poskusa vodi do nekega intuitivnega sklepa. Cilji so lahko različni, toda v filozofiji gre običajno za nek jasno določen spoznavni smoter – potrditev ali ovržbo neke teorije ali hipoteze. Pravi eksperiment je pogosto neizvedljiv, razlogi so lahko fizikalni tako kot v naši zgodbi (pošiljanje signala v času *nazaj*), 'banalno' finančni ali pa tehnološki in etični.

Vzemimo morda najbolj znani miselni eksperiment s področja svobode in odgovornosti. Splošno sprejeto načelo je načelo alternativnih možnosti (NAM): »Oseba je moralno odgovorna za to, kar je storila samo, če bi lahko ravnala drugače.« Če ste zapadli v stanje akinetičnega mutizma in nimate več sposobnosti premikanja, ne morete biti odgovorni za to, da se niste udeležili sestanka hišnega sveta. Frankfurt (1969) je predstavil nazoren scenarij, ki naj bi naše intuicije preusmeril do sklepa, da NAM ne drži. V malo posodobljeni različici: oseba A smrtno sovraži osebo B in načrtuje njen umor. Toda tudi oseba C sovraži B in si želi njene smrti, a želi ostati skrita v ozadju. Recimo še, da je C spreten nevrokirurg in manipulator, ki vstavi v možgane osebe A neko *napravico*, ki spremlja A-jevo možgansko aktivnost. Če znaki kažejo, da bo oseba A po svoji presoji izpeljala prvotno namero, naprava miruje. Če pa znaki kažejo, da bo oseba A odstopila od svoje namere, naprava sproži intervencijo (morda stimulira ustrezni del možganov), ki privede do tega, da se A spet odloči za umor. No, izkaže se, da intervencija ni potrebna, saj se A zaradi svojih razlogov odloči in izpelje načrtovani umor. Intuicija o tem konkretnem scenariju nam pravi, da je A ravnal svobodno, po lastni volji in je zato moralno odgovoren, čeprav ni mogel ravnati drugače.

Če to posplošimo – začeli smo s hipotezo (NAM), zamišljeni scenarij pa nas privede do tega, da jo zavrnemo. Hipotetični scenarij je zaradi tehnoloških in etičnih razlogov neizvedljiv. Pogled v 'drobovje' tega scenarija pa nam razkrije še nekaj pomembnih značilnosti miselnih eksperimentov. Nazorni scenarij sproži nekakšno mentalno 'destilacijo', na ognju domišljije se prvotna pojmovna 'zmes' razloči na komponente. Ko odstranimo možnost alternativ, preostali 'kondenzat' še vedno zadošča, da osebi pripišemo moralno odgovornost za njeno dejanje. Navsezadnje gre za *namerno* dejanje – oseba je izvor dejanja, z njim se poistoveti, če bi *njeni* razlogi govorili proti, bi se mu odpovedala. Miselni eksperiment nas vodi do tega, da v domišljiji 'vidimo' in razločimo pojmovne komponente tam, kjer jih običajno ne (prim. tudi Miščević, 2012: 202).

Poskusimo zdaj razčleniti zgodbo po fazah miselnega eksperimenta, kot jih predlaga Miščević (2017). V prvi fazi tvorec miselnega eksperimenta predstavi zamišljeni scenarij in zastavi vprašanje. Spomnimo so: Prediktor je enostavna igralna naprava, sestoji iz gumba in svetila, ki vedno zasveti sekundo *preden* pritisnete na gumb. Pojasnilo je znanstveno-fantastično: v Prediktorju je krogotok z negativnim časovnim zamikom, ki pošilja signal v času *nazaj*. *Hipoteza*, ki je, morda ironično, predstavljena precej apodoksično, se skriva v tezi, da »Prediktorji *dokazujejo*, da ni take stvari kot je svobodna volja.« Zato bo začetno vprašanje: »Ali Prediktor dokazuje, da

svobodna volja ne obstaja?» V naslednji fazi si bralec predloži situacijo in zgradi model. Razumeti mora, zakaj, glede na zgradbo Prediktorja, ni načina, kako bi ga pretentali: »Ne glede na to, kaj naredite, luč vedno zasveti *preden* pritisnete na gumb.« Ko scenarij razumemo, začne delovati naša mentalna 'mašinerija' (vživetje, razmišljanje, izpeljava posledic itd.): »implikacije nespremenljive prihodnosti prodrejo v zavest.« Očitno te implikacije načnejo naše izkustvo svobodne volje, saj ugotovimo, da »naše izbire niso pomembne.« V zadnji fazi potem sledi naš intuitivni odgovor: »Prav res, Prediktor dokazuje, da ni svobodne volje.«

Tudi Chiangov miselni eksperiment temelji na *ločitvi* pojmovnih komponent, za katere se nam zdi samoumevno, da sodijo skupaj. Filozofsko jedro je razmislek o pojmovanju časa, vzročnosti in našega delovanja. Vzročnost ima neko značilno smer – dogodek vzrok ima vzročno prioriteto, vzročnost nekako 'teče' od vzroka (pritisk gumba) do učinka (blisk), vzroki učinkujejo in vodijo do učinkov, nikdar obratno (prim. Tooley, 1987: 179). Naravno se nam zdi, da je smer vzročnosti smer časa: vzročnost je že za Huma relacija, v kateri so vzroki časovno *pred* učinki. Relacija med vzrokom in učinkom je asimetrična, ker je relacija »biti časovno predhoden« asimetrična, čas pa 'teče' samo naprej. Prediktor izzove domnevno apriorni status te povezave: ali ni (vsaj logično) možno 'vnazajšnje' povzročanje, v katerem učinek (blisk) nastopi časovno *pred* svojim vzrokom (pritisk gumba)?

Kar nekaj filozofskih teorij poskuša razložiti usmerjenost vzročne relacije, ne da bi jo takoj pojmovno utemeljili na časovnem zaporedju dogodkov. V starejših teorijah nastopa pojem moči – vzrok poseduje neko moč, ki privede do spremembe v drugi stvari (učinek). Drugi predlagajo bolj antropomorfna merila – nastop vzroka *pojasni*, zakaj nastopi učinek, ne pa obratno (npr. Horwich, 1987). Po teoriji manipulacije ima dogodek *V* vzročni primat pred dogodkom *U*, če lahko manipuliramo z dogodkom *V* in na ta način privedemo ali preprečimo nastop dogodka *U*. Po teh pojasnilih je ujemanje časovnega in vzročnega reda samo kontingentna značilnost dejanskega sveta ali naše pojmovne sheme. In Chiang s pomočjo izmišljene naprave, ki pošilja signal v času nazaj, preobrne in s tem *razloči* obe smeri: vzročni red signala teče od gumba do bliska, časovni red pa od bliska do gumba.

Nekateri so dokazovali, da je to nemogoče, ampak dovolimo si ta miselni izlet. Chianga ne zanima, kako je možen krogotok z negativnim časovnim zamikom, ampak raziskuje vprašanje, kako mi v svojem delovanju *doživljamo* to ločitev. Kaj se zgodi z našim doživljanjem svobodnih izbir, ko sta si smer časa in smer vzročnosti v nasprotju? Pripovedovalec pravi, da »Ne glede na to, kaj naredite, luč vedno zasveti, preden pritisnete na gumb« in zato, domnevno, naše izbire niso pomembne. Pravil igre ne moremo prekršiti, ne moremo pritisniti gumba, ne da bi se pred tem pojavil blisk,

točno eno sekundo pred tem. Blisk nas, tako rekoč, 'prisili' v premik prsta. Naš prst je kot marioneta na vrvici bliska, svoboda volje pa je iluzija.

Ampak *zakaj*, natanko? Vzemimo neko enostavno dejanje: poštar stoji pred vašimi vrati in pritisne na gumb zvonca. Zakaj? Ker vam *želi* izročiti pošto pošiljko, saj je *prepričan*, da ste doma (prej sta se dogovorila po mobilnem telefonu) in vam *namerava* pošiljko izročiti (to je njegov *cilj*, navsezadnje je to njegova služba). Po tej, zelo vplivni Davidsonovi analizi, poštarjevi *razlogi* (želje, prepričanja, namere, cilji) *povzročijo* njegovo dejanje. Nekateri takšni vzročni analizi sicer oporekajo, ampak v razpravi o Prediktorju bom privzel, da je namera pritisniti na gumb nujni del tistega, kar *povzroči* dejanje. Poštarjevo dejanje je *svobodno*, če bi se lahko odločil drugače, oblikoval drugačno namero, zaradi katere bi izbral drugačno ravnanje (vsi imamo izkušnje z 'lenimi' poštarji, ki iz takšnih ali drugačnih razlogov ne pozvonijo in vam samo vržejo obvestilo v nabiralnik).

Zdaj se počasi pokažejo razpoke v domnevnem 'dokazu', da nimamo svobodne volje. Ali zares nimamo *nadzora* nad tem, kaj se bo zgodilo? Naša namera pritisniti na gumb je nujni del tistega, kar *povzroči* dejanje. Kadar nekaj naredim, ker to *nameravam* storiti, potem je moje dejanje odvisno od mojih razlogov (želja in prepričanj, s katerimi se poistovetim) in v tem smislu ga (vsaj) doživljam kot svoje svobodno dejanje. Moj prst ni samo marioneta na vrvici Prediktorja, njegovi premiki so vzročno odvisni od *mojih* odločitev. Lahko bi se odločil drugače in oblikoval drugačno namero, zaradi katere bi izbral drugačno ravnanje. Preberimo še enkrat besedilo:

Če *poskusite* pritisniti gumb, ne da bi videli blisk, se blisk svetlobe takoj pojavi in ne glede na to, kako hitro se premikate, nikdar ne pritisnete gumba, preden mine ena sekunda. Če čakate na blisk z *namero*, da bi se kasneje odpovedali pritisku na gumb, potem se blisk nikoli ne pojavi.

Toda *poskus* pritiska gumba brez bliska vendar vključuje tudi enostavno *namero* pritisniti gumb. In če čakate na blisk z *namero*, da bi se pritisku odpovedali, potem očitno *ne* nameravate pritisniti gumba. Z drugimi besedami: pritisk gumba in z njim povezan blisk svetlobe sta vzročno odvisna od vaših namer. Res je, časovni in vzročni red sta obrnjena: blisk je časovno *pred* pritiskom na gumb in vzročno *za* njim. Ampak v opisanem scenariju je namera vzročno *pred* obema, morda tudi časovno. To se vsaj zdi smiselna interpretacija: blisk je v časovno negativnem zamiku, povezan s stanji gumba, psihološka stanja igralca pa ne. Namere igralca ne *registrirajo* položaja gumba, ampak ga povzročajo na klasičen način: učinek časovno sledi vzroku.

No, človeška psihologija je včasih zapletena in način oblikovanja namer ni vedno transparenten. Denimo, da je igralec v stanju negotovosti, potem se pojavi blisk, igralec (zelo hitro) oblikuje

namero in pritisne na gumb. Morda to ni nemogoče. V znamenitih Libetovih eksperimentih se oseba zave želje po premiku približno 200 milisekund pred aktivacijo mišice, ki sproži premik prsta (prim. Šuster: 2007, 307). Toda tudi v tem primeru je pritisk gumba pod mojim nadzorom, določajo ga moja volja in moji razlogi. Morda se res odločim zaradi bliska, ampak blisk me ne *prisili* v to odločitev. Zakaj potem ne gre za svobodno dejanje? Ali ne gre za zmotno sklepanje *post hoc ergo propter hoc*: »Izbira ne vodi (časovno) do bliska. Torej izbira ne vodi (vzročno) do bliska?« Premagati 'sistem' bi pomenilo *nemogoči* upor proti svoji *lastni* volji: nameravati pritisniti na gumb (to naznači blisk, ki registrira *učinek* naše namere), ampak *obenem* tega tudi ne nameravati (saj želimo pretentati Prediktorja).

Lahko bi se odločil drugače in oblikoval drugačno namero? Toda, ko se pojavi blisk, ne moreš ravnati drugače, kot da pritisneš na gumb! Nobene izbire nimaš. In dokler bliska ni, ne moreš pritisniti na gumb, nam prišepetava pripovedovalec. Ali ni to strašljivo?

Morda med pojavom bliska in pritiskom na gumb zares ni alternativ, ampak, prvič, imeli smo jih *pred* tem in drugič, kot kažejo Frankfurtovi primeri, še vedno ravnamo po lastni (svobodni) volji. Zakaj nas blisk ne 'prisili' v pritisk gumba? Ker gumb v vsakem primeru nameravamo stisniti, frustriran je le naš poskus privedi do gumba *brez* bliska. Tak poskus pa je obsojen na neuspeh, ne glede na to, kdaj se pojavi blisk: *pred* pritiskom na gumb ali *po* pritisku na gumb. Vzemimo namesto Prediktorja zelo navadni *Retrodiktor*: luč zasveti eno sekundo *po* tem, ko pritisnete na gumb. Retrodiktor je izjemno dobro narejen, s pomočjo novih materialov, supersonične tehnologije in toka nevtrinov, tako da je nemogoča kakršnakoli zunanja intervencija. Ko enkrat pritisnete na gumb, je glede na zakone narave fizikalno nemogoče, da se ne bi čez eno sekundo pojavil blisk svetlobe. Oziroma nemogoče vsaj v tistem smislu, v katerem je nemogoče, da ne bi krogotok v Prediktorju, ko enkrat pritisnete gumb, vodil do tega, da se sekundo *pred* tem pojavi blisk (tudi tu izključimo možnost napak v materialu in 'vnazajšnjih' intervencij, ki bi blokirale signal). Retrodiktor poroča o preteklosti, pove vam, da ste sekundo *pred* bliskom pritisnili na gumb. Izredno nezanimiva naprava za igranje, morda bi pritegnila šestmesečne dojenčke. Ampak tudi Retrodiktorja ne moremo pretentati: ne moremo pritisniti gumba, ne da bi se *za* tem pojavil blisk, točno eno sekundo kasneje. In nemogoče je, da se pojavi blisk, ne da bi *pred* tem pritisnili na gumb. Pravil igre ne moremo prekršiti, vendar mislim, da nihče ne bi trdil, da Retrodiktor dokazuje, da nimamo svobodne volje! Ne v enem, ne v drugem primeru vaša svoboda ne more biti v tem, da z neko magično močjo prekinete vzročno nujno zvezo med pritiskom na gumb in bliskom, ampak v tem, da po *lastni* izbiri pritisnete na gumb.

Prediktor (signal 'nazaj') v paru z Retrodiktorjem (signal 'naprej') ponazarja še eno od značilnosti miselnih eksperimentov. Miselni eksperiment, posebej to velja za 'močne' in prelomne, zelo pogosto

vodijo do variacij, majhnih ali večjih sprememb v scenariju, ki sklep okrepijo, razširijo na druga področja ipd. ali pa ga ovržejo (prim. Miščević, 2017: 121). Retrodiktor je proti-eksperiment, ki nam kaže, da je intuicija o nesvobodi, ki temelji na nujni zvezi med (časovno predhodnim) bliskom in (časovno kasnejšim) gumbom, zavajajoča. Še ena variacija bi bila 'Časovna turistka': denimo, da nekoč v daljni prihodnosti odkrijejo in izkoriščajo možnost potovanja v času. V naključnem pogovoru vas zgovorna turistka iz prihodnosti preseneti z napovedjo, da boste, danes še zakrneli samec in velik ljubitelj poznega vstajanja, čez šest let in pol vsako jutro sprehajali labradorca po imenu Maks. Tako kot je blisk signal iz prihodnosti (ena sekunda), je ta napoved poročilo iz prihodnosti (šest let in pol). Ali je lahko takšno resnično poročilo razlog za depresivno ugotovitev, da vaše izbire niso pomembne? Zaradi napovedi ne morete ravnati drugače kot da spremenite svoj življenjski slog? Pa saj boste vendar *sami* izbrali življenjsko pot pasjeljubca med različnimi alternativami, ki jih prinaša življenje.

3.

Prediktor zavaja, ker usmerja naše intuicije k napačnemu odgovoru na začetno vprašanje. Pripovedovalec predpostavlja, da naše izbire niso pomembne, a to bi moral biti šele naš intuitivni odgovor. No, v obrambo lahko navedemo, da so tudi najbolj slavni filozofski miselni eksperimenti deležni podobnih očitkov (prim. Miščević, 2017: 115). In tu gre, navsezadnje, za umetniško prozo, ne pa filozofsko razpravo. Pripovedovalec spretno manevrira med opisovanjem domnevnih dejstev in poročanjem o *učinku* Prediktorja na igralce. Miselni eksperiment črpa prepričljivost iz našega *doživljanja* vzročnosti in časa v našem vsakdanjem delovanju. Smer našega delovanja v našem izkustvu sovпада s smerjo časa, ki je smer vzročnosti in tega, kar pod našim nadzorom. Preteklost je vzročno izolirana, ne moremo je spremeniti. Celo v filozofski literaturi najdemo obrambo intuicije, da časovno *predhodni* blisk, nad katerim igralec domnevno nima nadzora, nekako *izsili* kasnejši pritisk na gumb.

Gre za standardni filozofski ugovor proti sami zamisli o vnazajšnjem povzročanju, t. i. 'argument preklica' (angl. *bilking argument*, prim. Horwich, 1987: 91–92). Uporabimo kar Prediktorja: blisk (B) je v zgodbi pred pritiskom na gumb (G), čeprav gre za vzročni učinek pritiska na gumb. Potem, pravi ta argument, je vsaj v načelu mogoče, da *interveniramo* v toku dogodkov in po nastopu bliska preprečimo kasnejši nastop G. S tem uničimo zvezo med B in G: blisk nastopi brez gumba, zato (kasnejši) G ne more biti vzrok za (predhodni) B. Če pa, ko enkrat nastopi B, ne moremo več *preprečiti* kasnejšega pritiska na gumb, potem to dokazuje (?), da je prvi dogodek (blisk) vzrok drugega – časovni in vzročni red sovpadeta, zamisel o vnazajšnji vzročnosti pa je nekoherentna.

A zakaj le? Chiang se s tem ne ukvarja, žanr znanstvene fantastike mu omogoča, da prekrši nek (domnevni) zakon narave in potem v zgodbi razdeluje posledice. Ampak zamisel sploh ni nekoherentna. Če ne moremo preprečiti pritiska na gumb, zakaj naj bi to pomenilo, da je potem blisk časovno predhodni vzrok in ne obratno? To je sicer čudno, ne more pa biti izključeno *a priori*. Faye (2018) lepo opozori, zakaj se nam to zdi *čudno*: ker razmišljamo o vnazajšnji vzročnosti kot pojavu, ki ga lahko neposredno nadziramo na običajen način s svojimi vsakdanjimi dejanji. Ampak to se nam zdi nemogoče, naše izbire vplivajo in določajo prihodnje dogodke, nikdar preteklih. Preteklosti je izven vzročnega 'dosega' naših dejanj, v tem smislu je fiksna (angl. settled) in neogibna. Prediktor privede do kolizije našega doživljanja časa in dojemanja sebe kot vzročnega dejavnika z obratom časovne in vzročne smeri. Dejanski red našega delovanja je vzročni (od namere do gumba, od gumba do bliska), naš izkustveni red pa je časovni (od bliska do gumba). Končni proizvod teh pojmovno-kemičnih reakcij pa je *fatalizem*.

Fatalizem v vsakdanji rabi označuje neko življenjsko *držo*, v katero zapademo, ko začnemo verjeti, da ničesar, kar lahko danes naredimo, ne more spremeniti tega, kar se *bo* zgodilo. Spomnimo se nesrečnega Ojdipa – usojeno je bilo, da bo ubil svojega očeta in se poročil z lastno materjo in ne glede na vse njegove izbire in poskuse, da bi to preprečil, se je prerokba uresničila. Tako kot preteklost tudi dogodki, ki šele prihajajo, niso pod našim nadzorom in so za nas neogibni, pravi fatalist. V našem doživljanju se lahko takšno prepričanje manifestira kot stoična 'vdanost' v usodo. Chiang dramatično opisuje bolj pogubne odzive, ko domnevne »implikacije nespremenljive prihodnosti prodrejo v zavest.« Nekateri ljudje, ki zaradi Prediktorja ugotovijo, da njihove odločitve niso pomembne, se ne lotevajo več nobenih spontanosti in tretjina igralcev konča »v stanju akinetičnega mutizma, v neke vrste budni komi.«

Filozofi danes uporabljajo izraz *fatalizem* drugače, kot oznako za *stališče*, da nimamo moči, da bi storili karkoli drugega od tistega, kar dejansko storimo: »Fatalizem /.../ je teza, da je logična ali pojmovna resnica, da nihče ne more ravnati drugače kot dejansko ravna, gola zamisel dejavnika, ki ima pred sabo alternativne poteke dejanj je protislovna« (van Inwagen, 1983: 23). Ker gre za logično resnico, je stališče znano tudi kot *logični* determinizem. *Navadni* determinizem je teza, da v danem trenutku iz dejstev o preteklosti sveta po *zakonih* narave sledijo vsa dejstva o prihodnosti sveta (prim. Šuster, 2007). V vsakem trenutku univerzuma obstaja natanko ena fizikalno možna prihodnost. Determinizem dopušča, da bi preteklost lahko bila drugačna, prav tako zakoni narave (fiksnost enega ali drugega ni logična in pojmovna resnica). Za logičnega determinista pa je vsak dogodek v preteklosti, sedanjosti in prihodnosti vedno že 'fiksiran' samo zaradi zakonov logike ter narave pojmov resnice in časa.

Od kod ta zamisel? Pri Aristotelu (*De interpretatione*: 18b23) najdemo znameniti primer jutrišnje pomorske bitke. Do nje bo prišlo ali pa ne, to nam pove 'čista' logika (zakon izključene tretje možnosti). Propozicija, da bo jutri pomorska bitka, je tako (danes, dan pred tem) resnična ali pa neresnična. Ampak, če je, denimo, (danes) resnično, da se bitka zgodi, potem je bilo to resnično tudi že včeraj. Propozicije v času ne spreminjajo svoje resničnostne vrednosti in resnice ne moremo spremeniti v neresnico. To velja tudi za resnične propozicije o bodočih dogodkih, zato v *nobenem* trenutku časa ni resnično, da bitke ne bo. Toda potem je tudi v preteklosti vedno že bilo resnično, da bitka bo. Preteklost pa je fiksirana, nihče ne more narediti, da bi bila drugačna. Torej tudi (pretekle) resnice tega, da jutri bo pomorska bitka, ne moremo spremeniti in je zato bitka *neogibna*. Enak razmislek pokaže, da je tudi propozicija, da jutri pomorske bitke ne bo, če je *ta* resnična, neogibna. Ali eno ali drugo: bodoči potek dogodkov je neodvisen od naših namer in odločitev.

Mislím, da strašljivost Prediktorja temelji na logiki argumenta za fatalizem. Zamislímo si igralca pred Prediktorjem, ki razmišlja o tem, kaj naj naredi in kaj sploh lahko naredi. Po vsakdanji fenomenologiji naše izbire vplivajo in določajo prihodnje dogodke, nikdar preteklih, ti so izven 'dosega' našega delovanja. Za poljubni trenutek *t* v prihodnosti igralca potem velja: ali bo pritisnil na gumb v tem trenutku, ali ne. Ampak, če bo pritisnil na gumb, se bo sekundo pred tem pojavil blisk in *napovedal* njegovo dejanje. In če ne bo pritisnil na gumb, potem bliska ne bo. Toda sekunda pred *t* je v preteklosti glede na njegov gib, pred trenutkom njegovega delovanja, zato v trenutku *t* ne more več vplivati na to, kaj se bo zgodilo *pred* tem. Resnične napovedi 'iz preteklosti' ne moremo spremeniti v neresnično, zato je njegovo dejanje neogibno.

Privlačno, ampak zmotno. Za aristoteljanskega fatalista ima resnična *napoved* (pred bitko) enak metafizični status kot resnično *poročilo* (po bitki). Oboje je fiksno, ničesar ni, kar bi kdorkoli lahko naredil, da bi eno ali drugo spremenil. Podobno pri igralcu: v trenutku pritiska na gumb ne more narediti ničesar, kar bi vplivalo na preteklost (pojav bliska). Pojav bliska je resnična *napoved* pritiska na gumb, zato lahko razmišljanje igralca prikažemo v skladu z logiko pomorske bitke:

Če je napovedano, da bo igralec pritisnil na gumb, potem je neogibno, da bo pritisnil na gumb. Če pa ni napovedano, da bo pritisnil na gumb, potem je neogibno, da ne bo pritisnil na gumb. Ampak pritisk na gumb je napovedan ali pa ni napovedan (blisk se pojavi ali pa se ne pojavi). Torej je neogibno, da bo igralec pritisnil na gumb ali pa je neogibno, da ne bo pritisnil na gumb.

Igralec v resnici nima izbire, svobodna volja pa je iluzija, pravi Pripovedovalec. Naj bo 'Np' oznaka za »Neogibno je, da *p*.« V najbolj splošnem smislu gre za neko nujnost, ki označuje našo nemoč: nezmožnost vplivanja, nesposobnost spreminjanja, odsotnost izbire ali kak podoben pojem, ki označuje nesvobodo glede tistega, kar izraža stavek '*p*' ali o čemer je propozicija, da *p*. Naj '*B*'

označuje »Napovedano je, da bo igralec čez eno sekundo pritisnil gumb,« 'G' pa »Igralec pritisne na gumb.« Razmišljanje lahko potem logično prikažemo na dva načina (prim. recimo Priest, 2017: 39-44):

PRVI	DRUGI
1. $B \rightarrow N G$	1. $N (B \rightarrow G)$
2. $\sim B \rightarrow N \sim G$	2. $N (\sim B \rightarrow \sim G)$
3. $B \vee \sim B$	3. $B \vee \sim B$
Torej,	Torej,
4. $N G \vee N \sim G$	4. $N G \vee N \sim G$

V PRVEM primeru je sklepanje veljavno, ampak prva in druga premisa sta sporni. V DRUGEM primeru pa sta premisi sprejemljivi, le da sklep *ne* sledi.

Zakaj je neresnična prva premisa, v kateri trdimo, da iz resnice napovedi sledi njena nujnost? Razprava o bodočih kontingencah ima častitljivo zgodovino, h kateri pa tu ne bom nič dodajal in spoštovanja vredno 'tehnično' obravnavo v sodobnosti (večvalentne logike, supervaluacija), ki se ji bom v glavnem izognil (prim. Šuster, 2014). Mislim, da je iz zgornje razprave jasno, kaj je narobe s prvo premiso igralca (PRVI vzorec):

Če je napovedano, da bo igralec pritisnil na gumb, potem je neogibno, da bo pritisnil na gumb.

Ali smo zares nemočni glede tega, ali bomo pritisnili na gumb? Spomnimo se: pritisk gumba in z njim povezan blisk svetlobe sta odvisna od igralčevih namer. Zato je to dejanje pod njegovim nadzorom, določajo ga njegova volja in njegovi razlogi. Prva (in po analogiji druga) premisa v PRVEM vzorcu je neresnična. Ampak, ali ni tako, da blisku *nujno* sledi pritisk na gumb, gre za neogibno zvezo? Res je, ampak potem gre za DRUGI vzorec in premiso:

Neogibno je, da, če je napovedano, da bo igralec pritisnil na gumb, potem bo pritisnil na gumb.

Ta premisa je resnična, ampak sklepanje ni veljavno. Vzemimo spet Retrodiktorja, denimo, da pritisnem na gumb. Zato velja tudi: »Neogibno je, če pritisnem na gumb, potem sekundo kasneje blisne svetloba.« Ampak iz tega ne sledi sklep, da je blisk svetlobe neogiben: lahko bi se odločil drugače in na gumb ne bi pritisnil (to je, navsezadnje, odvisno od moje volje). Premisi (»Pritisnem na gumb«, »Neogibno je, če pritisnem na gumb, potem sekundo kasneje blisne svetloba«) sta resnični, sklep (»Neogibno je, da blisne svetloba«) pa ne. Kot bi sklepali: »Imam sina in nujno, če imam sina, potem imam otroka. Torej je nujno, da imam otroka.« *Non sequitur!*

Kaj pa če gre za drugačen vzorec prenosa nemoči?

1. N B

2. N ($B \rightarrow G$)

Torej,

3. N B

Gre za znameniti 'modalni katapult', ki 'izstreli' začetno neogibnost na sklep prek vezi, ki so same neogibne (prim. Šuster, 2009). Recimo, pri Retrodiktorju: Neogibno je, da pritisnem na gumb in neogibno, če pritisnem na gumb, potem sekundo kasneje blisne svetloba. Torej je neogibno, da blisne svetloba. Razprava o tem načelu (znanem kot *Beta*) je razvejana, a se ji lahko izognemo, saj je jasno, da je prva premisa tako pri Retrodiktorju kot pri Prediktorju sporna. Zakaj bi bilo neogibno, da se pojavi blisk? To je neogibno samo, če je pritisk na gumb nujen glede na njegovo vzročno zgodovino, torej, če se igralec ne more odločiti drugače. Ampak o tem sploh ni bilo govora – rekli smo, da je to dejanje odvisno od igralčevih namer. Če bi *te* bile neogibne, potem bi nas res lahko začeli skrbeti za našo svobodo.

Prav to, po mnenju nekaterih, dokazuje znameniti *dejanski* eksperiment, še ena variacija Prediktorja, imenujmo jo Libet-Prediktor. Oseba naj bi svobodno izvedla neko enostavno dejanje – recimo premaknila prst. Meritve pokažejo, da pred dejanjem nastopi električna aktivnost v tistem področju možganov, ki sproži aktivacijo mišice, ki potem vodi do premika prsta. Toda ta aktivnost nastopi že 550 ms pred samim dejanjem, oseba pa se, po teh meritvah, zave svoje namere šele 150-200 ms pred premikom prsta. Torej *pravi* vzroki dejanj niso naše namere, saj se vzročna veriga sproži, preden se sploh zavemo odločitve, da dejanja izvedemo (prim. Libet, 2007)? Tudi ta *pravi* eksperiment je vodil do trditev, da je odsotnost svobodne volje zdaj *znanstveno* dejstvo, ne samo filozofska teorija (Harris 2012) in smešnih naslovov v popularnih medijih: »Moji možgani so me prisilili, da sem to storil« (angl. *My Brain Made Me Do It*). Kot da so moji možgani nek sadistični trener, ki me priganja k delanju sklec, ne pa substanca moje istovetnosti. In daleč od tega, da je odsotnost svobodne volje zdaj znanstveno dejstvo (prim. Šuster, 2007: 318–321). Eksperimenti pokažejo samo, da obstaja nek vzorec aktivnosti pred odločitvijo, ne pa, da je bila *odločitev* sprejeta preden smo se mi tega zavedali (prim. Mele, 2014). Interpretacija obeh eksperimentov, fiktivnega Prediktorja in dejanskega Libet-Prediktorja zahteva nekaj filozofskega pojmovnega dela!

4.

Pretekla resnica je del preteklosti (tudi, če je ta samo eno sekundo nazaj) in preteklost je fiksna, se glasi *mantra* fatalizma. Toda resnična (čeprav pretekla) napoved je rezultat naših odločitev in blisk

je samo 'senca', ki jo *naše* namerno dejanje meče v času nazaj. Podobno, mislim, velja za glavni argument fatalizma: resnica je odvisna od sveta in ne obratno, resnična napoved je samo senca, ki jo tisto, o čemer je, 'meče' v času nazaj (prim. Šuster, 2014). Torej *ne*, Prediktor ne dokazuje, da svobodna volja ne obstaja, glavni miselni eksperiment je neuspešen.

Pa vendar, miselni eksperiment nas opozarja, da poleg smeri časa in smeri vzročnosti obstaja tudi smer našega delovanja, ki je v našem doživljanju istovetna s smerjo časa. Prihodnost je odprta in pod našim nadzorom, preteklost je fiksna in vzročno izolirana. To je vir psihološke prepričljivosti zgodbe. Predstavljajte si, da se vrata odprejo, šele potem jih odklenete; da lonček s kavo zavre, šele potem ga pristavite na ogenj; da se najprej pojavijo črke na računalniškem monitorju, šele potem jih natipkate itd. Najmanj, kar lahko predvidimo, je to, da takšno časovno obrnjeno zaporedje vodi do psiholoških težav, ki jih Chiang tako dramatično prikaže. Če veš, kaj se bo zgodilo, zakaj ne poskusiš tega preprečiti, to je naša naravna reakcija, ki je tudi jedro argumenta 'iz preklica'. In če ne moreš, potem ti je *usojeno*, nimaš kaj. Chiang je prepričljiv v fenomenologiji fatalizma, a tudi psihologiji zasvojenosti – vsi poznamo razne viralne naprave, da ne govorimo o največjem sodobnem viru obsedenosti s pritiskanjem na gumb. Danes bi Prediktor bil aplikacija za pametne telefone, ki jih leta 2005 še ni bilo.

Zdaj se počasi izriše drugačna začetna hipoteza. Ne več, ali Prediktor dokazuje, da nimamo svobodne volje, ampak vprašanje, ali nas *doživljanje* ločitve smeri časa in smeri vzročnosti v našem delovanju vodi do fatalizma kot življenjske drže? Ločitev obeh smeri je za nas vir nelagodja, celo filozofi niso imuni, kot kažejo argumenti za domnevno pojmovno nemožnost vnazajšnjega povzročanja. Po argumentu preklica, denimo, moramo 'nenadni' pojav črk na monitorju interpretirati kot vzrok za to, da jih kasneje natipkate. Ali ni to tako, kot da bi bili naši prsti samo marionete na vrvicah usode? Za preverjanje te druge hipoteze bi moral biti scenarij bogatejši, lučka, ki zasveti sekundo, preden pritisnete na gumb ni dovolj.

Vzemimo zato še eno različico Prediktorja, imenujmo jo *Probadiktor*. Tokrat gre za napravo, kjer je vgrajeni krogotok probabilističen: brez pritiska na gumb ni bliska, toda pritisk na gumb samo v polovici primerov povzroči blisk svetlobe, mehanizem pa je povsem naključen in nepredvidljiv. Dodajmo še negativni časovni zamik – kadar se pojavi blisk, sekundo kasneje igralec 'nujno' pritisne na gumb. Še vedno velja: če čakate na blisk z namero, da bi se kasneje odpovedali pritisku gumba, potem bliska ni. Ampak, če poskusite pritisniti gumb, ne da bi videli blisk, vam to v polovici poskusov uspe, v polovici pa ne. Torej ni več res, da ne glede na to, kaj naredite, luč vedno zasveti, preden pritisnete na gumb. Kako bo igralka doživljala takšno igro?

Ne pri Prediktorju ne pri Probadiktorju igralka ne more privedi do situacije, v kateri se pojavi blisk, ona pa po eni sekundi ne pritisne na gumb. Kot rečeno gre za blokado, ki je *notranja* samemu scenariju: nameravati pritisniti gumb (blisk registrira učinek namere), ampak obenem tega tudi ne hoteti. Igralka bo frustrirana, ampak mislim, da je v ozadju tipična psihološka razklanost: nekaj nameravamo, čeprav tega nočemo. Če prevlada občutek nemoči in blokade, lahko konča v fatalizmu. Ampak tako kot pri Prediktorju: vedno ji uspe pritisk na gumb, kadar to (enostavno) namerava. V polovici primerov ji pri Probadiktorju uspe tudi bolj zapletena namera pritisniti gumb brez predhodnega bliska, pri Prediktorju ji to nikdar ne uspe. Popolnega nadzora pa še vedno nima: nemogoče je napovedati, kdaj bo njena (zapletena) namera uspela in kdaj ne. Tudi tu lahko prevlada občutek, da gre za igro usode ali neko 'silo', ki se z njo samo poigrava. Prediktor in Probadiktor kažeta, da obrat časovne in vzročne smeri lahko privede do naše *izkušnje* nemoči pred tem, kar prihaja. A to še ne pomeni, da naše izbire niso odvisne od naših namer in odločitev.

Predictor, Fatalism, and Thought Experiments

A short story by Ted Chiang (2005) is analysed as an example of a philosophical thought experiment allegedly proving that there's no such thing as free will. I explore the structure and stages of thought experiments as proposed by Mišćević (2012: 2017): design, the question, understanding, tentative conscious production, immediate spontaneous answer, variations, and generalizations. Another characteristic is a certain mental "chemistry": in imagination we separate conceptual traits that normally go together. Predictor separates the causal order from the temporal order of events. This inversion leads to dramatic collision between our experience of time and self-agency. The actual order of our action is causal, but our experiential order is temporal. Nevertheless, the thought experiment fails in its demonstration that our choices don't matter. Its logical core is a fallacious argument for fatalism.

Keywords: free will, fatalism, causality, time, thought experiment.

Literatura

- Chiang, T. (2005). »What's expected of us.« *Nature*, 436, str. 150.
- Clark, T. (2015). »The Perfectionist: Ted Chiang's science fiction wins piles of awards. When he publishes, which is hardly ever«. *The California Sunday Magazine* (20. julij 2020). URL = <https://story.californiasunday.com/ted-chiang-scifi-perfectionist/>.
- Faye, J. (2018). »Backward Causation«. V Zalta, E. N. (ur), *The Stanford Encyclopedia of Philosophy* (izdaja poletje 2018). URL = <https://plato.stanford.edu/archives/sum2018/entries/causation-backwards/>.
- Frankfurt, H. (1969). »Alternate Possibilities and Moral Responsibility.« *Journal of Philosophy*, 66, str. 829–839.
- Harris, S. (2012). *Free Will*. New York: Free Press.
- Haynes, J. (2011). »Beyond Libet: Long-Term Prediction of Free Choices from Neuroimaging Signals«. V Dehaene, S., Christen, Y. (urd), *Characterizing Consciousness: From Cognition to the Clinic? Research and Perspectives in Neurosciences*. Berlin: Springer, str. 161–174.
- Horwich, P. (1987). *Asymmetries in Time*. Massachusetts: The MIT Press.
- van Inwagen, P. (1983). *An Essay on Free Will*. Oxford: Clarendon Press.
- Libet, B. (2007). »Ali imamo svobodno voljo?« V Šuster, D. (ur), *O svobodni volji: od Leibniza do Libeta*. Maribor: Aristej, str. 303–317.
- Mele, A. (2014). *Why Science Hasn't Disproved Free Will*. Oxford: Oxford University Press.
- Miščević, N. (2012). »Political Thought Experiments from Plato to Rawls«. V M. Frappier, L. Meynell, R. Brown (urd.), *Thought Experiments in Science, Philosophy, and the Arts*. New York: Routledge, str. 191–206.
- Miščević, N. (2017). »In Defense of the Twin Earth – The Star Wars Continue«. *European Journal of Analytic Philosophy*, 12(2), str. 111–130.
- Priest, G. (2017). *Logic. A Very Short Introduction*. 2nd. Oxford: Oxford University Press.
- Šuster, D. (ur). (2007). *O svobodni volji: od Leibniza do Libeta*. Maribor: Aristej.
- Šuster, D. (2009). »'Modalni katapult' in argument za nezdružljivost«. *Analiza*, 13, str. 5–43.
- Šuster, D. (2014). »Med resnico in neresnico.« *Analiza*, 18, str. 19–56.
- Tooley, M. (1987). *Causation*. Oxford: Clarendon Press.

