

НИСАМ МОГАО ДА УЧИНИМ ДРУГАЧИЈЕ – ПА ШТА?¹

Кад год је прогрес поводом неког филозофског проблема у застоју, тактика вредна покушаја је наћи неку заједничку (и стога претежно неиспитану) претпоставку и побити је. Проблем слободе воље је такав проблем, и, као што Питер ван Инваген (Peter van Inwagen) примећује:

„...скоро сви филозофи се слажу да је нужан услов за сматрање агента одговорним за неко дело веровање да је агент могао да се уздржи од извршења тог дела“.²

Можда је Ван Инваген у праву; можда се већина филозофа слаже са овим. Ако је тако, ова заједничка претпоставка, коју ћу ја назвати МУД („моћи учинити другачије“), добар је кандидат за побијање, нарочито зато што се испоставља да има тако мало тога што би могло бити речено у прилог њој, једном када је доведена у питање. Ја ћу аргументисати да ова претпоставка, баш као они људи који су познати једино зато што су познати, дугује свој традиционално високи углед ничему другом до свом традиционално високом угледу. Готово да никада није доведена у питање. А сама традиција је, тврдим, испрва мотивисана нечим мало знатнијим од немарне екстраполације из сличних случајева.

Да заоштрим став, ја тврдим да је потпуно неважно за моралну одговорност да ли је агент могао да учини другачије у датим околностима. Како се, пак, поборник МУД-а сада припрема да покаже да ја очигледно грешим? Не тако што ме подсећа, непотребно, на широки консензус у филозофији у прилог МУД принципа, или понављајући га, непоколебљиво и вешто. Инерција традиције је сама по себи оскудна препорука, и ако се тврди да претпоставка није доведена у питање јер је очигледна, или самоевидентна, ја могу затражити макар неку илустрацију која би подржала самоевиденцију претпоставке у примени на сличне случајеве. Може ли ми ико навести пример некога ко се уздржава од суда о одговорности пре него што утврди (на сопствено задовољство) да ли је агент могао да учини другачије?

Можда ће изгледати да ја морам бити невероватно неупућен у предмете свакодневних разговора ако могу то питање поставити озбиљно. Истакну-

¹ Текст је преведен према изворнику: Daniel C. Dennett, *I could not have Done Otherwise – So What?*, *The Journal of Philosophy*, Volume 81, Issue 10, Eighty-First Annual Meeting American Philosophical Association, Eastern Division (Oct., 1984), 553–565.

² „The Incompatibility of Free Will and Determinism“, *Philosophical Studies*, xxv, 3 (March 1975): 185–99, p. 188, прештампао у Gary Watson, ed., *Free Will* (New York: Oxford, 1982): 46–58, p. 50.

та одлика многих актуелних расправа о одговорности појединачних агената је постављање питања „да ли је могао да учини другачије?“. Ово питање се поставља у парницама, како у грађанским, тако и у кривичним, али много чешће у ретроспективним расправама између појединаца које се тичу кривице или оправдања поводом појединачних дела због којих жалимо, а која настају услед пропуштања или из дужности.

Пре детаљнијег испитивања ових случајева, ваља приметити да поменуто питање не игра готово никакву улогу у расправама о похвали или награди за подесна дела због којих не жалимо – осим у уобичајеним учтивим опаскама некога ко је издвојен за похвалу или признање: „шта друго сам могао да урадим?“ („Било ко други би учинио исто“; „Ах; није то било ништа“). У овим примерима ми не сматрамо да агент пориче одговорност уопште, већ само изјављује да није било тешко бити одговоран под овим условима.

Можда је један разлог зашто не питамо „да ли је могао да учини другачије?“ када покушавамо да проценимо одговорност за добра дела и успехе тај што ми дајемо агентима олакшицу сумње (захваљујући великодушности нашег духа) када су урадили нешто добро, по нама, радије него да се задубљујемо превише скрупулозно у чињенице крајњег ауторства. Такав добронамеран импулс може одиграти одређену улогу, али постоје бољи разлози, као што ћемо видети. Но сасвим сигурно постављамо то питање када дело заслужује осуду. Али и када то учинимо, никада не искористимо познато питање да уведемо такву врсту испитивања које би бацило светло на традиционалну филозофску тему, за коју се очекује да би требало да буде покренута поменутиим питањем. Уместо тога ми настављамо потрагу за доказом онога што ја називам жаристем *локалног фатализма*: посебне околности у релевантном делу прошлости која је обезбедила да агент не учини другачије (током периода локалног фатализма) *без обзира шта је покушао, или желео, да уради*. Стандардни пример локалног фатализма је бити закључан у соби.³

Ако је агент закључан у соби (или је на неки други начин његова воља онеспособљена), онда се независно од истинитости или лажности детерминизма и невезано за то која врста узрочности влада унутар агентовог мозга (или душе, речено картезијански), слажемо да „он није могао да учини другачије“. Унапред одредљива емпиријска чињеница да је агент био под утицајем локалног фатализма прекида испитивање о узрочности. (Али, то не решава увек питање одговорности, како Хери Франкфурт (Harry Frankfurt) показује⁴; под специфичним околностима агент би ипак могао бити сматран одговорним.) А ако наше истраживање не успева да открије било какав доказ о таквом локалном фатализму, испитивање се такође завршава. Ми разматрамо завршену ствар:

³ Погрешна употреба овог стандардног примера – у екстраполацији до теме да, ако је детерминизам истинит, цео свет је затвор – описана је у мом *Elbow Room: The Varieties of Free Will Worth Wanting* (Cambridge, Mass.: Bradford/MIT Press, 1984), из кога су извучени делови садашњег аргумента.

⁴ „Alternate Possibilities and Moral Responsibility“, *The Journal of Philosophy*, LXV, 23, (Dec. 4, 1969): 829–833.

агент је ипак био одговоран; „могао је да учини другачије“. Доказивање негативне егзистенције не сматра се проблемом овде. Углавном можемо да рачунамо на то да би агент сам могао да нам скрене пажњу на било какав доказ локалног фатализма којег је свестан (ако може да докаже да су „му руке биле везане“, онда ће ово послужити да га изузме од кривице), тако да се његови неуспеси да изађе са таквим доказом узимају као поуздан (али не и потпуно сигуран) знак да је случај затворен.

Прва поента коју желим да истакнем је да ако поборници МУД-а погледају свакодневну праксу тражећи доказ за тврдњу да се *обични људи* „слажу да је нужни услов за сматрање агента одговорним за неко дело веровање да је агент могао да се уздржи од извршења тог дела“, они заправо неће наћи такву подршку. Када је дело које се разматра за похвалу, људи очигледно игноришу то питање и било би чудно када не би. А када се процењује неко дело за осуду, иако људи заиста питају „да ли је могао да учини другачије?“, они не показују никакав интерес да прате то питање даље од тачке где су задовољили своју знатижељу о егзистенцији или одсуству локалног фатализма – феномена који је потпуно неутралан између детерминизма и индетерминизма. На пример, људи се никада не уздржавају од суда о одговорности пре него што су консултовали физичаре (или метафизичаре или неуронаучнике) поводом њиховог мишљења о коначном статусу – детерминистичком или индетерминистичком – неуралних или менталних догађаја који су управљали агентовим понашањем. И, колико ја знам, ниједан адвокат одбране никада није ушао у судницу да изгради одбрану засновану на покушају да установи, путем експертског сведочења, да је оптужени био детерминисан да донесе одлуку која је довела до страшног дела и да стога није могао да учини другачије, те да стога („очигледно“) не би требало да буде сматран одговорним.

Стога МУД принцип није нешто што „сви знају“ чак и ако се већина филозофа слаже поводом њега. Принцип захтева аргумент који би га подржао. Моја друга поента је да се сваки такав подржавајући аргумент мора суочити са широком заступљеношћу крајње познатог доказа којим се сугерише да је често, када изгледа да смо заинтересовани за питање о томе да ли је агент могао да учини другачије, то зато што желимо да извучемо супротан закључак о одговорности у односу на онај који изводи филозофска традиција.

„Стојим овде“, рекао је Лутер. „Не могу да учиним другачије“. Лутер је тврдио да није могао да учини другачије, да му је савест *онемогућила* да порекне то. Он је могао, наравно, да не буде у праву, или да намерно пренагласи истину, али чак и ако јесте, можда нарочито ако јесте – његова објава је сведочанство о чињеници да ми просто не ослобађамо некога од осуде или похвале за неко дело зато што мислимо да није могао да учини другачије. Шта год да је Лутер радио, свакако није покушавао да изврда од одговорности.

Постоје случајеви у којима је тврдња „не могу да учиним другачије“ признање слабости: претпоставимо да је оно што треба да урадим да уђем у авион и одлетим њиме на сигурно место, али ја стојим укопан у месту и признајем да не могу да учиним другачије – због мог ирационалног и паралишућег страха

од летења. У таквом случају ја не могу учинити другачије, тврдим, зато што је моја способност рационалне контроле умањена. То је заиста оправдавајуће стање. Али у осталим случајевима, какав је на пример Лутеров, када кажем да не могу да учиним другачије, мислим да не могу зато што јасно видим каква је ситуација и зато што моја способност рационалне контроле *није* умањена. Ис-увише је очигледно шта треба чинити; разум диктира то; требало би да будем луд да учиним другачије, а, пошто нисам луд, ја не могу да учиним другачије.

Надам се да је истинито – и мислим да је велика вероватноћа да је истинито – да би било немогуће навести ме да мучим недужну особу нудећи ми хиљаду долара. „Ах“ – долази приговор – „али шта ако би неки зли свемирски гусари држали читав свет као таоца и обећали да га неће уништити само ако би ти мучио недужну особу? Да ли би и тада сматрао да је то нешто што је немогуће учинити?“ Вероватно не, али шта с тим? То је веома различит случај. Ако је оно за шта је неко заинтересован то да ли сам *иод наведеним околностима* могао да учиним другачије, онда је други поменути случај потпуно ирелевантан. Тврдио сам да би било немогуће навести ме да мучим недужну особу *за хиљаду долара*. Они којима је МУД принцип близак увек инсистирају да треба размотрити да ли је неко могао да учини другачије у *иошћуно* истим околностима. Ја тврдим нешто јаче; ја тврдим да не бих могао да учиним другачије чак ни у приближно сличном случају. *Никада* не бих пристао да мучим недужну особу за хиљаду долара. Не би правило разлику, тврдим, какав тон је подмићивач користио, или да ли сам био уморан или гладан, или да ли је потенцијална жртва била добро осветљена или делимично скривена у сенци. Ја сам, надам се, имун на све такве понуде.

Зашто би сада било чија интуиција сугерисала, ако сам у праву, да ако и када ја имам прилику да одбијем такву понуду, моје одбијање неће да се рачуна као одговорно дело? Ово је можда оно што неки људи мисле: они мисле да ако сам ја био у праву када сам тврдио да не бих могао да учиним другачије у таквим случајевима, ја бих био нека врста зомбија, „програмиран“ да увек одбијем подмићивања од хиљаду долара. Истински слободан агент, мисле они, мора бити променљивији некако. Ако би требало да будем у стању да саслушам разлоге, ако би требало да будем флексибилан на прави начин, мисле они, не смем бити исувише догматичан. Чак и у најбесмисленијим случајевима, онда, морам бити у стању да увидим да „свако питање поседује две стране“. Морам бити у стању да станем и извагам разлоге за и против поводом ове предложене уносне тортуре. Али једини начин на који бих могао да будем сачињен тако да увек могу да видим „обе стране“ – без обзира колико је бесмислена једна од њих – јесте да будем сачињен тако да сам у *сваком иојединачном случају* „могао да учиним другачије“.

То би било погрешно расуђивање. Виђење обе стране питања не подразумева да неко не буде надмоћно убеђен, на крају, једном страном. Флексибилност коју желимо да има одговорни агент је флексибилност да препозна један у милијарду случајева у коме, захваљујући тих хиљаду долара, који се не могу другачије добити, свет може бити спасен (или било шта друго). Али општа

способност да се флексибилно одговори у таквим случајевима уопште не захтева да је неко могао да учини другачије у неком појединачном случају, или у било ком појединачном случају, већ само да би под извесним варијацијама у околностима – варијацијама које су битне – неко учинио другачије. Филозофи су често запажали, нерадо, да случајеви тешких моралних проблема, одлуке које „могу отићи било на једну или на другу страну“, нису једине, нити чак најчешће, врсте одлука због којих људе сматрамо одговорним. Али они су ретко узимали за озбиљно тај наговештај и питали се да ли је МУД принцип једноставно погрешан.

Да наша одговорност заиста зависи, као што већинска филозофска традиција инсистира, од питања да ли икада можемо да учинимо другачије него што заиста чинимо у *ѿошћуно исћим околносћима*, били бисмо суочени са најчуднијим проблемом незнања: било би крајње невероватно, с обзиром на тренутну ситуацију у физици, да би ико икада знао да ли је ико икада био одговоран. Данас прихваћено схватање је да индетерминизам влада на субатомском нивоу квантне механике; дакле, у одсуству било ког општег и прихваћеног аргумента за универзални детерминизам, могуће је, колико ми знамо, да су наше одлуке и акције увећане, макроскопске последице индетерминистичких узрока у нашим мозговима на квантном нивоу. Али исто тако је могуће, колико ми знамо, да и поред тога што индетерминизам влада у нашим мозговима на субатомском квантно-механичком нивоу, наше макроскопске одлуке и дела буду самима собом одређени; квантне последице би такође могле бити самоукидајуће, а не појачане (као на пример органским Гајгеровим бројачима у неуронима). Такође, крајње је невероватно, с обзиром на комплексност мозга чак и на молекуларном нивоу (комплексност због које је реч „астрономски“ веома непрецизан израз), да бисмо икада могли да развијемо добар доказ да је било које појединачно дело било ефекат тако великих размера критичног субатомског индетерминизма. Дакле, ако је нечија одговорност за одређено дело зависила од тога да ли је, у тренутку одлуке, та одлука (већ) била детерминисана претходним стањем света, онда су, осим услед тријумфалног повратка универзалног детерминизма у микрофизику (који би отклонио сваку одговорност поводом овог гледишта), изгledi веома велики да ми никада нећемо имати *никакав* разлог да верујемо за било које појединачно дело да јесте или да није било одговорно. Критична разлика била би потпуно загонетна из било које макроскопске повољне позиције и практично загонетна из било које префињене микрофизичке повољне позиције коју је могуће замислити.

Већ смо видели да обични људи, када су заинтересовани за приписивање одговорности, заправо не упућују своје испитивање на то да ли су њихови пријатељи могли да учине другачије. Сада видимо и разлог зашто не би било мудро да покушају то: због чисте немогућности да се спроведе било какво значајније испитивање по овом питању – осим у случајевима у којима је откривен макроскопски локални фатализам. Шта онда људи мисле да раде када постављају МУД питање у појединачним случајевима? Да ли они заиста постављају метафизичко питање филозофа о томе да ли је агент био детерминисан

да учини оно што је учинио, с тим да одмах одустају чим испитивање постане тешко а изгледи за срећно проналажење негативног одговора (открићем неког локалног фатализма) постану нејасни? Не, јер постоји боље питање које су они можда све време постављали, питање које се *из принципских разлога* зауставља са закључком да је локални фатализам одбачен (или прихваћен). Оно је боље из два разлога: на њега је најчешће могуће емпиријски одговорити, и његов одговор је битан. А на традиционално метафизичко питање не само да није могуће одговорити; његов одговор, чак и када бисте га знали, био би бескористан.

Какве користи бисмо имали од знања, о неком агенту, да је у некој прилици (или у свакој прилици) могао да учини другачије него што је учинио? Или да није могао да учини другачије него што је учинио? Размотримо најпре потоњи случај. Претпоставимо да сте знали (рецимо зато што вам је Бог рекао) да када је Џонс повукао обарач и убио своју жену у времену t , он *није* могао да учини другачије. То јест, узимајући у обзир Џонсово микростање у времену t и целокупно микростање Џонсовог окружења (укључујући и гравитациона дејства удаљених звезда, итд.) у времену t , ниједна друга Џонсова путања није била могућа осим путање којом је кренуо. Ако би Џонс икада био стављен у потпуно исто стање, у потпуно истим околностима, он би поново повукао обарач – и ако би био стављен у то стање милион пута, повукао би обарач исто толико пута.

Ако сте сазнали ово, да ли сте тиме сазнали било шта значајно о Џонсу? Да ли сте сазнали било шта о његовом карактеру, на пример, или о његовом вероватном понашању у сличним приликама? Не. Иако су људи физички објекти који, као атоми или лагери или мостови, подлежу законима физике, они нису само компликованији од било чега у универзуму са чим смо упознати; они су такође осмишљени тако да буду веома осетљиви на пролазну представу тако да не могу бити никада у истом микростању двапут. Неко уопште не мора да се спусти на атомски ниво да би установио ово. Људи уче, памте, постаје им досадно, преусмеравају своју пажњу и непрестано мењају своја интересовања, тако да је готово бесконачно невероватно да је било која особа два пута у истом (целокупном) *психичком односно когнитивном* стању. Ово би било истинито чак и када бисмо произвели „потпуно исте“ околности у различитим приликама – ако ни због чега другог, онда зато што би други пут агент без сумње мислио о нечему што претходног пута није било мишљено, као на пример „О, све ово изгледа тако познато; шта ли сам оно урадио прошлог пута?“

Сазнање (од Бога, поново) да неки агент *није* био тако детерминисан да дела било би сазнање нечег једнако узалудног, са тачке гледишта процене карактера или планирања будућности. Истински недетерминисан агент није више флексибилан, прилагодљив, осетљив за нијансе, или променљив него што би његов детерминистички пандан био.⁵

Дакле, ако је било ко уопште заинтересован за питање о томе да ли је неко могао да учини другачије у *пошћуно* истим околностима (и унутрашњем стању) то би морала да буде чисто метафизичка знатижеља – то јест, знатижеља

⁵ Ово је показано у „Designing the Perfect Deliberator“, у *Elbow Room*.

толико чиста да би морала бити без икаквих скривених мотива, пошто одговор не би направио убедљиво никакву приметну разлику у односу на то како су се ствари одвијале.

Ако је мало вероватно да је од икаквог значаја да ли је особа могла да учини другачије – када погледамо микроскопски детаљно у узрочност која је на делу – које је друго питање за које смо (и за које би требало да смо) заинтересовани када питамо „али, да ли је могао да учини другачије?“? Размотримо слично питање које се може поставити о роботу, одређеног (по претпоставци) да живи читав свој живот као детерминистичка машина на детерминистичкој планети. Иако је овај робот, по претпоставци, потпуно детерминисан, он може бити контролисан путем „хеуристичких“ програма који изазивају „насумичне“ изборе – стратегија, тактика, приоритета, или било чега другог – у различитим приликама. Све што му је потребно је псеудонасумични генератор бројева, или унапред одабрана листа односно табела псеудонасумичних бројева које треба консултовати детерминистички када то прилика захтева, или алгоритам који продукује псеудонасумични низ цифара. У сваком случају он може имати неку врсту машине за бинго која би га снабдевала нешаблонском и арбитрарном серијом цифара на основу којих би управљао неким од својих активности.

Шта год да овај робот ради, он није могао да учини другачије, ако мислимо то у стриктном и метафизичком смислу тих речи на које су филозофи сконцентрисани. Претпоставимо затим да једног финог марсовског дана робот начини грешку вредну жаљења: он кује и извршава план којим уништава нешто вредно – неког другог робота, рецимо. Ја не претпостављам, ни за тренутак, да *он* може да зажали због било чега, већ само да његови ствараоци, доле на Земљи, жале због онога што је учињено и налазе се у чуду које би се природно могло изразити речима: *да ли је могао да учини другачије?* Претпоставимо да се они најпре задовољавају тиме да нема очигледног локалног фатализма (закључане собе, покварене батерије) који је утицао на њиховог робота. Али они и даље инсистирају на свом питању: да ли је могао да учини другачије? Они знају да је у питању детерминистички систем, наравно; стога они поступају паметније од постављања метафизичког питања. Њихово питање тиче се дизајна робота; јер ће с обзиром на овај догађај вредан жаљења можда пожелети да га редизајнирају, да би овакву *врсћу* догађаја учинили мање вероватном у будућности.⁶ Оно што они желе да знају, наравно, јесте информација којом се робот руководио, какво расуђивање или планирање је вршио, и да ли је употребио „довољно“ праве врсте расуђивања или планирања.

Наравно, у једном смислу „довољног“ они знају да робот није употребио довољно праве врсте расуђивања; да јесте, учинио би праву ствар. Али може бити да дизајн робота у овом случају заправо не може бити поправљен. Јер

⁶ „Ми смо једва икада заинтересовани за учинак комуникационо-инжењерске машине за један унос. Да би функционисала адекватно она мора показати задовољавајући учинак за читаву класу уноса, а то значи статистички задовољавајући учинак за класу уноса за коју се статистички очекује да буде добијен“ Norbert Wiener, *Cybernetics* (Cambridge, Mass: Technology Press; New York: Wiley, 1948), p. 55.

може се десити да је он на оптималан начин користио оптимално дизајниране хеуристичке процедуре – али овог пута, на несрећу, хеуристичке шансе које је опробао нису се исплатиле. Ставите робота у *сличну* ситуацију у будућности, и, само захваљујући чињеници да је генератор псеудонасумичних бројева у другачијем стању, он ће учинити нешто другачије; заправо он би углавном учинио праву ствар. Примамљиво је додати: *могао је* да учини праву ствар у тој прилици – мислећи тиме да је био довољно добро дизајниран, у датом тренутку, да учини праву ствар (његов „карактер“ није споран); његов неуспех није зависио ни од чега другог до од чињенице да се десило нешто *неплановано дизајном* (и непредвидиво) што је интервенисало у процесу на начин који је омогућио несрећну разлику.

Хеуристички програм не доноси загарантовано „прави“ или тражени резултат. Неки хеуристички програми су бољи од других; када један не успе, могуће је установити да се неуспех може приписати некој карактеристичној слабости у његовом дизајну, али чак и најбољи нису потпуно сигурни, и када они закажу, као што некада и морају, и даље уопште не мора да постоји *разлог* за неуспех; као што би Кол Портер (Cole Porter) рекао, то је била само једна од оних ствари.

Такви неуспеси нису једини случајеви неуспеха који би се „рачунали“ за дизајнере као случајеви у којима је систем „могао да учини другачије“. Ако открију да је роботов неуспех, у овој ситуацији, узрокован „несташном“ трунком прашине која је некако залутала на место где би могла да поремети систем, они могу да одлуче да је ово био тако мало вероватан догађај да нема потребе да се редизајнира систем да би се одбранио од понављања истих.⁷ Они ће приметити да на нивоу микрочестица (као и увек) њихов робот није могао да учини другачије; штавише, ако би се (путем неке веома слабе могућности) икада нашао у потпуно истим околностима поново, поново би доживео неуспех. Али дизајнери ће схватити да немају никакав рационалан интерес да било шта учине да побољшају дизајн свог робота. Он није успео у датој ситуацији, али је његов дизајн ипак беспрекоран. Постоји разлика између тога бити оптимално дизајниран и бити непогрешив.

Размотримо, пак, још једну врсту случаја. Робот поседује ласерску пушку која пуца са 99,9% прецизности. То јест, понекад, на великој удаљености, она не успева да погоди мету на коју је циљала. Кад год промаши, инжињери желе да знају нешто о промашају; да ли је он последица неке *систематске* грешке у контролама, нека мана или дефект који ће се понављати, или је то само једна од оних ствари – једно од тих „Божјих дела“ у коме је, упркос беспрекорном извршавању оптимално дизајнираног поступка циљања, ствар једноставно за мало промашила? Увек ће постојати такви случајеви; циљ је свести их на минимум – у складу са исплативошћу, наравно. Преко извесне границе није вредно

⁷ Строго говорећи, од понављања догађаја *овог ошћег типца*; нема потребе бранити се од понављања појединачних догађаја (то је логички немогуће) или од понављања догађаја *пошћуно* истог типа (то је номолошки немогуће).

бринути о грешкама. В. В. Квајн (W. V. Quine) примећује да инжењери имају појам о више него пролазном филозофском интересу: појам „занемарљивог“ – тј. случајева које је рационално игнорисати.⁸ Када су задовољени тиме да неки промашај спада у оно занемарљиво, они могу слегнути раменима и рећи: „па, то је могао бити погодак“.

Оно о чему инжењери воде рачуна када се сусретну са грешком у функционисању свог робота је то да ли је грешка у функционисању информативна: да ли открива нешто о схеми систематске слабости, за коју постоји вероватноћа поновног дешавања, или о неадекватној и несрећној вези између врста околности и врста реакција? Да ли је ова *врсћа* ствари подложна томе да се дешава поново, или је то само последица случајне конвергенције фундаментално независних фактора, за које готово не постоји вероватноћа поновног дешавања? Докле год њихов робот *не* функционише погрешно већ доноси „праве“ одлуке, поента постављања питања да ли је могао да учини другачије јесте у томе да потврде себи да подесно понашање није случајност или пука коинциденција већ пре резултат доброг дизајна. Они се надају да ће њихов робот, налик Лутеру, бити непоколебљив у својој мисији.

Да би добили доказ за ово они игноришу микродетаље, који никада неће бити исти у било ком случају, рачунају само просек између њих, анализирајући робота до коначног низа *макроскопски* одређених стања, организованих на такав начин да постоје везе између различитих степена слободе унутар система. Питање које они потом могу да поставе је: да ли су ове везе праве везе за овај задатак?

Ово образложење за игнорисање микродетерминизма (где год да он „у принципу“ постоји) и гледање кроз прсте таман толико да се замагле такве fine дистинкције у пробабилистички повезаним стањима и регионима, који се могу *шрећирати* као хомогени, је чисто, сигурно и непроблематично у науци, нарочито у инжењерингу и биологији. То не значи, наравно, да је ово такође прави начин да се мисли и о људима, када се питамо да ли су делали одговорно. Али у прилог томе има доста тога да се каже.

Зашто постављамо питање „да ли је могао да учини другачије?“? Питамо зато што се десило нешто што желимо да интерпретирамо. Извршено је извесно дело и ми желимо да схватимо како је дошло до њега, зашто је до њега дошло, и које значење бисмо могли да му припишемо. То јест, желимо да знамо какве закључке о будућности можемо извући из њега. Да ли нам оно нешто говори о агентовом карактеру, на пример? Да ли предлаже критику агента, која би могла, ако би била представљена на прави начин, да наведе агента да поправи своје понашање у извесном смислу? Да ли можемо да сазнамо из овог догађаја да ово јесте или да ово није агент за којег се може веровати да ће се понашати *слично* у *сличним* ситуацијама у будућности? Ако би неко одржавао свој карактер константним, али би мењао околности у незнатном – или чак у знатнијем – смислу, да ли би готово увек учинио исту врсту жалосне ствари?

⁸ *Word and Object* (Cambridge, Mass.: MIT Press, 1960), pp. 182, 259.

Да ли је оно што смо посматрали пука случајност, или је испољавање снажне тежње – тежње која истрајава, или је константна, упркос прилично великој разноликости услова?⁹

Када се агент пита о самом себи, онда је ово образложење још јасније уочљиво. Претпоставимо да ја откријем да сам учинио нешто страшно. *Кога је брига* да ли сам, у потпуно истим околностима и стању духа у коме се налазим, могао да учиним нешто друго? Нисам учинио нешто друго, и исувише је касно да поништим оно што сам учинио. Али када покушам да интерпретирам то што сам учинио, шта сазнајем о себи? Да ли би требало да вежбам ту врсту поступка који сам упропастио неспретношћу, у нади да ћу га учинити поузданијим, мање осетљивим на сметње, или би то био узалудан труд? Да ли би била добра ствар за мене, колико за сада могу да оценим, да покушам да прилагодим своје навике мишљења у таквим врстама случајева у будућности? Знајући да ћу увек донекле бити у немилости разматрања за која се испоставља да једва да се одвијају у мојој глави док време журбано пролази, знајући да не могу у потпуности да контролишем овај процес одлучивања, можда предузнем кораке да смањим вероватноћу извесних врста разматрања која обично „падају на памет“ у извесним критичним ситуацијама. На пример, ја могу да тежим да изградим навiku да избројим до 10 у себи пре него што кажем било шта о Роналду Регану (Ronald Reagan), научивши да се време одлучивања тако стечено добро исплати јер умањује нежељене испаде у виду неумерених коментара. Или могу да одлучим да, без обзира колико сам удубљен у разговор, морам научити да запитам себе колико сам чаша вина попио сваки пут када видим некога са флашом како гостољубиво обиграва око моје чаше. Овог пута сам испао будала; да је ситуација била прилично другачија, ја бих свакако учинио другачије; ако би ситуација била практично иста, ја бих можда учинио другачије, а можда не бих. Главна ствар је увидети да бих у будућности врло радо учинио другачије у ситуацијама које су (само) сличне овој.

То је, свакако, здрав однос који неко треба да изгради према деловима своје прошлости који су вредни жаљења. То је самопримењена верзија инжењерског става према постојаним слабостима у дизајну робота. Наравно, ако ћу радије наћи изговоре него поправити себе, ја могу да се задржим на чињеници да не *морам* да „преузем“ одговорност за моја дела, пошто увек могу да замислим уситњеније становиште са кога моја неприлика изгледа већом од мене. (Ако учините себе довољно малим, можете екстернализовати практично све.) Али ми мудро обесхрабрујемо ово бекство на уситњеније видове уклапања у свет, готово из истог разлога из кога је избегавано од стране инжењера: оно што можемо да схватимо из оваквог испитивања није ни од каквог значаја. Једноставно је небитно да ли је неко могао да учини другачије.

За робота то није битно, може неко узвратити, јер робот не може да *служи* казну или осуду за своје прекршаје. За нас је то битно зато што смо

⁹ Заинтересовани смо за тежње и случајности у оба смисла (и похвалном и оном вредног жаљења); ако бисмо имали доказ да се Лутер само шалио, да је његов наизглед чврст став био само нека врста комичне коинциденције, наш осећај поводом његове моралне снаге био би прилично умањен. „Није он тако одлучан“, могли бисмо рећи. „Могао је лако да учини другачије.“

кандидати за осуду и казну, не пуки редизајн. Не можете *кривити* некога за нешто што је учинио, ако није могао да учини другачије. Али ово је, ипак, само поновно навођење МУД принципа, а не ново разматрање, а ја оспоравам тај принцип од самог почетка. Зашто заиста не бисте кривили некога за чињење нечега од чега није могао да се уздржи? На крају крајева, ако је он то учинио, каква је разлика ако је био детерминисан да учини то?

„Разлика је у томе да ако је био детерминисан, онда *није имао шансе да не учини то*“. Али ово једноставно *не следи*, осим ако неко не заступа крајње сујеверно гледиште о томе шта је шанса. Упоредите следеће две лутрије по питању праведности. У Лутрији А, након што су сви тикети продати, њихови одсечци су убачени у одговарајући бубањ, и, након одговарајућег мешања (укључујући неку врсту изворног – квантно-механичког – насумичног мешања, ако тако преферирате), победнички тикет је извучен на слепо. У лутрији Б, мешање и извлачење тикета одигравају се *пре* него што су тикети продати, али што се осталог тиче лутрије су спроведене на исти начин. Многи људи мисле да је друга лутрија неправедна. Она је неправедна, мисле они, зато што је победнички тикет одређен пре него што људи уопште купе своје тикете; један од тих тикета је *већ* победнички; остали тикети су ништа друго до безвредан папир, и њихово продавање несумњичавим људима представља неку врсту преваре. Али у ствари, наравно, ове две лутрије су подједнако праведне: *свако има шансу да добије*. Време одабирања победника је потпуно небитна карактеристика. Обично је разлог за одлагање извлачења лутрије након продаје тикета само у томе да се јавност снабде очигледним доказом из прве руке да није било никакве преваре. Никакав тајни агент са знањем изнутра није манипулисао расподелом тикета, зато што знање о томе који је тикет победнички није било (и није могло бити) доступно ниједном агенту пре него што су тикети продати.

Занимљиво је да не прате све лутрије ову праксу. Издавачи Клерингхаус (Clearinghouse) и Ридерс Дајџест (Reader's Digest) шаљу милионе коверти сваке године на којима болдованим словима пише „МОЖДА СТЕ ВЕЋ ОСВОЈИЛИ“ – милион долара или неку другу награду. Наравно, ове скупе кампање засноване су на истраживању тржишта које показује да људи углавном мисле да су лутрије са унапред изабраним победницима праведне све док су поштено спровођене. Али можда људи одобравају ове лутрије зато што добијају бесплатно своје тикете. Да ли би већина људи *купила* тикет за лутрију у којој се за победнички одсечак, запечаћен у посебној коверти, зна да је од почетка био депонован у трезор банке? Сумњам да би већина обичних људи била безбрижна с обзиром на овакав сценарио, и да би сматрали да имају сасвим изгледну прилику да добију. То јест, сумњам да је већина обичних људи мање сујеверна од оних филозофа (враћајући се до Демокрита и Лукреција) који су убедили себе да, без сталног прилива истински насумичних *скокова* којим би се прекинула нит узрочности, не могу постојати никакве праве прилике односно шансе.

Ако је наш свет детерминисан, онда ми имамо генераторе псеудонасумичних бројева у себи, а не Гајгерове противнасумичне генераторе. То јест, ако је наш свет детерминисан, сви наши тикети за лутрију извучени су одјед-

ном, еонима уназад, стављени у коверте за нас, и раздељени када су нам били потребни у животу. „Али то није фер!“, неко ће рећи, „неким људима ће бити подељено више добитних тикета него другима“. Заиста, приликом било које поделе, неки људи имају јаче карте од других, али треба запамтити да се срећа упросечњава. „Али, ако се сва извлачења одигравају пре него што смо рођени, неки људи су *предодређени* да имају више среће од других!“ Но, то би било тачно чак и ако извлачења не би била одржана пре него што смо рођени, већ периодично, током наших живота.

Још једном, нема разлике – овог пута по питању праведности и, стога, по питању заслуге – да ли је агентова одлука била детерминисана еонима уназад (путем судбоносног тикета за лутрију смештеног у његовој можданој кутији за одлучивање, чекајући да буде употребљен), или је била индетерминистички одређена нечим налик квантном ефекту у, или непосредно пре, тренутка коначне одлуке.

Отворено је за присталице МУД принципа да покушају да пруже друге разлоге за верност овом принципу, али пошто тренутно не видим ништа што би могло да подржи ту верност сем навике верности саме, принуђен сам да закључим да би принцип требало одбацити као ништа друго до дуговечну филозофску илузију. Можда ћу погрешити ако закључим ово, наравно, али под датим околностима не могу да учиним другачије.

(Превод са енглеског: Марко Перић)