ANαliZA

časopis za kritično misel

Letnik 13, številka 1-2, junij 2009 · ISSN 1408-2969

MORALNA FILOZOFIJA R. M. HAREA

Vojko Strahovnik

Moralna filozofija R. M. Harea: univerzalni preskriptivizem in utilitarizem dveh ravni

Filozofska fakulteta Univerze v Ljubljani

Citiranje: Vojko Strahovnik, »Moralna filozofija R. M. Harea: univerzalni preskriptivizem in utilitarizem dveh ravni«, *Analiza: časopis za kritično misel* 13, št. 1-2 (2009): 47–62.

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

Vojko Strahovnik

Moralna filozofija R. M. Harea: Univerzalni preskriptivizem in utilitarizem dveh ravni

Članek poda osnovne značilnosti moralne teorije Richarda Mervyna Harea (1919-2002). Predstavljeno je mesto njegove teorije znotraj polja drugih moralnih teorij v dvajsetem stoletju. Hareova moralna teorija gradi na dveh osrednjih značilnostih logike moralnega jezika, in sicer na preskriptivnosti in univerzabilnosti moralnih sodb. Na koncu se izteče v utilitarizem, in sicer v t.i. utilitarizem dveh ravni, ki presega posamezne omejitve tako utilitarizma dejanj kot tudi utilitarizma pravil ter ponudi odgovor na večino klasičnih ugovorov, ki jih lahko postavimo utilitarizmu.

Ključne besede: etika, utilitarizem, preskriptivizem, dve ravni moralne misli, R. M. Hare, G. E. Moore

Richard Mervyn Hare (1919-2002) je bil eden najpomembnejših in najbolj znanih moralnih filozofov dvajsetega stoletja. V svoji filozofsko avtobiografijo, ki je izšla malo pred njegovo smrtjo leta 2002, je svoje filozofske poglede na izjemno zanimiv in barvit, tu in tam pa tudi duhovit način, vnesel tudi zgodbo svojega življenja. To ga je vodilo od meščanskega življenja na angleškem podeželju, smrti njegovih staršev in vstopa v vojsko ter bojevanja v drugi svetovni vojni na bojiščih v Indiji, čemur so sledila leta japonskega vojnega ujetništva v Burmi, in nato po povratku kariera univerzitetnega profesorja[1] na Balliol College v Oxfordu, nato na Corpus Christi, Oxford in nazadnje selitev na univerzo na Floridi (Gainseville, ZDA), kjer je službo sprejel tudi zato, ker v pravilih univerze niso imeli določila o prisilni upokojitvi, ki mu je zaradi njegove starosti grozila v Oxfordu. Avtobiografijo pričenja z naslednjimi besedami:

> »Pred kratkim sem imel čudne sanje, nekakšno vizijo napol v snu. Znašel sem se na vrhu gore, v gosti megli, zelo zadovoljen s sabo – ne samo zaradi tega, ker mi je uspelo splezati na vrh gore, temveč tudi zato, ker sem uresničil svoj življenjski cilj in našel način, kako razumno odgovoriti na moralna vprašanja. Toda medtem ko sem si domišljal ta dosežek, se je

[1] Hare je Univerza v Lundu na Švedskem leta 1991 podelil častni doktorat, kar je bil tudi njegov prvi in edini doktorski naziv, ki ga je pridobil.

> megla začela razkrajati in videl sem, da sem na tem vrhu gore obkrožen z grobovi vseh drugih filozofov (velikih in majhnih), ki so imeli isti cilj in za katerega so mislili, da so ga uresničili. In od tega trenutka dalje sem na podlagi razmisleka o mojih sanjah uvidel, da so prizadevni filozofski črvi pridno grizljali sisteme teh filozofov in razkrivali, da so bili njihovi dosežki zgolj utvara. Res je, da so njihovi skeleti, ki jih črvi niso mogli prebaviti, ostali nedotaknjeni, a bili so presenetljivo podobni drug drugemu. Vse to me je pripeljalo do tega, da še enkrat premislim o tem, kaj je na tem področju sploh mogoče doseči.« (Hare, 2002: 269)

Kot dve največji spodbudi za to, da se je pričel ukvarjati z moralno filozofijo, navaja kot prvo dilemo, s katero se je srečal ob bližajoči se vihri druge svetovne vojne, ko je tehtal med tem, ali naj sprejme pacifistične poglede ali pa naj vstopi v vojsko. Druga spodbuda je bilo nelagodje oz. občutje krivde, ki ga je kot mladostnik doživljal zaradi dejstva, da je bila njegova družina razmeroma premožna oz. se ji ni bilo treba ukvarjati z vprašanjem dnevnega preživetja, medtem ko je okoli sebe lahko opazoval množice brezposelnih delavcev, ki so se komaj prebijali skozi življenje. Del opisa iz zgornjega odlomka avtobiografije za Harea nedvomno drži; v vsem svojem filozofskem razvoju je sledil temeljni ideji zagotovitve *razumnega iskanja odgovorov na moralna vprašanja*. Ob tem je razvil svojstveno moralno teorijo, ki si – kljub umanjkanju velikega števila zagovornikov oz. naslednikov, zato pa toliko večjega števila Hareovih učencev (Bernard Williams, Peter Singer, Brian McGuinness, John Lucas, Richard Wollheim), na katere je pomembno vplival – zasluži mesto med osrednjimi moralnimi teorijami prejšnjega stoletja.

I. Metaetika

Kako Harea umestiti v kontekst moralnih teorij dvajsetega stoletja, ki bi ga lahko po mnogo čem označili tudi za stoletje metaetike? Hare je prvenstveno razvijal svojo moralno teorijo kot alternativo dvema, do tedaj prevladujočima moralnima teorijama, tj. *moralnemu kognitivizmu in realizmu* na eni strani, katerih najbolj vneti zagovorniki so bili britanski *intuicionisti* (Moore, Prichard, Ross in Broad) in pa *moralnemu nekognitivizmu* na drugi strani, ki sta ga utelešala Ayerjev in Stevensonov *emotivizem* in ki je temeljil na kritiki opisnega pomena moralnih sodb, ki so jo podali Carnap in drugi logični pozitivisti in ki ga je že skoraj dve desetletji prej razvijal tudi švedski filozof in pravnik Axel Hägerström. Prvi so zagovarjali sklop raznolikih trditev, med katerimi so najbolj temeljne: teza opisnosti moralnih izjav oz. sodb (semantična teza); teza, da moralne izjave oz. sodbe so oz. izražajo prepričanja (psihološka teza) in teza, da obstaja objektivna moralna resnica, na katero lahko s svojimi moralnimi sodbami zadevamo (ontološke teza) ter imamo o njej moralno vednost (epistemološka teza). Na drugi strani je nekognitivizem vse omenjene teze zavračal, vsaj v takšni neposredni obliki. Emotivisti so trdili, da moralne sodbe oz. izjave niso izjave o moralnih dejstvih ali moralnih kvalitetah dejanj, oseb in stanj stvari,

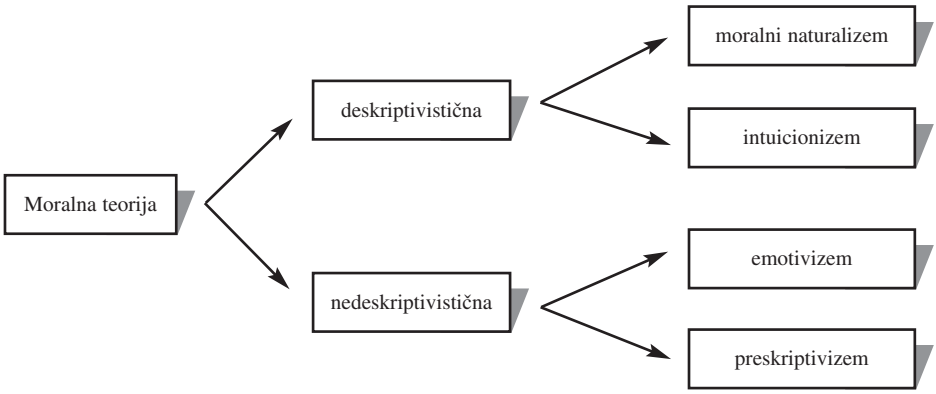
temveč zgolj odraz našim moralnih čustev in naravnanosti do teh predmetov sodb. Ob prelomu dvajsetega stoletja v njegovo drugo polovico, ko je s svojo teorijo nastopil Hare, je bil emotivizem izrazito v prevladi in skoraj splošno sprejet kot pravilna teorija moralnosti.

Hare je ponujal pomembno tretjo možnost, ki naj bi presegla omejitve prvotnih dveh alternativ (Hare, 1993: 451). Predvsem je emotivizmu nasprotoval iz razloga, ker ni mogel razumeti, kako bi takšna moralna teorija lahko ponujala način, da bi nam pomagala razumno odgovoriti na moralna vprašanja (Hare, 2002: 288). Namesto uvida moralnih dejstev, ki so ga ponujali intuicionisti in namesto golega sklicevanja na čustva in naravnanosti pri emotivizmu, je Hare poskušal razviti *logiko moralnega jezika in moralne misli*, ki bi nam omogočala razumno odločati o naših dejanjih ter tudi razumno vrednotiti ter kritizirati ravnanja drugih. Uvedba možnosti razumske razprave nazaj v našo moralno misel in torej nadgraditev emotivistične zavnitve kognitivizma, je tudi eden največjih dosežkov Hareove moralne misli, kakor ocenjuje tudi Peter Singer (Singer, 2002: 310). In če so mnogi drugi filozofi sledili Ayerjevemu napotku, da *»od moralnih filozofov ne moremo pričakovati nobenih vodil za delovanje«*, je Hare sedaj na novo odprl tudi prostor *praktične etike*, kateri se je prav tako pogosto vneto posvečal, začenši z vprašanjem rasizma. Za Harea je etika tako v prvi vrsti logika moralnega jezika oz. logika moralnih argumentov. Tej logiki pa se je uvodoma poskušal približati z raziskovanjem logike predpisov oz. imperativov, kar je osrednja vsebina njegove prve knjige *Jezik moralnosti* (*The Language of Morals*, 1952).

Preden se posvetimo logiki moralnega jezika, kakor jo vidi Hare, in normativni strukturi njegove moralne teorije, si poglejmo še natančnejšo umestitev njegove teorije znotraj polja metaetike. Zanj sta z vidika metaetike dve v temelju različni moralni teoriji *moralni deskriptivizem* in *moralni nedeskriptivizem*. Ob tem Hare zavrne običajne razmejitve med obema stališčema kot zavajajoče. Takšna je npr. razmejitev, da moralni deskriptivizem trdi, da so moralne sodbe lahko resnične ali neresnične, medtem ko nedeskriptivizem to zavrača. Prav tako nam ni v pomoč razumevanje deskriptivizma kot moralnega kognitivizma, ki dopušča vednost glede moralnih propozicij, medtem ko nedeskriptivizem kot nekognitivizem možnost vednosti zavrača. Enako je, če poskušamo razliko zarisati na področju metafizike, kjer naj bi deskriptivizem zatrjeval obstoj moralnih lastnosti in moralnih dejstev v svetu, nedeskriptivizem pa takšne vrste bitnosti zavračal. Hare prva kriterija zavrne iz razloga, ker lahko v pomembnem smislu tudi nedeskriptivist trdi, da so moralne sodbe resnične in da imamo moralno vednost. Zadnji, metafizični razlikovalni kriterij pa zavrača, ker je preveč nejasen in se pravzaprav ne ve, o čem sploh govorimo, če rečemo, da neka moralna lastnost obstaja v svetu. Tudi nasploh je Hare zavračal pomembnost ontologije in metafizičnih razprav za moralno teorijo (npr. v eseju »Ontologija v etiki«, 1985).

Razliko med obema osnovnima stališčema je najbolje podati v okviru *teorije pomena*, ki se opira na *pogoje za resničnost stavka*. Glede na to teorijo pomeni razumeti pomen stavka kot izjave razumeti resničnostne pogoje te izjave, tj. kaj mora biti res, da bi jo lahko označili za resnično. Moralni deskriptivisti tako

trdijo, da je pomen moralnih stavkov (stavkov, ki izražajo moralne sodbe) mogoče razumeti na tak način in da je torej njihov pomen mogoče določiti preko resničnostnih pogojev. Moralni nedeskriptivizem pa to zanika in trdi, da je poleg morebitnega opisnega pomena moralnih stavkov v njih navzoč tudi drug gradnik, npr. vrednostni, emotivni ali predpisni gradnik, ki ga ne moremo določiti z resničnostnimi pogoji, ampak izraža naše vrednote, naravnanosti ali pa predpise, katere sprejemamo oz. nanje pristajamo. Npr. stavek »Avto je parkiran pred hišo«, je moč razumeti preko tega, kaj mora biti res, da je ta izjava resnična, medtem ko moralne sodbe npr. »Mučenje zapornikov ni moralno sprejemljivo« ne moremo opredeliti zgolj preko opisnega pomena te izjave, ampak tudi naših naravnanosti do mučenja.



Nadalje, Hare deli deskriptivizem na *moralni naturalizem* in *intuicionizem*, nedeskriptivizem pa na *emotivizem* in *preskriptivizem*. Razlika med naturalizmom in intuicionizmom je v tem, da prvi trdi, da lahko resničnostne pogoje moralnih sodb podamo s pomočjo opredelitev, ki vključujejo zgolj ne-moralne oz. zunajmoralne resnice ali lastnosti, intuicionizem pa vse to zanika. Za Harea se nazadnje obe vrsti deskripitivističnih moralnih teorij iztečeta v moralni *relativizem*, ki se mu sicer deskriptivizem v izhodišču želi izogniti. Naturalizem se izteče v relativizem, ker če pomen moralnih izrazov razumemo preko resničnostnih pogojev, potem bo nenazadnje določalo resničnost ali neresničnost moralnih sodb dejstvo, kakšne resničnostne pogoje sprejema posamezna skupnost. Podobno je pri intuicionizmu, ki kot osnovo moralnosti razume splošno sprejeta temeljna moralna prepričanja (moralne intuicije), ki jih delijo vsi izobraženi oz. moralni zreli ljudje. Toda dejstvo je, da so takšna prepričanja od skupnosti do skupnosti različna (Hare, 1993: 452-454).

Nedeskriptivizem predstavlja odgovor na deskriptivizem. Prvo izmed njegovih vej predstavlja emotivizem. Emotivisti so po zavrnitvi naturalizma in intuicionizma pričenjali z izhodišča, da so *moralne sodbe* izraz iracionalnih ali vsaj neracionalnih *naravnanosti naklonjenosti ali zavračanja*; odobravanja ali neodobravanja. Tako npr. Ayer: *»Če sedaj posplošim to, kar sem do sedaj dejal in rečem 'Kraja denarja je moralno nepravilna', potem sem izrekel stavek, ki nima*

dejstvenega pomena – tj. ne izraža nobene propozicije, ki bi bila lahko resnična ali neresnična. To je enako, kot če bi zapisal 'Kraja denarja!!', kjer oblika in poudarjenost klicajev kažeta (glede na izbrano konvencijo), da je ta posebna vrsta moralnega neodobravanja čustvo, ki je tako izraženo« (Ayer, 1952: 107). Zatorej o moralnih vprašanjih ne moremo razumno razpravljati; sogovorca, ki se ne strinja z našo moralno sodbo, ne moremo prepričati ali mu dokazati naš prav, kar pa pravzaprav pomeni, da se *»naposled zatečemo h golemu zmerjanju«* (Ayer, 2004: 20). Čeprav je bil emotivizem eden izmed pomembnejših vplivov na Harea, pa ga v takšni obliki ni mogel sprejeti predvsem zaradi njegovih neracionalističnih izhodišč. Začetek preskriptivizma kot nove oblike nedeskriptivizma je torej v tem, da kot zmotno zavrne predpostavko, da je mogoče racionalno oz. razumno razpravljati zgolj o dejstvenih oz. opisnih vprašanjih. Loti se iskanja logike oz. pravil mišljenja, ki zaobsega tako opisne kot neopisne izjave oz. jezikovna dejanja.

Glede vprašanja moralne motivacije, je bil Hare *internalist*. Internalizem je stališče, ki zagovarja tesno povezanost moralnosti oz. moralnih sodb in motivacije za delovanje. Moralni premisleki (sodbe, izjave, ...) nas motivirajo oz. priskrbijo razloge za delovanje, in to neodvisno od dejstev o vsebini morale in dejstev o samem delovalcu (njegovih siceršnjih želja, interesov, teženj). V tej povezavi je motivacijski internalizem blizu razmislekom, ki jih podpira moralna fenomenologija oz. fenomenologija moralnega delovanja ter naše vsakdanje izkustvo. Pri slednjem lahko izhajamo iz tega, da nas naše moralne sodbe vedno ali pa vsaj pogosto motivirajo, da ravnamo v skladu z njimi ali pa (v primeru konflikta motivacij) vsaj, da bi jim sledili ob odsotnosti drugih premislekov; da nad nami izkazujejo nekakšno racionalno avtoriteto v smislu, da nam dajejo dobre razloge, da z svojim delovanjem tem sodbam pogosto sledimo, tudi če to delovanje ni v našem interesu oz. tudi če pred tem v sebi ne prepoznamo neke določene želje, da bi tako ravnali. Internalizmu nasprotno stališče je eksternalizem (glede moralne motivacije in razlogov za delovanje), ki trdi, da sta motivacijska moč in racionalnost moralnosti odvisni od dejavnikov, ki so tej moralnosti oz. moralnim sodbam zunanji. Povezava med sodbami in motivacijo je zgolj naključna in je nadalje lahko odvisna od same vsebine moralnih sodb ali dejstev o delovalcu (njegovih želja, interesov, ...). Če zastavimo vprašanje »Čemu biti moralen?« in ga na primer razumemo kot to, ki vprašuje »Čemu naj bi storil to, za kar sodim, da je moralno pravilno ali dobro?«, potem je za motivacijskega internalista to vprašanje absurdno, nesmiselno, ali pa (za internalista, ki je kognitivist, na primer Kant) vsaj neracionalno. Kdor ga bi zastavil, očitno ne bi poznal uporabljenih pojmov oz. ne bi vedel, o čem govori. Za motivacijskega eksternalista pa to isto vprašanje ni nesmiselno, temveč kvečjemu nemoralno; kdor ga zastavlja je kvečjemu moralno nezrel, nerazvit, neobčutljiv ali sprijen. Hare je na strani internalizma, saj trdi, da sprejeti neko moralno sodbo (kot predpis) že pomeni tudi biti motiviran za delovanje v skladu z njo. Internalizem ima kar nekaj problematičnih točk, izmed katerih je največja zanka za Harea bila figura akratika, tj. osebe z akrazijo oz. šibko voljo. Poleg takšne osebe sta ga begali še dve, in sicer fanatik ter amoralist. Zakaj in v kakšnem smislu vse te tri figure predstavljajo težave za Harea, bomo videli ob koncu naslednjega razdelka.

II. Logika moralnega jezika: predpisnost in univerzabilnost moralnih sodb

Poglaviten odmik od emotivizma se pri Hareu zgodi s tem, ko moralnih sodb ne razume več kot izraza naših čustev, teženj ali drugih naravnanosti, ampak kot vrsto *predpisov*. Moralne sodbe in tudi vrednostne sodbe nasploh so vrsta predpisov, podobno kot so to tudi preprosti *ukazi*. In že pri ukazih lahko vidimo, da so podvrženi določeni logiki oz. zahtevam konsistentnosti, kljub temu, da nimajo opisne vloge. Če mi oseba A npr. najprej reče: »Odpri vhodna vrata!« in takoj zatem oseba B: »Ne odpiraj nobenih vrat!«, je s tem očitno nekaj narobe. Ne morem se resno zavezati obema zahtevama oz. ukazoma. Emotivizem bi obe izjavi razumel kot takšni, ki izražata nekaj o njunima govorcema, ter bi jih analiziral kot »A želi, da odprem vrata« in »B želi, da ne odprem vrat«, pri čemer med obema ni prav nobenega nasprotja. Toda:

> Na nek način ni nič narobe, če zastavimo stvar tako, vendar je lahko to v filozofskem smislu zelo zavajajoče. Za posledico ima namreč, da če jaz rečem »Zapri vrata« in ti (isti osebi) rečeš »Ne zapri vrat«, potem si drug drugemu ne nasprotujeva. Zagovornik takšne teorije bi lahko odvrnil, da čeravno ne gre za nasprotovanje oz. protislovje, nahajamo nestrinjanje v željah, ter da je to dovolj, da pojasnimo občutek, da sta stavka vseeno na nek način nezdružljiva med sabo… Toda še vedno ostaja težava, da se za stavek »Zapri vrata« zdi, da govori o vratih, in ne o govorčevem stanju duha. Podobno so navodila za pripravo palačink (»Vzemite 4 jajca…«) o jajcih, ne pa introspektivna analiza stanja duha oz. duše sestre Vendeline (Hare, 1952: 1.3).

Sprejeti oz. pristajati na nek predpis pomeni za Harea biti iskreno *zavezan k delovanju* (ob pogoju, da lahko oz. sem zmožen tako delovati), vsebovanem v tem predpisu, podobno kako iskreno sprejemanje neke opisne izjave pomeni biti prepričan, da je resnična. Naravnanost ob iskrenem sprejemanju moralne sodbe je torej ta, da bomo stremeli k temu, da v skladu z njo tudi ravnamo. Tukaj je tudi jedro Hareovega internalizma, ki je bil opisan že zgoraj.

Moralno mišljenje sledi logiki, ki jo srečamo tudi pri drugih predpisih in ukazih, kjer lahko iz predpisne in opisne premise izpeljemo praktičen sklep. Npr.

Odnesi vsa ta pisma na pošto.
To je pismo
∴ Odnesi to pismo na pošto.

Glede prehoda od opisnih do predpisnih izjav poda Hare dve omejitvi. Prvič, nobenega opisnega sklepa ne moremo izpeljati iz danih premis, če ga ni mogoče izpeljati tudi zgolj iz opisnih premis. Torej iz »Zapri vsa okna« in »Tole je okno« ne moremo izpeljati »Zaprl bom tole okno« Drugič, nobenega predpis-

nega sklepa ni mogoče izpeljati iz množice premis, ki ne bi vsebovala vsaj ene predpisne premise. Gre za znano načelo, da iz dejstev ne moremo izpeljati najstva, na katerega je opozarjal že Hume.

Moralne sodbe so torej predpisne na zgoraj opisan način. Od preprostih ukazov pa se razlikujejo v tem, da so *univerzabilne*. Slednje pomeni, da se predpis, vsebovan v njih, z njihovim sprejetjem nanaša na vse podobne delovalce v enakih okoliščinah. Ne morem npr. iskreno reči »Moraš mi vrniti denar, ki si si ga od mene izposodil«, ne da bi hkrati sprejemal tudi tega, da bi moral vsakdo dovolj podoben tebi in v podobni situaciji vrniti denar (vključno z menoj samim). Sprejetje takšne partikularne moralne sodbe tako pomeni tudi sprejetje njej podobnih sodb, ki se ne nanašajo samo na delovalca, ki ga zadeva ta sodba, ampak tudi vse druge, njemu podobne delovalce v podobnih okoliščinah ter zaobsegajo podobna dejanja in predmete kot prvotna sodba, in to ne glede na prostor in čas teh primerov. Če npr. tako ocenjujemo določeno dejanje neke osebe kot moralno napačno, s tem sprejemamo enako sodbo glede vseh pomembno podobnih dejanj, ne glede na to, kdo bi jih izvršil, kje in kdaj.

Moralna teorija tako določa pomene in funkcije moralnih besed in s tem podaja nekakšna pravila moralne igre. Tako pripravlja teren, na katerem se odvija moralno mišljenje, sama pa je nevtralna med posameznimi moralnimi stališči. Videli smo, kako sta glavni predvsem dve pravili, ki sta vezani na predpisnost in univerzabilnost naših moralnih sodb. Ko se odločamo za to, kaj bi v določeni situaciji morali storiti, je to, kar iščemo neko dejanje, h kateremu se bomo lahko zavezali (*predpisnost*) in ki bo hkrati takšno, da bi ga lahko kot predpis sprejeli tudi za druge, ki bi bili v podobni situaciji (*univerzabilnost*) (Hare, 1998: 51). Moralno misel tako v pomembni meri tvori to, da na podlagi dejanj, ki se nam ponujajo kot moralno obvezna, prepovedana ali dopustna oblikujemo moralna načela ter njih sprejemljivost nato preverjamo tako, da sledimo njihovim posledicam in premislimo, ali bi jih lahko sprejeli. Ta logika moralnega jezika je tako eden izmed bistvenih delov moralne razprave. Za slednjo je pomembno še poznavanje dejstev o primerih oz. situacijah in pa obstoj določenih teženj, interesov, želja in stremljenj, ki jih v naši sodbi upoštevamo.

Premislimo to logiko moralnega jezika in moralne misli tudi ob primeru, ki ga uporabi Hare sam. A je dolžan denar B-ju, B pa je dolžan C-ju, ob tem pa zakon dopušča, da lahko upnik izterja dolgovani denar od dolžnika s tem, da ga proti njegovi volji zapre. Sedaj se B vpraša: »Ali naj zaprem A-ja in na ta način uveljavim svojo zahtevo?« Predpostavimo, da B to želi storiti in se nagiba k temu. Če v igri ne bi bilo zahteve univerzabilnosti, bi torej brez težav sprejel partikularen predpis »A bom dal zapreti« Toda ko želi ta svoj predpis razumeti kot moralno sodbo in se vpraša »Ali naj bi zaprl A-ja, ker mi ni povrnil tega, kar mi dolguje?«, se izkaže, da bi to pomenilo, da sprejema obče načelo »Vsakdo, ki bi bil v moji situaciji, bi moral dati zapreti svojega dolžnika, če mi ni vrnil posojenega denarja«. Toda v naslednjem koraku uvidi tudi to, da je C nasproti njemu v enaki situaciji, kot je sam nasproti A-ju. Vse to pomeni, da mora sedaj sprejeti tudi sodbo, da bi moral tudi C dati zapreti njega. Sprejemanje slednjega pomeni, da torej sprejema predpis, da naj ga C da zapreti, česar pa si – predpostavljamo – ne želi. Zaradi tega ne more sprejeti niti prvotne sodbe, da bi

moral sam dati zapreti A-ja. Za takšen celoten argument sta značilnosti predpisnosti in univerzabilnosti sodb ključni, kajti ravno ti omogočata prehode med sprejemanjem posameznih predpisov. Videli smo tudi, da je v primeru predpostavljeno, da si B ne želi biti poslan v zapor. Toda to je zgolj predpostavka.

Kaj pa če je B povsem mirno pripravljen biti poslan v zapor? Vzpostavi se problem t.i. *fanatika*. Fanatik je figura, ki je pripravljen zagovarjati nek ideal, tudi če le-ta nasprotuje njegovim interesom in ga torej tudi sklicevanje na interese drugih ne bo prepričalo. Podobno sta za Hareovo teorijo težavni figuri akratika in amoralista. *Akratik* je oseba, ki ji umanjka vsakršna moralna motivacija; sprejema torej moralne sodbe kot univerzalne predpise, a jim v dejanjih ni pripravljen slediti. *Amoralist* je po drugi strani tisti, ki ni pripravljen sprejeti nobenega univerzalnega predpisa oz. iskreno ne sprejema nobene moralne sodbe; uporablja lahko sicer moralni besednjak, vendar mu pomeni nekaj drugega kot ostalim. Npr. ko reče: »Moralno napačno je trpinčiti ljudi za zabavo« s tem npr. želi reči, da drugi ocenjujejo, da je to moralno napačno. Nadaljni razvoj Hareove moralne teorije je mogoče razumeti prav v okviru tega, da poskuša najti kar najbolj prepričljive odgovore za te tri težave.

Zgoraj opisana struktura predpisnosti in univerzabilnosti moralnih sodb je na področju normativne etike Harea privedla do *utilitarizma*, natančneje *utilitarizma zadovoljitve preferenc*. Najprej, ker gre za vrsto posledicizma, ta teorija razume moralno pravilnost in napačnost dejanj kot odvisno od posledic teh dejanj. Moralno pravilno je dejanje, ki v dani situaciji privede do najboljših možnih posledic. Utilitarizem zadovoljitve preference vrednost posledic dejanja tehta glede na zadovoljitev preferenc udeleženih v situaciji, tj. *moralno pravilno* je tisto dejanje v situaciji, ki v *kar največji meri zadovolji želje, stremljenja in teženja udeleženih oseb*.

III. Utilitarizem dveh ravni

Videli smo, kako se Hareova moralna teorija izteče v utilitarizem kot vrsto posledicistične etike, glede na katero so moralno relevantne posledice naših dejanj vezane na izpolnitev preferenc. V zgodovini moralne misli sta se oblikovali dve različni vrsti utilitarizma, in sicer utilitarizem dejanj in utilitarizem pravil. *Utilitarizem dejanj* trdi, da je moralno pravilno oz. naša dolžnost tisto dejanje izmed vseh alternativ v dani situaciji, ki bo privedlo do najboljših posledic oz. največ koristi. Ko se v situaciji odločamo o tem, kaj moramo storiti, moramo najprej premisliti o vseh dejanjih, ki so nam na voljo, ter s primerjanjem izbrati tistega, ki bo privedel do največ koristi. *Utilitarizem pravil* po drugi strani trdi, da moramo sprejeti mnoštvo moralni načel in pravil, ki bodo v celoti gledano privedla do največ koristi, če jim bomo sledili. Relevantna primerjava v tem slednjem primeru je torej med različnimi množicami moralnih načel in pravil. Vsaki od teh dveh vrst utilitarizma lahko postavimo resne ugovore.

Osrednji ugovor utilitarizmu pravil je v tem, da pogosto privede do neintuitivnih posledic in nam nalaga, da sledimo pravilu, ne glede na to, da takšno ravnanje očitno ne bo privedlo do najboljših posledic. Npr. če sprejmemo pravilo »Ne laži!«, potem mu moramo glede na osnovno različico utilitarizma pravil

slediti v vseh situacijah, ne glede na to, da lahko s precejšnjo verjetnostjo predvidevamo, da bo tu privedlo do manj vrednih posledic kot pa alternativno dejanje, s katerim bomo lagali in kršili to pravilo.

Zdi se, da je torej v prednosti utilitarizem dejanj, ki nam bo vedno zapovedoval prav to, tehtanje vrednosti posledic vsakega posamičnega dejanja. Toda po drugi strani je takšna zahteva precej nerealistična oz. neuresničljiva, kajti spregleduje dejstvo, da smo povprečni moralni delovalci kognitivno omejena bitja. Da imamo torej človeške omejitve. Težko si v vsaki situaciji zamislimo vsa alternativna dejanja, ki bi jih lahko storili. Še težje je oceniti posledice vseh teh alternativ, ki gredo v neskončnost. Hare tudi dodaja, da je naša vednost omejena, da smo pogosto pristranski, da se ne zmoremo v popolnosti postaviti v situacijo drugega ter da je veliko situacij, kjer nimamo na voljo dovolj časa časa, da bi opravili takšno kalkulacijo zadovoljitve preferenc.[2]

Hare je uvidel te težave in svojo moralno teorijo zasnoval tako, da se jim lahko izogne. To stori na način, da uvede dve ravni moralnega mišljenja. To razlikovanje (ali vsaj zametke takšnega razlikovanja med ravnmi moralnega mišljenja) sicer pripisuje že Platonu in Aristotelu. Platon v *Menonu* (98b) tako razlikuje med resničnim mnenjem in razumevanjem ter nadalje (preko pogovora med Sokratom in Menonom) ugotavlja, da sta obe lahko vodili oz. temelj za naše delovanje, pri čemer so resnična mnenja manj vredna od razumevanja, vendar – kolikor ostajajo resnična – vseeno dobra podlaga našega delovanja. Aristotel pa podobno temo odpre v 6. knjigi *Nikomahove etike*, kjer govori o razliki med mnenjem in pametnostjo ter modrostjo. Tudi klasični utilitaristi, npr. Mill in Sidgwick so se zapletali v težave glede (ne)sprejemljivosti utilitarizma dejanj in pravil. Pomen Harea je v tem, da je to razliko med dvema ravnema moralnega mišljenja najbolj jasno pojasnil in razdelal. Za njim je različico utilitarizma dveh ravni oz. posrednega utilitarizma prevzela večina sodobnih utilitaristov (Peter Singer, David Brink, Brad Hooker). V naslednjem razdelku bomo podrobneje predstavili, kako je to težavo na podoben način poskušal razrešiti G. E. Moore.

III.a. G. E. Moore in vprašanje dolžnosti

Moorova moralna teorija je posledicistična, saj trdi, da »*izjaviti, da je določeno ravnanje v danem trenutku brezpogojno pravilno ali obvezujoče, očitno pomeni*

[2] Podobno, a bolj sistematično, je tudi Mackie (Mackie, 1977: 155–6) podal šest razlogov, zakaj utilitarizem dejanja ne more biti uspešen model mehanizma neposrednega moralnega presojanja: (1) Pomanjkanje časa in zmožnosti močno omejuje preračunavanje vseh posledic možnih alternativ. (2) Tudi če bi imeli čas in zmožnosti, bomo le redko imeli na voljo vse pomembne informacije. (3) Delovalčeva sodba o posamičnih zadevah bo zelo verjetno izkrivljena zaradi njegovih posebnih interesov in udeleženosti v odnosih. (4) Tudi če bi razumsko dospeli do določitve pravilnega dejanja, bo šibkost naše volje zelo verjetno ogrožala njeno udejanjenje. (5) Odločitve, ki so sicer same po sebi pravilne, in dejanja na njihovi podlagi, bodo lahko izrabljene kot zgled, tako da bodo napotovale druge, da bodo na ta način upravičevali dejanja, ki so tem podobna, a v temelju napačna. (6) Glede na človeško naravo, ki je takšna, kakršna je, praktična in delujoča morala ne sme biti prezahtevna; nekoristno je postaviti standarde tako

izjaviti, da bo v primeru takega ravnanja v svetu več dobrega in manj zla, kot pa bi ga bilo v primeru, da bi namesto tega storili kaj drugega« (Moore, 2000: 90). *Moralna pravilnost* oz. *dolžnost* je opredeljena s pojmom dobrega. Ta pojem oz. pojem neodnosne vrednosti ima tako znotraj normativne strukture Moorove moralne teorije osrednje mesto kot temeljni normativni pojem, kar pomeni, da nam to, da je neka stvar neodnosno dobra, priskrbi razlog za delovanje. Prav tako je glede na pojem dobrega opredeljen pojem *vrline* – vrlino Moore definira kot *»trajno dispozicijo za izvrševanje določenih dejanj, ki praviloma proizvedejo najboljše možne izide«* (Moore, 2000: 243). Vrline so za Moora nekaj, kar nas postavlja v odnos do dobrega, nam omogoča njegovo doseganje; opredeljene so funkcionalno in same po sebi nimajo neodnosne vrednosti.

Vse omenjeno vodi Moora do stališča, ki bi ga lahko označili kot *neosebni posledicizem* oz. osebno-nevtralni (angl. *agent-neutral*) posledicizem, tj. stališča, da je naša dolžnost vedno ravnanje, ki bo privedlo do največ dobrega, neodvisno in nepristransko tehtano glede na vsakega posameznika enako. Hkrati je Moorova teorija vrednosti *nehedonistična*, saj za Moora moralna vrednost ni zgolj ali primarno v sreči, ugodju ali zadovoljevanju želja, temveč v določenih stanjih duha oz. zavesti, katerih vrednost je neodvisna od posameznikove naravnanosti do njih. Moore tako trdi, da so *»daleč najbolj dragocene stvari, ki jih poznamo ali ki si jih lahko zamislimo, določena stanja zavesti, ki bi jih lahko v grobem opisali kot uživanje v medčloveških stikih in veselje nad lepimi predmeti«* (Moore, 2000: 260).

Pri sodbah o našem ravnanju (»Kaj bi morali storiti?«) moramo glede na zgoraj podana izhodišča tako po Mooru upoštevati dvoje stvari: (i) katera stanja stvari imajo *neodnosno vrednost* oz. katere stvari so dobre po sebi in (ii) *vzročne sodbe* o tem, katere stvari so v vzročnem razmerju s prvimi in na kakšen način, saj *»vprašati, katere vrste dejanje bi morali storiti ali katera vrsta ravnanja je pravilna, je isto kot vprašati, katere vrste učinkov bo tako dejanje ali ravnanje proizvedlo.* […] *Vsa ta vprašanja seveda vključujejo* tudi *pravo etično sodbo – sodbo, da so določeni učinki po sebi boljši od drugih. Vendar pa* dejansko *zatrjujejo, da so te boljše stvari učinki – da so vzročno povezane z zadevnimi dejanji«* (Moore, 2000: 217).

Na podlagi te analize odnosa med dobrim in dolžnostjo, Moore trdi, da so intuicionisti v zmoti, ko zagovarjajo samorazvidnost moralnih zakonov. Nobena trditev o tem, da je določena vrsta dejanja brezpogojno naša *dolžnost*, ne more biti *samorazvidna*. Samorazvidne so zgolj trditve o tem, kar je *dobro samo po sebi* (Moore, 2000: 219). Moralne sodbe o dolžnostih *niso neizpeljane* intuicije, saj jih lahko dokažemo, potrdimo ali ovržemo prav z raziskovanjem tega, kaj je dobro samo po sebi in hkrati raziskovanja vzrokov in učinkov dejanj, ki jih sodbe zadevajo.[3] Predvsem slednje je zaradi pogosto zapletenih okoliščin in omejitev pri poznavanju celotne verige vzrokov in posledic izjemno *težko dosegljivo znanje*; prav tako pa nas omejujejo razmisleki o vseh možnih alter-

[3] Cf. *»Etični zakon nima značaja znanstvenega zakona, prej je podoben znanstveni* napovedi*: slednja pa je ob še tako veliki verjetnosti vedno zgolj verjetna«* (Moore, 2000: 226).

nativah. Zato je naša (praktična) moralna vednost nujno pomanjkljiva in iz tega Moore zaključi, *»da nimamo nikoli nobenega razloga za domnevo, da je kako dejanje naša dolžnost: nikoli namreč ne moremo biti gotovi, da bo kako dejanje proizvedlo največjo možno vrednost«* (Moore, 2000: 220). Zaradi tega se morata etika in moralni delovalec zadovoljiti s skromnejšim ciljem, tj. določitvijo tega, katera od verjetno razpoložljivih alternativ, ki so na voljo komur koli, bo po vsej verjetnosti proizvedla največ dobrega. Odgovori na vprašanja te slednje vrste pa se skrivajo v *konvencionalni morali* in tako Moore sklene, da lahko posamezniku *»torej z gotovostjo priporočimo, da naj se* vedno *ravna po pravilih, ki so splošno koristna in se po njih ravna večina ljudi«* (Moore, 2000: 235).

Moore tako pristane v t. i. *posrednem posledicizmu* oz. *posledicizmu dveh ravni*. Ko se moralni delovalci odločajo o tem, kaj storiti oz. kako ravnati, ne smejo tehtati vrednosti posamičnih alternativnih dejanj, ki so jim v situaciji na voljo, oz. tehtati vrednosti njihovih posledic, temveč morajo slediti splošno sprejetim moralnim načelom, ki jih nahajamo v konvencionalni morali.[4] To »na dolgi rok« zagotavlja več dobrega v svetu. Podobno sta pred Moorom zagovarjala že Mill in Sidgwick, toda Moore je bil v tej točki izredno tog; njegova *»etika delovanja se tako nazadnje izteče v najhujši moralni konzervativizem«* (Klampfer, 2000: 340). Tem običajnim moralnim načelom naj bi sledili tudi v primeru, ko gre za kompetentne moralne presojevalce in ko imamo sicer dobra dokazila o posledicah posameznih alternativnih dejanj oz. o tem, da bo kršenje teh načel v določenem primeru privedlo do več dobrega (Moore, 2000: 234–235).

Ta Moorov konzervativizem postavlja v čudno luč sam posledicizem oz. utilitarizem dveh ravni, kar v grobem pomeni, da poskušamo združiti utilitarizem pravil in utilitarizem dejanj, pri čemer naše moralno presojanje in odločanje sledi enkrat enemu, drugič drugemu glede na okoliščine. Največja težava je seveda razumevanje tega prehoda iz ene na drugo raven. Vsaj na prvi pogled se zdi, da omenjeni Moorov skrajni konzervativizem onemogoča vsak tak prehod. (To je tako, vsaj kar zadeva naše praktične odločitve, kjer je udeleženo vzročno znanje.) Na slednje kaže tudi Moorovo zavračanje tega, da bi lahko kdajkoli upravičeno domnevali, da imamo opravka z izjemo glede na izbrano splošno sprejeto in koristno pravilo, in slednjih zatorej ne smemo nikoli kršiti[5] (Moore, 2000: 233). Po drugi strani dva premisleka kažeta v nasprotno smer.

Prvič, Moore v *Principia Ethica* problematike moralnih dilem ne omenja posebej. Čeprav zatrjuje, da znotraj običajne moralnosti nahajamo splošno koristna

[4] Moralna načela in pravila konvencionalne morale, na katere meri Moore, delijo naslednji dve značilnosti: (i) njihovo splošno spoštovanje bi bilo v vsaki družbeni skupnosti dobro kot sredstvo, saj so okoliščine, od katerih je odvisna njihova koristnost, univerzalne; in (ii) ker so sredstvo za ohranitev civiliziranе skupnosti, ki pa je sama nujni pogoj vsake velike dobrine, njihov zagovor ni odvisen od pravilnih stališč v zvezi z vprašanjem o dobrem po sebi (Moore, 2000: 228–9).

[5] Moore tukaj svoje stališče dodatno podpre z argumentom, da bi takšna (sicer morebiti upravičena) kršitev dajala slab zgled drugim (cf. Moore, 2000: 233–4). Slov. prevod *Principie Ethica* je na tem mestu zmoten in zavajajoč, kajti pravi, da bi takšna upravičena kršitev *»prav gotovo spodbudila druge kršitve nekoristnih pravil«* (Moore, 2000: 234); pravilneje pa bi bilo to mesto prevesti kot *»spodbudila druge, nekoristne kršitve pravila.«*

načela, kot so prepoved ubijanja (oz. škodljivost umora) in načela povezana s spoštovanjem lastnine, marljivostjo, zmernostjo in izpolnjevanjem obljub (Moore, 2000: 227–228), ne uvidi, da je tako zasnovana običajna moralnost lahko tudi polna moralnih *konfliktov* med temi načeli in iz njih izhajajočih moralnih dilem. Tako ni povsem jasno, kako se moramo odločati v takšnih primerih. Vprašanje je, ali bi Moore tukaj odgovoril, da se mora naš razmislek v takšnih primerih dvigniti na višjo raven in da se moramo odločiti glede na tehtanje vrednosti učinkov alternativnih dejanj? Če obstajata zgolj ti dve ravni in če na spodnji ravni konflikt ni razrešljiv, potem moramo preiti na zgornjo raven. Konflikt na spodnji ravni tudi ne bo neredek, sploh glede na načela običajne moralnosti, ki jih izpostavi Moore. Drugič, Moore govori tudi o t. i. *nedoločenih dejanjih*, tj. o dejanjih, ki jih ne uravnava nobeno splošno koristno načelo ali pravilo, a so vseeno moralno pomembna in katerih splošne koristnosti se ne da dokazati. Najprej ni povsem jasno, kakšen je sploh *obseg* takšnih dejanj. Po eni strani se *prima facie* ponuja odgovor, da je delež »nedoločenih dejanj« majhen, kajti običajna morala pokriva precejšen del naših moralnih življenj. Moore po drugi strani pravi:

> V skladu s prejšnjimi sklepi bo taka razprava zadevala skoraj vsa dejanja z izjemo tistih, ki *jih* v danem stanju družbe opravlja večina ljudi. Poudarili smo namreč, da je dokazovanje splošne koristnosti tako težavno, da je lahko dokaz dokončen le v majhnem številu primerov. Vsekakor ni mogoč v zvezi z vsemi dejanji, ki *jih* opravlja večina ljudi; pa čeprav v tem primeru, če so kazni le dovolj stroge, sama po sebi [verjetno so tukaj mišljene kazni in ne dejanja, torej »same po sebi« – op. VS] zadoščajo za dokaz, da je za posameznika uklonitev običaju praviloma koristna (Moore, 2000: 236).

Takšna dejanja lahko presojamo le *»tako, da pokažemo na njihovo neposredno težnjo k proizvajanju po sebi dobrih ali k preprečevanju po sebi slabih stvari«* (Moore, 2000: 236). In potem sledi še glede na zgornja izvajanja presenetljiv sklep, *»da se mora posameznik v primerih dvoma v svojih odločitvah ravnati raje po neposrednih premislekih o intrinzični vrednosti ali nizkotnosti učinkov, ki jih njegovo dejanje lahko proizvede, namesto po pravilih, katerih dobre učinke je v svojem posameznem primeru nezmožen videt.«* (Moore, 2000: 237). Morda se tu skriva tudi odgovor na vprašanje o moralnih dilemah, saj so le-te bistveno zaznamovane s takšnim dvomom v odločitev.

Kot se je pokazalo, so Moorovi praktični vidiki njegove moralne teorije v napetosti sami s sabo. Čeprav je njegova misel, kar se tiče temeljnega merila moralne pravilnosti dejanj, jasna, je toliko več težav, ko se obrnemo k moralnim načelom kot vodilom pri odločanju. Moore preskakuje z ravni posledicizma dejanja k posledicizmu pravil in konvencionalni morali, pri čemer pa ni jasnih razmejitev med enim in drugim, predvsem glede obsega dejanj, ki jih nahajamo na vsaki izmed ravni.

III.b. Hare o dveh ravneh moralnega mišljenja

Utilitaristi se zaradi zgoraj omenjenih ugovorov tako pogosto znajdejo v zagati, ko morajo po eni strani pristajati na utilitarizem dejanj, po drugi strani pa na utilitarizem pravil. Pri tem različni misleci to zagato razrešijo različno. Ena izmed rešitev je – kot smo videli zgoraj – da nekako znotraj moralne teorije sprejmemo obe vrsti utilitarizma. Sidgwick je to storil na način, da je ravnanje v skladu z utilitarizmom načel rezerviral za poseben razred ljudi, ki so takšnega zahtevnega moralnega premisleka sposobni v vsaki situaciji. Vsi preostali pa naj sledijo preprostim moralnim načelom, ki jih vsebuje zdravorazumska moralnost (Sidgwick, 1907). Po drugi strani Moore ni razlikoval med razredi ljudi, ampak je poskušal to razliko najti na strani dejanj. Preprostim moralnim načelom sledimo pri tisti vrsti dejanj, ki je dovolj pogosta in jo izvršuje večina ljudi, kar nam posledično omogoča oceno koristnosti takšnega sledenja pravilom. Pri nedoločenih dejanjih pa moramo njihove posledice tehtani neposredno, tj. za vsak posamezen primer.

Hare ubere tretjo pot. Razlika med obema vrstama utilitarizma se ne odraža niti v dveh različnih skupinah ljudi niti v različnih vrstah dejanj. Razliko moramo razumeti kot razliko v *dveh različnih ravneh moralnega mišljenja*, pri čemer se lahko ena in ista oseba zdaj nahaja na eni, nato pa na drugi ravni. Prav tako lahko na obeh ravneh govorimo o enakih vrstah dejanj oz. o vseh dejanjih. To je pomemben Hareov dosežek, ki so ga večinoma prevzeli vsi utilitaristi, ki zagovarjajo to dvonivojsko različico utilitarizma, npr. Hooker (Hooker, 2000).

Moralno mišljenje je torej razdeljeno na dve ravni.[6] Prva je *intuitivna*, zdravorazumska ali običajna raven, ki zadeva splošna pravila, ki jim sledimo, ter hkrati zajema naša moralna občutja, vest, naravnanosti, čustva ipd. Druga raven je *kritična* oz. reflektivna raven, na kateri lahko kritično premislimo načela ali odločitve s prve ravni. Nanjo nas lahko »prestavi« npr. soočenje z moralno dilemo (situacijo, kjer nam eno načelo narekuje določeno ravnanje, hkrati pa nam drugo načelo narekuje drugačno ravnanje in se nam zdi, da imamo dve dolžnosti, ki jih ne moremo izpolniti hkrati) ali pa iskanje upravičenja za načela, ki jim sicer sledimo.

Na *intuitivni ravni* nahajamo preprosta moralna načela, ki jih lahko z moralno vzgojo ponotranjimo oz. se jih naučimo, hkrati pa so takšna, da so uporabna pri odločanju v različnih situacijah, kjer je pomembno, da se odločimo hitro ali kjer je nevarnost, da bomo neupravičeno dali prednost našim lastnim interesom. V tem smislu lahko ta načela razumemo kot zelo uporabna vodila pri odločanju za ravnanje. Prav tako na tej ravni nahajamo npr. ponotranjene odzive, motivacijo, dispozicije, nagnjenja, teženja in naravnanosti, ki spremljajo določene vrste dejanj, npr. da občutimo obžalovanje, ko moramo zaradi druge, bolj pomembne dolžnosti prelomiti našo dano obljubo, čeprav smo – gledano v celoti – ravnali prav. Ta raven je tako osrednja za moralno vzgojo in osrednja tudi za to, čemur bi lahko rekli moralno zrela oseba.

[6] Poleg omenjenih dveh Hare govori tudi o tretji ravni, ki pa je metaetična raven in se ukvarja predvsem z logiko moralnega jezika.

Po drugi strani je *kritična raven* moralnega mišljenja ta, ki sledi osnovnemu načelu zasledovanja najboljših posledic ravnanja (v smislu zadovoljitve preferenc), ki ga postavi utilitarizem dejanj. Kadar nam za neko situacijo intuitivna raven moralnega mišljenja ne ponuja odgovora ali pa nam ponuja več nasprotujočih si napotil za delovanje, lahko takšen konflikt razrešimo na kritični ravni.[7] Prav tako lahko na tej ravni kritično ovrednotimo načela (npr. *prima facie* dolžnosti pravičnosti, neškodovanja, iskrenosti, zvestobe idr.), ki jih sprejemamo kot del intuitivne ravni. Na podlagi premisleka jih lahko preoblikujemo ali zavrnemo. Sledimo tudi koristnosti, ki je vezana na njih sprejemanje, torej temu, ali bi gledano v celoti sprejemanje izbrane množice moralnih načel privedlo do boljših posledic kakor pa katere od drugih, alternativnih množic. Kritična raven je torej tista, ki sledi izbranemu moralnemu načelu in ga razume kot merilo (in ne kot uporabno vodilo) moralnega statusa dejanj.

Razliko med ravnema Hare slikovito oriše tudi s pomočjo dveh figur. Prva figura je bitje, ki ima nadčloveške moči ter zmožnosti. Ima neomejeno znanje, sposoben je izjemnega kritičnega premisleka, nikoli ne zapade dajanju prednosti samemu sebi, ni egoističen ali kako drugače pristranski. Hare to bitje imenuje tudi *nadangel*. Nadangel je v situaciji, ki lahko povsem opusti intuitivno raven svojega moralnega mišljenja in v vsaki situaciji sledi zgolj utilitarizmu dejanj. Na drugi strani si zamisli *pripadnika raje* oz. navadnega delavca (angl. *prole*, po pripadnikih najnižjega razreda v Orwellovem romanu *1984*). Gre za osebo, ki praktično nima sposobnosti kritičnega mišljenja in vse njegovo moralno ravnanje zgolj sledi naučenim ali pridobljenim *prima facie* načelom intuitivne ravni moralnega mišljenja. Nikoli se ne dvigne na kritično raven.

Vse to pa moramo razumeti kot ilustracijo razlike med tema dvema ravnema, ne pa kot dejansko delitev ljudi v ta dva razreda. Vsi mi smo v izbranih ozirih podobni tako raji kot nadangelom, posamezne razlike med nami pa seveda obstajajo glede izgrajenosti našega moralnega značaja in sposobnosti za kritične moralne premisleka. Gotovo je, pravi Hare, da ne moremo ves čas delovati kot nadangeli. Po drugi strani pa je gotovo tudi to, da se ne moremo ves čas vesti kot raja, če naj naš moralni sistem sploh deluje, saj je kritična raven tista, ki sploh izbere *prima facie* načela intuitivne ravni. Slednja torej ne more podpirati same sebe, medtem ko prva lahko. Zato je za Harea kritična raven primarna raven moralnega mišljenja.

Zaključek

Videli smo, kako je Hare dokončno oblikoval najbolj sprejemljivo različico utilitarizma dveh ravni. Kot tak se utilitarizem lahko izogne tradicionalnim ugov-

[7] Sam pojav navidezno nerešljivih moralnih dilem gre po Hareu pripisati prav naši zavezanosti »naivni« zdravorazumski oz. intuitivni moralnosti, ki jo sestavljajo preprosta, uporabna in naučljiva moralna načela, ki pa seveda v izbranih primerih lahko vodijo do nasprotujočih si sodb. Ko se dvignemo na kritično raven moralne misli in uporabimo utilitaristično načelo kot moralno merilo moralne pravilnosti, potem ti konflikti in te dileme izginejo. Ker pa smo vseeno še tesno povezani z intuitivno moralnostjo, ki ji sicer običajno sledimo in smo je vajeni, še vedno občutimo čustveno nelagodje ob tem, ko moramo prekršiti eno izmed preprostih intuitivnih moralnih pravil. Toda razumsko vzeto tu ni nobenega konflikta in nobene dileme. Da ne bi bili podobni podgani v labirintu brez izhoda, se moramo odpovedati intuicijam in uporabiti razum (Hare, 1978). Za več o tem glej tudi Strahovnik (Strahovnik, 2005).

orom, ki jih lahko naperimo zoper utlitarizem dejanj in utilitarizem pravil. Vseeno pa tudi utilitarizem dveh ravni oz. posredni utilitarizem ni povsem odporen za nadaljnje ugovore. Pomembnega je podal Bernard Willams, ki trdi, da tako zamišljena moralna teorija spodkopava odnos med *teorijo* in *prakso* ter s tem posledično spodkopava tudi samo sebe. Tovrstni utilitarizem razume vrednost načel, pravil, zavezanosti, vrednot in teženj, ki jih ponotranjimo in jim sledimo na intuitivni ravni moralnega mišljenja kot zgolj instrumentalno vredne. Toda nesmiselno bi bilo npr. trditi, da je vrednost prijateljstva ali pravičnosti zgolj instrumentalna in da moramo to uvideti na kritični ravni, z oddaljenim in hladnim pogledom. Kajti vse zgoraj našteto bo na intiutuvni ravni dobro delovalo zgolj, če vsega ne bomo imeli le za instrumentalno vredno. Samo tako bodo ta ponotranjena načela, vrednotenja, stremljenja in naravnanosti opravljala svoje vlogo. Vse to pa moralno teorijo preveč oddalji od prakse.

Nekatere različice posrednega utilitarizma sploh ne dovoljujejo nobenega prostora za teorijo. Razumejo jo kot transcendentalno glede na življenje; nahajajočo se v prostoru onkraj prakse, ki naj bi jo upravičevala in usmerjala. V psihološki različici [tj. različici, ki jo zagovarja Hare – op. VS] to skušnjavo najdemo v posebnemu razumevanju časa za teorijo: gre za čas, ko delovalec nekako zapusti samega sebe in vidi vse, vključno s svojimi dipozicijami, z vidika vesolja, in se potem vrne nazaj v svojo prako oz. življenje. Toda vsak stvaren proces takšnega teoretiziranja bi moral biti del življenja in torej sam del prakse. Ne moremo namreč – razen z vsiljeno in iluzorno disociacijo – razločiti teoretika v sebi od drugega dela sebe, katerega dispozicije so pod vprašajem. Tako v posrednemu utilitarizmu ta disociacija zgolj pomaga prekriti to težavo, tj. konflikt med pogledom teoretika na te dispozicije in pogledom na svet, ki jih ima glede oz. na podlagi teh dispozicij (Williams, 1985: 110).

Hare je s svojim preskriptivizmom in utilitarizmom priskrbel pomembno izbiro na polju etiških teorij. Utilitarizem je vztrajno poskušal ubraniti pred različnimi ugovori, hkrati pa ga je prikazal v novi luči, tj. kot mnogo bližje kantovski etiki, kakor pa ga običajno razumemo. Izjemno pomemben je tudi njegov prispevek k t.i. praktični oz. uporabni etiki, kjer je odpiral prostor za to raven moralne filozofije.

R.M. Hare's moral philosophy: Universal prescriptivism and two-level utilitarianism

The article presents the basic characteristics of R.M. Hare's (1919-2002) moral theory. His theory is presented as occupying an important place within the field of other metaethical positions of 20th century. Hare's moral theory is based upon two central features of the logic of moral thought, i.e. prescriptivity and universalizability of moral judgments. Hare elaborates a two-level utilitarianism, which surpasses limitations of act-utilitarianism and rule-utilitarianism, offering replies to the majority of traditional objections against utilitarianism.

Keywords: ethics, utilitarianism, prescriptivism, two levels of moral thought, R.M. Hare, G.E. Moore.

Literatura

Ayer, A. J. (1946 [2004]). *Language, Truth and Logic*, London: Gollancz; slov. prevod pomembnejših odlomkov Ayer, A. J. »Etične sodbe so zgolj izraz čustev«, *Filozofija na maturi*; 11 (1-2), 2004, str. 18-20.

Hare, R. M. (1952). *The Language of Morals*. Oxford: Oxford University Press.

Hare, R. M. (1963). *Freedom and Reason*. Oxford: Oxford University Press.

Hare, R. M. (1967a). »Geach: Good and Evil«. V Foot. P. (ur.), *Theories of Ethics*, Oxford: Oxford University Press, str. 74-82.

Hare, R. M. (1967b). "The promising Game". V Foot, P. (ur.) *Theories of Ethics*, Oxford: Oxford University Press, str. 115-127.

Hare, R. M. (1971). *Practical Inferences*. London: Macmillan; University of California Press.

Hare, R. M. (1978). »*Moral Conflicts: The Tanner Lecture on Human Values*«. Utah State University, October 5, 1978, str. 169-193.

Hare, R. M. (1981). *Moral Thinking: Its Levels, Method, and Point*. Oxford: Oxford University Press.

Hare, R. M. (1989). *Essays in Ethical Theory*. Oxford: Oxford University Press.

Hare, R. M. (1993). »Universal Prescriptivism«. V Singer, P. (ur.), *A Companion to Ethics*, Blackwell, str. 451-463.

Hare, R. M. (1998). »A Moral Argument«. V Rachels, J. (ur.) *Ethical Theory 1,* Oxford: Oxford University Press, str. 51-57.

Hare, R. M. (2002). »A Philosophical Autobiography«. *Utilitas*, let. 14. št. 3, str. 269-305.

Hooker, B. W. (2000). *Ideal Code, Real World*. Oxford: Oxford University Press.

Klampfer, F. (2000). »George E. Moore in začetki teoretične etike v analitični filozofiji«. V Moore, G. E., *Principia Ethica*, Ljubljana: Študentska založba, str. 317–44.

Mackie, J. L. (1977). *Ethics: Inventing Right and Wrong*. New York: Penguin.

Moore, G. E. (1903/2000). *Principia Ethica,* Cambridge: Cambridge University Press; slov. prevod *Principia Ethica*, (prev. F. Klampfer) Ljubljana: Študentska založba, 2000.

Sidgwick, H. (1907/1981). *Methods of Ethics*, 7th ed. London: Macmillan and Company; ponatis Indianapolis: Hackett, 1981.

Sidgwick, H. (1907/1981). *Methods of Ethics*, 7th ed. London: Macmillan and Company; ponatis Indianapolis: Hackett, 1981.

Singer, P. (2002). "R.M. Hare's Achievements in Moral Philosophy". *Utilitas,* 14 (3), str. 309-317.

Stevenson, C.L. (1944). *Ethics and Language.* New Haven: Yale University Press.

Strahovnik, V. (2005) "Intuicionizem, moralni konflikt in pluralizem dolžnosti". *Analiza* 4/2005, str. 106–29.

Williams, B. A. O. (1985). *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.