ANαliZA

časopis za kritično misel

Letnik 2, številka 1, junij 1998 · ISSN 1408-2969

PREVOD IN KOMENTAR

Anton Mohorčič

Ali je inteligenca sploh možna?

Filozofska fakulteta Univerze v Ljubljani

Citiranje: Anton Mohorčič, »Ali je inteligenca sploh možna?«, *Analiza: časopis za kritično misel* 2, št. 1 (1998): 141–148.

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

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

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

Anton Mohorčič
Filozofska fakulteta Univerze v Ljubljani

ALI JE INTELIGENCA SPLOH MOŽNA?

Dretske v prispevku z naslovom *Ali je inteligenca lahko umetna?* sicer zastavi vprašanje možnosti umetne inteligence in na koncu nanj odgovori, vendar je predmet njegove raziskave prav inteligenca kot taka. Pravzaprav je cela razprava nekakšno izpeljevanje in pojasnjevanje pojma inteligence, tako kot ga pojmuje Dretske. Iz pojma inteligence samodejno sledi odgovor na vprašanje o možnosti nastanka umetne inteligence. V tem prispevku se bomo miselno sprehodili skozi Dretskejevo razpravo in tudi preskusili trdnost stebrov, na katerih gradi svojo izpeljavo.

Ključne besede: (umetna) inteligenca, misel, delovanje, F. Dretske.

Inteligenca je več kot misel

> Na začetku svojega prispevka Dretske omeni dva pogleda na inteligenco. O njej se lahko razmišlja kot o denarju, o nečem, kar lahko skoraj vsakdo ima, ampak eni več kot drugi. Na drugi strani pa je inteligenca lahko kot bogastvo - nekaj, kar imajo le tisti, ki posedujejo več kot povprečno vsoto denarja. Na prvi način naj bi na inteligenco gledali filozofi, na drugi pa računalniški strokovnjaki. Dretskeja kot filozofa zanima, v prenesenem pomenu, kdo ima denar in ne kdo ga ima veliko.
>
> Če neumni pezdirk [*Lebistes retikulatus*], ko je lačen, plava na drugi konec bazena zato, ker misli, da je tam hrana, lahko njegovo obnašanje ocenimo kot inteligentno. Inteligentno pa je zato, ker pezdirk počne nekaj, za kar misli, da je način, da dobi tisto, kar si želi. Vsak organizem, čigar obnašanje je na ta način vodeno z mislijo, je inteligenten.
>
> Tako to pojasni Dretske in pove, da njega pri vprašanju umetne inteligence zani-

ma to, ali je možno v tem smislu zgraditi inteligenten stroj. Ne zanima ga visok IQ pri strojih, temveč, ali je ta IQ sploh večji od nič - ali je pri strojih prisotno karkoli od miselnega.

Da pa stvar ni tako enostavna, vidimo v nadaljevanju. Najprej pove, da je vprašanje možnosti umetne inteligence pravzaprav vprašanje možnosti zgraditve takšnih umetnih sistemov, katerih obnašanje bo razumljeno v mejah tistega, kar mislijo. In takoj dodaja, da za to ni dovolj zgraditi stroje, ki bi mislili, kajti inteligenca zahteva več kot misel.

Misliti in delovati še ne pomeni biti inteligenten, temveč inteligenca zahteva delovanje, ki ga vodi misel. Za primer navede nepričakovan gib proti očem - mežiknemo. Nihče si ne želi, da mu kaj prileti v oko. Ne mežiknemo zato, ker imamo tako misel, temveč zato, ker imamo ta refleks že vgrajen. Naša misel ni zaslužna za mežikanje.

Do tega mesta ni Dretskeju kaj pripomniti. Marsikdo bi se strinjal z njegovim izvajanjem in z definicijo inteligentnega obnašanja kot obnašanja, vodenega z mislijo. Sedaj pa zadevo poglobi tako, da pravi, da je inteligentno obnašanje tisto, ki je pod nadzorom “tistega, kar nekdo misli” - vsebine misli in ne nosilca te vsebine. To ponazori z naslednjim. Če mikrofonu rečemo “Zatresi se!”, se ta zatrese. In če mu rečemo “Bodi tiho!”, se prav tako zatrese. Ne trese se zaradi pomena besed, temveč zaradi zvoka, nosilca besed. Prav tako bi lahko bilo z mislijo. Celo če bi neka žival imela misli in bi te misli povzročile neko njeno delovanje, to še ne pomeni nujno, da dela to, kar dela, zato, ker misli to, kar misli. To, kar misli, vsebina ali pomen njenih misli, je lahko povsem postransko.

Potem navede primer robota, ki se imenuje Clyde. Ta je opremljen z umetnimi čutili, ki ga vodijo skozi njegovo okolje. Ima baterije, ki jih mora polniti vsakih osem ur. Clydeovi ustvarjalci so nekatera njegova interna električna stanja spremenili v misli. Interna stanja nadzirajo njegovo obnašanje. Tako da je stanje, ki skrbi za priključitev na priključek za polnjenje, sedaj misel z vsebino, da priključitev polni baterije. Clyde se sedaj priključi na priključek, ker misli, da s tem polni svoje baterije. Ali njegovo obnašanje lahko imenujemo inteligentno? Dretske odgovarja, da ne nujno. Obnašanje ni inteligentno, če ni povzročeno z vsebino misli. Clydeovo obnašanje je lahko povzročeno z mislijo, ki ima pravo vsebino, ne pa zaradi te vsebine! To je podobno kot besede s pomenom “zatresi se”.

Po Dretskeju je nauk te zgodbe, da je lažje narediti stroj, ki misli, kot pa takega, ki se obnaša inteligentno.

Nosilec misli in vsebina misli

Potem se sprašuje, kako vsebina misli sploh lahko nadzira obnašanje. Če smo vsi mi fizični sistemi, v temelju torej stroji - kot pravi, nekateri mokri, drugi suhi -, potem morajo biti misli materialna stanja našega živčnega sistema. Vse naše obnašanje, ki je pod nadzorom misli, je torej pod nadzorom fizičnih stanj našega živčnega sistema.

Preden nadaljujemo s spremljanjem Dretskejeve razlage obnašanja, si podrobneje poglejmo nekaj njegovih trditev.

Misel naj bi bila sestavljena iz nosilca in vsebine misli. Nosilec je tisti, ki lahko povzroča to in ono delovanje. Ne smemo pozabiti, na kar Dretske posebej opozarja v svoji prvi opombi, da z delovanjem ni mišljena samo telesna dejavnost. To je lahko, tako pravi, tudi reševanje nekega povsem miselnega problema. Kaj pa je vsebina misli? Dretske večkrat omenja to kot nekaj, "kar nekdo misli", drugje pa spet kot "pomen". Na eni strani imamo nosilce, ki vstopajo v medsebojna razmerja in ustvarjajo nove nosilce ali pa povzročajo določene dejavnosti. Ali je to mišljenje ali pa je mišljenje tisto, ko vstopi določena vsebina misli v zavest? Ali pa oboje? Na eni strani bi torej imeli "učinkujoče mišljenje" z nosilci, na drugi strani pa "pasivno mišljenje" za zavest.

Zakaj bi sploh obstajalo neko zavestno mišljenje, če se s tem pri celi stvari nič ne spremeni? Lahko da je že sama zavest odveč.

Za ponazoritev razlike med nosilcem in vsebino misli je Dretske uporabil že prej omenjen primer z mikrofonom, ki mu rečemo "Zatresi se!". To je na prvi pogled zelo posrečen primer, kajti nosilec misli (zvok) povzroči učinek ne glede na svoj pomen (tisto, kar rečemo). Pri vseh različnih primerih je učinek vedno isti: mikrofon se trese. Kaj pa, če bomo hoteli različen učinek? Verjetno bi morali zvok temu primerno oblikovati. Prav tako je pri mislih. Različne misli učinkujejo različno zato, ker so nosilci misli različni, in ne zato, ker imajo različno vsebino. Nekdo od zunaj, ki bi lahko nekako pogledal "obliko" nosilcev, bi vedel, kako te misli učinkujejo, kar je pravzaprav njihov pomen, ne da bi "odčital" vsebino iz nosilcev! Gledano od zunaj in na predhodno povedano, ima ena misel lahko dva pomena. Ali sta ta dva pomena nujno enaka? Mogoče celo obstajajo misli brez vsebine, ker nosilci misli lahko povsem nemoteno sprožajo razne dejavnosti, pri čemer je sama vsebina postranska, kot pravi Dretske - irelevantna. Mislim, da na takšna pota zaide vsako mišljenje, ki deli misel na nosilca in vsebino.

Ni možnosti (pri Dretskeju), da vsebina misli ni nekako "napisana" na nosilcu misli, ampak se samo pojavi v zavesti takrat, ko se ta misli. Tako bi vsebino povzročal in pogojeval nosilec, kar pa ne sledi iz Dretskejevega primera z mikrofonom, ker je pomen "zapisan" v zvoku in neodvisen od učinka, ki ga ta povzroči. Prav tako pa sam Dretske jasno pove, da mora biti obnašanje, da bi bilo inteligentno,

pod nadzorom tistega, kar nekdo misli: vsebine misli in ne nosilca te vsebine.

Kaj vse je možno pri konceptu “nosilec in vsebina”, si lahko pogledamo pri računanju. Imamo troje misli. Učinek in vsebina prve je “dve”, druge “štiri” in tretje “seštevanje”. V medsebojnem učinkovanju dobimo misel z učinkom in vsebino “šest”. Mišljenje je tukaj potekalo na dveh ravneh: kot učinkujoči nosilci in ozaveščena vsebina le-teh. Če bi nam nekdo (recimo Clydeovi oblikovalci) spremenil vsebino druge misli na “pet” (nosilec bi imel še vedno isti učinek!), bi rezultat ostal enak in spremembe ne bi opazili - le kako naj bi jo!

Kolibriji ali hrošči?

Zakaj je potrebna vsebina misli in s tem zavest, če vse deluje tudi brez tega?

Dretske vsekakor pravi, da so misli materialna stanja v telesu. Pri tem pa dodaja, da ni vsaka lastnost teh notranjih (internih) stanj in še posebej ne njihova vsebina, dejstvo, ki ima opraviti izključno (ali celo primarno) z notranjim povezovanjem.

Kaj s tem misli, ponazori z rastlino škrlatno gilijo [*Gilia aggregata*], ki sredi poletja spremeni barvo iz rdeče v belo. Pojasnilo te spremembe je, da to naredi zato, da privabi opraševalce. Njeni glavni opraševalci so kolibriji, ki imajo raje rdeče cvetove. Sredi poletja kolibriji odidejo in osnovni opraševalci postanejo vešče, ki so jim ljubši beli cvetovi.

Neka druga rastlina, fizično povsem enaka giliji, imenuje jo dvojčica, pa se je razvila v nekem drugem okolju. Prav tako spreminja barvo sredi poletja. To pa stori zato, da se obrani pred požeruškimi hrošči, ki jih privlači rdeče.

Fizično dogajanje pri obeh rastlinah je enako. Pojasnilo za obnašanje posamezne rastline moramo poiskati zunaj rastline, v zgodovini rastline.

Dretske doda še tretjo rastlino, ki je povsem enaka prvima dvema, vendar je narejena v laboratoriju po vzoru škrlatne gilije. Ta rastlina sredi poletja prav tako spremeni barvo. Ali naredi to zato, da privabi opraševalce? Ali pa, da prežene hrošče? Ne. Za takšno obnašanje nima nobenega razloga.

Tako zadevo postavi Dretske. Da bi pojasnili obnašanje nekega sistema, moramo pogosto poiskati vzroke zunaj njega. Imenuje jih strukturne vzroke, za razliko od prožilnih.

Nemožnost umetne inteligence

Za razumevanje nastanka naravne inteligence, ko je sistem upravljan od tistega, kar misli, nam Dretske pove zgodbo o živali z imenom Buster. Buster naj bi se

naučil novega načina inteligentnega obnašanja preko nečesa, kar Dretske imenuje proces učenja. V istem okolju živijo tudi kožuhovinasti črvi ali kožrvi. So različnih velikosti in barv. Skupno jim je to, da imajo kožuh (K) in obliko črva (Č). Buster nima pojma KOŽRV. Za njega je to le ena izmed mnogih nekategoriziranih stvari v njegovem okolju. Nekega dne po naključju povoha velikega rdečega kožrva. V nosu začuti pekočo bolečino in od takrat se izogiba velikim rdečim kožrvom. Podobno se mu zgodi z malim zelenim kožrvom. Sčasoma se začne obnašati na način, za katerega bi se vsi strinjali (tako pravi Dretske), da je inteligenten: izogiba se kožrvom.

Tudi če Bustru podtaknemo ponaredek kožrva, se bo umaknil. To naredi zato, ker misli, da je tisto, kar vidi, kožrv, njega pa je strah pred njimi. Skozi te boleče izkušnje si je Buster izoblikoval pojem kožrva.

Na tem mestu Dretske zatrdi, da se inteligentnega obnašanja ne da ustvariti umetno v laboratoriju. Če bi naredili umetnega Busterja, seveda povsem enakega originalu, bi se ta tudi povsem enako obnašal. Prav tako bi se izogibal kožrvom, ampak to ne bi počel zaradi istega razloga (če sploh zaradi kakšnega) kot original. Ima povsem enaka notranja stanja, vendar se zaradi pomanjkanja zgodovine, njegovo (te konkretne živali) obnašanje ne da pojasniti iz vsebine pojma.

Dretske pride do zaključka, da skoraj ni možnosti za nastanek umetne inteligence.

Definicija inteligence

Z naslednjim stavkom Dretske zaokroži bistvo svoje celotne razlage: “Če z inteligenco razumemo upravljanje delovanja z mislijo, potem moramo za to, da bi ustvarili sistem, ki se ne obnaša samo na enak način kot inteligenten sistem, ampak se obnaša na ta način zaradi razlogov in s tem kaže inteligenco, temu sistemu dati zgodovino, ki takšno upravljanje naredi možno.” Umetnemu Busterju ne moremo dati pojasnjevalnih odnosov med vsebino misli in obnašanjem.

Dretske zaključi s trditvijo, da bi bili mogoče zmožni zgraditi umetne sisteme, ki bi se obnašali kot pravi, ampak se ne bi obnašali tako zaradi istih razlogov. Takšni sistemi ne bi bili inteligentni. Prava inteligenca potrebuje pravo zgodovino, ki bo vzpostavila pojasnjevalno povezavo med vsebino in obnašanjem, ki ga ta vsebina pomaga pojasniti.

Spomnimo se Dretskejeve omembe Fodorjevih H-delcev. Delci postanejo H-delci, če so obstajali takrat, ko je kovanec v Fodorjevi roki ležal s številko navzgor. Nič, kar H-delci počnejo, ni pojasnjeno s tem, da so H-delci. Prav tako je z Dretskejevo inteligenco. Če bi opazovali dejavnost ali obnašanje nekega bitja in pri tem imeli vpogled v vsako njegovo misel in v vsak njegov delec, še vedno ne bi

mogli ugotoviti, ali se to bitje obnaša inteligentno ali ne. Dretskejeva inteligenca ni lastnost bitja, ali pa je lastnost tako, kot je lastnost svinčnika to, da je moj.

Zakaj torej razlaga in razumevanje nekega obnašanja pri Dretskejevih inteligentnih bitjih ostane zunaj njih?

Dretske to doseže z izbiro svojih primerov. Čeprav pravi, da rastlinam pripisuje namene samo zaradi dramatičnih učinkov, pa te pripisovane lastnosti uporablja v izpeljevanju. Samo tako lahko pride do ponaredka škrlatne gilije, ki spreminja barvo "brez razloga". Če pojasnilo, da gilija spreminja barvo zaradi opraševalcev, razumemo po vzročni verigi, pridemo do kolibrijev in vešč, ki so v daljni preteklosti nekako povzročili to spremembo. Ker pa je ponarejena gilija nastala po vzorcu iz prave, je v isti vzročni verigi kot prava in spreminja barvo iz istega "razloga". Če ji pa pripišemo namen, potem ima ta namen tudi ponarejana.

S primerom z gilijami si Dretske poskuša pomagati nekako v prenesenem smislu. Zgodba o Busterju pa naj bi bila nekako bolj resnična. Tukaj nam je postopno prikazana pot pridobivanja pojma skozi "proces učenja". Na koncu se Buster obnaša tako, da se izogiba vsemu, za kar misli, da so kožrvi. Njegovo obnašanje pa je inteligentno le takrat, če je "on osebno" (točno ta) šel skozi proces učenja in ima sedaj zgodovino, ki se ujema z vsebino pojma. Ponarejeni Buster se obnaša povsem enako, vendar neinteligentno, ker nima ustrezne zgodovine.

Čeprav je Dretske večkrat označil inteligentno obnašanje kot obnašanje, ki ga upravlja vsebina misli, pa na koncu jasno vidimo, da ne gre za nikakršno upravljanje, temveč le za ujemanje vsebine misli in osebne zgodovine.

Zakaj pa Dretske tako opredeli inteligenco?

Glavni razlog vsekakor tiči v Dretskejem razumevanju misli in mišljenja. Misel, kot smo že povedali, deli na nosilca in vsebino. Nosilec je notranje stanje, ki v medsebojnem učinkovanju v resnici upravlja delovanje. Z delovanjem so mišljene tako fizične kakor tudi t.i. miselne dejavnosti, kot so reševanje problemov in podobno. Vsebina misli je potem nekaj, kar ne vpliva neposredno na nič. Vsebina misli je to, kar mislimo, ko mislimo. Iz tega sledi, da mišljenje, kot odčitavanje vsebine misli, ni nikakršno delovanje, ampak postransko dogajanje. Tako postavljeno mišljenje lahko enačimo z zavestjo, v kateri se pojavljajo vsebine misli kot posledica medsebojnega učinkovanja nekaterih notranjih stanj.

Kaj je tisti ključni pojem, ki povezuje Dretskejevo opredelitev inteligence in njegov pogled na misli in mišljenje?

Ta pojem je razumevanje. Dretske "čuti potrebo" po razumevanju, vendar ga zaradi svojega pogleda na misli uvrsti izven mislečega bitja. To vidimo v njegovi zahtevi po povezovanju vsebine misli in zgodovine dogajanja. Razumevanje ob-

staja, ampak samo za nekoga zunaj, in to zunaj pomeni, da ni nikogar, ki bi razumel.

Zaključek

Vprašanje umetne inteligence pri Dretskeju ni vprašanje v pravem smislu te besede, kajti že v izhodišču opredeli vsa bitja kot stroje, ene mokre, druge suhe, kot sam pravi.

Dretske v svojih primerih uporablja rastline in nekakšne živali, zato si mi na koncu lahko privoščimo še človeka - in to prav Dretskeja. S pomočjo njegovih znanstvenikov naredimo njegov ponaredek, ki mu je v vsem enak. Dodamo mu še znanje o njegovem nastanku. Ponarejeni Dretske bi na vprašanje o lastni inteligenci odgovoril, da ni inteligentno bitje. Vprašanje je, ali je to v Dretskejevem smislu inteligenten zaključek ali ne. Če je, potem je ponarejeni Dretske postal inteligenten. Enako velja za nadaljnje pridobljene pojme ponarejenega Busterja in za obnašanja, ki iz teh sledijo. Opirajoč se na dosedanje Dretskejevo izvajanje, bi se to imenovalo inteligentno, ker vsebine novih pojmov imajo ustrezno zgodovino. Ali bi se s tem strinjal tudi Dretske?

Is Intelligence Possible at All?

Dretske in the article *Can intelligence be artificial?* raises the question about the possibility of artificial intelligence. But the main subject of his research is, nevertheless, the notion of intelligence as such. The answer to the question: What is artificial intelligence is derived from the answer to the question: What is intelligence as such. In this commentary we analyze Dretske's discussion and test out the solidity of foundations on which he builds his conclusions.
Keywords: (artificial) intelligence, thought, action, F. Dretske.

Literatura

Dretske, F. (1993). “Can Inteligence Be Artificial”. *Philosophical Studies*, 71, str. 201-216.