

STROJE A MYSLENIE

Metodologické poznámky

WŁADYSŁAW KRAJEWSKI

Z filozofických problémov kybernetiky sú nepochybne najspornejšie otázky, či stroje môžu myslieť, či stroje môžu mať vedomie. Vyvolali už mnoho živých diskusií. Tieto diskusie sa však nezriedka vyznačujú nejasnosťou, týkajúcou sa samých pojmov myslenia či vedomia. Preto by som tu chcel sformulovať isté návrhy, ako treba tieto pojmy chápať, a na tomto základe naznačiť smery riešenia samého problému.

1. Známe sú rôzne pokusy vymedziť pojem *myslenia*. Zdá sa, že najpriliehavejšia je definícia, podľa ktorej myslenie je *procesom riešenia úloh* (otázok). Pre mnohé účely je táto definícia celkom uspokojivá. V našom prípade však nestačí, lebo — ako sa ukazuje — môžeme ju ponímať dvojakým odlišným spôsobom, čím dostaneme dva pojmy myslenia. Jeden možno nazvať logickým, druhý psychologickým.

Logický pojem myslenia získavame vtedy, keď riešenie úloh berieme ako čisto logický proces. Z hľadiska logiky spočíva úloha v tom, že sa sformuluje otázka a určité tvrdenia, ktoré majú charakter „premis“. Úloha sa rieši tak, že sa vysloví úsudok, v ktorom by tieto „premisy“ vystupovali ako predpoklady, kým záver by bol odpoveďou na položenú otázku.

Na otázku, či stroje môžu myslieť, treba pri takomto ponímaní veci odpovedať kladne. Sú stroje, ktoré môžu riešiť rozsiahle triedy úloh. Dá sa skonštruovať stroj na riešenie každej úlohy, pre ktorú existuje určitý algoritmus (schéma uvažovania). Nejde tu však o myslenie v akomkoľvek zmysle, ale práve len v zmysle logickom. Inakšie to vyzerá, keď k mysleniu pristúpime z psychologického hľadiska.

Psychologický pojem myslenia získavame vtedy, keď riešenie úloh berieme ako psychologický proces, predovšetkým ako proces charakterizovaný nasledujúcimi etapami: polozenie otázky v dôsledku želania uspokojiť určité potreby (takouto potrebou môže byť aj zvedavosť v oblasti poznania), snaha vyriešiť ju, späť s určitou námahou a napokon samotné riešenie spojené s uspokojením a súčasne s určitou únavou atď. Použili sa tu teda také pojmy ako: želanie, zvedavosť, snaha, uspokojenie a pod. To sú všetko psychologické pojmy, označujúce isté stavy vedomia. Aby sme mohli odpovedať na otázku, či môžu byť vlastné stroju, musíme dať odpoveď na generálnu otázku — čo je vedomie.

2. Vymedzenie pojmu *vedomia* naráža na omnoho väčšie ťažkosti než vymedzenie pojmu myslenia.

Boli pokusy definovať vedomie čisto behavioristicky. Vyskytol sa napr. návrh, aby sa vlastnosť vedomia pripisovala takému predmetu, ktorého správanie je dostatočne komplikované, dostatočne „dôvtipné“, ktorý je, povedzme, dostatočne dlhý čas schopný podvádzať človeka, sám sa vydávajúc za človeka.¹ Hoci vezmeme do úvahy akokoľvek komplikovaný typ ľudského správania, je pravdepodobné, že vždy bude možné skonštruovať stroj, ktorý ho napodobní (čím ešte nie je rozhodnutá otázka, či bude možné, aby jeden stroj reprodukoval *všetky* typy ľudského správania). Znamená to však, že stroje tohto druhu majú vedomie? Nechávam stranou fakt úplnej ľubovoľnosti behavioristických definícií (môžeme si predsa celkom nezáväzne vybrať stupeň komplikovanosti a „dôvtipnosti“ v správaní stroja, od ktorého začneme pripisovať stroju vedomie). Dôležitejšie je, že keď hovoríme o vedomí, a to či už v živote, alebo vo filozofii (napr. keď kladieme otázku vzťahu vedomia a bytia), máme na mysli niečo viac ako určitý typ správania.

Čo treba uznať za vedomie v takom prípade?

Keď hovoríme o vedomí, ide nám o schopnosť pociťovať, prežívať (dojmy, emócie, akty vôle a pod.), teda o to, čo často nazývame *subjektívnym svetom*. Vzniká, samozrejme, otázka, odkiaľ vieme o jestvovaní tohto subjektívneho sveta. Odpeveď je jednoduchá: poznáme ho z introspekcie.

Ako je známe, introspekcia ako metóda poznania má ten veľký nedostatok, že jej chýba hodnota intersubjektívnej verifikovateľnosti. Práve preto psychológia a spoločenské vedy snažia sa ju, pokiaľ možno, eliminovať a nahradiť takými metódami, ktoré by spomínanú hodnotu mali. Je to nepochybne správna tendencia, podmieňujúca pokrok v týchto oblastiach vedy. No napriek tomu nie som tej mienky, že by bolo možné introspekciu *úplne* vylúčiť z psychologického bádania² i z filozofických rozborov. Nad všetku pochybnosť patrí k určitým druhom skúsenosti, a to najbezprostrednejšej, v istom zmysle najistejšej. Descartovské *cogito* sa mi vidí zrejmu tézou; Descartov omyl spočíval iba v tom, že z tejto tézy chcel vyvodiť poznatky o vonkajšom svete. Z tohto hľadiska mal pravdu skôr Locke, ktorý rozlišoval medzi dvoma druhmi skúseností — vonkajšími a vnútornými. Treba len dodať, že *vnútorná* skúsenosť hovorí o subjektívnych stavoch, ktoré sú určitými aspektmi stavov nášho tela a koniec koncov sú podmienené aj *vonkajším* prostredím, v ktorom žijeme. Subjektívnym idealizmom nie je zrejme samotné uznanie vnútornej skúsenosti, ale redukovanie *každej* skúsenosti na ňu.

Pochopiteľne, uznávame nielen vlastné vedomie, ale aj vedomie iných ľudí. Na akom základe to robíme?

3. O subjektívnom svete iných ľudí nemáme priame dôkazy. Robiť závery o jeho existencii môžeme preto iba na základe nepriamych dôkazov.

Medzi tieto dôkazy patrí predovšetkým správanie iných ľudí. Keď iní, podobne

¹ Pozri napr. A. Turing, *Can the Machine Think?*

² Toto stanovisko presvedčivým spôsobom odôvodnil prof. M. Kreutz — pozri *Studia Filozoficne* (1960), č. 5, 20.

ako ja, rozprávajú, smejú sa, vzdychajú a pod., domýšľam sa, že majú tiež vedomie a pociťujú rovnako ako ja sám. No behavioristické dôkazy nie sú jediné. Ľudia mimo mňa sa mi podobajú aj po iných stránkach. Majú analogickú morfológickú a anatomickú stavbu, skladajú sa z tých istých prvkov a chemických zlúčenín, v ich tkanivách prebiehajú rovnaké fyziologické procesy. Všetci ľudia prechádzajú rovnakými štádiami embryonálneho vývoja a v podstate tými istými postnatálnymi procesmi rastu, zrelého života, starnutia, smrti. Pritom rodičia všetkých ľudí sú tiež ľuďmi, patria k druhu *Homo sapiens*, ktorý — ako nás pouča veda — je odvodený z iných zvieracích druhov.

Spomínané vlastnosti možno teda zhrnúť do nasledujúcich skupín: 1. vlastnosti chovania, 2. vlastnosti stavby a zloženia (napr. chemického), 3. vlastnosti pôvodu a individuálneho vývinu.

Vo všetkých týchto troch skupinách máme do činenia s vlastnosťami, ktoré sú intersubjektívne verifikovateľné. O ich podobnosti u všetkých ľudí nám hovorí buď priama skúsenosť, alebo veda opierajúca sa tiež o skúsenosť.

A keďže všetci ľudia sú si podobní z hľadiska týchto dokázateľných základných vlastností, musíme uznať, že aj jediná neoveriteľná vlastnosť — vlastnosť vedomia — je im spoločná. Myšlienka, že iní ľudia majú vedomie podobné môjmu, je teda nadmieru pravdepodobná, ba prakticky istá hypotéza, takže nikomu nenapadne pochybovať o jej pravdivosti.

Aj o zvieratách môžeme tvrdiť, že sa nám podobajú vlastnosťami všetkých uvedených skupín, hoci táto podobnosť nesiahá tak ďaleko ako medzi bytosťami nášho druhu; preto usudzujeme, že zvieratá majú tiež subjektívny svet, ale ten sa určitým spôsobom od nášho líši, je oveľa primitívnejší. Je to tiež dobre odôvodnená hypotéza.

Ako vyzerá v tomto svetle hypotéza, že stroje môžu mať vedomie?

4. Stúpenci náhľadu, podľa ktorého stroje môžu myslieť, odvolávajú sa výlučne na behavioristické vlastnosti. Ako som už spomenul, treba počítať s eventuálnou pravdivosťou tézy, že aj najkomplikovanejšie vlastnosti ľudského správania sa podarí skôr či neskôr reprodukovať pomocou strojov. Nevidím však dôvod, prečo by sa malo porovnávanie človeka a strojov obmedzovať na vlastnosti náležiacie do prvej z uvedených skupín.

Pokiaľ ide o druhú skupinu, tak isto možno nájsť isté spoločné vlastnosti u ľudí i kybernetických strojov. Sú to štrukturálne vlastnosti, ktorými sa zapodieva kybernetika: prítomnosť prvkov, ktoré preberajú informáciu a energiu z vonkajšieho sveta, prvkov, ktoré skladujú energiu a informáciu atď., ako aj systém spätných väzieb týchto prvkov. Tým sa však podobnosť vyčerpáva. Chemické zloženie strojov (kov, sklo a pod.) je úplne iné než zloženie živých organizmov, fyzikálne procesy, ktoré prebiehajú v strojoch, nepodobajú sa fyziologickým procesom prebiehajúcim v organizme (s výnimkou určitých štruktúrálnych vlastností týchto procesov) a pod.

Napokon, v prípade tretej skupiny medzi strojmi a ľudskými organizmami nemožno zistiť vonkoncom žiadnu podobnosť. Stroje, aspoň tie, ktoré už existujú, vznikajú celkom inakšie než organizmy, neprechádzajú takými štádiami indivi-

duálneho vývinu, akými prechádza organizmus. Tým menej sa dá vo vzťahu k nim hovoriť o filogenéze.

Pretože sa stroje líšia od organizmov toľkými intersubjektívne verifikovateľnými vlastnosťami, nemáme žiaden dôvod pripúšťať, že sa podobajú ľudským organizmom v takej intersubjektívne neverifikovateľnej vlastnosti, akou je vlastnosť vedomia.

Niekedy sa môžeme stretnúť s tvrdením, že fakt neexistencie vedomia u strojov nemožno dokázať. Takto klásť otázku je však rovnako nesprávne, ako keď sa na adresu ateistov vyslovujú námietky, že nemôžu dokázať nejestvovanie boha. Onus probandi je na strane tých, ktorí postulujú existenciu nejakého hypotetického objektu, a nie na strane tých, ktorí súdia, že takýto objekt nejestvuje.

5. Otázka, či stroje môžu mať vedomie, nemá výlučne filozofický charakter. Z tej či onej odpovede na túto otázku môžu časom vyplývať podstatné praktické dôsledky. Ak niekto tvrdí, že určité stroje majú vedomie, že sa po tejto stránke ničím nelíšia od človeka, musí z toho vyvodiť závery týkajúce sa zaobchádzania so strojmi takéhoto druhu. Ako je známe, pri styku s ľuďmi (a do istej miery aj so zvieratami) platia určité etické zásady. Máme sa týmito zásadami riadiť aj v styku so strojmi? Známe je aj to, že ľudia sú zodpovední za svoje činy a v určitých prípadoch sú trestne stíhaní. Treba brať na zodpovednosť aj stroje? Pravda, autori čisto behavioristického zamerania neradi hovoria o svedomí, morálke, zodpovednosti. Podľa nich sú to nevedecké pojmy. Avšak v živote musia uznávať isté etické zásady, zaiste schvaľujú aj existenciu súdництва, teda nemôžu utiecť ani pred vyššie položenými otázkami.

6. Zo všetkého, čo bolo povedané, samozrejme nevyplýva, že podľa mňa akékoľvek vedomé bytosti sa nám musia podobáť čo do stavby, chemického zloženia atď. Je celkom možné, že na niektorých planétach žijú rozumné bytosti, ktoré sa nám vôbec nepodobajú. Keby sa s nimi niekedy stretli budúci kozmonauti, mohli by mať nemalé ťažkosti pri ich identifikovaní. Zrejme by najprv upriamili pozornosť na ich správanie. Aby však získali istotu, že sú to živé a vedomé bytosti, a nie napr. stroje skonštruované zatiaľ neznámymi vedomými bytosťami, museli by oblasť svojich pozorovaní rozšíriť. Začali by sa zaujímať, z čoho sa tieto tvory skladajú. Pravdepodobnosť, že sú to živé bytosti, by sa zväčšila, keby sa ukázalo, že sú zostrojené z plastickej substancie, ktorá sa ostatne vôbec nemusí nevyhnutne skladať z bielkovín alebo nukleínových kyselín, ako je to na Zemi. Ešte viacej faktov by prinieslo dlhšie pozorovanie týchto tvorov. Pravdepodobnosť, že sú to živé bytosti, by nesmierne vzrástla, keby sa potvrdilo, že rastú, že sa rozvíjajú a množia, teda — že ich „populácia“ podlieha analogickému procesu, akým je biologický vývin pozemských živých bytostí.

Pravda, nejde o to, aby okolnosť, že ktorákoľvek z týchto vlastností chýba, mala rozhodnúť o veci v negatívnom zmysle. Ide iba o rozšírenie oblasti pozorovania, o to, aby sme sa neobmedzili na pozorovanie správania, ale venovali pozornosť aj ostatným dvom skupinám vlastností.

7. Treba rozlišovať medzi otázkou, či stroje môžu mať vedomie, a otázkou, či človek môže umelým spôsobom vytvoriť vedomú bytosť. Osobne myslím, že sa na túto druhú otázku musí odpovedať kladne. Treba pripustiť, že sa raz biológii

podarí syntéza zvieracích a potom aj ľudských zárodkov a že bude možné docieľiť, aby sa z takéhoto zárodku in vitro rozvinul človek, ktorý bude, samozrejme, vedomou bytosťou. Bude to však práve človek, bytosť z krvi a kostí, a nie z kovu a zo skla. Bude sa podobat' ostatným ľuďom všetkými svojimi vlastnosťami okrem jednej — okrem pôvodu, svojej „histórie“ pred syntézou. Avšak od okamihu, keď nastane syntéza, bude sa už normálne rozvíjať, bude prechádzať tými istými štádiami ontogenézy ako každý človek, bude schopný zanechať svet potomstvo atď. Spolu s podobnosťou vo všetkých vlastnostiach prvej a druhej skupiny to bude stačiť, aby sme ho uznali za človeka obdareného vedomím a podľa toho zaobchádzali s ním tak, ako zaobchádzame so všetkými ľuďmi.

[Studia Filozoficzne (1963), č. 1]