

A kibernetikai „belső cél” fogalom filozófiai tartalmáról

RÓZSAHEGYI EDIT

A kibernetikához — éppúgy mint minden szaktudományhoz — a filozófia két szempontból közelíthet. Vagy azzal a fő céllal, hogy aktuális elméleti problémáinak megoldásához a maga eszközeivel módszertani-elvi segítséget nyújtson, vagy azzal a fő céllal, hogy a szaktudomány már elért eredményeiből kihámozza azt, ami őt illeti, azt, ami — akarva-akaratlanul — filozófiai kérdésekhez szól hozzá. E két szempont ugyan sohasem választható el teljesen egymástól, de nem is azonos. A következőkben a kibernetikához — a teleológia vonatkozásában — kifejezetten ez utóbbi céllal fordulunk.

A teleológia komplex, igen sokrétű filozófiai problémája korunkban sok szempontból új megvilágításba került, „átstrukturálódott”, nem utolsósorban azért, mert a tudományos fejlődés, a régi idealista természeti teleológia kódét eloszlátva, új sikon, reális tartalommal veti fel a célszerűség kérdését. Természeti célok ugyan nincsenek, bizonyos területeken azonban mégis szükségesnek és megalapozottnak bizonyult egy *metodológiai célszerűség-elv* megtartása; milyen értelemben, hol és milyen tartalommal jogosult ez az egyes illetékes területeken? azonos-e, mondjuk, a biológiában és a kibernetikában? — ezekre, és még számos kérdésre ma a filozófiának, ha a teleológia-problémát ki akarja dolgozni, felelnie *kell*.

Ezt a „kell”-t reprezentálja különös energiával, friss, szokatlan hangan szólva a kibernetika. A szakember legyinthet erre: a kibernetika központi fogalma lehet a negatív visszacsatolás vagy az információ, de a teleológia, a „belső cél” semmi esetre sem. Igaza van, és még sincs igaza. *Amennyiben a kibernetika elveit általános, filozófiai tartalmuk szintjén vizsgáljuk*, a „belső cél” igenis lehet elvi jelentőségű, központi fogalma. A szaktudománynak természetesen nem az az égető gondja, hogy a belső cél fogalom tartalmát tisztázza. Sőt még akkor is, amikor saját legáltalánosabb elméleti problémáin dolgozva, mégiscsak eljut ehhez a fogalomhoz — ahogyan, mint az alábbiakban látni fogjuk, pl. Wienerék tették —, akkor is mint a kibernetika *speciális* fogalmát vagy elvét kívánja megragadni, és korántsem filozófiai érvényű kijelentéseket akar tenni. De változtat-e ez azon, hogy ezeknek az elveknek és ezeknek a speciális tartalmaknak *van* filozófiai aspektusuk, *van* filozófiai érvényű mondanivalójuk? A filozófus teljes joggal igyekezhethet *ezt* kihámozni s tisztázni, azon túl, hogy a maga speciális problémáinak megoldásához is szüksége van rá. Wienerék egy *kibernetikai* célszerűség-elv tartalmát próbálták körvonalazni, de amikor akadt, aki elképzeléseikkel filozófiai szemszögből szállt vitába, ezt nem utasították vissza azzal, hogy, mondjuk, ez speciális, szaktudományos igényeik félremagyarázása. Ellenkezőleg, úgy reagáltak, hogy megpróbálták pontosabbá, világosabbá tenni elképzeléseiket. Egyébként nyíl-

ván nemcsak azért jártak el így, mert tudatában voltak elvi nézeteik filozófiai jelentőségének, hanem mert e szinten is megláthatták egy „negatív visszacsatolás” ígérétét; lehet, hogy igen széles feladatsávon közömbös a szaktudomány számára, milyen „belső cél” értelmezéseken vitatkozik a filozófia, de a (szaktudományos) elméleti kérdések egy bizonyos szintjén nem árt, sőt szükséges, hogy egy nagyobb probléma-egész összefüggéseiben el lehessen helyezni a speciális elveket, és akkor nem mindegy, képes-e és mennyire képes a filozófia ennek az átfogó probléma-egésznek a területén egy bizonyos „fogalom-szabályozással” segíteni a szaktudomány „működését”.

*

Az az új nézőpont, amelyből a kibernetika az anyagi rendszerekhez közeledik, önmaga kifejezi a célszerűség-probléma és a fiatal tudományág szoros kapcsolatát. A kibernetika köztudomás szerint elvonatkoztat a különböző rendszerek anyagi tartalmától, sajátos szubsztrátumától, és azoknak csak funkcionális vonatkozásai, működésük „hogyan”-ja érdekli. Ebből a szempontból közös nevezőre hozza az élő és az élettelen. Organizációs elvei, működési törvényszerűségei oldaláról közeledve az élő organizmushoz, nemcsak ismerni, hanem *alkalmazni* is kívánja azt, amit pl. a fiziológiától tanul. Az új elveket élettelen automatákon reprodukálja, és valóban bámulatos eredményeket ér el. A célkereső szerkezettel ellátott légelhárítótól a feltételes reflexeket modellező „állatokig” és a „gondolkodó” gépekig önvezérléses mechanizmusok egész sorát alkotja meg, és az, ami ezekben különbségeiken, igen eltérő komplexitási fokokon túl közös, úgy tűnik, éppen működésük célszerűsége, egy „belső célszerűség”, amely nélkül viselkedésüket nem lehet megérteni.

Nem csoda ezért, hogy a célszerűség-probléma a szó szoros értelmében az új tudománynak már a hajnalán színre lép. Rosenblueth és Wiener az új nézőpontot elméletileg megalapozó cikkükben már 1943-ban kifejtik, hogy a teleológia eszméi nemcsak hasznosak és fontosak, hanem feltétlenül *szükségesek a viselkedés, illetve a működés bizonyos fajtáinak interpretálásához, akár élő szervezetről, akár élettelen gépekről van szó.*¹

A célszerűség a kibernetika alapjáról tehát mint szükséges, élő szervezeteiket és élettelen szerkezeteket egy bizonyos szempontból egyaránt interpretálni képes *módszertani elv* bukkan fel — és éppen mint ilyen igen érdekes a *filozófia* szempontjából is. A kibernetika nem utolsósorban az élő szervezet viselkedésének célszerű organizációját ragadta meg és vitte át élettelen rendszerek működésére, forradalmian új nézőpontjának ez legalábbis az egyik elvi alapja. Milyen célszerűség-fogalomra épül ez az elv? Mi a tartalma, a jelentése itt a célszerűségnek? Lehet-e egyáltalán egy olyan célszerűség-értelmezést kidolgozni, amely magyarázataival valóban át tudja fogni a kibernetika érdeklődési körébe tartozó minőségileg annyira heterogén területet?

Rosenblueth és Wiener említett cikkükben megpróbálják jellemezni és meghatározni a célszerűség tartalmát és kritériumát. Évekkel később a R. Taylorrel kipattant vitában fenntartják és pontosabbá teszik álláspontjukat.²

¹ Rosenblueth and N. Wiener: *Behavior, Purposes and Teleology* (*Phil. of Science*, vol. 10. 1943. 18—24. o.)

² A vitát lásd: R. Taylor: *Comments of a Mechanistic Conception of Purposefulness* (*Phil. of Science*, vol. 17. 1950. 310—317. o.); A. Rosenblueth and N. Wiener: *Purposeful and Non-Purposeful Behavior* (*Uo.* vol. 17. 1950. 310—317. o.); R. Taylor: *Purposeful and Non-Purposeful Behavior: a Rejoinder* (*Uo.* vol. 17. 1950. 327—332. o.)

Ezek a „korai” fejtegetések és máig is gyűrűző filozófiai visszhangjuk mindenképpen érdekes tanulságokat kínálnak, különösen ha számításba vesszük, hogy a kibernetika filozófiai tanulságait elemző irodalmunkban a célszerűség kérdése nem szerepel túlságosan nagy súllyal.

Wienerék szerint a célszerűség viselkedés-kategória. A viselkedést igen tág értelemben határozzák meg, mint „egy objektum bármely modifikációját, amely külsőleg felfogható”, azaz egy élettelen rendszerre, egy járó órára, egy működő autóra éppúgy vonatkoztatható, mint egy élőlényre. Ezek után szerintük célszerű minden olyan viselkedés, „amely egy cél, azaz egy végső állapot elérésére irányul, egy olyan végső állapotra, amelyben a viselkedő objektum meghatározott időbeli és térbeli korrelációba jut más tárgyakkal vagy eseményekkel”.³ A célszerűség tehát mindenekelőtt célirányultságot jelent. A célként felfogott végső állapot azonban nem valami vaktában elért, külső, esetleges kölcsönviszony a viselkedő objektum számára. Viselkedésükre nézve valóban célszerűeknek csak azokat az entitásokat nevezhetjük, amelyeknek belső organizációjához ez a végső korreláció szorosan hozzátartozik, működésük elemzésének szerves része a cél. A falióra 11 órakor megállt. Mutatóinak ez a végső helyzete volna az a cél, amelyre működése irányul, az az idő- és térbeli korreláció, amelyet az óra működése végül is elért? Nyilván nem. Az időpont, a térbeli konfiguráció, melyet a szóban forgó állapot képvisel, teljesen közömbös, külső az óra előző viselkedésének megértése szempontjából, a falióra működése ezen az alapon nem nevezhető célszerűnek. A közönséges pisztollyal célba lehetek, de ugyanígy lőhetnek vele vaktában, bármire, a célpontok külsők, tetszőlegesen a pisztoly belső mechanizmusa szempontjából. A pisztoly viselkedése nem célszerű. Célszerű viszont minden olyan mechanizmus (par excellence a szervomechanizmusok), amelyben a működésnek egy bizonyos önellenőrzése oly módon megy végbe, hogy a rendszer működése során produkált reális modifikációkat egy alkalmas érzékelő berendezés folyamatosan összehasonlítja a kívánt végeredménnyel, az eltéréseket képviselő „hibajelek” pedig automatikusan a korrekció irányában befolyásolják a rendszer működését. Az ilyen mechanizmus tehát a kívánt cél elérése érdekében egy sajátos működési elv, a negatív visszacsatolás segítségével folyamatosan önmagát korrigálja. Rosenblueth és Wiener a negatív visszacsatolás elvét célszerűség-fogalmuk igen lényeges mozzanatának tartják. Tulajdonképpen ez, a célra irányultságot, a cél regulatív szerepét realizáló mechanizmus az, amely egy rendszer viselkedésének és a viselkedés céljának *belső* egységét megvalósítja. A negatív visszacsatolás így szerintük a célszerű viselkedés kritériuma lesz. Ime tehát a Wienerék által javasolt célszerűség-fogalom: *a célszerűség minden olyan viselkedés belső jellemzője, amely egy meghatározott végső állapot elérésére irányul: ezt a célirányultságot a viselkedés sajátos szervezettsége képviseli, amely lényegében azt jelenti, hogy a cselekvő objektum állandó irányító jelzéseket kap a céltől, s segítségével viselkedését önmaga modifikálja.*

Kétségtelen, hogy az automatikus önszabályozó berendezéssel rendelkező gépek működését a negatív visszacsatolás elve nélkül nem lehet megérteni. Nem vitás az sem, hogy ez az elv „természetes” megvalósulási formájában az élő szervezet magasán szervezett működésének igen fontos funkcionális eleme. A testhőmérséklet-szabályozás, az egyensúlytartás stb. mechanizmusán túl a biológia egyre mélyebb szinteken, az elemi és a legbonyolultabb anyagcsere

³ Rosenblueth — Wiener: *Behavior, purposes . . .* 18. o.

folyamatok alapján felismeri ennek a regulatív elvnek az érvényesülését.⁴ És, végül, tagadhatatlan, hogy a modern bonyolult kibernetikus gépekben, amelyeknek egyik alapvető működési elve a negatív visszacsatolás, több-kevesebb sikerrel igyekeznek instrumentálni az ember pszichés működésének bizonyos folyamatait, illetve vonásait is. Úgy tűnik tehát, hogy a negatív visszacsatolásban a célszerűségnek valóban valami közös, általános kritériumát sikerült megragadni.

De mennyire mély, hol és meddig használható ez a kritérium?

Mindenekelőtt valamit tisztázni kell: tulajdonképpen mi is az, amit Wienerék célként határoznak meg?

Emlékeztetünk rá, hogy az imént idézett definícióban a célt kifejezetten mint a viselkedő objektum egy bizonyos végső állapotát jelölték meg, amelyre ez az objektum törekszik. I. Scheffler szerint azonban ezzel kapcsolatban a következő nehézség merül fel:⁵ a célszerű viselkedés egy végső állapot elérésére irányul. Ezt az irányulást a célból folyamatosan érkező jelzések regisztrálása biztosítja. A cél tehát az, ami a viselkedés folyamán jelzéseket küld az objektumnak. A végső állapot azonban nyilván nem lehet azonos a céllal, hiszen időben későbbi, mint a viselkedés, és így nem vezérelheti az utóbbit. Mi hát akkor a cél? Ebben a felfogásban nem lehet más — mondja —, mint az az entitás, amellyel a korrelációnak — a végső állapotnak — ki kell alakulnia, de nem maga a végső állapot. A cél tulajdonképpen a viselkedés során végig egzisztáló *cél-objektum*. — Wienerék néhány példája, amelyen illusztrálni és magyarázni kívánják elképzelésüket, valóban megerősíteni látszik ezt a felfogást. Hangsúlyozzák, hogy egy kibernetikus berendezés célszerűségének kérdését nem oldja meg a tervező ember céljaira való hivatkozás. A szervomechanizmus célszerűsége mindig az ő saját, belső célját tételezi fel. Milyen célt? Egy szonikus célkereső apparátussal felszerelt torpedó *célja* pl. — mondják — az irányító *hanghullámokat kibocsátó fókusz*. Egy fototropikus szervomechanizmus *célja a fény*. Egy nyulat kergető mű-kutya *célja a nyúl*. Ezekben a példákban a megjelölt célok valóban mind maguk az egzisztáló entitások, *célobjektumok*, amelyekkel egy bizonyos végső korrelációnak ki kell alakulnia.

Pontatlanság tehát a cél „végső állapotként” való meghatározása az adott felfogáson belül is? Úgy gondoljuk, a példák és a definíció eltérése bizonyos tisztázatlanságot tükröz, de ezt ilyen egyszerűen nem lehet elintézni. Valójában arról van itt szó, hogy bár Wienerék kifejezetten „metodológiai” célszerűségről beszélnek, az ontológiai és a metodológiai szempontok viszonya „belső cél”-magyarázatukban keveredik egymással, ezért azután a metodológiai elv tartalma sem világos és problémamentes.

Nézzük csak meg mindenekelőtt egy, a negatív visszacsatolás elvén alapuló viszonylag egyszerű berendezés, pl. egy szervomotor segítségével kormányozott hajó működését. A hajó irányát a hajó hossz tengelye reprezentálja, ezzel, amíg a hajó helyes irányban halad, a kormánylapát síkja egybeesik. Az esetleges eltéréseket egy érzékelő berendezés fogja fel (a mi szempontunkból közömbös, hogy ez a kormányos vagy egy mesterséges automatikus érzékelő apparátus-e), és a kormánykerék elfordításával reagál rá. Erre egy villamos berendezés a kerék elfordulási irányát és nagyságát jelző feszültséget állít

⁴ О сущности жизни. Ред. Г. М. Франк, А. М. Кузин. Москва 1964.

⁵ I. Scheffler: *The Anatomy of Inquiry. Phil. Studies in Theory of Science*. New York 1963.

elő, egy elektromos berendezés ugyancsak feszültséggel jelzi a kormánylapát tényleges helyzetét. A két feszültséget egy összehasonlító egység kivonja egymásból, és a szervomotorra visszavezetett (felerősített) *különbség* segítségével a kormánylapátot megfelelő irányban és mértékben elforgatják.⁶ A vezérlést tehát itt a *meghatározott irányú útvonalat* jelző optikai jelek regisztrálásával egy érzékelő berendezés — mondjuk a kormányos — valósítja meg. Van cél-objektum: ez a bizonyos meghatározott útvonal, és irányító jelzéseket is küld a szerkezetnek. De mondhatjuk-e, hogy a cél-objektum a kormányzó szerkezet működésének célja? Ha meggondoljuk, hogy az irányító jelzéseket kibocsátó objektum *nem* határozza meg az automata konkrét viselkedését, sőt konkrétan azt a viszonyt sem, amelybe végül is a szerkezet kerül vele, azaz a viselkedés konkrét eredményét sem, akkor, természetesen, úgy tűnik, nem, nem mondhatjuk. Elvileg minden további nélkül elképzelhető, hogy, mondjuk, speciális körülmények között a hajót úgy kormányozzák, hogy *ne* haladjon a hossz tengely képviselte egyenes irányban. Ez a berendezés ugyanolyan elvek alapján működik, ugyanaz marad a cél-objektum, éppúgy az irányító jelekre reagál, éppen csak ellenkező előjelű viselkedéssel, mint az előző esetben. „Pozitív” és „negatív” célok lehetőségére Wienerék is utalnak. (Egy fototropikus szervomechanizmus célja lehet az, hogy *keresi*, de ugyanúgy az is, hogy *kerüli* a fényt — mondják.)⁷ A „pozitív”, illetve „negatív” megkülönböztetés itt nyilvánvalóan nem a célobjektumra, hanem a kialakuló végső állapotra vonatkozik. Mindezek alapján a cél a viselkedő automata és a célobjektum egy bizonyos végső kölcsönviszonya lesz. A célobjektumból érkező jelek „értelmét”, jelentését az automata belső szerkezete adja meg. A viselkedés magyarázatához ezért nem az egyik vagy a másik mozzanatnak, hanem a kettő kölcsönviszonyának megértése szükséges. Természetes azonban, hogy a viselkedés során egzisztáló célobjektum mint a kölcsönhatás egyik pólusa valóban szükségszerű összetevője lesz az automata célszerű viselkedésének.

Ezen az úton tehát egy relációs cél-fogalomhoz jutottunk el, egyébként egybehangzóan a wieneri definícióval. De metodológiai, „belső” cél ez? A metodológiai célszerűség-felfogásnak azt kell jelentenie, hogy képesek vagyunk teleologikus módszerrel, a célként felfogott eredményből magyarázni a viselkedést, anélkül hogy ehhez szellemi (emberi vagy emberfeletti) tényezők bevonására lenne szükség. — Vegyük most, számba a Scheffler által említett értelmezési problémát. Ha a cél-korreláció végső állapot, hogyan határozhatja meg a korábbi viselkedést? Erre a kérdésre az *ember* céljaira való utalás nélkül nem lehet felelni. Az automata belső szerkezete nyilvánvalóan potenciális hordozója a cél-korrelációnak, a „képességnek”, a lehetőségnek, hogy a megfelelő célobjektummal kapcsolatba kerülve a gép meghatározott állapotba jusson. A szerkezet is, a kapcsolat is az ember konstruktív munkájának eredménye, s így azt, hogy *mi* az automata célja, *milyen* a belső szerkezete, *mi* az irányító célobjektum és *így mi lesz a végső korreláció tartalma*, az ember céljaira való hivatkozás nélkül nem lehet megmagyarázni. A végső állapot regulatív cél-elvként az ember fejében alakul ki, az automata viselkedése egészében *ezt* realizálja, s így viselkedésének célszerűsége, céljai csak az emberi tevékenység-folyamaton belül nyernek értelmet. *Ebben az értelemben* igaz pl. V. N. Szvincsikijnek az a megállapítása, hogy a mesterséges automaták viselkedését

⁶ Tarján R.: *Kibernetika*. Gondolat, 1964. 110—111. o.

⁷ Rosenblueth—Wiener: *Purposeful and ...* 325. o.

tekintve a szubjektummal való kapcsolaton kívül nincs értelme a célnak.⁸ Amint láttuk, ez abból következik, hogy ha a viselkedésüket, s így viselkedésük eredményét is *tartalmasságukban* próbáljuk megragadni, ez egy szükségszerűen relációs célértelmezésen át elkerülhetetlenül az emberi tevékenységgel való kapcsolat kérdéséhez, tehát az ontológiai alapok tisztázásának szükségességéhez vezet. Wienerék ezt nem látják, célmeghatározásukat „a” belső, metodológiai célszerűség-felfogás alapjának tekintik, amellyel kapcsolatban — mint a Taylorrel folytatott vitában hangsúlyozzák — irreleváns az emberi célokra való minden hivatkozás.

Abban persze igazuk van, hogy az emberi célokra hivatkozni *nem elég*, mert ezzel még nem tisztázzuk a szóban forgó viselkedés *specifikumát*.

Egy közönséges vadászfegyver is emberi célokat valósít meg. Emberi alkotás, adekvát egy céllal, azzal, hogy lelőjje a vadat, a vadász e cél érdekében hozza működésbe. A fegyverre való vonatkozást kiemelve a cél itt így fogalmazható meg: az ember célja az, hogy a célobjektum és a fegyver (pontosabban egy része: a golyó) meghatározott kölcsönviszonyba kerüljön egymással. Világos, hogy ez a kölcsönhatás, a végső állapot, reálisan nem létezik a fegyver működése *közben* vagy *előtte*. De létezik ideálisan a tervező, illetve a vadász fejében. Éppen ez a céljuk tartalma. Sőt, bizonyos értelemben nem is pusztán ideálisan létezik ez az állapot. Amikor a tervező, illetve a készítő létrehozta a fegyvert, az eszköz szerkezete, ezzel kapcsolatban várható (mert „előre kiszámított”) viselkedése tulajdonképpen éppen az ember fejében előzetesen meglevő kölcsönhatást reprezentálja. Ezért, amikor a vadász működésbe hozza a fegyvert, célja nemcsak azt a kívánságot vagy hitet fejezi ki, hogy a fegyver segítségével elejtheti a vadat, hanem azt a tudást is, hogy a puská meghatározott feltételek között alkalmas arra, hogy a cél-korrelációt realizálja. Ezeket a meghatározott feltételeket (legalábbis részben) ismét az ember teremti meg. Megközelíti, célbaveszi a vadat, és működésbe hozza a fegyvert. Mint az egész tevékenységfolyamat láncszeme, mint az emberi cél reprezentánsa, világos, hogy a fegyver működése is célszerű lesz. De csak mint ilyen. Hogy ez a célszerűség „külső” marad a fegyver viselkedése számára, azt az bizonyítja, hogy a működését mintegy körülfogó emberi tevékenységtől elvonatkoztatva nincs értelme vele kapcsolatban célszerűségről beszélni. A célobjektum a fegyver működését nem befolyásolja. Ha valaki a fegyver tökéletes nem-ismeretében véletlenül meghúzza a ravaszt, a fegyver ugyanúgy fog működni (amennyiben működik), függetlenül attól, hogy a golyó a levegőbe repül-e vagy vak-tában el is talál valamit. A fegyver viselkedésének magyarázatához nem kell elemezni a célobjektum viselkedését.

Egy automatikus célkereső berendezéssel ellátott mechanizmus (pl. egy korszerű torpedó) viselkedése ebből a szempontból nyilvánvalóan minőségileg más. A célobjektum megjelenése megváltoztatja a rendszer egész működését, és ha a tipikus viselkedését önmagában, az emberi tevékenységtől elvonatkoztatva akarjuk elemezni, sajátosságát ekkor is csak mint cél-meghatározottságot, célrairányultságot tudjuk megnevezni. Ebben az értelemben beszélhetünk a mechanizmus „belső célszerűségéről”. A törekvés, hogy ezt a *belső* célszerűséget mint ilyet értelmezzük és tisztázzuk, teljesen jogos, sőt éppen ez a metodológiai probléma. Az előbb elmondottakból következik, hogy ez a probléma nem

⁸ В. Н. Свинцинский: *Понятие целесообразности и функционирование кибернетических систем (О сущности жизни)*. Москва 1964. 192—200. o.).

azonos azzal a feladattal (amelyre gyakran le szokták szűkíteni), hogy bebizonyítsuk a célszerű emberi tevékenység és a gépi „célszerűség” ontologikus szétszakíthatatlanságát, de nem is oldható meg nélküle.

Először is az előbbieik alapján lesz világos, hogy amikor bármely gép működéséről önmagában beszélünk, mindig absztrahálunk. Egy összetett tevékenységfolyamatból kiemelünk bizonyos szakaszt vagy szakaszokat. Az iménti egyszerű példák esetében aligha lehet kétséges, hogy ezek a szakaszok az (emberi) cél megvalósítási folyamatának részei. Csakhogy: a célmegvalósítás folyamata maga is összetett. És nemcsak mechanisztikus értelemben, azaz amennyiben egymástól megkülönböztethető működésfázisok egymásutánja, amelyekben az eszközt kauzálisan hatni engedjük. Jóllehet „gyakorlati” (pl. termelési) célok esetében a cél megvalósítása egészében gyakorlati folyamat, azaz megfigyelhető, „külső” változásokban manifesztálódó operációk sora, ez a kauzális sor célt megvalósítani csak akkor fog, ha — mint a vadászfegyver példáján láttuk — *meghatározott feltételek között* kapcsolódik a tevékenységfolyamathoz, illeszkedik az egészbe. Az egyes manipulatív tevékenység-mozzanatokat, műveleteket, bonyolultságuktól függően több-kevesebb nehézséggel, de elvileg minden további nélkül formalizálni és instrumentálni lehet, azaz emberi funkciókból élettelen eszközök, gépek funkciójává lehet őket tenni. Hogy a „közönséges” gépek működését ezek után „önmagában” vizsgálva nem nevezhetjük célszerűnek, ez csak az „elidegenült” gépi formában nyilvánvalóvá vált kifejezése annak, hogy ezek az egyszerű vagy összetett műveletek emberi formájukban is viszonylag önálló mozzanatai voltak a célszerű tevékenységnek, azaz bizonyos értelemben közömbösek voltak e céllal szemben, különböző célok „szolgálatára” egyformán alkalmasak lehettek.⁹ De lehet-e instrumentálni a célmegvalósítás folyamatát ezen túlmenően, nevezetesen magát a célkeresés, a célhoz való alkalmazkodás aktusát? Ha csak a tényeket nézzük, akárcsak az olyan viszonylag egyszerű automatikus berendezéseket, amelyeket az előbbiekből említettünk, a kérdés feleslegesnek látszik. A kibernetika ilyen, sőt ezeken ma már messze túlmenő eredményei gyakorlatilag látszanak bizonyítani, hogy, legalábbis bizonyos határok között, igen. Hát nem éppen a (belső) célirányultságában, a célokhoz való alkalmazkodásban nevezhető-e meg pl. a szervomotoros kormányzó berendezés működésének specifikuma? — A nehézség ott van, hogy noha szinte önként kínálkozik ez az értelmezés, mihelyt a belső célszerűséget elemezni kezdjük, kiderül, hogy nagyon nehéz megragadni a *belső cél* fogalmát úgy, hogy valóban belső is legyen és cél is legyen. Wienerék szerint a „belső” jelleg azt jelenti, hogy a gép „célja” nem azonos a tervező, az ember céljaival. Láttuk viszont, hogy pl. az iménti kormányzó apparátus célját az egyenes útvonal megtartásában, tehát egy konkrét, tartalmas végső állapotban jelölve meg, kiderült, hogy ez éppen nem „más”, mint az ember célja, sőt, célmivolta éppen hogy *csak* az emberi tevékenységre vonatkoztatva érthető meg. Ha viszont azt állítjuk, hogy a cél egyszerűen egy meghatározott minőségű *objektum*, ezzel a célfogalmat tulajdonképpen viszony, állapot jellegétől, tehát éppen a viselkedésre való vonatkozásától fosztottuk meg. A tartalmasságában megragadott cél tehát nem szakítható ki az emberi tevékenységből anélkül, hogy értelmét ne

⁹ A. N. Leontyev, D. J. Panov: *Az ember pszichológiája és a technikai haladás. (Filozófiai problémák a magasabbrendű idegműködés fiziológiájában és a pszichológiában. Akadémiai Kiadó, 1965. 514—519. o.)*

veszítené, azaz nem tekinthető „belső” célnak. Világos ezek után, hogy ha a kibernetikus berendezések működési specifikumát célszerűségükben állapítjuk meg, ezen csak egy *formalizált* és valamiképpen *objektívált* célt és ennek a gép viselkedésére gyakorolt hatását, pontosabban e hatás hogyanját érthetjük. A cél formalizálása azonban nem egyszerűen azt jelenti, hogy a rendszer működésébe be kell kapcsolni, magában a rendszerben képviseltetni kell azt a bizonyos objektumot, amellyel meghatározott kölcsönviszonynak kell végül is kialakulnia, hanem formalizálni és „objektíválni”, képviseltetni kell magát a végső állapotot, a cél tulajdonképpeni tartalmát is. A rendszerben valamiképpen eleve egzisztálnia kell a végső állapotnak, enélkül a célszerűsége nem lehet „belső”, de ugyanakkor potenciálisnak, „megvalósítandónak” is kell lennie, enélkül nem lehet cél. Az egyszerű negatív visszacsatoláson alapuló szervomechanizmusokban az a bizonyos célobjektum tulajdonképpen kettős szerepet tölt be: nemcsak a „környezetet” képviseli, az entitást, amellyel a viselkedő objektum kapcsolatba kerülve, minden időpillanatban valamilyen kölcsönviszonyban van, hanem előre meghatározott jel-nagyságok, a *referencia szintet szolgáltató fizikai jel-értékek formájában egyúttal magát a végső állapotot is*. Sőt e kettős funkciójában ez az utóbbi a döntő, a gép működési specifikuma szempontjából nemcsak az objektum konkrét fizikai tartalma lesz mellékes, hanem tulajdonképpen maga ez az objektumként való egzisztálás is csak azért lesz jelentős, mert hordozója lehet a rendszeren belül a végső állapotot reprezentáló jel-értékeknek. A gépi célszerűségben a cél tehát lényegében kvantitatív formát ölt, fizikai jelek meghatározott mennyiségi értékévé transzformálódik. Ennek megfelelően, természetesen, magának a célszerűségnek az értelme is megváltozik. A cél elérése, a célnak megfelelő viselkedés kialakítása egyszerűen annyit jelent, mint a viselkedés során ugyanezen jelek ugyanilyen értékeit produkálni. Ehhez, csak a belső célszerűség szempontjából szemlélve a gép viselkedését, működésének konkrét tartalmától is el lehet, sőt el is kell tekinteni. A lényeg az lesz, hogy az egész viselkedést éppúgy „mérhetővé” kell tenni, formalizálni kell, mint a célt, olyan értelemben, hogy eltérő viselkedések a referenciaszinttel összehasonlítható fizikai jelek különböző értékeiben legyenek kifejezhetők. Ezzel az értelmezés legnagyobb nehézségét, a cél és az eredmény elvi-minőségi heterogenitását sikerül kiküszöbölni, s így a kérdés most már „csak” a „hogyan” lesz. Hogyan alakulhat a gép viselkedése úgy, hogy végül is a két érték — a referenciajel és a kimenőjel értéke egybeessék? Erre válaszol a kibernetika a negatív visszacsatolás elvével. Világos ezek után, hogy az *organizációnak* ez a sajátos módja a kibernetikában a célszerűség központi elve, csomópontja lesz, de világos az is, hogy a tartalomtól való alapos elvonatkoztatás, egy sajátosan formalizált cél objektíválása talaján lesz az. Ennek megfelelően mint „belső” kibernetikai elv a célszerű organizációnak is egy elvont modelljét adja csupán, amely nem zárja ki a struktúra nem is lényegtelen eltéréseit.

Mindezek alapján pl. a szervomotoros hajókormányzó berendezés belső célszerűségéről a követhető értelemben beszélhetünk: Az egyenes útvonal nem „belső cél”, de meghatározott értékű (pl. optikai) jeleket „küld”, amelyeket más alkalmas „fizikai nyelvre” lefordítva (azaz, mint esetünkben, pl. elektromos feszültséggé átalakítva) a szerkezet működtetésére lehet felhasználni. E „másik” nyelv konkrét minőségét (elektromos hatás-e vagy valami más) a kiinduló objektum természete nem határozza meg, de az irányító jelek meghatározott értékét (pl. a feszültség előjelében és nagyságában) ki kell fejeznie.

Ezért a cél-objektumra vonatkoztatva úgy fogható fel, mint ennek bizonyos paraméter-értékeit reprezentáló jel. Ha ebből a jelértékből, a referenciajelből indulunk ki, a rendszer viselkedését a következőképpen írhatjuk le: Ha a referenciajelet nem közvetlenül használjuk fel a szerkezet működtetésére, hanem folyamatosan összehasonlítva a működés reális eredményét azonos módon kodifikáló értékekkel, a kettő *eltérését*, a „hibajelet” alkalmazzuk vezérlésre, a szerkezet mindaddig fog működni, míg az eltérés nulla nem lesz, azaz, amíg a szerkezet olyan állapotot nem ér el, amely a referenciajellel megegyező értékkel fejezhető ki. E *belső* működést vizsgálva ezek után a referenciajelet úgy tekinthetjük, mint a rendszer végső állapotát előzetesen képviselő és a viselkedést emberi beavatkozás nélkül irányító szerkezeti elemet, tehát mint *belső célt*, és a viselkedést ezen az alapon belsőleg célszerűnek nevezhetjük. Az, hogy mikor *mi* szolgáltatja a referenciajelet, ill. a referenciaszintet, hogy a konkrét célkitűzéseknek megfelelően milyen anyagi szubsztrátumok miféle konkrét szerkezete realizálja a célszerű organizációt, ez, természetesen, korántsem közömbös a megoldandó feladatok, tehát a gép konkrét viselkedése szempontjából, de mihelyt ezeket is bevonjuk a vizsgálatba, ez a célszerűség elvi tisztázása szempontjából azt jelenti, hogy túllépünk a szorosan vett belső kibernetikai, metodológiai síkon és a problémát visszahelyezzük a célszerű emberi tevékenység reális összefüggéseibe. Amikor Wienerék metodológiai célnak tüntették fel a fényt, a hajót vagy a célpontul szolgáló nyulat, amikor célszerűségfogalmukat bizonyos *konkrét* viselkedések elemzéséhez hasznos és szükséges módszertani elvnek nevezték, függetlenül attól, hogy élőről vagy élettelen gépekről van-e szó, ilyen „túllépést” hajtottak végre anélkül, hogy ezt észrevették volna. Ezzel pedig tulajdonképpen, minden metodológiai igazságmag ellenére is, amelyet új nézőpontjuk képviselt — kettős elvi félreértés számára teremtetek alapot.

Először is, joggal úgy tűnhetett, hogy a módszerük elvi alapját képező célszerűség-felfogás nem egyszerűen egy sajátos tudományág speciális metodológiai elve, hanem általában a célszerű viselkedés, sőt „a” célszerűség kritériuma kíván lenni. Jellemző, hogy Taylor is, Scheffler is így kezelik és erről az alapról bírálják elképzeléseiket. Úgy látják, hogy Wienerék álláspontjának megfelelően a célszerűség kritériuma a jeleket produkáló célobjektum. De ha ezt elfogadjuk, mondják, ezzel eleve kizárunk a célszerűség köréből minden olyan viselkedésformát, ahol ilyen célobjektum nincs. Márpedig a — legalábbis a tudatos emberi — viselkedés számos esetben ilyen. Az ember célul tűzheti ki pl. a bölcsek kövének megtalálását, és mindent megtehet, hogy ezt el is érje — jelenti ki Taylor. Sőt, az embernek lehetnek pl. olyan céljai is, hogy, mondjuk napról napra alázatosabb legyen — így Scheffler. Vajon lehet-e tagadni, hogy mindezekben az esetekben célirányuló, célszerű a viselkedése? Aligha. Márpedig itt irányító jeleket küldő célobjektumról nem igen beszélhetünk. Mi következik ebből? Taylor szerint az, hogy ez is bizonyítja az új célszerűség-elv hasznának és értékének kétséges voltát. Scheffler szerint az, hogy a kibernetika által javasolt kategóriákkal le lehet írni és meg lehet magyarázni egyes viselkedésformákat, másokat viszont nem. Ahol célobjektum van, ott a viselkedés célszerűsége a kibernetikai értelmezésben ragadható meg (legalábbis az esetek nagy részében), függetlenül attól, hogy élettelen gép vagy élő entitás viselkedéséről van-e szó, ahol azonban nincs, ott ez az értelmezés nem kielégítő.¹⁰ —

¹⁰ I. Scheffler: *id. mű.*

Az előzők alapján azt kell mondanunk, hogy *így* ez a bírálat rossz vágányon halad. Wienerék belső célszerűsége — függetlenül attól, hogy ők maguk ezt elvileg mennyire tisztázták — valójában nem „a” célszerűség értelmezése valamilyen általános metodológiai értelemben. Legelőször ezt kell meglátni — tisztázni kell a mesterséges szervomechanizmusok belső célszerűségének *speciális* metodológiai jellegét, amely belső célszerűséghez csak a szervomechanizmusok reális alapját képező emberi tevékenység számos tartalmi és struktúrális sajátosságának — a célszerűséget érintő sajátosságának — absztrahálása révén jutunk el. Az eredmény — a kibernetika speciális módszertani elve — természetesen tartalmazza a célszerű viselkedés bizonyos általános jellemzőit. De amit mint *általános* jellemzőt megragad és felhasznál, az éppen nem a cél-objektum szükségessége, hanem általában a *célrairányultság* mint sajátos organizációs elv, amelyet a negatív visszacsatolás elve fejez ki. Ahhoz, hogy a célra-irányultságot mesterséges szervomechanizmusokon modellezni lehessen, a célt objektívnálni kellett, mégpedig a működést reguláló referenciaszint formájában. A Wienerék által felhozott konkrét példákban ezeknek az irányító jeleknek különböző objektumok voltak a forrásai, a hordozói. Az előbbieket alapján azonban világos, hogy ebben éppen a célrairányultság *speciális* megjelenítésének hogyanja fejeződik ki, az a speciális út, amelyen keresztül az ember „belső” céljából kiindulva el lehetett jutni élettelen szerkezetek „belső” célszerűségéhez, rájuk ruházva a célmegvalósítás bizonyos funkcióit. Ezen az alapon nem elvi ellentmondás és nem értelmezési nehézség, ha „tagadhatatlanul” célszerű viselkedések esetében az irányító célt objektum-formában nem találjuk. Ez csak azt bizonyítja, hogy a kibernetika célszerűség-elvét kell az emberi célszerű tevékenységből kiindulva magyarázni, nem pedig megfordítva. Ez az előbbi speciális érvényességét nem szünteti meg, sőt éppen nem zárja ki azt sem, hogy „a” célszerűség (elsősorban az emberi célszerűség) bizonyos jellemzőit segíthessen alaposabban megérteni.

A másik félreértés szorosan összefügg ezzel. R. Taylor szerint a gépi teleologizmus megoldhatatlan értelmezési nehézségekhez vezető alapvető hibaforrása az, hogy a célszerűséget viselkedés-kategóriának tekinti. A viselkedésből ugyanis, írja, korántsem lehet mindig egyértelműen következtetni a célra. (Ha pl. egy autó követ egy gyalogost, sőt egyre közeledik hozzá, esetleg el is üti, a „viselkedéséből” egyáltalán nem derül az ki, vajon a vezetője célja az volt-e, hogy elgázolja az embert, vagy csak ijesztgetni akarta, vagy esetleg egyszerűen csak elromlott a kocsí fékje (s így a viselkedés az adott vonatkozásban nem is volt célszerű). Ebből viszont az következik, hogy a célszerűség nem szűkíthető le a viselkedés ismérveire, hanem szélesebb fogalom: a viselkedést megelőzően a jövő bizonyos állapotaira vonatkozó kívánságot és véleményt is feltételez.¹¹ — A tudatos emberi tevékenység célszerűségére vonatkozóan Taylornak nyilvánvalóan igaza van. De ha ezt az igazságot egyszerűen szembe-állítjuk a kibernetika belső célszerűségével, ez azt jelenti, hogy egy tudományág speciális *módszertani* célszerűség-elvének lehetséges tartalmát nem is vizsgáljuk, az ontológiai alapját képező összefüggéseket nem tisztázzuk, s így, a jelen esetben kétségtől vitatható megoldási javaslatokkal együtt ab ovo a problémákat is elvetjük. Valójában, mint láttuk, a szóban forgó belső célszerűség mint speciális metodológiai elv viselkedés-kategória, de anélkül, hogy mint ilyen általánosítható lenne, anélkül, hogy *általában* igaz lenne Wieneréknek

¹¹ R. Taylor: *Purposeful and ...* 328. o.

bármely efajta kitétele: „a célszerűséget felismerni a cselekvés természetéből kell, nem a cselekvő objektum struktúrájának és természetének vizsgálatából és a rajta való spekulációból”.¹²

Úgy gondoljuk, az eddigiekben felvázolt problémák filozófiai szempontból a kibernetikai belső célszerűség gerincét érintik. Ez pontosabban a következőket jelenti: 1. ha azok a válaszok, amelyeket egyrészt az emberi tevékenység és a gépi célszerűség ontológikus összefüggésére vonatkozóan, másrészt ennek alapján a metodológiai teleológia-elv filozófiai tartalmára vonatkozóan adni próbáltunk, alapjában véve helyesek, a viszonylag egyszerű berendezések elemzésén túl a bonyolultabb kibernetikus gépek működésére vonatkozóan is *lényegében* érvényeseknek kell lenniök; 2. ez nem zárja ki azt, hogy a kibernetika fejlődése a célszerűség-kérdésre vonatkozóan is ne gazdagítaná tudásunkat, s ne változtatná, mélyítené saját metodológiai elvének tartalmát is.

A negatív visszacsatolás mint műszaki elv köztudomásúan nem a kibernetika felfedezése. J. Watt golyósszabályozója gyakorlatilag már 1876-ban megvalósította. Mint fiziológiai jelenséget, mint az organizmus és a környezet kapcsolatának ciklikus organizációját szintén nem a kibernetika ismerte fel elsőnek. De az a belátás, hogy itt a viselkedés szempontjából jelentősen közös vonásról van szó élő és élettelen között, a kibernetika megalapítóinak érdeme. Jelentősége pedig abban van, hogy az élő organizmus célszerű viselkedésének valamiféle belső vázát adja meg, s így perspektívát nyújt egyes eddig csak az élő anyag speciális funkcióinak tartott viselkedésformák gépi instrumentálására, elvileg új, az eddigieknél összehasonlíthatatlanul komplexebb feladatok automatizálására. Minél újszerűbbek és minél komplexebbek az automatizálendő feladatok, természetesen annál kevésbé jelentik a visszacsatolás elvének műszakilag viszonylag egyszerű, közvetlen demonstrálását. Az ezen túlmenő feladatok pedig igen sokszor elvileg jelentős és nehéz problémák megoldását igénylik, amihez a kibernetikának pl. az információelmélet, vagy a matematikai logika eredményeit is messzemenően be kell vonnia érdeklődési körébe. Világos, hogy ennek megfelelően a célszerű organizáció megteremtése is egyre bonyolultabb és nehezebb dolog. A referencijel — a „belső” cél produkálása —, a szabályozandó változók optimális értékeinek előállítása a szabályozandók számától és egymáshoz való viszonyától függően igen bonyolult számítási műveleteket igényelhet (amelyek esetleg csak elektronikus számológépek mint vezérlőberendezések segítségével oldhatók meg), a gépet a feladat egész megoldási menetét programozó, esetleg több ezer utasítással kell ellátni stb. A kibernetikai célszerűség *elvi tartalmát* azonban mindez nem érinti. Az automata célszerűsége ezekben az esetekben is ugyanúgy illeszkedik be az emberi teleologikus tevékenységfolyamatba, és ugyanolyan jellegű, a negatív visszacsatolással jellemezhető metodológiai önállósággal rendelkezik, mint az egyszerűbb esetekben.

Van azonban a dolognak egy másik oldala is. Ez pedig a következő: az önszabályozó berendezések célkereső működésükkel a teleologikus viselkedés *bizonyos funkcióit* modellezik. Az előbbiekben ez azt jelentette, hogy mivel a gép reálisan nem tűz ki célokat, hanem végeredményben csak az emberi célok megvalósítási folyamatába kapcsolódik be, a belső célszerűség fogalmát csak úgy érthetjük meg, ha a célt a gép működése számára valamilyen adekvát formában eleve adottnak vesszük. Az egyszerű negatív visszacsatolással működő

¹² Rosenblueth — Wiener: *Purposeful and . . .* 323. o.

automaták esetében viszonylag könnyű kimutatni, hogy ez valóban így van. Ezek viselkedése azonban ennek megfelelően viszonylag merev. Működésüket az ember által produkált, többé-kevésbé „kész cél” irányítja és „véletlen” körülményekhez, megváltozott külső vagy belső működési feltételekhez nem tudnak alkalmazkodni. Mármint nemcsak az ember, hanem általában az élő célszerű viselkedését éppen az jellemzi, hogy adaptálódik a környezet változásaihoz. Felmerül tehát a kérdés, vajon nem lehet-e plasztikusabb, „értelmesebb”, „tanulékony” gépekkel a viselkedésnek ezt a magasabb típusát is modellezni.

Nem nehéz meglátni, hogy a kibernetika ebből a szempontból vett fejlődése új kérdéseket vet fel a célszerűség-problémával kapcsolatban. Hiszen nem arról van-e szó, hogy amennyiben az új feladatokat egyre tökéletesebben sikerül megvalósítani, ez szükségessé teszi a belső célszerűség *elvéleg új értelmezését*? A plasztikusabb viselkedés instrumentálására vonatkozó kutatások, a gyakorlati eredmények már ma sem terméketlenek. Az *elméleti lehetőség* problematikája pedig már jónéhány évvel ezelőtt felvetődött (pl. W. R. Ashby érdekes könyvében, amelynek első kiadása 1952-ben jelent meg).¹³ Mindez azonban valójában azt bizonyítja, hogy az iménti kérdésre nemmel kell felelni.

Ashby abból indul ki, hogy az idegrendszer adaptív viselkedésének ténye és az a hipotézis, hogy működése a lényegét tekintve mechanisztikus (azaz nem teleologikus), nem kibékíthetetlen egymással. Tehát, mondja, az agyműködés „logikai mechanizmusát” kidolgozva bizonyíthatjuk, hogy az élő szervezet viselkedése lehet egyszerre szelektív és mechanisztikus, következésképpen adaptív voltában is hozzáférhető a gépi modellezés számára. Mindabból, amit erre épít, a mi szempontunkból két dolgot kell kiemelni:

1. mivel meg akarja magyarázni, *hogyan jön létre* az a viselkedés, amely de facto célra irányult — mondja —, nem indulhat ki a cselekedet hasznosságának teleologikus fogalmából, mert így bűvös körbe jutna. Abból a feltevésből kell kiindulnia, hogy a viselkedés minden pillanatban a viselkedő entitás fizikai természete által meghatározott, nem pedig finálisan. E feltevéssel összhangban szigorúan „objektív módszerrel” tanulmányozza az adaptív viselkedést. Ehhez szorosan hozzátartozik, hogy csak olyan fogalmakat használ, amelyeknek objektíve megragadható értelmük, objektíve demonstrálható tartalmuk van. De megvalósítható-e ez minden esetben? Ashby elismeri, hogy az *emberi tevékenységet* tekintve az ilyen törekvés hallatlan gyakorlati és filozófiai nehézségekbe ütköznék, az állati és a gépi viselkedésre nézve azonban megvalósítható. Ami számunkra ebből fontos, az az emberi tevékenységre vonatkozó, csöppet sem lényegtelen korlátozás. Az tudniillik, hogy a tanulás, az alkalmazkodó viselkedés automatizálása az elvi lehetőségeit tekintve éppúgy *nem* az emberi célszerűség maradéktalan lefordítása a gép nyelvére, legfeljebb bizonyos vonásainak modellezése, mint ahogy a célkeresés egyszerű aktusának automatizálása sem volt az;

2. arra a kérdésre, hogy mi az adaptív viselkedés ilyen módszerekkel megragadható „logikai mechanizmusának” lényege, Ashby a *kettős visszacsatolás* elvével felel. A következőképpen: az organizmus fennmaradása belső reakcióinak kiegyensúlyozott voltát, működésének *stabilitását* feltételezi. Ez a stabilitás egyet jelent azzal, hogy az organizmus számos lényeges jellemzőjé-

¹³ R. Ashby: *Design for a Brain. The Origin of Adaptive Behavior*. London 1960.

nek (funkcionális szempontból: változóinak) kvantitatív értéke bizonyos határokron belül marad. A környezet változásai azonban hatnak a lényeges változókra is, kritikus értékeiket ezért a kedvezőtlen hatásokkal szemben *fenn kell tartani*. Az egyéni viselkedés célszerűsége éppen az e hatások kiegyensúlyozását eredményező alkalmazkodást jelenti. Az állat azonban nyilvánvalóan nem rendelkezhet a lehetséges összes változáskombinációkra megfelelően reagálni kész öröklött mechanizmusokkal, a célszerű viselkedést tehát meg kell tanulnia. Ezt biztosítja a kettős visszacsatolással jellemezhető organizáció. Ez lényegében azt jelenti, hogy a környezet nemcsak az organizmus külső viselkedését kiváltó „jelzéseket” küld ez utóbbi érzékszervein keresztül, hanem közvetlenül is hat a szervezet lényeges változóira, amelyeknek a „normálist” meghaladó értékváltozása viszont a rendszer bizonyos működési feltételeinek (meghatározott külső paramétereknek) megváltoztatásán át modifikálja a viselkedést. (Ashby a macska és a tűz viszonyával szemlélteti ezt a kapcsolattípust. A macska, ha nincsenek még tapasztalatai a tűzzel kapcsolatban, pl. meglátva a lángokat, egészen közel mehet hozzájuk, körüljárhatja, megszagolni igyekezik őket stb., azaz a legkülönbébb viselkedéssel reagálhat a látványra. Így azonban adott esetben megéghet, rágurulhat egy parázs, ami azt jelenti, hogy a szervezete működési egyensúlyát biztosító bizonyos belső viszonyok, lényeges változók értéke az egyensúlyt megbontó mértékben megváltozik. Ennek hatására meg kell változnia bizonyos feltételeknek, pl. a parázs vagy a lángok és a macska közötti térbeli távolságnak, ami ismét meghatározott hatással van a macska viselkedésére.) A viselkedés adekvát voltáról akkor beszélhetünk, ha ez a rendszer (a szervezet és a környezet) egyensúlyi állapotát biztosítja. Ehhez viszont a rendszer minden részének, tehát az organizmusnak a dinamikus stabilitása is szükséges. Ezért mindaddig, amíg az egyensúlyi állapot ki nem alakul (illetve helyre nem áll), amíg tehát az organizmus meg nem találja a változó feltételeknek megfelelő, a lényeges változók megfelelően stabil értékeit biztosító viselkedésformát, az egyes láncszemek automatikusan modifikálják egymást a fenti organizációnak megfelelő rendben, azaz mindenfajta céltudat nélkül a rendszer állandó mozgásban van.

Arra a kérdésre, amit az imént feltettünk (hogy ti. adaptív viselkedésre képes automaták építése nem teszi-e szükségessé a kibernetikai belső célszerűség fogalmának elvi átértékelését), úgy gondoljuk, az állati tanulás „logikai mechanizmusának” ez a felfogása világosan nemet mond. Ennek alapján tanuló gépek létrehozása azt jelenti, hogy mindenekelőtt a fentivel analóg dinamikus egyensúlynak, egy „ultrastabil” rendszernek a feltételeit kell élettelen szubsztrátumon instrumentálni, ehhez viszont itt is elsősorban a szervezet lényeges belső viszonyait modellező határértékeket, a szabályozás alapját képező belső célt kell „előállítani”. A belső célszerűség értelme sem e cél kibernetikai tartalmát és funkcióját tekintve, sem a belőle kiinduló organizáció *lényegét* tekintve nem fog megváltozni. Ez azonban nem jelenti azt, hogy az organizáció itt a metodológiai célszerűség tartalmára vonatkoztatva csak jelentéktelen mértékben változik. A visszacsatolás kettős hurka ebből a szempontból jelzése a tartalmi változásnak. A gépnek most éppúgy önmagának kell a célhoz vezető viselkedést (tehát a működés helyes programját) kidolgoznia, mint ahogyan az állatnak is meg kell tanulnia, hogyan „fordítsa hasznára” a tüzet. A kettős visszacsatolással jellemzett organizáció éppen azért lehet az állat és a gép számára közös tanulás-mechanizmus alapja, mert lehetővé teszi nem-célszerű, „véletlen” viselkedés-aktusok felhalmozódásából végül is egy adaptív viselkedésmód

kibontakozását. Ennek a hogyanja valóban új mozzanat az immanens célszerűség tartalmában.

Az állat nem látja előre viselkedése eredményét, a környezet várható visszahatását saját aktivitására, ezért, ha valamely cselekvése inadekvátnak bizonyul (a lényeges változóknak a tűrési határokat meghaladó értékváltozáshoz vezet), nem tudja, milyen irányban kell megváltoztatnia magatartását. A környezettől információkat csak saját cselekedetei eredményében kaphat — mondja Ashby. Minden sikertelen kísérlet után vaktában változtatja viselkedését, próbálkozik, mindaddig, amíg szervezete működési egyensúlya helyre nem áll. A célszerűtlen viselkedésmódokat a kettős visszacsatolással „dolgozó” organizáció utólag kiszűri, az eredményt minden esetben mintegy lemérve a lényeges változók értékalakulásaiban. Ilyenformán Ashby szerint a próbaszerencse módszer az állat adaptív viselkedésének szükséges eleme, a cél eléréséhez szükséges információk megszerzésének objektív módszerekkel tanulmányozható útja. Mint ilyen, közös eleme lehet az élő organizmus és a gép viselkedésének. — Kétségtelen, hogy a tanulógépek bizonyos típusai valóban igyekeznek ezt az elvet felhasználni. Az egyik leglényegesebb új mozzanat működésükben a „próbálkozás”, a „tapasztalatgyűjtés” és a „szelekció” alkalmazása — lényegében az Ashby által megrajzolt mechanizmus logikája szerint.

Tulajdonképpen ennyi az, amit el akartunk mondani. Mégsem tehetünk itt pontot, ha el akarjuk kerülni azt a hamis látszatot, hogy szerintünk a filozófia ezeknek a problémásikoknak az elemzésével „lényegében” megoldottnak nyilváníthatja a kibernetikai cél-elv értelmezését, legalábbis azt, ami ebből órát tartozik. Szó sincs róla, a problémahorizont még csak most tágul igazán. Ezért választjuk befejezés helyett bizonyos további problémák jelzését.

Az imént azt mondtuk, hogy az adaptív viselkedésre képes automaták szintje valami jelentősen újat hoz a metodológiai célszerűség tartalmában. Mit? Röviden úgy fogalmazhatnánk meg, hogy a céltevékenységnek egy újabb jellemző vonását kíséri meg gépi funkcióvá tenni: a célnak azt a sajátosságát, hogy mivel csupán többé-kevésbé általános, elvont „tükrözése”, ideális képe a kívánt eredménynek, önmagában nem határozza meg mereven és egyértelműen az utat, amelyen elérhető, annál is inkább, hiszen az út mindig változó környezeti feltételekkel való kölcsönhatást feltételez, s így a viselkedő entitás szelektív magatartását követeli meg.

Ezek után azonban máris tökéletesen érthető, ha egy újabb, filozófiailag is izgalmas kérdés merül fel: az „egyszerűbb” élő organizmus, a növény és az állat célszerű élettevékenysége és az alkalmazkodásra is kiterjesztett gépi célszerűség viszonyának kérdése. Bármennyire kezdetiek is ma még az eredmények, nem jelzik-e éppen azt, hogy az aktív alkalmazkodás funkcionális megragadásával az élet lényeges vonásait „sajátítja ki” a kibernetika, s ezen az úton továbbhaladva egyre teljesebben (sőt végül teljesen) sikerül a maga eszközeivel produkálnia azt, amit életnek nevezünk.¹⁴ Azaz a mi szempontunkból: ha alapvetőnek kell is tekintenünk a kibernetikai célszerűség tisztázásában ennek a céltudatos emberi tevékenységhez való viszonyát (és a kettő elvi különbségét), ezen belül esetleg még jogos lehet feltenni a kérdést, van-e egyáltalán bármiféle elvi különbség a kibernetikai és az ún. szerves célszerűség között.

Semmiképpen sem intézhetjük el könnyedén ezt a kérdést. És nemcsak azért nem, mert neves tudósok meggondolandó érvekkel alátámasztott fel-

¹⁴ A. H. Колмогоров: *Жизнь и мышление с точки зрения кибернетики*. Москва 1962.

tevése, hogy esetleg — nincs. Nem is csak azért nem, mert a kibernetika már eddig is tagadhatatlanul viharos lendülettel söpört el nem egy eleve eléje állított elvi korlátot, amelyen túl — állították — nem mehet. (Ami persze nem azt jelenti, hogy megfordítani a dolgot, és eleve korlátlanak nyilvánítani a kibernetika lehetőségeit, kevésbé komolytalan „segítség” lenne, mint az előbbi.) Hanem azért sem, mert magától értetődő, hogy ebben a kérdésben a szerves célszerűség „meghallgatása” nélkül még hipotéziseket felállítani sem lehet. A szerves célszerűség „belső célja” is metodológiai elv, immanens, anélkül, hogy tudatos cél lenne. Ha viszont ennek alapján a szerves célszerűséget filozófiailag úgy értelmezzük, hogy ez csupán formális, szerkezeti analógia az emberi teleológiával, akkor tulajdonképpen nem igen lehet elvi különbséget kimutatni a kibernetika belső célszerűsége és a szerves teleológia között. De nyilván itt is bonyolultabb a dolog, és a kérdés éppen az, hogy mennyivel több a szerves célszerűség formális analogikus jegyeknél, mi a szerves teleológia elvének *speciális tartalma*. Ezen a szinten a kibernetika belső célproblémájához a filozófia nem igen nyúlhat anélkül, hogy ezzel a másik cél-elvvel is ne foglalkoznék alaposan.

ON THE PHILOSOPHICAL CONTENT OF THE CONCEPT OF “INNER PURPOSE” IN CYBERNETICS

E. Rózsahegyi

Starting from N. Wiener's and A. Rosenbluth's conception of teleology and its contemporary and later philosophical reception, the paper examines the philosophical content of the cybernetic “inner purpose”. Stressing the ontological inseparability of mechanical functionalism and purposeful human activity, it considers within the frame of this relationship the plane of so called “methodological functionalism” a relatively independent problem. The special content of the cybernetic principle of methodological functionalism is outlined by the interpretation of the concept of “inner purpose”.

О ФИЛОСОФСКОМ СОДЕРЖАНИИ ПОНЯТИЯ КИБЕРНЕТИЧЕСКОЙ «ВНУТРЕННЕЙ ЦЕЛИ»

Э. Розахеда

Отправляясь от телеологического взгляда Н. Винера и А. Розенблюта, а также современных им и позднейших философских откликов, данная статья исследует философское содержание кибернетической «внутренней цели». Подчеркивая онтологическую неотрывность машинной целесообразности от целенаправленной сознательной человеческой деятельности, в пределах их отношений автор выделяет как относительно самостоятельную проблему плоскость т. н. методологической целесообразности. Специальное содержание методологического принципа целесообразности кибернетики автор старается определить при помощи толкования понятия «внутренней цели».