BÁTFAI NORBERT

A játékok és a mesterséges intelligencia mint a kultúra jövője

A cikk célja a mesterséges intelligencia kutatásokat az emberi önmegismerés szolgálatába állítani. Ehhez egyrészt filozófiai hátteret biztosítani, másrészt a mesterséges intelligencia társadalmi elfogadottságát megalapozni. Tézisünk, hogy az emberi kultúra fenntartásához és fejlesztéséhez a játékokon és a mesterséges intelligencián keresztül vezet az út. E tézis alátámasztásnak támogatására kísérletet teszünk a szubjektivitás elméletének megalapozására.

Kulcsszavak: *mesterséges intelligencia, bonyolultság, entrópia, mém, számítógépes játékok, e-sport*

Szerzői információ:

Bátfai Norbert, PhD 1972-ben Salgótarjánban született. Kitüntetéses programtervező matematikus oklevelét a Kossuth Lajos Tudományegyetemen 1998-ban szerezte. 1999-ben megnyerte a Java szövetség (Sun, IBM, Oracle, Novell, IQSoft) Java programozási versenyét. Mobil információtechnológiai cége megnyerte 2004-ben a Sun és a Nokia Magyarország mobil Java programozási versenyét. 2008-ban megkapta a Vezető Informatikusok Szövetsége Év Informatika Oktatója címét. 2011-ben szerzett doktori fokozatot informatikából a Debreceni Egyetemen. 2012-ben megkapta a Hírközlési és Informatikai Tudományos Egyesület Pollák-Virág Díját. Kutatási területei: a játékfejlesztés és a robotpszichológia. A Debreceni Egyetem Informatikai Kara Információ Technológiai Tanszékének adjunktusa. A DEAC-Hackers e-sport szakosztály kutatási vezetője.

Így hivatkozzon erre a cikkre:
Bátfai Norbert, „A játékok és a mesterséges intelligencia mint a kultúra jövője”. *Információs Társadalom* XVIII, 2. szám (2018): 28–40.
https://dx.doi.org/10.22503/inftars.XVIII.2018.2.2

Bátfai Norbert

A játékok és a mesterséges intelligencia mint a kultúra jövője – egy kísérlet a szubjektivitás elméletének kialakítására

Bevezetés

A közelmúlt mesterséges intelligencia kutatásának áttörő eredményei nem csupán ígérték az emberi szintű gépi intelligenciát, hanem adott esetekben – ilyen például a számítógépes játékokkal (Mnih et al. 2013, Mnih et al. 2015) vagy a Góval (Silver et al. 2016) való játék – el is hozták azt: a programok tudása immár összemérhető az emberekével vagy meg is haladhatja azt. Ez az oka, amiért érdekes a tudatosságot programozóként vizsgálni. Miközben nem a ma már lehetségesnek látszó programozó ágensek[1] (Reed és de Freitas 2015, Bhupatiraju et al. 2017, Denil et al. 2017) témakörrel foglalkozunk, hanem magának az embernek a behatóbb megismerésével. Kérdésünk, hogy egyáltalán miért szeretünk játszani és miért van szükségünk mesterséges intelligenciára? A tapasztalat azt mutatja, hogy az általános mesterséges intelligencia kora közeleg, a játékokkal szimbiózisban közeleg. Vajon miért alakul így?

A programozás

A programozás az a folyamat, amely során a programozó előállítja a számítógépen futni képes kódot. Ez a kód lehet közelebb (gépi kód) vagy távolabb (magasabb szintű programozási nyelv) egy adott gép saját nyelvéhez, ami számok szisztematikus sorozata. Ezek a mesterséges nyelvek mérnöki alkotások, elméleti modelljük és vizsgálatuk alapja a Turing-gép. A Turing-féle (matematikai) „gép" egy betűket tartalmazó szalag, betűkkel jelzett állapotok és olyan szabályok összessége, amelyek megmondják, hogy abban az esetben, ha a gép beolvas egy adott betűt a szalagról egy adott állapotban, akkor mely betűt írja vissza, mely betűvel jelölje a következő belső állapotát, és melyik irányba lépjen a szalagján. Egy ilyen szabályt tekintünk a Turing-féle gondolkodás egy absztrakt alaplépésének, amelyről azóta úgy véljük – erről szól a Church-tézis –, hogy az emberi algoritmikus gondolkodás egyik elképzelhető alaplépésének is tekinthető. Neumann János már ismerte és alkalmazta (Neumann 1951: 198) Turing eredményeit az univerzális gépről (Turing 1937) nevezetesen, hogy létezik olyan Turing-gép, amely bármely más Turing-gép és inputjának működését képes leutánozni, azaz szimulálni. Önreprodukáló automatája megalkotásakor hivatkozik is erre, de a későbbiekben Neumannt inkább McCulloch és Pitts neurális jellegű munkái (McCulloch és Pitts 1943) ihlették meg, amely gyümölcsét a megbízhatatlan alkatrészekből épített megbízható számításokat végezni képes hálózat megalkotása (Neumann 1956) jelenti, amelyben egyértelmű az analógia a természetes számító egység (agy) és az akkori elektroncsöves mesterséges technológia között. A mesterséges neurális hálózatok viszont nem ebbe az irányba fejlődtek, sőt Minsky és Papert félreértett kritikája (Minsky és Papert 1969) miatt sokáig egyáltalán nem is fejlődtek. Napjaink új lendületét a többrétegű és „mély" neurális hálózatok adják, ezek állnak az első bekezdésben említett Google DeepMind[2] munkák sikerei mögött is.

[1] Értsd: majdani robot programozók (programozó programok).
[2] https://deepmind.com/

Az élet

Neumann saját munkája akkor még nem ért össze Schrödingernek az életről alkotott entrópia alapú (Schrödinger 1944) tézisével, miszerint az élő rendszerek szemben a termodinamika második főtételével, képesek belső entrópiájuk nem növelésére. Később már igen: vannak olyan tudósok, akiket már e kettő (Neumann önreprodukáló automatája és Schrödinger életről alkotott elképzelése) közösen ihletett meg.[3]

Az élet és a programozás

A növények és az állatok viselkedését célszerűnek és adott cél elérése érdekében tett tevékenységnek tekintjük. Tesszük ezt annak ellenére, hogy számos olyan kísérletet ismerünk, amely kimutatja, hogy ez a célszerűség nem belátáson alapul, hanem szimpla (előre programozott vagy előre programozottan tanult) algoritmikus alapú viselkedés. Például Tinbergen (Tinbergen 1976) kísérletei azt mutatják, hogy a tüskés pikó (Tinbergen 1952) vagy az ásódarázs (Tinbergen 1984) nem ismeri fel (nincs tudatában), hogy tevékenysége nem célszerű, az adott cél elérése szempontjából hasztalan.[4] Mennyire bonyolult az ilyen viselkedés? Segíthet erről intuitív képet alkotni, ha tanulmányozzuk például a kopoltyús vízicsiga védekező reflexét vezérlő habituációs-szenzitizációs tanulási folyamatának biokémiai leírását (Klix 1985: 71–76), miközben a bonyolultságot a leírás bonyolultságával azonosítjuk. Mi, emberek úgy véljük magunkról, hogy célszerű viselkedésünk belátáson alapul, jelentsen ez bármit: „értjük a dolgokat", működik bennünk a donaldi homunkulusz[5] (Donald 2001: 316). Mi nem pusztán előre programozott vagy tanult komplex tevékenység sorozatokat hajtunk végre, de szabad akaratunk létezése legalábbis vita tárgyát képezheti.

Nyilvánvalóan az emberi test működésének is léteznek adott célt szolgáló működési mechanizmusai, például az autonóm idegrendszer. Vajon az autonóm vagy a szomatikus vezérlés a komplexebb, „bonyolultabb"?

Bonyolultság

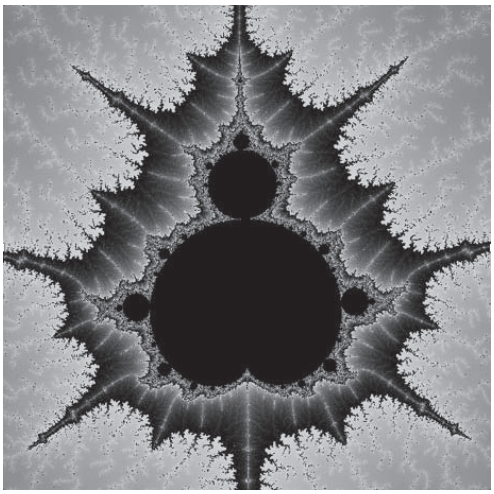
Neumann (Neumann 1951) munkájának zárógondolata a bonyolultság elméletének kifejlesztéséért kiáltott. Hite szerint adott bonyolultsági szint fölött adott kreatúra önmagával megegyező vagy még bonyolultabb kreatúrákat tud létrehozni. Ugyanebben a dolgozatában korábban (Neumann 1951: 200–201 oldal átmenete) előzetesen meglepőnek tartotta Turingnak az univerzális gépről szóló tételét, miszerint egy adott Turing-gép bármely más Turing-gépet képes leszimulálni, legyen az „dupla olyan nagy vagy bonyolult" – fogalmaz Neumann – mint a szimulált gép. Nekünk már nincs ilyen előzetes megérzésünk, hiszen

[3] Ilyen például a genetikai kódot „jegyző", Nobel-díjas Sydney Brenner (lásd például a http://scarc.library.oregonstate.edu/coll/pauling/dna/quotes/john_von_neumann.html vagy a https://www.cell.com/current-biology/pdf/0960-9822(93)90337-N.pdf lapot).

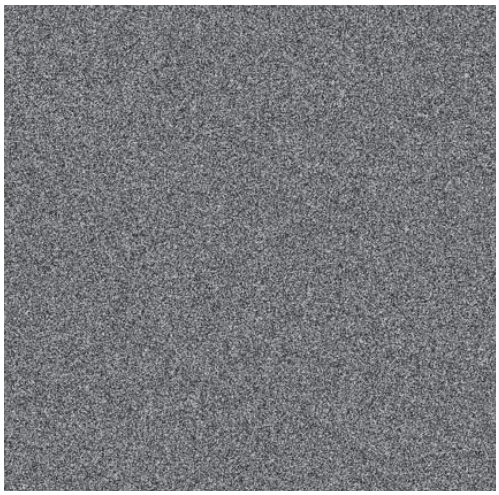
[4] Amikor a tüskéspikó a vörösre festett hasú fadarab modellel harcol, vagy az ásódarázs a korábban a fészkét jelölő, éppen elmozgatott tobozok között keresi a bejáratot.

[5] Hitem szerint szép lenne, ha az embert megkülönböztető donaldi „molekuláris újítás" (Donald 2001: 21) mondjuk, Penrose–Hameroff OrchOR tudatmodellje (Hameroff és Penrose 1996), azaz egy direktben kvantum-mechanikai hátterű folyamat lenne.

egyrészt azon nőttünk fel, hogy gépünk alapszoftvere (korábban egy interpreter, aztán az operációs rendszer) bármilyen más programot tud futtatni, illetve a hatvanas évek elején megjelent a Solomonoff–Kolmogorov–Chaitin-bonyolultság (Solomonoff 1960, Kolmogorov 1968, 1998, Chaitin 1969). Ez azt mondja ki, hogy valaminek a bonyolultsága a valamit kiszámoló program és e program inputjának az együttes hossza. Neumann előzetes aggályát ez azzal oszlatja el, hogy amikor az univerzális gép megkapja a nála „kétszer bonyolultabb" bemenetet szimulálás céljából, akkor az ő bonyolultságának kalkulációjába az inputként kapott „kétszer bonyolultabb" gép leírását is számításba kell venni. Így a Solomonoff–Kolmogorov–Chaitin-bonyolultsággal a Neumann áhított bonyolultsági elmélete vélhetően a kezünkbe került.



1. ábra: Egy Mandelbrot-halmaz nagyítás a komplex sík téglája fölött (saját programmal)



2. ábra: Az iménti Mandelbrot-halmaz színértékeinek egy pszeudó- (szoftverből generált) véletlen sorozata (saját programmal)

Tekintsünk meg egy Mandelbrot-halmaz (Mandelbrot 1983) képét (1. ábra)! Ezt „bonyolult" (és nem véletlen[6]) képződménynek tartjuk, pedig a „kódja" csupán egyetlen iterációs képlet. Tehát a Solomonoff–Kolmogorov–Chaitin-komplexitás fogalma szerint nem bonyolult.

Tekintsünk meg egy számítógéppel generált véletlen számsorozatot (2. ábra)! Ezek ránézésre valóban véletlennek látszanak – még statisztikai tesztekkel vizsgálva is –, amelyen nem lepődhetünk meg, hiszen eleve ez volt előállításuk kritériuma. Valójában persze nem véletlenek, hiszen erre a célra kifejlesztett véletlenszám-generáló algoritmusokkal vannak generálva. Ezért a Solomonoff–Kolmogorov–Chaitin-megközelítés szerint ez sem bonyolult.

[6] Ebben a példában a bonyolult és a véletlen ellentétesnek érzett fogalmak, ezzel szemben a Solomonoff–Kolmogorov–Chaitin-bonyolultság szerint viszont közeli rokonok: a véletlenek a legbonyolultabbak.

[7] Ez a LEGO-hoz hasonló építőjáték, de sokkal kisebb az elemek mérete és az építőelemek homogénabbak.

Első példánk tehát bonyolultnak tetsző, de valójában nem az, második pedig véletlennek tetsző, de valójában nem az.

Vegyünk alapul egy csomag Barbár királyos MICROBlocks építő játékot[7], amelyet éppen kiöntünk az asztalra! Ezt nem összetettnek vagy bonyolultnak, hanem véletlenszerűnek látjuk, ha pedig megépítjük belőle a Barbár királyt (3. ábra), azt rendezettnek és komplexnek látjuk. A Solomonoff–Kolmogorov–Chaitin-bonyolultság és az entrópia itt is korrigálhatja az intuíciónkat, miszerint a felépítettet tekintenénk komplexnek. Tehát ezzel szemben azt mondhatjuk: a kiöntött kevésbé rendezett és bonyolultabb, a felépített rendezettebb és kevésbé bonyolult.[8] De van egy másik szint is, ahol intuíciónk megcsalhat, hiszen ha – Schrödinger (1970: 253) észrevételének mintájára – a Barbár király felépített és a kiöntött alkatrészei is egy-egy konkrét elrendezést alkotnak, akkor miért rendezettebb az egyik a másiknál? Mert a Barbár királyban vannak kisebb „szabályos" részek, például a szeme, a kardja? Emlékezzünk a véletlen sorozatra, ott nem látszódnak ilyen szabályosságok, pedig mégis „szabályos". Az entrópia oldaláról fognánk meg, mely szerint sokkal több „rendezetlen" elrendezés van, mint „rendezett"? Ez már felhasználná a „rendezettség" fogalmát… Nézzük meg a Barbár király építési előírását! Ez minden tégla helyét és pozícióját rögzíti. Ám a kiöntött állapotot is ugyanúgy tudjuk leírni: minden tégla helyének és pozíciójának megadásával. A Solomonoff–Kolmogorov–Chaitin-megközelítés azért mondja egyszerűbbnek a téglák Barbár király elrendezését, mert annak leírását lehet tömöríteni, a véletlen elrendezés leírásán meg nem lehet.

3. ábra: A felépített Barbár király[10] (saját fotó)

Ezt a tömöríthetőséget jobban kidomborítja, ha a végtelen határátmenetben vizsgálódunk tovább.[9] Képzeljünk el egy olyan futószalagot, amelyen felépítendő Barbár királyok követik egymást! Mennyi építési előírásra van szükségünk az első 'n' darab megépítéséhez? Mivel minden építendő tétel ugyanaz, így egyetlenre. Viszont ha a Barbár király helyett minden tétel esetén egy adott véletlen elrendezést szeretnénk, akkor a szükséges előírások száma a tételek számával egyezik meg. Határátmenetben, ha a futószalag tart a végtelenbe, a Barbár királyos esetben az egyetlen leírás jelentősége elfogy, a véletlen tételek esetén viszont állandóan megmarad.

A tételenkénti Barbár király és véletlen elrendezések helyett gondoljunk most egy tartályban mozgó sokrészecskés rendszerre! Egy jövőbeli időpont leírását a Barbár király építési előírásához hasonlóval adhatjuk meg, amelyben a részecskék helyét és mozgását kell jellemeznünk. Ha két állapot között van törvény, amely előírja, hogy hol lesz a részecske, akkor az 'n' állapot megadásához elég a kezdeti állapot, a törvények leírásának és

[8] Összhangban azzal, hogy a bonyolultság és az entrópia között erős kapcsolat van, lásd Lovász (1994).
[9] A (Rónyai, Iványos és Szabó 2008: 237) véletlensorozat-definíciójának mintájára.
[10] Köszönet kisfiamnak Bátfai Nándor Benjáminnak, hogy felépítette és fotózásra a rendelkezésemre bocsájtotta.
[11] Például n=1000 esetén 4 jegy kell az 'n' leírásához, ami ránézésre is látszik, de általánosan úgy fogalmaznánk, hogy 4=log(1000)+1.

az 'n' értékének a megadása. Határátmenetben ez ugyanúgy „elfogy", mint az iménti Barbár királyos sorozatra számolt érték, hiszen a kezdeti állapot és a törvények leírása konstans hosszú, az 'n' leírásának hossza pedig az 'n' logaritmusa[11], amely nem nő egyformán az 'n' értékével.

Szubjektivitás

Amikor hitet teszünk az objektív valóságban, hallgatólagosan feltételezzük annak objektív érzékelését is. Intuíciónk lecsendesítésére[12] vonjuk be az iménti példák vizsgálatába a megfigyelést. Tekintsük magunkat egy olyan mesterséges neurális hálózatnak, amelynek egy klasszifikációs (például megmondani, mit látunk) feladatot kell megoldani. Ránézek a Barbár királyra, ha még sose láttam a Barbár királyt és hasonlót sem, akkor kis túlzással bármi lehet, amit látok, azaz közel egyforma valószínűséggel kell kiválasztanom a döntésem eredményét, hogy mit is látok. Ez nagy entrópiát – és a (Lovász 1994: 69) értelmében bonyolultságot – jelent, mondhatjuk, hogy a (rám vonatkoztatott) bonyolultsága nagy. Ha viszont jól ismerem a Barbár királyt, akkor a Barbár király kiválasztására nagy valószínűséget kapok. Ez kis entrópiát (és bonyolultságot) jelent, tehát a (rám vonatkoztatott) bonyolultsága kicsi. Ez a szemlélet a bonyolultság egy lokális (vagy még inkább szubjektív[13]) megközelítése, amit szubjektív bonyolultságnak nevezünk majd a következőkben. Ez úgy válik globálissá (látszólag objektívvé), hogy mindannyiunkat hasonló mesterséges neurális hálózatnak kell tekintenünk a példában. Tekintsük megfordítva a „sose láttam a Barbár királyt és hasonlót sem" esetet! Ekkor ezt az elrendezést olyannak érezném, mint bármely más „véletlen"[14] elrendezést? Vagy nem is látnám?[15]

Képzeljünk el egy olyan gépet[16], amely bármely más gép szubjektív bonyolultságát meg tudja tanulni.[17] Mennyire lenne ez a gép bonyolult?[18] Turing univerzális tételének bizonyítása nem a neumanni bonyolultság alapú megközelítésen alapul (mondván, hogy adott bonyolultság felett létezne az univerzális gép), hanem egy direkt konstruktív bizonyítás, egy adott univerzális gép konstruálása. A bonyolultsággal úgy köthető össze, hogy

[12] Mert az eddigi példákban első benyomásunk mást súgott a bonyolultságról, egyszerűt bonyolultnak láttatott és megfordítva, hátterében a véletlen és a bonyolult különbözőségének feltételezésével.
[13] A klasszifikációs feladatban a szóba jöhető a hálózat kimenetén megjelenő valószínűség-eloszlás entrópiáját tekintjük itt a szubjektív entrópia származtatási alapjának. Megemlíthetjük, hogy ettől a megközelítéstől ugyan függetlenül, de egy korábbi munkában (Bátfai et al. 2017) a hálózat pontosságát információs pontosság néven már próbáltuk bitekben kifejezni az entrópia felhasználásával.
[14] Itt a véletlen alatt azt értjük, hogy mondjuk „kidobjuk kockával".
[15] Lásd a „What the #$*! Do We (K)now!?", http://www.imdb.com/title/tt0399877 filmben a történetet arról, hogy a karibiak eleinte nem látták Kolombusz hajóit, illetve hasonló jelenséget vélhetünk felfedezni „Westworld" című, http://www.imdb.com/title/ AI sorozatban, például az első évad hetedik epizódjában, 00:46-nál.
[16] Például egy számítógép programot vagy egy mesterséges neurális hálózatot. Vagy akár az embert?
[17] Például egy CIFAR-10-es, https://www.cs.toronto.edu/~kriz/cifar.html, tesztben ugyanazok az értékek jelennek meg idővel a „tanuló" hálózat kimenetén, mint amelyeket a „tanulandó" hálózat ad.
[18] Vélhetően – annak ellenére, hogy faji korlátaink is lehetnek a megismerési képességekben, lásd Wigner (1972) – az emberi gép van ilyen bonyolult, mert feltételezzük, hogy bármit meg tudunk tanulni.
[19] Ugyanez a gondolatmenet Neumann automatáira is igaz. Mi tudjuk érteni a kutya viselkedését, fordítva és a kutya-kutya relációban ez nem valószínű, általános értelemben biztosan nem.

a konstruált univerzális gép bonyolultsága felülről becsüli azt a bonyolultságot, amely az univerzalitáshoz szükséges (de persze nem elégséges).[19] Ez várhatóan a „szubjektív gép" esetén is így lenne: konstruálnunk kellene egy olyan univerzális jellegű gépet, amely várhatóan létezik, hiszen a megfelelő neurális háló univerzális approximáló és osztályozó (Altrichter et al. 2006), így elvben az éppen megtanulandó gép leképezését is meg tudja tanulni, ezáltal tetszőleges gépet meg tudna „tanulni". Egy ilyen szubjektivitási tétel megalkotása megnyithatná az utat a szubjektivitás elméletének kidolgozásához. Neumann automatája lemásolja magát (Neumann 1951), Turing univerzális gépe (Turing 1937) megkap egy másik gépet és szimulálja azt, egy korábbi a munkában (Bátfai 2016) hasonló Turing-féle gépet konstruáltunk, amely „figyeli" a megkapott Turing-gép működését és így „másolja le", adja vissza azt.[20] Utóbbi közelebb áll a szubjektív bonyolultságot megtanuló, lemásoló, leszimuláló fejlesztendő gépünk intuitív elképzeléséhez.

Végezzünk el egy Barbár királyos gondolatkísérletet Einstein speciális relativitáselméletével! A zacskóba most Einstein 1905-ös cikkét tegyük, de úgy, hogy a cikk szövegét betűire vágjuk! Kiöntjük a betűket, az anyanyelvi beszélő jó eséllyel rak össze belőle anyanyelvi szövegeket, Einstein, illetve az elméletet értő és az eredeti cikket olvasó tudósok jó eséllyel reprodukálni is tudnák az eredeti cikket. Tekintsük sorba rendezve a betűket! A betűk véletlenszerű sorba rendezésénél az anyanyelvi szöveg rendezettebb, mert a kódfejtő matematikus ki tudja mutatni, hogy a szöveg redundáns (természetes nyelvű, nem véletlen). De ezt azért teheti meg, mert jártas az információelméletben.[21] Az eredeti cikket visszaállítók jártasak a relativitáselméletben. Nekik az eredeti (vagy ahhoz lényegében hasonló) cikk szövege még rendezettebb, mondhatni, mert nem csak olvasni tudják, hanem értik is azt, ám az előző lábjegyzetet is figyelembe véve az „értést" definiáljuk úgy, hogy az érti, akinek kicsi szubjektív bonyolultság adódik rá. Tehát számukra a cikknek még kisebb a szubjektív bonyolultsága, mintha pusztán természetes nyelvű szövegként olvasnák. Ezek olyan elméletek, amelyeket gondolatok nagyon letisztult, igen rendezett összességének tekintünk. Megérthetném én a relativitáselméletet? Úgy ahogyan egy fizikus érti, vagy akár maga Einstein értette?[22] Mi itt az „elmélet" szerepe? Ez a kérdés lenne a „szubjektív gép" definiálásának alapja. Mert az intuitív válasz az, hogy igen, meg.

Gondoljunk arra a gépre, amely „érti"[23] az imént „barbarizált relativitáselméletet"![24] Tehát amely számára ez a cikk „input" kis szubjektív entrópiájú, rendezett. Kapja meg ugyanezt a bemenetet egy másik gép, amely számára viszont ez az input nagy szubjektív entrópiájú, rendezetlen. Meg tudná tanulni az utóbbi gép az előző gép „szubjektív entrópiáját", azaz, hogy számára ugyanolyan is kis szubjektív entrópiájú, rendezett legyen ez az input?

Kicsit konstruktívabban: legyen a szubjektív gép (aki egy „másikat másol") most egy többrétegű neurális hálózat, melynek tanító pontjai a megtanulandó gép bemenete és az

[20] Ebben a készülő kéziratban az univerzális tanuló gép megtekinthető:
https://github.com/nbatfai/AlgorithmicFractals/blob/master/manuscript/ULM.tex

[21] Szem előtt tartva a Rényitől származó (1976: 73) elvet, hogy az információelmélet nem foglalkozik az információ értelmezésével.

[22] Turing-teszt (Turing 1950) jelleggel ugyanazon eldöntendő kérdésekre tudnék olyan szubjektív bonyolultságú (klasszifikált) igen-nem válaszokat adni?

[23] Megint csak egy számítógép programot, egy mesterséges neurális hálózatot vagy akár az embert.

[24] Az eredetit vagy egy lényegében hasonló cikket nézünk most.

[25] http://yann.lecun.com/exdb/mnist/, kézzel írt számjegyek felismerésének sztenderd alapfeladata.

[26] https://www.cs.toronto.edu/~kriz/cifar.html, 10 kategóriába (mint autó, madár stb.) tartozó fotók felismerésének sztenderd alapfeladata.

ezekhez tartozó kimenetei alkotta párokból álljanak. (Például, ha a megtanulandó gép egy MNIST[25] gép lenne, akkor a szubjektív gép tanító pontjai az MNIST gép bemenő képeiből és az ezekre adott 10 dimenziós kimenetek alkotta párokból állnának. Ha a megtanulandó gép egy CIFAR-10[26] gép lenne, akkor a szubjektív gép tanító pontjai a CIFAR-10 gép bemenő képeiből és az ezekre adott ugyancsak 10 dimenziós kimenetek alkotta párokból állnának.) A szubjektív gép bármely (megfelelő) leképezést megtanulna, így a megtanulandó hálózat leképezését is abban az értelemben, hogy tudja reprodukálni azt az említett példák értelmében. Ekkor szembe is kerülünk a már említett neumanni rácsodálkozás egy analógiájával, hogy egy sokkal „egyszerűbb" (de még arra alkalmas) gép megtanul adott bemenetekre ugyanúgy viselkedni, mint egy akár nála sokkal „bonyolultabb". Intuitíven egy univerzális szubjektivitási tétel egy olyan neurális hálózatot konstruálna meg architektúrástól, neuronszámostól stb., amely bármely más neurális háló leképezését képes megtanulni. A dilemmát pedig itt is az oldaná fel, hogy a tanuláshoz át kell adnunk vagy a fent említett tanító pontokat, vagy magát a bonyolultabb, megtanulandó hálózatot és inputját (csak abból a célból, hogy a szubjektív gép architektúrája saját maga elő tudja állítani a szubjektív gép tanító pontjait).

Ha a szubjektív gépet az univerzális géppel hasonlítanánk össze, akkor azt láthatnánk, hogy az univerzális gép „ismeri" a szimulálandó gépet, és pontosan ugyanúgy működik (a szimulálandó gép működését teljesen pontosan „másolja-utánozza le"). Ellenben az intuitív gép nem ismerné a megtanulandó gépet, hanem csak annak az adott bemenetekre adott válaszait és ezeket tudná reprodukálni, „lemásolni". Tehát az intuitív gép esetén egyfajta univerzális lemásolásról beszélhetnénk. Ez alapvetően a felügyelt tanítás alapesete, amelyet ha a memetikai (Dawkins 2005) környezetben értelmezünk, akkor a szubjektív gépet memetikus gépnek is nevezhetnénk, hiszen „lemásolja" egy másik gép működését. Susan Blackmore (Blackmore 1999: 29) meghatározásában a mém az, ami másolódik. Vizsgáljunk meg egy CIFAR-10 gépet! Ekkor a szubjektív gép azt másolja le, ahogyan ez a gép osztályozza az inputját. Végső soron a bemenetén megjelenő képekről megmondja, hogy az autót, repülőt, madarat stb. ábrázol. Ezt (a hálózatot, súlyait) tekinthetjük az autó, a repülő, a madár mémjének (amelyet ha a szubjektív gép megtanul, akkor reprezentációjában egy másik hálózat más súlyai határoznak meg, de a CIFAR-10 klaszterezése ugyanaz). Mivel mém az, amit a szubjektív gép lemásol, így ebből a nézőpontból a mém nem lenne igazán tartalmas fogalom (mert a szubjektív gép intuitíven „bármit" lemásol), viszont utat mutat a pontosabb meghatározásához. Maradjunk a CIFAR-10 példánál, ez véges sok tanító autó, repülő, madár… megtanulása után tetszőlegesen nagyszámú további bemenő képről tud nyilatkozni, hogy az autó, repülő, madár stb. kép-e. A neumanni dilemmát is figyelembe véve véges sok tanító pontot alkalmazva a hálózat „bármennyi" további más autó, repülő, madár stb. képet képes (elegendően jó pontossággal) beklaszterezni. Mi másolódna a CIFAR-10 szubjektív gépes példánkban? A kocsi, madár, repülő… felismerésének képessége, melyet a kocsi, madár, repülő… mémjével azonosítunk. S itt vehetünk észre egy lényeges dolgot, éppen a neumanni dilemma mentén: a mém rövidebb, mint az általa „jelölt dolgok". Nem csak intuitíven, hanem praktikusan is, ez következik az iménti neumanni dilemma feloldása kapcsán említettből, mivel véges sok adattal klaszterezünk be potenciálisan végtelen sokat. A génekkel szokásosan bemutatott szoros analógiára is rámutathatunk a szubjektív gépek nézőpontjából: ahogyan az élőlényekben a DNS az adott környezetbeli rendezettség fenntartásának alapja, úgy a mentális működés tekintetében (a kultúrában) ez a szerep a mémeké.

Tételezzük fel, hogy a tanító (megtanulandó gép) neurális architektúráját nem ismerjük, a tanulóét (szubjektív gép) viszont igen. Ez lehetővé teszi, hogy a tanítás után, a

tanuló ágens súlyainak vizsgálatával megmérjük a tanító tanítottra vonatkozó, ezen adott másolásbeli intelligenciáját (bonyolultságát). Ez egyfajta intelligencia vagy szubjektivitás „hőmérés", amelyben az intelligenciát, a szubjektivitás egymásra vonatkoztatott mértékét a hőhöz hasonlóan mérnénk. Itt a tanítottra vonatkozó kikötéstől el is tekinthetünk, ha az intuitív gép univerzalitását feltesszük (s persze, ha hasonlóan viselkedik, mint a Kolmogorov-bonyolultságban az univerzális gép, azaz ha nem nagyon különbözik két intuitív gép tanulása egymástól).

Egy szoftver termék akkor mesterséges intelligencia, ha meg tudja tanulni a természetes intelligencia szubjektív bonyolultságát (a tanulás során a szubjektivitás egymásra vonatkoztatott mértéke lecsökken), azaz olyan szubjektív gép, amely ember módjára tud gondolkodni. Ez a Turing-tesztnél erősebb kritérium lenne, az következne belőle. Mivel az univerzális szubjektív gép bármit megtanul, így a mesterséges intelligencia elméleti modelljének a szubjektív gépet tekinthetnénk. Tehát egy szoftver akkor lenne intelligencia, ha bármely szubjektív bonyolultságot meg tudna tanulni. A tanuláshoz idő kell. Embernél a kultúra ezt az időt rövidíti le. Hogyan? Talán mert biztosítja a bemenő, éppen megfelelően alacsony entrópiát (az alacsony entrópiájú szellemi táplálékot). A kultúra az alacsony entrópiájú (tehát nagy rendezettségű) mémek „forrása"?

Szellemi táplálék

Térjünk vissza a bevezetőnek a bonyolultságot felvető „Vajon az autonóm vagy a szomatikus vezérlés a komplexebb, „bonyolultabb"? kérdésére, melyre válaszolni persze nem tudunk, de megpróbálhatjuk alkalmazni a schrödingeri tézist – miszerint az anyagcsere[27] során alacsony entrópia felvételével és magas entrópia leadásával tartjuk fent szervezetünk rendezettségének állandó szintjét, folyamatos küzdelemben a termodinamika második főtételével – mentális működéseinkre, a kultúrára. Ha ezzel a tézissel vonnánk analógiát a szellemi működések tekintetében, akkor azt mondhatnánk, hogy rendezett bemenetet veszünk magunkhoz (például meséket hallgatunk, könyveket olvasunk, filmeket nézünk, iskolába járunk), ezekkel tartjuk fel „agyi szoftverünk" rendezettségét. A schrödingeri tézis magasabb entrópiájú kimenetével nehezebb analógiát vonni. Lehetne az autonóm, a szomatikus idegrendszer vezérlő kimenete vagy a gondolkodás. Ezen belül agyunk szomatikus vezérlő kimenete lenne a kisebb entrópiájú, a gondolkodásunk egy nagyobb entrópiájú kimenet. Hiszen a szerveknek és zsigereknek csak nem lehet akármit mondani, ellenben nyugodtan beszélhetünk összevissza, mondjuk, handabandázva is (profán határátmenetben még nagyobb entrópiájú kimenet lenne az érzelemdús veszekedés vagy a pletyka).

A schrödingeri értelmezésben az anyagcsere az energiát pótolja, a „szellemi bemenet" az információt pótolná? Schrödinger csipkelődő megjegyzésével (Schrödinger 1970: 195. oldal közepe és lásd még a fejezet végi jegyzetet) összhangban, akkor a könyveken, filmeken azok információ tartalmát kellene feltüntetni? Várhatóan egy rajzolt gyerekmese kevesebb információt hordoz (kisebb az entrópiája), mint egy valódi képfelvételekből álló film, egy gyerekkönyv meg még kevesebbet. Egy számítógépes játék meg várhatóan több információt hordoz, mint egy valódi film. Mivel a klasszikusan értelmezett objektív szemlélet szerint (tehát a mi olvasatunkban például a Kolmogorov bonyolultsággal számoló in-

[27]Állati és növényi eredetű táplálékot és oxigént veszünk fel, széndioxidot és mást adunk le (Penrose 1993: 346).

formatikus kitüntetett szubjektív entrópiája alkotta vonatkoztatási rendszerben) úgy okoskodhatunk, hogy az említett könyvet lehet legjobban tömöríteni. A rajzolt mesét tipikusan jobban lehet, mint a valódi filmet. A játék meg nem azért bonyolultabb a rajzolt mesefilmnél, mert már fotórealisztikusabb, hanem azért, mert ott bele kell számolni az elvbe az akár véletlennek is tekinthető inputot a játékos részéről. Ezzel a számítógépes játékot bonyolultságban a spektrum legvégén lévő teljesen véletlen képkockákból álló film felé tolja (minél több a „random" input, annál tovább[28]). Ennek az objektív skálának az esetlegességére mutathat rá az az intuitív határeset, amikor maga „a Kolmogorov-bonyolultsággal számoló informatikus" idéz fel olyan példát, melyben délelőtt egy programot még bonyolultnak látott, délután meg ugyanazt egyszerűnek. Bizonyára több 100 programozóval esik meg nap mint nap ez a használati eset: kiderül, hogy egy korábban írt függvénye bugos! Nézi a kódot, de bonyolult, nem jön rá a hibára. Ellene dolgozik, hogy a függvénynek az elvártnál magasabb a ciklomatikus komplexitása[29], ezért nem vezet gyorsan célra a függvényen belüli vezérlés esetleges lefolyásának kitalálása néhány „véletlen" bemenettel. De a programozó annak ellenére, hogy Kernighan–Plauger után tudja, hogy tervezni és kódolni élvezet, nyomkövetni és hibát keresni „büntetés" (Kernighan és Plauger 1982), mégis élvezi a bugot! Vajon miért? Talán mert mivel maga írta és látja a problémát, biztos benne, hogy hamarosan ki tudja adni a bugfixet? Avagy a jelen terminológiánkban biztos benne, hogy a képernyőn bámult forrás jelen pillanatban magas szubjektív entrópiáját hamarosan alacsony szubjektív entrópiás élményre tudja majd változtatni, azaz tudja majd csökkenteni a szubjektív entrópiát. A szubjektív megközelítést támogatja ugyanazon könyvek magyar és kínai fordításának tanulmányozása. Van várakozás az információ ilyen irányú (nem „bitenkénti") értelmezésére, lásd például Dömölki Bálint arc poeticáját az utolsó előtti kérdésre Kömlődi Ferenc kötetében (Kömlődi 2007: 44). Jelen munkánk a szubjektivitást tenné egyetemessé, az objektivitás pedig ennek szimpla következménye, amely a tanuló szubjektív gépek hasonló architektúrájából és hasonló tanításából adódik. A magyar-kínai vonalról általánosabban közelítve özönlenek az érdekes kérdések: megérthetek egy kutyát[30]? Két kutya megértheti egymást? Megértheti-e a kutya az embert? Mert mikor beszélhetünk megértésről? Ha szubjektív gépként tudnak működni.

Játékok

Praktikusabban közelítve, Schrödinger és Penrose hivatkozott (Schrödinger 1970, Penrose 1993) munkái nyomán tudjuk, hogy testünk rendezettségének fenntartásához (kvázi az élethez) az energia alacsony entrópiájú formáit fogyasztjuk és magas entrópiájú formában adunk le energiát. A test entrópiájának alacsony szinten tartásához perces nagyságrendű periódusidővel kell felvennünk az oxigént, napival a táplálékot. Van intuitív megfelelő mentális működéseinkre? Lehetne ilyen a gyerekeknél tipikusan jól megfigyelhető játék iránti vágy? „Játék szomj – Játék éhség"? Azért szeretnek játszani, mert a játékok „entrópiája" fogyasztásra éppen megfelel az agyuk rendezettségének? Ha a játék túl egyszerű

[28] Tehát egy 5v5 emberekkel játszott League of Legends (az e-sport egyik zászlóshajó játéka, röviden LoL, http://eune.leagueoflegends.com) meccs várhatóan bonyolultabb, mint egy olyan, amelyben botokkal játszunk.

[29] Megmutatja, hogy a függvény végrehajtása során a vezérlés hányféle ágon tud lefolyni benne (McCabe 1976).

[30] Emlékezzünk Wigner Jenő „kutyás-szorzótáblás" példájára is, amely az ember faji korlátait is felveti (Wigner 1972).

(túl rendezett – túl kicsi entrópia) vagy túl bonyolult (túl kicsi rendezettség – túl nagy entrópia), akkor nem játszunk vele, ha adott helyzetben éppen megfelelő, fogyasztjuk? Ilyen értelemben a felnőtté válás során tipikusan egyre komplexebb szellemi táplálékok magunkhoz vétele tud felkészíteni arra, hogy még összetettebb rendszereket tömegesen értsünk majd meg, például akár informatikusként a Boost C++ könyvtárának tervezésétől, a Zermelo–Fraenkel-féle axiomatikus halmazelméleten át egészen a kvantummechanika és a relativitáselmélet egyesítéséig. Ilyen értelemben a játékok szerepe felértékelődhet nemcsak a játékosok, hanem a társadalom szemében is. Ha a játékok, a kompetitív játékok (e-sport) szeretete mögött olyan elementáris ösztön működik, mint a légzés vagy táplálkozás esetében, akkor arról kell majd gondolkodnunk, hogy az e-sport elsöprő népszerűségével az egyetemi képzésben már direktben megjelenő – például (Bátfai 2017a) – játékok a közoktatásban hogyan tudnak majd szervezett formában is megjelenni.

Mégis mi végre vannak a játékok? Mi végre a mesterséges intelligencia? Mert a jelenkorban kibontakozó IKT katalizálta kulturális burjánzás-virágzás (előbbi tekintetében gondoljunk csak az aktuális mumusra, a közismert „fake news" jelenségére, utóbbi kapcsán pedig értsd úgy Wagner (Wagner 2006) e-sport definícióját, hogy a számítógépes játékot az IKT tette e-sporttá) nem tudja majd fenntartani a kultúra alacsony entrópiáját. Amelynek viszont a növekedése, ha természetesen nem is a wigneri kutya-szorzótábla szintjére vetheti vissza az általában vett egyéni embert, de felvillantja a lehetőségét egy új „sötét középkori" periódusnak a jelen információs kultúrában. Nyilván nem véletlen, hogy a meghatározó „tech cégek" is ott bábáskodnak[31] az MI mögött. De itt most nem csak az információs közösségi terek javításáról van szó! Megközelítésünkben ez csak a jéghegy csúcsa. Ennek a folyamatnak a háttere az egyedfejlődésben az öregedéssel lehet rokon. Vegyük azt az érzést, amikor szembesülök azzal a tapasztalattal, hogy a tizenéves gyerekeim sokkal több információt „felvesznek"[32] egy LoL meccs lejátszása során (olyan elvben közös tapasztalatokról számolnak be, amelyek bennem legalábbis nem tudatosultak). Mert a hatékonyabb működés miatt én ezeket az információkat már ignorálom? Amíg a „rendszereim" (például szervek) fejlettségi szintje nem homogén (analógia a játékkal: van a csapatban olyan is, aki maga lehozza az egész ösvényt), addig ezek az információk nem számítanak. Ha viszont a rendszerek már homogén teljesítményt nyújtanak (a játékban minden hős egyforma erős), akkor ezek az információk döntenek (győzelem a játékban = egészség az életben). Az analógiával ott kötjük össze a játékot az öregedéssel, hogy a játékbeli karakterek vezérlését összemossuk a való testünk szerveinek vezérlésével. Ebben az olvasatban az „öregedés" az a jelenség, amikor a fejlődésre programozott vezérlés a finomabb információkat már figyelmen kívül hagyja, pedig már fontosak lennének.

[31]Vagy még inkább szülik (lásd például a Google, Facebook, Amazon, Nvidia kapcsolódó eredményeit).

[32] Az ezzel kapcsolatos kutatásainkat lásd a szerző korábbi előadásban (Bátfai 2017a) és a téma motorjában a https://github.com/nbatfai/esport-talent-search mérőprogramban. Az alapvető mérés azt vizsgálja, hogy a játékos a képernyő milyen növekvő komplexitás (bit/sec) változása mellett veszti el a kontrollt a karaktere felett, majd milyen csökkenő komplexitás változása mellett veszi vissza azt.

Összefoglalás

A szubjektív gépet egyértelműen az emberi mentális működés absztrakt modelljének szánjuk. Ezért intuitíven azt is feltesszük, hogy tipikusan a szubjektív gépek tanulnának szubjektív gépeket, minden konkrét gép különböző, de architektúrájuk és működésük szervezésében megegyező. A bonyolultságot csak szubjektív értelemben definiáljuk.[33] A megtapasztalt objektivitást, az objektivitás látszatát, az adja, hogy mint szubjektív gépek ugyanabból a kultúrából másolunk. Másolt mémjeink sikerét az adja, ha tapasztalatainkat minél szélesebb körét meg tudják magyarázni, Occam borotvája ezt az igazság fogalmával kapcsolja össze. Ilyen értelemben a formális axiomatikus elméletek lehetnek a ma ismert legalacsonyabb entrópiájú mémek forrásai. Ám Neumann *„a logikáról és a matematikáról is feltételezhetjük, hogy azok is csak történeti eredetű, esetleges kifejezési formák"* (Neumann 2006: 88) ráérzése nyomán az is elképzelhető, hogy tapasztalatunk szerint majd két vagy akár még több olyan axiomatizált elméletünk is lesz, amelyeket igaznak tartunk, de egymással nem összeegyeztethetőek. Ez harmonizál azzal a képpel, hogy az objektívnek tartott tapasztalataink nem pusztán közös metszetei szubjektív egyéni tapasztalatainknak, hanem a tanult objektivitás (a kultúra, a józanész) visszahat, az egyéni tapasztalataink meghatározója, irányítója is.

Induló kérdéseink voltak, hogy miért szeretünk játszani, és miért van szükségünk mesterséges intelligenciára? A játékban a szubjektív entrópia könnyebben csökkenthető, mint a valóságban. Kvázi könnyebb megtanulni LoL-ozni, mint adott esetben osztani az általánosban, trigonometrikus egyenleteket megoldani a középiskolában, vagy komplex kalkulusozni a főiskolán. Pedig elvben mindegynek kellene lennie, hogy a szubjektív gép mit másol, ha az szisztematikusan ezekkel a példaként megnevezett burkolt célokkal lenne felépítve: játékokba, konkrétan számítógépes játékokba oltva, mert akkor tömegesen öröm lenne másolni (játszani) az osztást is, az egyenleteket is, a kalkulust is. Tézisünk, hogy azért szeretünk játszani, mert ahogyan a testünknek rendezett bemenetre van szüksége rendezettségének fenntartásához, úgy a lelkünknek (agyi szoftverünknek) is megfelelően rendezett bemenetre van szüksége rendezettsége fenntartásához. S ugyanez igaz az emberi kultúrára is: szükségünk van a szervezett oktatás következő lépcsőfokaként olyan személyi tanító ágensekre, akik mesterségesen intelligensek és ott lehetnek minden egyes tanuló (másoló) ember mellett, hogy tudjuk őket másolni, ezáltal fenntartani az emberi kultúrát!

Irodalom

Altrichter Márta, Horváth Gábor, Pataki Béla, Strausz György, Takács Gábor és Valyon József, *Neurális hálózatok*, Panem, Budapest, 2006.

Bátfai Norbert, "Theoretical Robopsychology: Samu Has Learned Turing Machines", *arXiv eprints*, 1606.02767, 2016. https://arxiv.org/abs/1606.02767

Bátfai Norbert, Besenczi Renátó, Bogacsovics Gergő and Monori Fanny, "Entropy Non-increasing Games for the Improvement of Dataflow Programming", *arXiv eprints*, 1702.04389, 2017. https://arxiv.org/abs/1702.04389

[33] Ebben az irányban kifejlesztendőek konkrét becslő-mérő programok, a mérésben a mi közös objektivitásunk (konkrétan a matematikánk) lenne az a vonatkoztatási rendszer, amelyben a mérési eredmények születnek. Ezt az irányt nem gondoltuk végig, de a végén felmerülhet egy alternatív válasz is Fermi „Where is everybody" kérdésére.

Bátfai Norbert, „Az e-sport egyetemi oktatása”, Országos Neveléstudományi Konferencia (Nyíregyházi Egyetem, 2017. november 9–11.), poszter, 2017. https://arato.inf.unideb.hu/batfai.norbert/NEMESPOR/ONK2017/ONK_esport_poster_BN.pdf

Bátfai Norbert, „Tehetségkutatás az e-sportban”, in Kerülő Judit, Jenei Teréz, Gyarmati Imre (szerk.) Országos Neveléstudományi Konferencia (Nyíregyházi Egyetem, 2017. november 9–11.), MTA Pedagógiai Tudományos Bizottság, Nyíregyházi Egyetem, Nyíregyháza, 2017a, 38. old. http://onk2017.hu/wp-content/uploads/2017/11/ONK_2017_november_20171110.pdf

Blackmore, Susan, A mém-gépezet, Magyar Könyvklub, Budapest, 2001.

Bhupatiraju, Surya, Rishabh Singh, Abdel-rahman Mohamed and „Deep API Programmer: Learning to Program with APIs”, *arXiv eprints*,1704.04327, 2017. https://arxiv.org/abs/1704.04327

Chaitin, Gregory J., ”On the Simplicity and Speed of Programs for Computing Infinite Sets of Natural Numbers”, *Journal of the ACM*, Vol. 16. (1969) Issue 3., pp. 407–422. https://doi.org/10.1145/321526.321530

Richard Dawkins, *Az önző gén*, Kossuth Kiadó, Budapest, 2005.

Donald, Merlin, *Az emberi gondolkodás eredete*, Osiris Kiadó, Budapest, 2001.

Hameroff, Stuart and Roger Penrose, ”Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness”, *Mathematics and Computers in Simulation*, Vol. 40. (1996) Issues 3-4., pp. 453–480. https://doi.org/10.1016/0378-4754(96)80476-9

Kernighan, Brian W. és Plauger, P. J., *A programozás magasiskolája*, Műszaki Kiadó, 1982.

Friedhart Klix, *Az ébredő gondolkodás*, Gondolat Kiadó, Budapest, 1985.

Kolmogorov, Andrej N., ”On tables of random numbers”, *Theoretical Computer Science*, Vol. 207. (1998) Issue 2., pp. 387–395. https://doi.org/10.1016/S0304-3975(98)00075-9

Kolmogorov, Andrej N., ”Three approaches to the quantitative definition of information”, *International Journal of Computer Mathematics*, Vol. 2. (1968) Issue 1-4., pp. 157–168. https://doi.org/10.1080/00207166808803030

Kömlődi Ferenc, *Mesterséges intelligencia és határterületei*, Akadémiai Kiadó, Budapest, 2007.

Lovász László, *Algoritmusok bonyolultsága*, Nemzeti Tankönyvkiadó, Budapest, 1994.

Mandelbrot, Benoit. B., *The fractal geometry of nature*, W. H. Freeman and Comp., New York, 1983.

McCabe, Thomas J., ”A Complexity Measure”, *IEEE Transactions on Software Engineering*, Vol. SE-2. (1976) Issue /4., pp. 308 –320. https://doi.org/10.1109/TSE.1976.233837

McCulloch, Warrwn S. and Walter Pitts, ”A Logical Calculus of the Ideas Immanent in Nervous Activity”, *Bulletin of Mathematical Biophysics*, Vol. 5. (1943) Issue 4., pp. 115–133. https://doi.org/10.1007/BF02478259

Minsky, Marvin and Seymour Papert, *Perceptrons: An Introduction to Computational Geometry*, MIT Press, Cambridge, 1969.

Misha, Denil, Sergio Gómez Colmenarejo, Serkan Cabi, David Saxton and Nando de Freitas, ”Programmable Agents”, *arXiv eprints*, 1706.06383, 2017. https://arxiv.org/abs/1706.06383

Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg and Demis Hassabis, ”Human-level control through deep reinforcement learning”, *Nature*, Vol. 518. (2015), pp. 529–533. http://dx.doi.org/10.1038/nature14236

Mnih, Volodymyr, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra and Martin Riedmiller, ”Playing Atari with Deep Reinforcement Learning”, *arXiv eprints*, 1312.5602, 2013. https://arxiv.org/abs/1312.5602

Neumann János, „A számítógép és az agy”, NetAcademia Oktatóközpont, Budapest, 2006.

Neumann János, „Az automaták általános és logikai elmélete” (1951), in: Ropolyi László és Szegedi Péter (szerk.), *Neumann János válogatott írásai*, Typotex, Budapest, 2010, 156–227. old.

Neumann, John von, ”Probabilistic logics and synthesis of reliable organisms from unreliable components”, in Claude Shannon and John McCarthy (eds.), *Automata Studies AM-34*, Princeton University Press, Princeton, 1956, pp. 43–98.

Penrose, Roger, *A császár új elméje*, Akadémiai Kiadó, Budapest, 1993.
Rényi Alfréd, *Napló az információelméletről*, Gondolat Kiadó, Budapest, 1976.
Reed, Scott and Nando de Freitas, "Neural Programmer-Interpreters", *arXiv eprints*,1511.06279, 2015. https://arxiv.org/abs/1511.06279
Schrödinger, Erwin, „A természettudományos világkép sajátosságai", in Tőrös Róbert (szerk.), *Válogatott tanulmányok / Erwin Schrödinger*, Gondolat Kiadó, Budapest, 1970, 221–280. old.
Schrödinger, Erwin, *What is life? : the physical aspect of the living cell*, Cambridge University Press, Cambridge, 1944.
Silver, David, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel and Demis Hassabis, "Mastering the game of Go with deep neural networks and tree search", *Nature*, Vol. 529. (2016), pp. 484–489. http://dx.doi.org/10.1038/nature16961
Solomonoff, Ray J., *A Preliminary Report on a General Theory of Inductive Inference*, Zator Co., Cambridge, 1960.
Tinbergen, Niko, *Curious Naturalists*, University of Massachusetts Press, Amhert, 1984.
Tinbergen, Niko, „The Curious Behavior of the Stickleback", *Scientific American*, Vol. 18. (1952) Issue 6., pp. 22–26. https://doi.org/10.1038/scientificamerican1252-22
Tinbergen Niko, *Az ösztönről*, Gondolat Kiadó, Budapest, 1976.
Turing, Alan, "On Computable Numbers, with an Application to the Entscheidungsproblem", *Proceedings of the London Mathematical Society*, Vol. s2-42. (1937) Issue 1., pp. 230–265. https://doi.org/10.1112/plms/s2-42.1.230
Turing, Alan, "Computing machinery and intelligence", *Mind*, Vol. LIX. (1950) No. 236., pp. 433–460. https://doi.org/10.1093/mind/LIX.236.433
Wagner, Michael. G., "On the Scientific Relevance of eSports", in Hamid R. Arabnia (ed.) *Proceedings of the 2006 International Conference on Internet Computing. Conference on Computer Games Development (ICOMP, Las Vegas, Nevada, USA, June 26-29, 2006)*, CSREA Press, 2006, pp. 437–442.
Wigner Jenő, *Szimmetriák és reflexiók*, Gondolat Kiadó, Budapest, 1972.