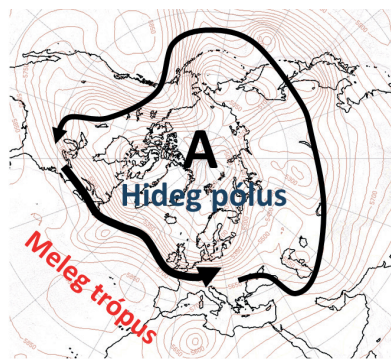


HORVÁTH ÁKOS

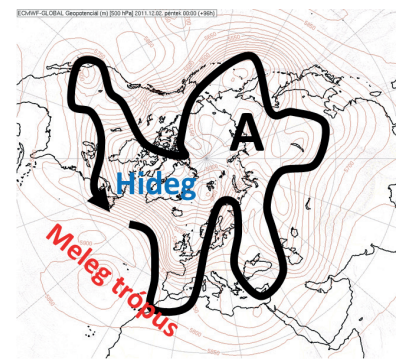
A viharjelzés bizonytalanságai

A légköri folyamatok egyik legfőbb sajátossága, hogy gyakran nehéz előre jelezni azok jövőbeli állapotát; időnként a legkisebb külső behatásra is gyökeresen más irányba alakul az időjárás, más szóval, a légkör sokszor kaotikusan viselkedik [1][2]. Ez a kaotikusság főként akkor jelenik meg, amikor nagy változásokat okozó, heves időjárási folyamatok indulnak be, melyek során jelentős energia koncentráció, erősen turbulens áramlások, és heves csapadéktevékenység indul meg. Egy stabil, „lamináris” anticiklonban sokkal áttekinthetőbbek, így jobban előre jelezhetőek a légmozgások, mint egy „turbulens” ciklonban. A veszélyes időjárási jelenségek, azaz a szélviharok, jégesők, tornádók, heves csapadékok azonban éppen a légkör turbulens viselkedéséhez köthetőek.

Napjainkban a mezőgazdaság, az ipar vagy a szabadidős tevékenységek, lényegében az egész modern társadalom egyre érzékenyebb a markáns időjárási eseményekre. A meteorológiával szemben egyre nagyobbak az elvárások a veszélyjelzés tekintetében, a veszélyes jelenségek viszont a legnehezebben előre jelezhető folyamatok közé tartoznak. Ezen írás célja, hogy áttekinthetően a ciklonoktól a tornádókig tartó széles meteorológiai skálán előforduló veszélyes



Zonális típus



Meridionális típus

1. ábra. A pólust körülölelő, meghatározóan nyugat-kelet irányú zónális áramlási típus, és az észak-déli (illetve dél-északi) irányú meridionális áramlási típus

jelenségeket és bemutassuk az előrejelzési módszereket és az ott fellépő bizonytalanságokat.

A ciklonok skálája

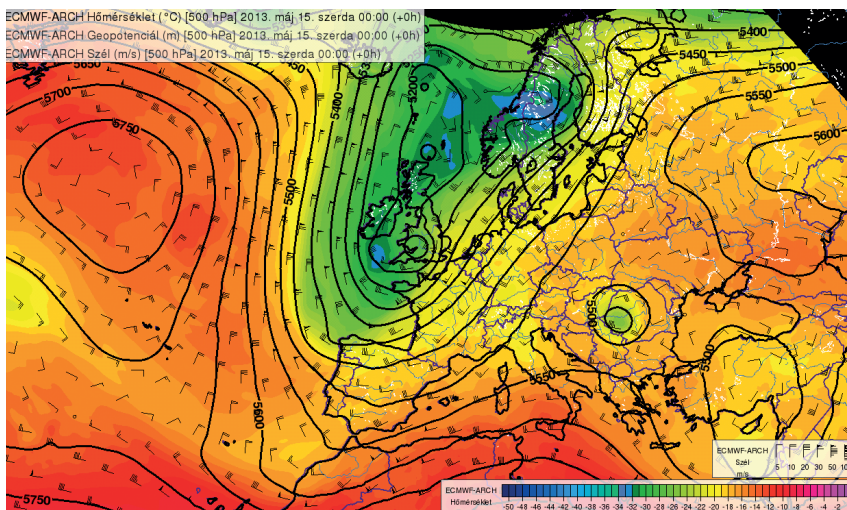
A ciklonok kialakulása szorosan kötődik a globális cirkulációhoz. A globális cirkuláció lényegében egy gigászi hőerőgépként ír-

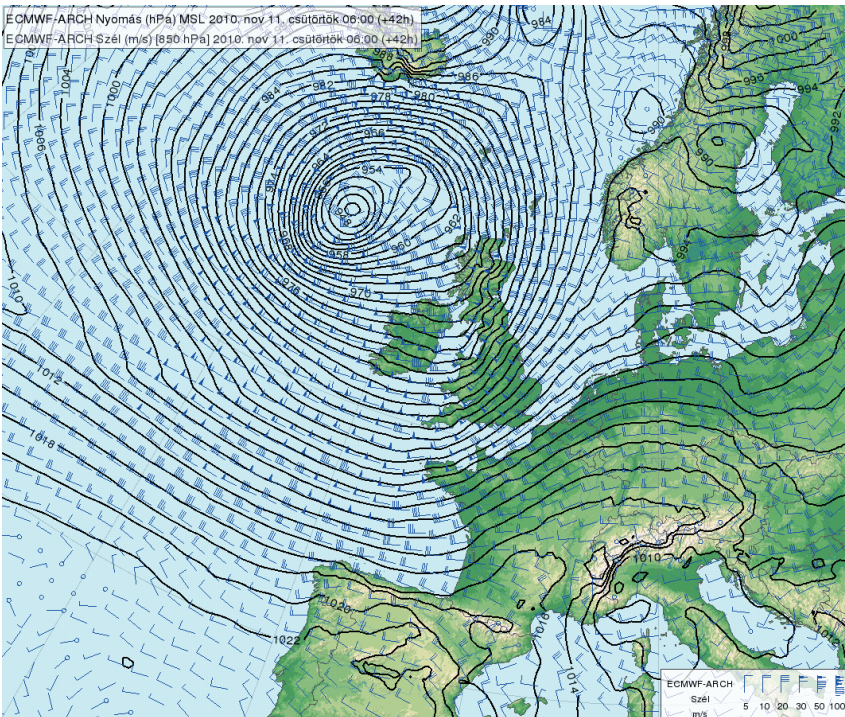
ható le, ahol a kazán szerepét az Egyenlítő vidéke, a hűtő szerepét pedig a sarki hideg pólus környéke játssza. Az így keletkező hőenergia-különbség, valamint a forgó Földön fellépő Coriolis-erő hatására létrejön a pólusokat körülvevő áramlási rendszer, a nyugati szelek öve. A nyugati szelek, vagyis a zónális áramlás azonban nem kedvez az észak és dél közötti hőcserének, így a trópus tovább melegszik, a pólus tovább hűl, a nyugati szelek pedig egyre erősebbek lesznek. Nyilvánvalóan egy idő után nem maradhat fenn ez az állapot és légköri instabilitások lépnek fel. Az ún. barotróp, majd baroklin instabilitások hatására a nyugati szelek övében egyre növekvő hullámok jelennek meg, amelyek egy idő után leszakadhatnak az áramlásról, illetve a nyugat-keleti zonalitás átalakul észak-déli (illetve dél-északi) meridionalitássá. A valódi légköri cirkuláció a két alaptípus között változik (1. ábra). A meridionális áramlásban a légköri hullámok, ciklonok hátoldalán hideg levegő zúdul dél felé, míg a hullám előoldalán meleg áramlik északra (2. ábra).

A viharciklonok

A ciklonok skáláján fellépő egyik legveszélyesebb időjárási jelenség éppen a nagyon erős zónális áramlás felbomlásához kötődik, amikor a légköri hullám robbanás-szerűen alakul ki. Ez a gyors ciklogenezis (rapid cyclogenesis), vagy „tengeri bomba” néven elhíresült folyamat legtöbbször az Atlanti-óceán északnyugati területei fő-

2. ábra. Meridionális áramlás kialakulása Európa nyugati partjainál a középső troposzférában. A folytonos vonalak az 500 hPa nyomás szint magasságát, a színezett területek a hőmérséklet eloszlását, a szélzónák pedig az 500 hPa nyomásfelületen uralkodó szelet mutatják. A kezdődő sarki hidegkitörés keltette azt a ciklont, amely egy héttel később a dunai árvíz okozója volt





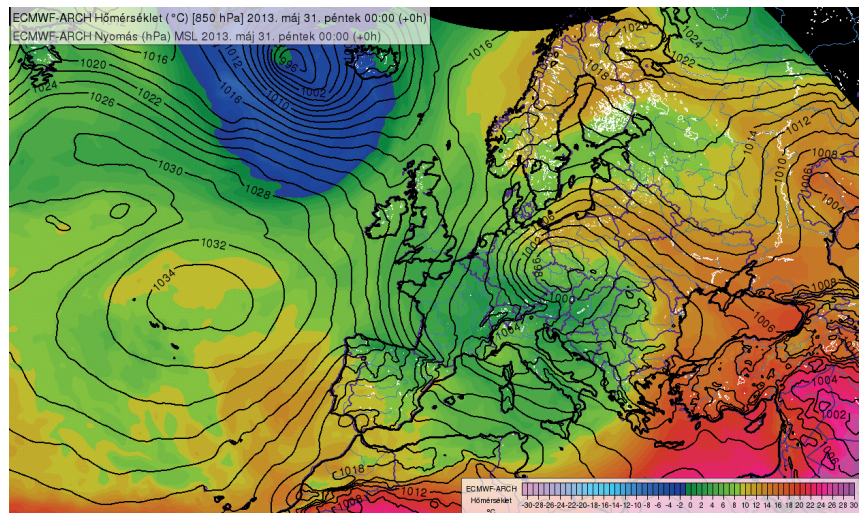
3. ábra. Viharciklon Nyugat-Európa partjainál 2010. november 11-én. A ciklon centrumában a tengerszintű légnyomás 945 hPa-ra süllyedt. A folytonos vonalak a tengerszintű légnyomást jelölik, a szélszálók pedig a felszín közeli szelet

lött alakul ki. Legfőbb sajátossága a rendkívül erős nyomáscsökkenés (24 óra alatt akár 20–30 hPa). A gyorsan mélyülő ciklonban forgó levegő sebessége hozzáadódva a ciklon gyors áthelyeződési sebességéhez nagyon erős viharokat képes létrehozni. Az egyik ilyen hírhedt vihar 1999. december 26-án csapott le Nyugat-Európára. A számítógépes előrejelzések alábecsülték a vihar erősségét, amely a valóságban például Párizsban 180 km/ó legerősebb szél-lökést okozott, emberéleteket követelve, óriási károkat okozva. Pár évvel később a 2004. november közepén lecsapott vihar-ciklont (amely többek között a Magas-Tátra erdőseit tarolta le) már sikerült nagy pontossággal előre jelezni. Az elmúlt években a téli fél évben gyakran jelentek meg hasonló jellegű viharok Nyugat- és Észak-Európa partjainál [3] (**3. ábra**).

A gyorsan fejlődő ciklonok előrejelzése nagyon nehéz feladat, az itt jelentkező bizonytalanság meglehetősen nagy. Ennek egyik oka magában a légkörben van: nagyon nehéz megmondani, hogy a több ezer km hosszú zonális áramlásban pontosan hol alakul ki az első kis légköri zavar (perturbáció), amely elindítja a hullám fejlődését. Viszont ha már egyszer sikerült analizálni a fejlődő ciklont, akkor a számítógépes modellek meglehetősen pontosan tudják követni a fejlődést és az áthelyeződést. Az előrejelzésben jelentkező bizonytalanság másik forrása a mérési pontok alacsony szá-

ma, amely főként az óceánok fölött jelent problémát. A bizonytalanság leírására az ún. ensemble modellek segítségével van lehetőség, vagyis amikor a hidrodinamikai modellek kezdeti mezőire apró kis változtatásokat, azaz perturbációkat helyeznek. Ilyenkor a perturbálatlan kezdeti feltételek-

4. ábra. Leszakadt légörvény Közép-Európa felett 2013. május 31-én. Az alig mozduló ciklon keleti oldalán rendkívül meleg levegő áramlik a Baltikum irányába, míg a nyugati oldalon hideg levegő zúdul Nyugat-Európára. Ez a ciklon volt a dunai árvíz közvetlen kiváltója. A folytonos vonalak a tengerszintre számított légnyomás-eloszlást, a színezett területek pedig a 850 hPa nyomási felület (kb. 1500 m magasság) hőmérsékletét mutatják



ból indított referenciamodell mellett több, perturbált kezdeti mezővel indított modell is futtatnak és az így kapott előrejelzés-együttes lehetővé teszi a valószínűségi előrejelzések készítését. Az ensemble modellezésnek a legnehezebb pontja annak meghatározása, hogy hol érdemes módosítani a kezdeti mezőt. A probléma megoldásának összetett matematikai eszköztársa van, és ezzel a kérdéskörrel a kiadvány egy másik írása foglalkozik részletesen [4]. A mérési hiányok csökkentésére ugyancsak a matematikai modellezés nyújthat segítséget, azzal, hogy kijelöli azokat a területeket, ahol a számítás legérzékenyebb a kezdeti feltételekre. Ez lehetővé teszi, hogy az adott térségben sűrűbben végezzenek méréseket, például repülőgépekről dobott ejtő szondák segítségével. A célzott mérések technikája (targeted observation) ugyan költséges, de a befektetett erőfeszítések bőven megtérülhetnek egy veszedelmes ciklon pontosabb előrejelzésével.

A lefűződött (cut off) ciklonok

Míg a viharciklonok gyors fejlődésükkel és nagy sebességükkel okozzák a pusztító szélviharokat, addig a lefűződött ciklonok esetében éppen a lassú áthelyeződés vált ki szélsőséges időjárást. A földi cirkuláció zonális áramlási típusával szemben áll a meridionális áramlás, amikor a mély hullámokban a ciklonok alig, vagy csak lassan helyeződnek át (**4. ábra**). Ilyen esetekben a ciklonok sokáig maradnak egy terület fölött, és ha elegendő meleg és nedves levegőhöz jutnak az előoldalukon (pl. a Földközi-tenger medencéjéből) és a hideg utánpótlás is rendelkezésre áll (pl. az észak-at-



5. ábra. Zivatarlánc közelít a Balatonhoz

lanti térségből), akkor jelentős csapadékot tudnak produkálni. A 2013-as, illetve a 2002-es dunai árvizek ilyen lefűződött ciklonokhoz köthetők [5]. Amikor a ciklon töltődésnek indul és a napsugárzás hatására az alsó légrétegek felmelegednek, akkor a magasban még fennmaradó hideg mag és az örvénylés együttesen kedvező körülményeket teremtenek a heves zivatarok kialakulásának.

A lefűződött ciklonok esetében a bizonytalanság nem annyira a ciklon környezetének állapotából, hanem sokkal inkább magában a ciklonban lezajló folyamatokból adódik. A lefűződött ciklonban a hideg- és melegfront már többszörösen is utolérte egymást és többszörösen körbefordult okklúziós frontokká alakult. Ezekben az összeáramlási zónákban koncentrálnak a nedvesség és a csapadéksávok is ilyen szalagokba rendeződnek. A sávban az egymás után mozgó felhőkből nagy mennyiségű csapadék hullhat viszonylag keskeny sávban. A műholdképeken ezek a csapadéksávok a ciklon jellegzetes karjaiként azonosíthatók. Ha az összeáramlási vonal a ciklon méretéhez képest csak egy kicsit is elmozdul, az komoly eltérést eredményezhet a csapadék eloszlásában. A lefűződött ciklonokban a viharciklonokhoz képest újabb hidrodinamikai instabilitási formák is megjelennek. Az előrejelzések bevalásának növelésére itt a finomabb mozgásformák és a csapadékképződési folyamatok fokozott figyelembevételével van lehetőség. Az alkalmazott hidrodinamikai modelleknek finomabb horizontális felbontással kell rendelkezniük és a csapadékképződési folyamatokat leíró moduloknak is részletesebbnek kell lenniük.

Az időjárási frontok és zivatarláncok skálája

A ciklonok szerkezetében, főleg a fiatalabb légyörvények esetén nagyon jól elkülöníthetőek a légköri frontok. A légkör állapotát leíró paraméterek tekintetében a frontok törésvonalakként, vagy szakadási felületekként jelennek meg. A frontokon az állapotváltozók értékei térben nagyot változnak, ott első vagy másodfajú szakadások jelennek meg. (Így pl. a hidegfrontok átvonulásakor a hőmérséklet nem folytonosan változik, hanem gyakran „ugrik”, vagy a légnyomás mezejében, az időjárási frontok mentén az elsőrendű deriváltakban figyelhető meg szakadás, vagyis az izobár „megtörik”). Az időjárási frontok tehát éles felületek, amelyek hullámok jelennek meg és jóval kisebb skálán, de hasonló instabilitási formák alakulhatnak ki, mint a ciklonok fejlődésekor. A frontokon kialakuló hullámok problémát jelentenek a szél és a csapadék előrejelzésben, hiszen mindkét elem erősen kötődik a frontfelület okozta függőleges mozgásokhoz, illetve a légnyomásváltozáshoz. Gyakori jelenség, hogy a hidegfronton keletkező hullám csak az alsó légkörben jelentkezik, a magasban a fronthoz tartozó hideg levegő zavartalanul áramlik a hullámban lévő meleg levegő fölé, így egy újabb markáns instabilitási forma jelenik meg, a zivatarokért felelős *konvektív instabilitás*. A konvektív instabilitás végső soron a felhajtóerőhöz köthető és legegyszerűbb megjelenési formája az, amikor a napsugárzás hatására felmelegedett talaj közeli meleg levegő sűrűsége kisebb lesz, mint a fölötte lévő levegőé, így az egy légbuborék formájában emelkedni kezd, majd a benne kicsa-

pódó nedvesség hatására a jól ismert kumulusz felhő formájában láthatóvá válik. Jórészt ez a konvektív instabilitás felelős a hidegfrontok előtt megjelenő vonalba rendezett zivatarokért (5. ábra). A meteorológusok által „squall line” néven emlegetett jelenség az egyik legmarkánsabb időjárási objektum. A csatárláncban sorakozó zivatarcellák nemegyszer orkán erejű széllel, felhőszakadással és jégesővel járnak. A kockázat mértékét is figyelembe véve a viharjelzés egyik legnehezebb feladata azt megmondani, hogy hol és mikor jönnek létre zivatarláncok, azon belül is melyik cella produkálja a legnagyobb szelet, lesz-e jégeső stb. A légköri skálákon lefelé haladva a konvektív instabilitás terén a zivatarláncok kifejlődésének és mozgásának előrejelzése valószínűleg az a határ, ahol a jelenlegi gyakorlatban alkalmazott modellek rövidtávú (+24 órára előre szóló) előrejelzése még felhasználható, bár a bizonytalanság már jelentős. Az egyedi zivatarfelhők, főként az első cellák kialakulásának dinamikus alapú előrejelzésére jelenleg is komoly erőfeszítések történnek [6].

A zivatarfelhők skálája

A veszélyes időjárási folyamatokat kiváltani képes jelenségek közül a zivatarcellák a skála alján helyezkednek el az 5–10 km-es átmérőjükkel. A zivatarok lehetnek egyediek (lokálisak), vagy valamilyen rendszerbe szerveződők. A lokális zivatarok legtöbbször nem jelentenek nagyobb időjárási veszélyt, bár vannak olyan területek, ahol az általuk keltett 60–70 km/ó széllelökések mégis komoly problémát okozhatnak, így például a Balatonon, a Velencei-tavon vagy a Tiszatavon fürdőzők, csónakázók számára. A lokális zivatarok legtöbbször a frontoktól távolabb alakulnak ki, ilyenkor légtömegben belüli, vagy hőzivataroknak is nevezik őket.

Ha labilisabb a levegő, vagyis nagyobb a konvektív instabilitás, akkor zivatarrendszerek, zivatargóccok, azaz többcellás zivatarok alakulnak ki, amelyek szinte mindig veszélyt jelentenek. Itt az egyes cellák sokszor egymást erősítve magas tetővel rendelkező (12 km) felhőket is képesek létrehozni. A felhőkben nagyon erős feláramlások (15–20 m/s) jönnek létre, amelyek elősegítik a nagy szemű jég megjelenését, amely a kihullás során nem olvad el, hanem a talajra érve pusztító jégveréseket okoz. A többcellás zivatarok nagyon erős szelet is tudnak produkálni, akár 100–120 km/ó-t meghaladó széllelökésekkel. Ennek esélye, hogy a zivatarokból többcellás zivatargóccok legyenek annál nagyobb, minél közelebb vannak az időjárási frontokhoz, illetve, mint láttuk, a zivatarláncok közvetlen kapcsolatban vannak a hidegfrontokkal [7]. Még finomabb felbontást tekintve látható, hogy egy cellán belül néhány kilométer távolság-

A viharjelzés bizonytalanságai



6. ábra. A keskeny csapadéksávok jelzik a zivatarfelhőből leáramló légzuhatagot

ban jelentős különbség van a csapadék mennyiségében, a jégverés eloszlásában vagy a szélvihar erősségében. Gyakran a határ egy részét elveri a jég, néhány száz méterrel arrébb pedig csak eső esik. A zivatarfelhőből kifutó heves szélben is nagy lehet a térbeli változékonyság, különösen a koncentrált leáramlási zónák, az ún. down-burst jelenségek esetén (6. ábra). A nagyon heves zivatarokban a levegő koncentrált beáramlása során tornádók is kialakulhatnak, amelyek legtöbbször csak néhány száz méteres pusztító sávval rendelkeznek, ott viszont mindent visznek magukkal. Éppen a kis méretüknél fogva, a tornádókat nagyon nehéz a hagyományos meteorológiai mérőhálózáttal azonosítani. Inkább csak az időjárási radarok segítségével mért paraméterek

alapján van mód a tornádót létrehozó ún. szupercellás zivatarcellákat detektálni, melyek jellegzetes formájú, baljós kinézetű felhőjét szabad szemmel is könnyű felismerni (7. ábra).

Ha a viharjelzés bizonytalanságát tekintjük, akkor a zivatarok a legnehezebben előre jelezhető folyamatok közé tartoznak. Kiskalájú, de nagyon gyorsan felépülő rendszerekről van szó, egy fejlett zivatarfelhő a „semiből” 20–30 perc alatt is képes kialakulni. Bár főként a többcellás zivatarok esetén a nagyobb skálájú rendszerek (pl. a ciklonok) meghatározzák, hogy hol lesz a legnagyobb a konvektív instabilitás, azonban az első cellák kirobbanását bármilyen kicsiny hatás is kiválthatja. Itt valóban nem túlzás a neves amerikai mete-

7. ábra. A szupercella jellegzetes lenyúló felhőzete a cella középpontjában feláramló örvénylő levegőre utal



orológusnak, Lorenznek tulajdonított pilangóhatás [8]. Egy apró fuvallat elindíthat egy kis feláramlást, amely az instabil lég-rétegződésben gyorsan termiké alakul. A felemelkedő levegő pótlására egyre intenzívebb módon újabb légmozgások indulnak el, a termikből gomolyfelhő lesz, abból zivatarfelhő, amelyből a kifutó szél újabb cellákat torlaszt fel. Ha a kezdeti apró hatás 60 kilométerrel arrébb indítja el a folyamatot, akkor a zivatarok annyival arrébb és annyival másabb körülmények között indulnak meg, a zivatar okozta változások visszahatnak a frontok mozgására, az pedig magára a ciklonra. Végző soron a gyenge fuvallat egy egész földrész időjárását megváltoztathatja.

Egy szög miatt a patkó elveszett.

A patkó miatt a ló elveszett.

A ló miatt a lovas elveszett.

A lovas miatt a csata elveszett.

A csata miatt az ország elveszett.

(Máskor verd be jól a patkószeget!)

(angol népköltés)

Károlyi Amy fordításában)

A légköri energiaátadás alapvetően nagyobb skálák felől a kisebb skálák felé történik. A ciklon energiáját főként a földi cirkulációtól kapja, az időjárási frontok mozgását jórészt a ciklon szerkezete határozza meg, a zivatarok skáláján lezajló fejlődések pedig erőteljesen függenek a ciklon frontjainak elhelyezkedésétől, mozgásától. A skálák közötti függőség azonban nem feltétlenül egyirányú. Mint az említettekben is látható, zivatarok skáláján lezajló folyamatok megzavarhatják az időjárási frontok, végző soron a ciklon fejlődését is, amely akár az egész nyugati szelek övében is elindíthat változásokat. A bizonytalanságok nagyban függenek az adott skálát jellemző hidrodinamikai instabilitások mértékétől. Ha minden skálán nagy az instabilitás, a gyorsan fejlődő veszélyes folyamatokat nagyon nehéz leírni, a veszélyjelzés ilyenkor az időjárás követésére szorítkozik. A veszélyes folyamatokat követése, és azok pár órára történő előrejelzése a meteorológia egyik speciális ágához, az ultrarövidtávú előrejelzéshez, szakszóval a *nowcasting*-hoz tartozik (1. blokk). Napjainkban terjedőben vannak a számítógépeken futó automatikus viharjelző rendszerek, vagy *nowcasting* rendszerek. Az időjárási veszélyjelzésben a bizonytalanság és a nagy felelősség miatt azonban még mindig a végző szó az adott terület időjárási sajátosságait ismerő meteorológusé.

Időjárási veszélyjelzés Magyarországon

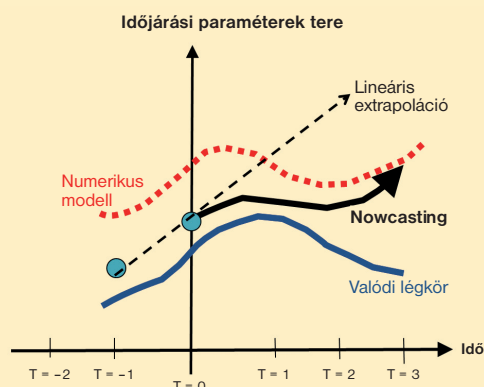
Hazánk időjárásában ugyancsak megjelennek a veszélyes időjárási jelenségek, bár kisebb gyakorisággal és hevességgel, mint

A veszélyjelzés módszertana, a nowcasting

A nowcasting kifejezés a pillanatnyi állapot (now) és az előrejelzés (forecast) szavakból összerakott mozaik szó. Ha feltételezzük, hogy a légköri folyamatok időbeli változását egy nagyon bonyolult függvény írja le, akkor az ultrarövidtávú előrejelzést felfoghatjuk úgy, mint ennek a függvénynek egy lineáris közelítését, (akár a matematikában a Taylor-sor első két tagját) a dinamikus modelleket pedig mint a függvényt valamilyen közelítéssel leíró, bonyolult eljárást. Nyilvánvalóan ezt a lineáris közelítést csak rövid időre előre lehet alkalmazni és az újonnan bejövő információk alapján nagyon gyakran ismételni kell. Ezt a megközelítést takarja a nowcasting kifejezés: a folyamatosan mért adatok alapján meg kell határozni a tendenciákat és azok alapján továbbvinni az időjárás változását rövid időre (maximum 3 órára) előre. A nowcasting feladata, hogy a következő egy-két órára előre minden adott pontra megmondja, hogy lesz-e valamilyen veszélyes időjárási jelenség, vagyis közvetlenül segítse a döntéshozók munkáját.

Az Országos Meteorológiai Szolgálatnál (OMSZ) alkalmazott nowcasting rendszer legelső lépéseként a sokféle meteorológiai mérésből meg kell határozni a pillanatnyi időjárást. A pillanatnyi időjárás alatt az adott területet lefedő rácsra kiszámolt meteorológiai paramétereket értjük. A térben rendezetlenül elhelyezkedő automata meteorológiai állomások 10 perces sűrűséggel szolgáltatják az adatokat, hasonlóan az időjárási radarokhoz illetve a meteorológiai műholdakhoz. A sokféle mérésből, valamint a korábbi kiindulási időből számított dinamikus alapú előrejelzések (ugyancsak rácspontri) adataiból készül el az adott időpont állapotát leíró ún. időjárási objektív analízis. Az objektív analízis a lehető legpontosabban írja le a légkört, azonban nem tökéletesen. Az ultrarövidtávú előrejelzés egyik lehetősége az, hogy az objektív analízisek sorozatából időben előre lineáris extrapolációt végeznek, amely a következő órára még sikeres is lehet, azonban a lineárisnak nem tekinthető légkör miatt ennek érvényessége az előrejelzési idővel erősen csökken. A három órára szóló előrejelzés azzal javítható, hogy időben előre haladva fokozatosan a jövő állapotát dinamikai alapon előrejelző numerikus modellhez közelítjük az előrejelzésünket (**8. ábra**). A dinamikus modellek hatalmas számítási kapacitást igényelnek, ezért azokat nem lehet túl gyakran futtatni (naponta 2 vagy legfeljebb 4 futást szoktak indítani) és a modell is hibával terhelt, azonban az ultrarövidtávú előrejelzendő időszak végén (tipikusan 3 órára előre) mégis jobb lesz a lineáris extrapolációnál. Az OMSZ-nál jelenleg minden tízpercenként indul egy nowcasting előrejelzés amely a fentiek alapján +3 órára előre készíti előrejelzést.

Vannak próbálkozások a nowcasting valószínűségi előrejelzések készítésére is. A legkézenfekvőbb megoldás az, hogy a néhány percnél bejövő új információk felhasználásával mindig újrafuttatják a nowcasting eljárásokat. Az így kapott előrejelzések 1–2 órára előre már egy előrejelzés együttest alkotnak, amelyből ismét valószínűséget lehet számolni. Mindez megköveteli, hogy a nowcasting előrejelző programok nagyon gyorsak legyenek, ami az eljárás lineáris voltából következik, szemben a nagyságrendekkel többet számoló, komplex dinamikus modellel.



8. ábra. Az OMSZ nowcasting rendszerének alapelve. A függőleges tengely a légkör állapotát reprezentálja, a vízszintes tengely az időt. A valódi légkör állapotának változását a sötét-kék vonal mutatja, a két világoskék korong a valódi légkört leírni próbáló objektív analíziseket jelzik. A piros szaggatott vonal azt mutatja, hogy a dinamikus modell hogyan jelzi előre a légkör változását. A fekete vastag vonal a nowcasting előrejelzés trajektóriája, a fekete szaggatott vonal pedig a lineáris extrapolációval készített közelítést mutatja

például az Egyesült Államok középső részén. „Meteorológiai védelmet” először a Balaton kapott, ahol közel 80 éve működik a viharjelző rendszer. Ennek elsődleges oka, hogy az erős kategóriába tartozó szél (40 és 60 km/ó között) már veszélyt jelent a kis csónakokra, az úszókra, a viharos kategória alsó határát jelentő 60 km/ó széllekek pedig már a nagyobb hajókat is igénybe veszik. A szárazföldön ugyanakkor 70 km/ó várható szélnél adják ki a legalacsonyabb (citrom színű) veszélyjelzést. A balatoni viharjelzésnél az első fok az erős (40–60 km/ó), míg a másodfok a viharos (> 60 km/ó) széllekek várható bekövetkezésére figyelmeztet. A másodfokú viharjelzés kiadása számos korlátozással jár, aminek a nyaralók érthetően nem örülnek. A jelzés elmaradása viszont közvetlenül emberéleteket veszélyeztethet. Mint láttuk, egyetlen helyi zivatar is bőven elegendő a viharos kategóriájú széllekek kiváltására, vagyis a legkisebb skálán fellépő instabilitások és bizonytalanságok is közvetlenül érintik a tavi viharjelzést. 2006 óta az egész országra készülnek figyelmeztető előrejelzések és riasztások a hazai meteorológia és a nowcasting széles eszköztárát felhasználva.

Az utóbbi években egyre nagyobb számban szerveznek tömegrendezvényeket, ahol százezrek jönnek össze szabadtéri fesztiválokra vagy koncertekre kitéve egy váratlan vihar okozta szélviharoknak vagy jégesezőnek. Van olyan nyári rendezvény, ahol tízezer ember ússza át a Balatont egyszerre és egy váratlanul feltámadó 40 km/ó szél is végzetes tömegszerencsétlenséghez vezethet. A szervezők, a rendezők, vagy a biztonságért felelős emberek egyértelmű választ várnak, hogy megtartható-e az adott esemény. Legyen szó egy szabadtéri koncertről, vagy egy, a történelem menetét meghatározó partraszállásról, döntési lánc legelőjén a meteorológia van, a sokféle instabilitással és bizonytalansággal, amely gyakran nem teszi lehetővé az egyértelmű és határozott választ.

IRODALOM

- [1] Götz Gusztáv (2001): *Káosz és prognosztika*. Országos Meteorológiai Szolgálat, Budapest.
- [2] Götz Gusztáv (1993): *A pillangó-effektus. A káosz felfedezése a meteorológiában*. Fizikai Szemle 1993/12.
- [3] Seres András Tamás, Fodor Zoltán, Horváth Ákos (2007) *Januári vihar Európában*. LÉGKÖR 52:(2) pp. 6–9. (2007).
- [4] Horányi András, Szépszó Gabriella, Szűcs M.: *Valószínűségi meteorológiai előrejelzések. Áldás vagy átok?* Természet Világa jelenlegi száma.
- [5] Horváth Ákos, Nagy Attila Simon André (2013). *A dunai árvíz időjárási háttere*. Természet Világa 2013. augusztus.
- [6] Horváth Á, I Geresdi, P Németh, K Csirmaz, F Dombai (2009): *Numerical modeling of severe convective storms occurring in the Carpathian Basin*. ATMOSPHERIC RESEARCH 93: pp. 221–237. (2009)
- [7] Horváth Ákos (2007). *A légköri konvekció és a budapesti vihar augusztus 20-án*. Természet Világa 2007. május.
- [8] Lorenz, Edward N., 1963: *Deterministic Nonperiodic Flow*. J. Atmos. Sci., 20, 130–141.

HASZPRA TÍMEA

Világjáró részecskék a légkörben

Az Eyjafjallajökull vulkán kitörésének és a fukushimai balesetnek a tanulságai

Miert volt szükséges Európa-szerte légtérzárat elrendelni 2010-ben az izlandi Eyjafjallajökull vulkán kitörése miatt? Hogyan lehetséges, hogy a vulkántól mindössze 150 km-re található keflavíki repülőtér mégis fogadhatta a gépeket? Hogyan juthatott el a vulkáni hamu egészen délre, még Spanyolország térségébe is, majd terjedt szét szinte az egész kontinens felett a légkörben? Miért voltak – ha más nem is – a meteorológiai körülmények viszonylag szerencsésnek mondhatók a fukushimai atomerőmű balesete idején? Hogyan lehet előre jelezni a légköri szennyeződések útját, és ezáltal idejében felkészülni a szükséges légiforgalmi korlátozásokra vagy kitelepítésekre? Milyen tényezők befolyásolhatják ezen előrejelzések bizonytalanságát? Hogyan jelenik meg a káosz az aeroszolrészecskék¹ sodródásában? A cikk ezekre a kérdésekre igyekszik válaszolni.

A különböző légköri szennyeződések (gázok, aeroszolrészecskék) terjedésének számítógépes vizsgálata, előrejelzése fontos feladat, hiszen akár az emberi tevékenységből származó, akár a természetes eredetű kibocsátások komoly hatással lehetnek az élővilágra, társadalomra, gazdaságra, és nem eleget, ha csupán aktuális mérésekből (pl. műholdképről) tudunk tájékozódni. Gondoljunk csak a fukushimai balesetet követő intézkedésekre, vagy idézzük fel például az említett vulkánkitörést: a hamu egy része a szigetre visszahullva okozott nehézségeket, a levegőben sodródván pedig, amellett, hogy megnehezítette az utazni vágyók életét, a légitársaságoknak is hatalmas gazdasági veszteségeket okozott.

A szennyeződések nyomában

A szennyeződések nagy távolságú mozgását légköri terjedési modellekkel lehet követni. Az ún. lagrange-i terjedési modellek

¹ Aeroszol: gáznemű közeg és benne finoman eloszlott szilárd vagy cseppfolyós részecskék együttes rendszere. A szilárd vagy cseppfolyós részecskéket aeroszolrészecskéknek nevezzük. Aeroszolrészecskékről csak vízgőzben telítetlen levegő esetén beszélünk, a köd-, felhő-, csapadékcseppeket nem tekintjük azoknak.

1. A határsebesség

A határsebesség kicsiny, m_r tömegű, r sugarú gömbnek feltételezett aeroszol részecskékre az ún. Stokes-féle közegellenállási erő ($6\pi r \rho_v w_{\text{határ}}$) és a nehézségi erő ($m_r g = 4r^3 \pi/3 \cdot \rho_r g$) egyensúlyából számítható (a részecskékre ható felhajtóerő elhanyagolható):

$$w_{\text{határ}} = \frac{2\rho_r r^2 g}{9\rho_v} \quad (2)$$

A határsebesség tehát függ a részecske r sugarától és ρ_r sűrűségétől. g jelöli a nehézségi gyorsulást, ρ és v a levegő sűrűsége, illetve kinematikai viszkozitása. A közelítés a légkörben általában előforduló $\rho_r \approx 1000\text{--}2000 \text{ kg/m}^3$ sűrűségű részecskékre $r \leq 10\text{--}15 \text{ }\mu\text{m}$ sugárral érvényes. Érdemes megjegyezni, hogy $r = 1\text{--}10 \text{ }\mu\text{m}$ -es részecskékre a határsebesség nagyságrendje $10^{-4}\text{--}10^{-2} \text{ m/s}$, amely összemérhető a levegőben jellemző függőleges áramlási sebességekkel, ezért a részecskék függőleges áthelyeződését a határsebesség mellett a fel- és leáramlások ténylegesen is befolyásolják.

A nagyobb részecskék határsebessége a négyzetes közegellenállási törvényből számolva ennél lényegesen nagyobbak, $1\text{--}10 \text{ m/s}$ nagyságrendűnek adódik. Ennélfogva ezek a részecskék hamar, akár a légkörbe való kikerülésük után néhány perccel visszahullnak a földre, és helyileg okozhatnak problémákat (pl. az Eyjafjallajökull kitörésekor visszahulló hamu).

A gázok $r = 0 \text{ }\mu\text{m}$ sugarú, $w_{\text{határ}} = 0 \text{ m/s}$ határsebességű részecskéknél tekinthetők, amelyek az (1) egyenlet alapján minden időpillanatban a környező levegő sebességével haladnak.

alkalmasak a szennyeződés minden egyes részecskéje útvonalának (trajektóriájának) meghatározására a részecskék mozgásegyenlete alapján. Ennek a módszernek előnye, hogy a különböző tulajdonságú (pl. eltérő sűrűségű vagy méretű) részecskék egyedileg kezelhetők. A mozgásukat a Newton-egyenlet írja le, amely a légköri terjedés skáláján a következő egyenlettel egyszerűsödik:

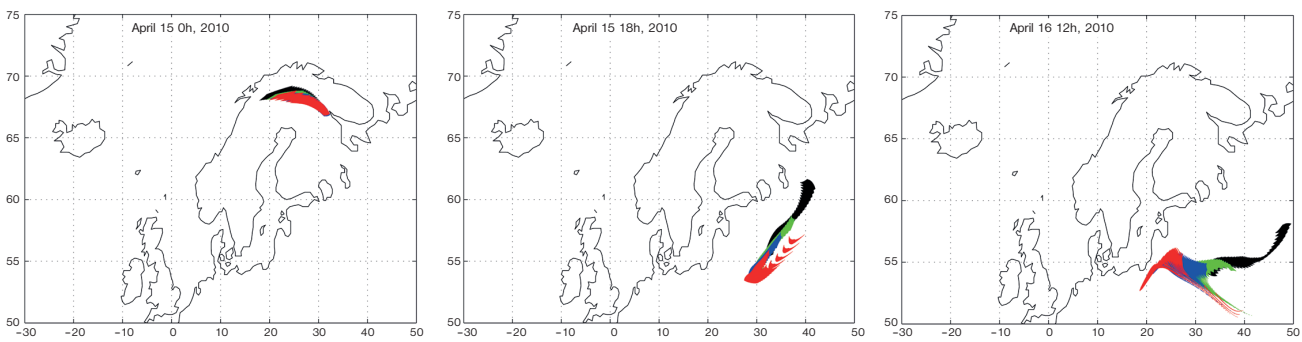
$$\mathbf{v}_r = \mathbf{v} - w_{\text{határ}} \mathbf{n}. \quad (1)$$

Itt \mathbf{v}_r jelöli a részecske sebességvektorát, \mathbf{v} a szélesebbésgvektor a részecske helyén, $w_{\text{határ}}$ pedig a részecskére ható nehézségi erő hatását figyelembevevő ún. határsebesség (a részecske ezzel a sebességgel esne álló levegőben), \mathbf{n} egy függőlegesen felfelé irányuló egységvektor. Parányi aeroszolrészecskékre a határsebesség négyzetesen növekszik a részecskemérettel (l. **1. blokk**). Az (1) egyenlet annyiban tér el Károlyi György írásában vagy a tavi és folya-

mi sodródást bemutató cikkeken használt mozgásegyenlettől, hogy megjelenik benne a részecskék süllyedését jellemző, $w_{\text{határ}}$ -t tartalmazó tag is.

A részecskék pályáinak numerikus számítása, vagyis a részecskék kis időközönkénti újabb és újabb helyzetének meghatározása alapvetően az (1) egyenletben megjelenő \mathbf{v} és $w_{\text{határ}}$ szél- és határsebesség ismeretében zajlik. A légkör felszínhez közeli néhány km-es rétegében a csapadék is jelentős szerepet játszik: felhő- vagy esőcseppel ütközve, azokba kerülve a részecske kimosódhat a felszínre, így a szimulációkban ezt a hatást is célszerű figyelembe venni.

A terjedési számításokhoz szükséges meteorológiai változók (pl. szélesebbésg, csapadék) múltbéli adatokat tároló adatbázisokból (pl. ERA-Interim [1]) vagy előrejelzési mezőkből származhatnak. Ezek a mennyiségek általában egy háromdimenziós rácsban állnak rendelkezésre néhány óras fel-



1. ábra. A Eyjafjallajökull vulkánból származó, különböző méretű aeroszolrészecskékből álló vulkáni hamufelhők szimulációja 18, 36 és 54 órával a „kitörés” után (piros: $r = 1 \mu\text{m}$, kék: $r = 2 \mu\text{m}$, zöld: $r = 5 \mu\text{m}$, fekete: $r = 10 \mu\text{m}$)

bontásban, így minden számítási időlépésben (aminek a nagysága az adatok időbeli felbontásának töredéke, csak néhány perc) a részecskék pillanatnyi sebességének (1) egyenlet szerinti meghatározása a megfelelő rácpontokból a részecskék helyére való térbeli és időbeli interpolációval történik.

Kicsiny méret – hosszú élet

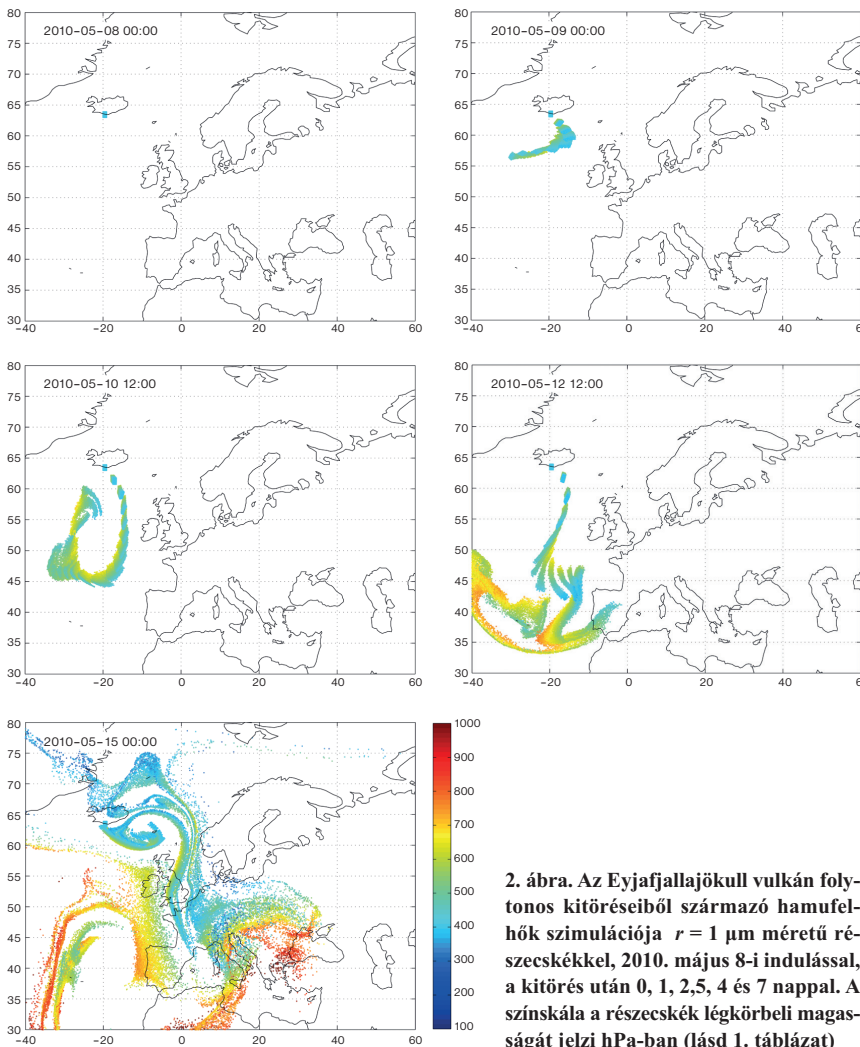
Tekintsünk meg néhány szimulációt, hogy képet kapjunk a vulkáni hamura jellemző sűrűségű ($\rho_r = 2000 \text{ kg/m}^3$), de különböző méretű aeroszolrészecskék levegőbeli vándorlásáról [2]. Az Eyjafjallajökull valóságban bekövetkezett, többé-kevésbé folyama-

tos pöfögésének számítógépes szimulációja helyett először érdemes csak egy-egy pillanatszerűen kibocsátott vulkáni hamufelhő útját nyomon követni. Az 1. ábra egy kezdetben $100 \text{ km} \times 100 \text{ km} \times 4 \text{ km}$ -es, 2010. április 14-én 6 UTC-kor az Eyjafjallajökull fölött kb. 5,5 km magasságban induló hamufelhő terjedését mutatja felülnézetből a következő néhány napra vonatkozóan. A hamufelhő útvonala és szétterjedése nagymértékben függ attól, hogy mekkora méretű szemcsékből áll, hiszen a különböző sugarú, de azonos kiindulási helyzetű részecskék az eltérő határsebességük miatt hamar más-más magasságban találják magukat, ahol az épp uralkodó különböző szélviszonyok hatására egyre jobban elkülönülnek eltérő méretű „párjuktól”, mind magasságban, mind vízszintes irányban. A nagyobb részecskék hamarabb közelítik meg a felszínt, és mivel a felszín közelében általában kisebb és eltérő irányú szélebségek lépnek fel, vízszintesen „lemaradnak” és más irányban terjednek apróbb társaikhoz képest. Ahogy a keretezett írásban említettük, az igazán nagy ($r > 100 \mu\text{m}$) hamuszemcsék már rögtön a kezdet kezdetén annyira hátramaradnak, hogy szinte azonnal kiülednek, és csupán 10–100 km-es távolságokat tesznek meg a levegőben.

Kedvező és kedvezőtlen áramlási viszonyok

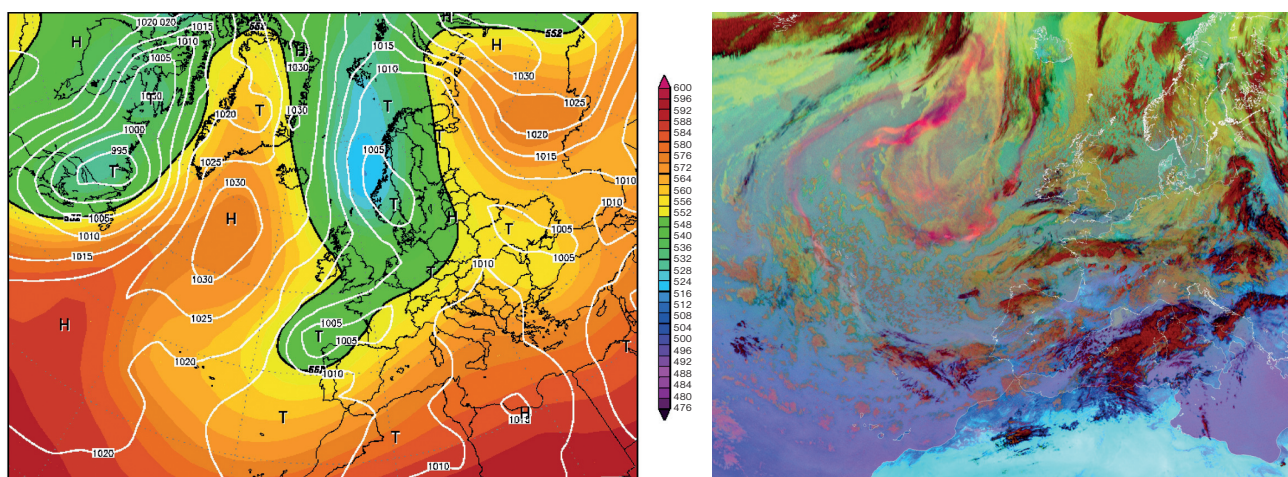
Az Eyjafjallajökull kitörései

A 2. ábrán az Eyjafjallajökull vulkán folyamatos kitöréseiből származó, kizárólag $1 \mu\text{m}$ -es részecskéket tartalmazó hamufelhők terjedésének szimulációja látható a 2010. május 8–15. közötti időszakra, a 3. ábra pedig a május 10-i meteorológiai helyzetet mutatja be. A részecskék mozgását mindig az éppen aktuális áramlási mező határozza meg, ezt követik nagyon rövid ideig. A 2. és 3. ábrát összevetve máris választ kapunk a cikk elején felvetett kérdések egy részére. Ezekben a napokban Izland térségében észak-

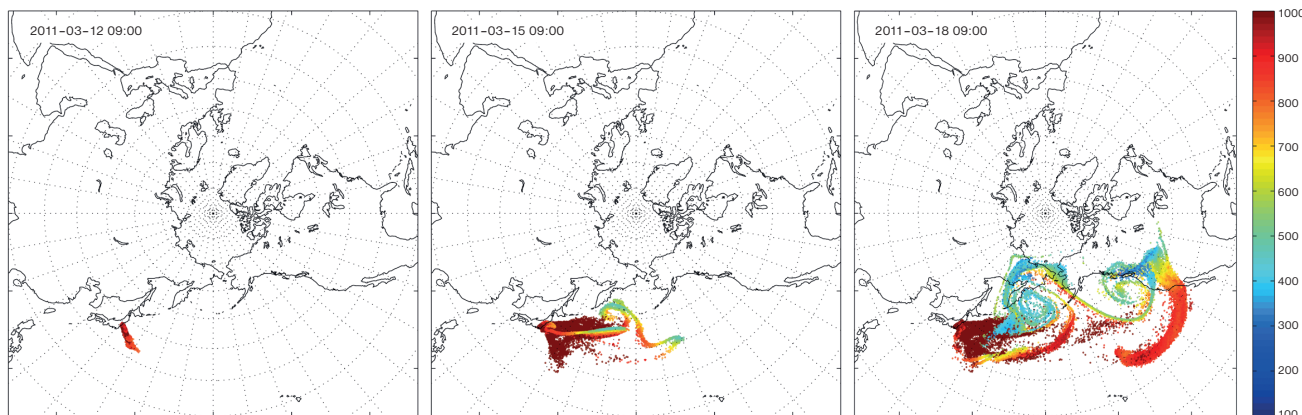


2. ábra. Az Eyjafjallajökull vulkán folyamatos kitöréseiből származó hamufelhők szimulációja $r = 1 \mu\text{m}$ méretű részecskékkel, 2010. május 8-i indulással, a kitörés után 0, 1, 2, 5, 4 és 7 nappal. A színskála a részecskék légköri magasságát jelzi hPa-ban (lásd 1. táblázat)

Világjáró részecskék a légkörben



3. ábra. 2010. május 10-i időjárási helyzet. Bal: felszíni nyomásmező [hPa] (fehér kontúr, H: magas nyomás – anticiklon, T: alacsony nyomás – ciklon) és az 500 hPa-os geopotenciális magasság³ mezeje [dekaméter] (színes kontúr) [3]. Jobb: műholdkép (rózsaszín: vulkáni hamu, sárga/drapp: alacsony szintű felhők, piros: vastag felhők) [4]



4. ábra. A fukushimai atomerőmű balesetekor kiáramló radioaktív szennyeződés terjedésének szimulációja a légkörben. A skála a részecskék légkörbeli magasságát jelzi hPa-ban (a barna felel meg a kiülepedett részecskéknek, lásd 1. táblázat) ($r = 1 \mu\text{m}$)

ki, északnyugati szél volt jellemző (amely dél, délkelet irányba fúj), így fordulhatott elő, hogy habár a keflavíki repülőtér meglehetősen közel fekszik a vulkánkitörés helyszínéhez, mégis nyitva tarthatott [5, 6], mivel a hamufelhő az uralkodó szelek miatt mindvégig elsodródott Izlandtól délre, délkeletre elnyúlva, és nem a sziget felett terjeszkedett egyre nagyobbra, mint egy festékpaca. A 2. ábra május 10-i képén jól megfigyelhető, hogy a hamu az Izlandtól délnyugatra található anticiklon (3. ábra bal oldali képén H-val jelölve) áramlási tartományába kerül, és az óramutató járásával egyezően visszatekeredik észak felé. A május 10-i szimuláció és az ugyanezen a napon készült műholdképen (3. ábra jobb ol-

² Az UTC a Coordinated Universal Time (egyezményes koordinált világidő) rövidítése.
³ A geopotenciális magasság körülbelül az adott nyomási szint tengerszint feletti magasságának felel meg.

p [hPa]	z [km]
1000	felszín
850	1,5
700	3
500	5,5
300	9

1. táblázat. A nyomási szintek átlagos magassága

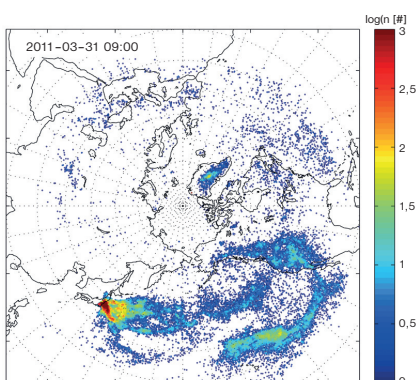
dali kép) rózsaszínben megjelenő hamufelhő elhelyezkedése között jó egyezés tapasztalható. Ugyancsak az atlanti térségben fellépő északias szelek eredményeképpen juthatott el a vulkáni hamu egészen az Izlandtól 2000 km-nyire lévő Ibériai-félszigetig is, ezzel akadályozva az Atlanti-óceán fölötti légi közlekedést [6]. A 2. ábra május 15-i képén a légkörbe kikerülő részecskék jól kirajzolják egy Izlandtól dél-

keletre felépülő ciklon, illetve egy Portugáliától nyugatra elhelyezkedő anticiklon áramlási viszonyait, valamint az is figyelemre méltó, hogy a részecskék a Skandináv-félszigettől eltekintve szinte egész Európa fölött megtalálhatók (a koncentráció nem mindenütt lépte át a légtérzár elrendeléséhez szükséges határértéket). A 2. ábra színeiből arra is következtethetünk, hogy ugyanakkora méretű részecskék ugyanannyi idő alatt egészen eltérő magasságokba kerülhetnek.

A fukushimai baleset

Míg 2010 tavaszán az Eyjafjallajökull kitöréssorozata meglehetősen rosszul érintette az európai repülőjáratokat, közel egy évvel később, a Fukushima Daiichi atomerőmű balesetekor a radioaktív szennyeződés levegőbe kiáramló része szerencsére nem a kontinens, hanem nagyrészt a Csendes-óceán felé sodródott. A 4. ábra a folyama-

tos, minden időlépésben a becsült kibocsátással arányos számú, $r = 1 \mu\text{m}$ sugarú részecske terjedését mutatja be, barnával jelölve a felszint elérő, kiülepedett részecskéket. A légköri szennyeződés többnyire délkelet-kelet-északkelet irányban hagyja el Japán térségét. A márciusi 15-i képen már észrevehető, ahogy az aeroszolrészecskék a Kamcsatka-félszigetnél ciklonális (az óramutató járásával ellentétes) irányban áramlanak, a 18-i kép pedig még szebben elénk tárja ezt, illetve az észak-amerikai partnál kifejlődött ciklont, amelyekbe belesodródva a részecskék a feláramlásokkal a felszínhez közeli légrétegekből nagyobb magasságokba (6–7 km) jutnak. A magasba került aeroszolrészecskék számára így lehetővé vált, hogy a szabad légköri áramlásokkal világszerte útra induljanak. A radioaktív szennyeződés néhány nap alatt – természetesen kisebb koncentrációban, felhígulva – megjelent, mérhető volt Észak-Amerikában, majd kis idő elteltével Európában is [7]. Az **5. ábrán** a felszint elérő, kiülepedett részecskék számsűrűsége (azaz az ülepedési „koncentrációmező”) igencsak inhomogén eloszlást mutat. A legjelentősebb

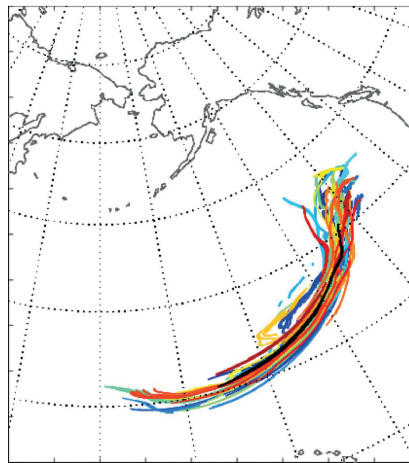


5. ábra. A kiülepedett radioaktív szennyeződés eloszlása. A színek a március 12–31. időszakban kiülepedett részecskék számát (n) jelölik 50 km×50 km-es területi felbontásban

kiülepedés Japán partjainál lép fel, de a féltekén máshol is található olyan területek, ahol sok részecske távozott a légkörből. Ez a fraktálszerű kiülepedési mintázat a vulkáni hamura és mindenfajta kisméretű szennyező részecskére jellemző, és a sodródás kaotikus jellegének a következménye.

Megbízhatunk-e egyetlen előrejelzésben?

Természetesen akár vulkánkitörésből, akár ipari balesetből származó aeroszolrészecskék vagy gázok útját próbáljuk nyomon követni, előre jelezni a légkörben, a szimulációk mindenképpen többféle hibával lesz-



6. ábra. Kicsit különböző kezdeti feltételekből induló meteorológiai előrejelzések felhasználásával kapott $r = 1 \mu\text{m}$ sugarú részecskékből álló szennyeződésh felhők 2,5 nap után. A szennyeződésh felhők 2011. március 12-én 0 UTC-kor kb. 5,5 km magasságban indultak Fukushima felett egy kb. 70 km×70 km-es terület-ről

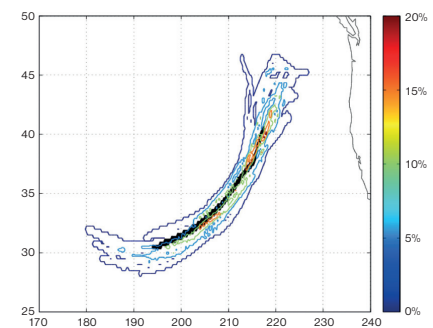
nek terheltek. Egyrészt, ha a számításhoz szükséges összes felhasznált adatot tökéletesen pontosan is ismernénk, a terjedési modell által figyelembe vett folyamatok, numerikus közelítések befolyásolnák a szimulációk eredményét. Ráadásul a legtöbb esetben csak meglehetősen nagy bizonytalansággal lehet megbecsülni a kibocsátott szennyeződésh felhők kezdeti jellemzőit, és mindemellett a felhasznált meteorológiai adatok is számos bizonytalanságot tartalmaznak, amelyek többek között a meteorológiai mezők kezdőfeltételeiben rejlő hibák következményei. Ez utóbbi hatásának feltárására az „alap”, nagyfelbontású előrejelzés mellett, feleakkora felbontással, kissé perturbált kezdeti állapotokat előállítva is lefuttatják a meteorológiai modellt. Így jön létre a több előrejelzési tagból álló valószínűségi (ensemble, sokasági) előrejelzés. Az ensemble előrejelzés szétartásából az előrejelzések bizonytalanságára következtethetünk, amint azt Horányi András és munkatársainak írása részletesen bemutatja [8].

A többi hibaforrástól ideiglenesen eltekintve, csupán a meteorológiai előrejelzésből fakadó, a terjedésszámításra hatással lévő bizonytalanságok tanulmányozása céljából tekintünk például egy feltételezett, pillanatszerű, kezdetben kb. 70 km×70 km-es kiterjedésű szennyeződésh felhőt Fukushima felett [9]. A **6. ábra** mutatja a kissé különböző kezdeti feltételekből induló meteorológiai előrejelzések felhasználásával készített szimulációkban a szennyeződésh felhők elhelyezkedését a Csendes-óceán fölött 2,5 nap után. A fekete a nagyfelbontású előrejelzés, a színek pedig a különböző ensemble tagok alkalmazásával nyert szenny-

yeződésh felhőket jelölik. Mi az, ami első ránézésre szembeötlő? A szennyeződésh felhők az összes szimuláció szerint erősen megnyúltak kezdeti kompakt alakjukhoz képest. Ezt az ezekben a napokban a Japántól keletre található futóáramlás⁴ segítette elő: a részecskék a futóáramlás erős, nyugatias szeleivel kelet felé sodródtak, és a térségben fellépő nagy szélnyírásnak köszönhetően erőteljesen széthúzódtak. A másik fontos észrevétel a különböző színű felhők eltérő elhelyezkedése, amelyek együttesen egy több száz kilométer széles sávon terülnek el.

Akár légtérzár, akár kitelepítés elrendelése kerül területekre, kockázatelemzésnél célszerű megbecsülni, hogy a szennyeződés merre sodródhat, hol haladhat meg a koncentrációja bizonyos határértéket és ennek mekkora a valószínűsége. Az előbbi példában a nagyfelbontású előrejelzés szennyeződésh felhőjéhez képest jóval nagyobb terület fölött jelennek meg a sokasági előrejelzés felhői. A **7. ábrán** a kontúrvonalak arra utalnak, hogy az ensemble tagok hány százaléka szerint lép túl az adott térségben a koncentráció egy bizonyos határértéket. Ebből képet nyerhetünk az adott területre az esemény bekövetkezésének valószínűségéről.

Láttuk tehát, ha pusztán csak az egyik bizonytalansági forrás hatását vizsgáljuk is,

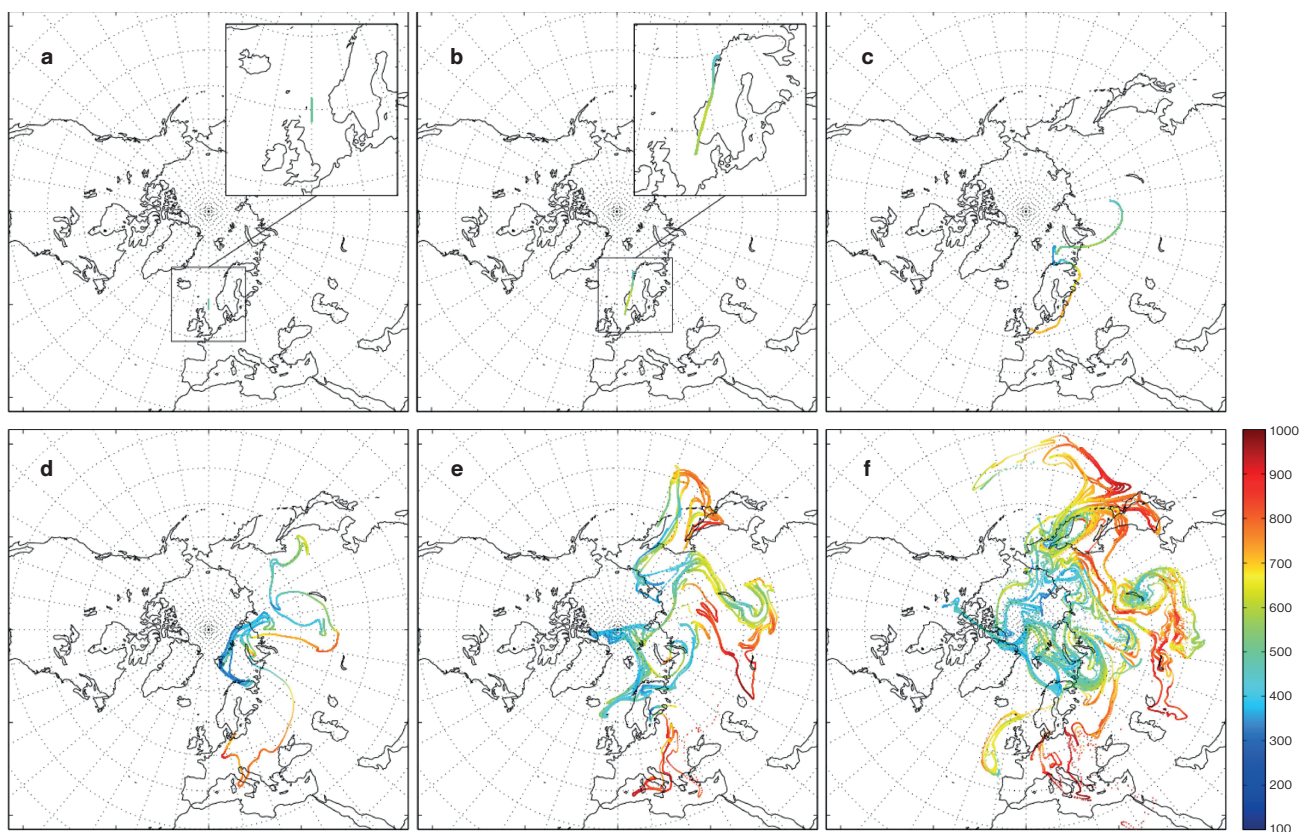


7. ábra. A térképen a kontúrvonalak (2, 5, 10, 15, 20%) mutatják, hogy az ensemble tagok hány százaléka szerint számíthatunk az adott terület fölött szennyeződésre. A fekete szín jelöli azt a területet, ahová a nagyfelbontású előrejelzés felhasználásával jutnak a részecskék ($r = 1 \mu\text{m}$)

a szennyeződésterjedési szimulációk között igen nagy különbségek alakulhatnak ki. Így, ha lehetőség van rá, érdemes minél több hibaforrásra tekintettel lenni, hogy minél teljesebb képet kapjunk az előrejelzés bizonytalanságáról, a modellezett esemény lehetséges kimeneteleiről.

⁴ Futóáramlás: erős, keskeny áramlás egy közel vízszintes tengely mentén, amit a szélsőséges erős nyírása jellemez.

Világjáró részecskék a légkörben



8. ábra. Egy kezdetben $L_0 = 3^\circ \approx 333$ km hosszúságú $r = 0$ μm sugarú, gáznak megfeleltethető részecskékből álló egydimenziós szennyeződéshő sodródási képe 0, 2, 4, 6, 8 és 10 nappal a kibocsátás után. A színskála a részecskék magasságát jelzi hPa-ban (lásd 1. táblázat)

Girbegurba, tekervényes fonalak – megmutatkozik a kaosz

Az előző esettanulmányokból kiderült, hogy egy-egy kezdetben kicsiny, kompakt szennyeződéshő alakja hamar eltorzul. A 8. ábra egy vonalszerű felhő fejlődését követi nyomon, amely a sodródás folyamán egyre jobban megnyúlik, széthúzódik, és a szál a különböző légköri képződményekbe kerülve újabb és újabb kacsokat növeszt, az áramlások hatására „összegyűrődik”. Egykét hét elteltével a megfigyelő a félteke tekintélyes részét behálózó, tekervényes, kacsoringós fonal képével találkozhat. A részecskék sodródása során a kaotikus visel-

kedés nyilvánul meg abban, hogy a részecskék felhő szálas, fraktál alakzatban terjed szét, és teríti be az egyik, majd előbb-utóbb a másik féltekét is, bonyolult mintázatot kialakítva. A szennyeződéshők nyúlásának mértéke az ún. topologikus entrópiával jellemezhető (1. 2. blokk).

A véges idejű (pl. 10 nap alapján számolt) topologikus entrópia jól megfigyelhető földrajzi és évszakos eloszlással rendelkezik (9. ábra). Ahogy az áramlási viszonyok alapján várható is, a közepes szélességeken (főként a féltekék téli évszakában) lépnek fel a legnagyobb értékek ($0,6 - 0,9$ nap⁻¹) a folyamatosan kialakuló, fejlődő és elhaló ciklonok erős keverési és nyírási hatásának

következtében, míg a legkisebb értékekkel ($0,2 - 0,5$ nap⁻¹) a nyugodtabb trópusokon találkozhatunk [11].⁵

A világszerte kalandnak vége szakad

Az aeroszolrészecskék a nehézségi erő, a levegő fel- és leáramlásainak, valamint a csapadék hatására, amint ez az előzőekből is kiderült, előbb-utóbb elhagyják a légkört, kiülepednek a felszínre. Az ülepedési mező, ahogyan ez az 5. ábrán is jól kitűnik, egyáltalán nem egyenletes, hanem térben és időben meglehetősen inhomogén eloszlást mutat. Mivel a felszínhez közel haladó szennyezőanyagokat alkalmasszerűen az éppen hulló csapadék is könnyen kimoshatja a légkörből, rövid időtartamot tekintve a csapadékeloszlás mintázata és a légkörből távozó részecskék helyzetének rajzolata között erős kapcsolat fedezhető fel (lásd pl. 10. ábra).

⁵ Néhány hónap elteltével egy szennyeződéshő részecskéi (amennyiben elég aprók, és ezalatt nem távoztak számottevő mennyiségben a légkörből) már többeképp jól elkeverednek a féltekék fölött, így a fonal hosszának növekedése, a topologikus entrópia értéke sem függ a kiindulási helyzettől.

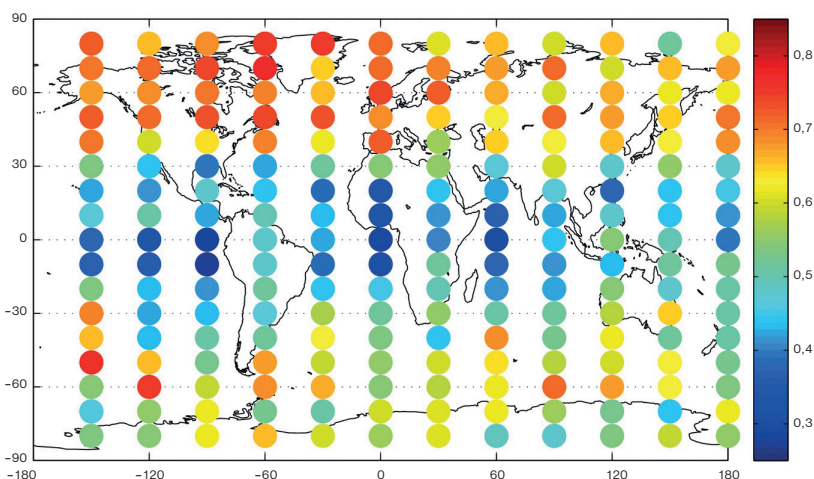
2. Topologikus entrópia

A tapasztalat szerint kaotikus rendszerekben a gyorsan girbegurba alakzattá fejlődő fonalak L hossza időben sebesen, nevezetesen exponenciálisan növekszik

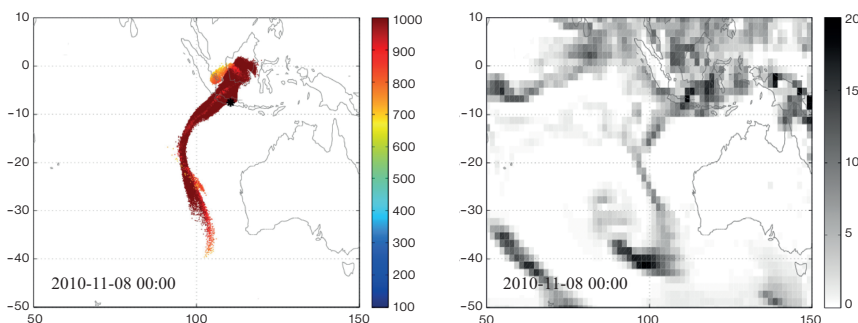
$$L(t) \sim e^{ht},$$

és a növekedés mértékét a topologikus entrópia (h) írja le [10]. Ez a mennyiség a bonyolultság, szabálytalanság mérőszámának is tekinthető. A topologikus entrópia általában nagyobb a Ljapunov-exponensnél, és bizonyos esetekben, így a légköri szennyeződésterjedés kapcsán is, meghatározása egyszerűbb, mint a Ljapunov-exponensé. Jelen esetben h viszonylag nagy értékének ($0,89$ nap⁻¹) köszönhetően a 10. napon a felhőfonál hossza már több ezerszeresére gyarapodott, és meghaladja a 10^6 km-t.

Haszpra Timea: Világjáró részecskék a légkörben



9. ábra. A körökkel jelölt földrajzi helyekre 10 napos intervallumokra számolt és időben átlagolt topologikus entrópia [nap^{-1}] a 2009. december – 2010. február időszakban



10. ábra. Bal: A Merapi vulkán kitéréséből származó hamufelhő szimulációja $r = 10 \mu\text{m}$ sugarú részecskékkel, 2010. november 1. 0 UTC indulással, a kitérés után 7 nappal (kezdeti méret: $100 \text{ km} \times 100 \text{ km} \times 4 \text{ km}$, kibocsátás magassága: 5 km). A színskála a részecskék légköri magasságát jelzi hPa-ban (lásd 1. táblázat). Jobb: Teljes csapadékmennyiség a november 8. 0 UTC-t megelőző 12 órában mm-ben

Mivel jellemezhető általánosságban a részecskék légkörből való kiürülésének üteme, és mi mondható el ennek méretfüggéséről? A tapasztalatok szerint a t idő után levegőben maradó részecskék $n(t)/n(0)$ aránya néhány nap után időben exponenciálisan csökken:

$$\frac{n(t)}{n(0)} \sim e^{-kt} = e^{-t/\tau}$$

A kiülepedési folyamat gyorsasága tehát k -val, a kaotikus jelenségek körében használatos mennyiséggel, az ún. szökési rátával számszerűsíthető (l. Gruiz Márton megfelelő keretezett írását). Ennek reciproka az elterjedtebben használt és a való életben jobban mérhető átlagos légköri

tartózkodási idő τ . Az exponenciális összefüggés csapadéktevékenység hiányában és annak jelenlétében is fennáll. Természetesen a várakozásoknak megfelelően az utóbbi esetben fokozottabb az ülepedés, nagyobb a szökési ráta, és rövidebb az aeroszolrészecskék átlagos légköri tartózkodási ideje.

Meglepő módon az átlagos tartózkodási idő rendkívül gyorsan (ismét exponenciálisan) csökken az r részecskemérettel:

$$\tau(r) \sim e^{-kr},$$

ahol $k \approx 0,21 - 0,28 \mu\text{m}^{-1}$. Parányi szennyezőkre τ a kibocsátás z magasságától függően elérheti a 20–60 napot is, míg az $r = 10 \mu\text{m}$ sugarú részecskék akár a nap töredékére, néhány óra alatt is kipotyoghatnak a levegőből⁶. Érdeemes utánagondolni, hogy ha azzal a naiv feltételezéssel élünk, hogy a részecskék csupán a saját határsebességükkel hullanak lefelé, az átlagos élettartamuk ($\tau = z/w_{\text{határ}}$) négyzete-

sen függne r -től a (2) határsebesség r -függéséből adódóan. Az exponenciális méretfüggés arra utal, hogy a légáramlások jelentősen hozzájárulnak a részecskék légköri kaotikus mozgásának kialakulásához.

A tanulság dióhéjban

A természetes eredetű és az emberi tevékenységből származó aeroszolrészecskék és gázok légköri mozgását mind környezet-, mind egészségvédelmi szempontból fontos modellezni. A terjedési modellek szimulációit mérések segítségével tesztelhetjük. Fontos megjegyezni azonban, hogy a modell által készített előrejelzések számos hibával terheltek, amiket érdemes szem előtt tartani az eredmények kiértékelése során. A részecskék kaotikus sodródásának, dinamikájának jellemzésére a nemlineáris rendszerek elméletéből ismert fogalmak (pl. a szennyeződésfelhők nyúlását jellemző topologikus entrópia vagy az ülepedési folyamat ütemét leíró szökési ráta) jól alkalmazhatók.

A munka az OTKA NK100296 pályázat támogatásával készült.

IRODALOM

- [1] ERA-Interim: http://data-portal.ecmwf.int/data/d/interim_daily/
- [2] Haszpra, T. T. (2011): *Volcanic ash in the free atmosphere: A dynamical systems approach*. Journal of Physics: Conference Series, 333, 012008.
- [3] Wetterzentrale: <http://www.wetterzentrale.de/topkarten/fsreaeur.html>
- [4] EUMETSAT: http://oiswww.eumetsat.org/WEB-OPS/medialib/medialib/images/2010_05_10_1200_m8_rgb_24hmicr.jpg
- [5] http://en.wikipedia.org/wiki/Air_travel_disruption_after_the_2010_Eyjafjallajökull_eruption
- [6] http://en.wikipedia.org/wiki/2010_eruptions_of_Eyjafjallajökull#Effects_of_the_ash_plume_on_air_travel
- [7] Stohl, A., P. Seibert, G. Wotawa, D. Arnold, J. F. Burkhardt, S. Eckhardt, C. Tapia, A. Vargas, and T. J. Yasunari (2012): *Xenon-133 and caesium-137 releases into the atmosphere from the Fukushima Dai-ichi nuclear power plant: determination of the source term, atmospheric dispersion, and deposition*. Atmospheric Chemistry and Physics 12, no. 5, 2313–2343.
- [8] Horányi A., Szépszó G. és Szűcs M. (2013): *Valószínűségi előrejelzések: áldás vagy átok?* Természet Világa, jelenlegi szám.
- [9] Haszpra, T., Lagzi, I., and Tél, T. (2013): *Dispersion of aerosol particles in the free atmosphere using ensemble forecasts*. Nonlinear Processes in Geophysics, 20, 759–770.
- [10] Tél T. és Gruiz M. (2002): *Kaotikus dinamika*. Nemzeti Tankönyvkiadó, Budapest.
- [11] Haszpra, T. and Tél, T. (2013): *Topological entropy: a Lagrangian measure of the state of the free atmosphere*. Journal of the Atmospheric Sciences. doi: 10.1175/JAS-D-13-069.1.

⁶ A szimulációinkban megbecsült tartózkodási idő nem veszi figyelembe az eltérő tulajdonságú részecskék kémiai és fizikai átalakulásait, ezért a különböző szennyeződések mért tartózkodási idejéhez képest előfordulhatnak eltérések.

HOMONNAI VIKTÓRIA

A légköri hosszú távú kapcsolatok titka

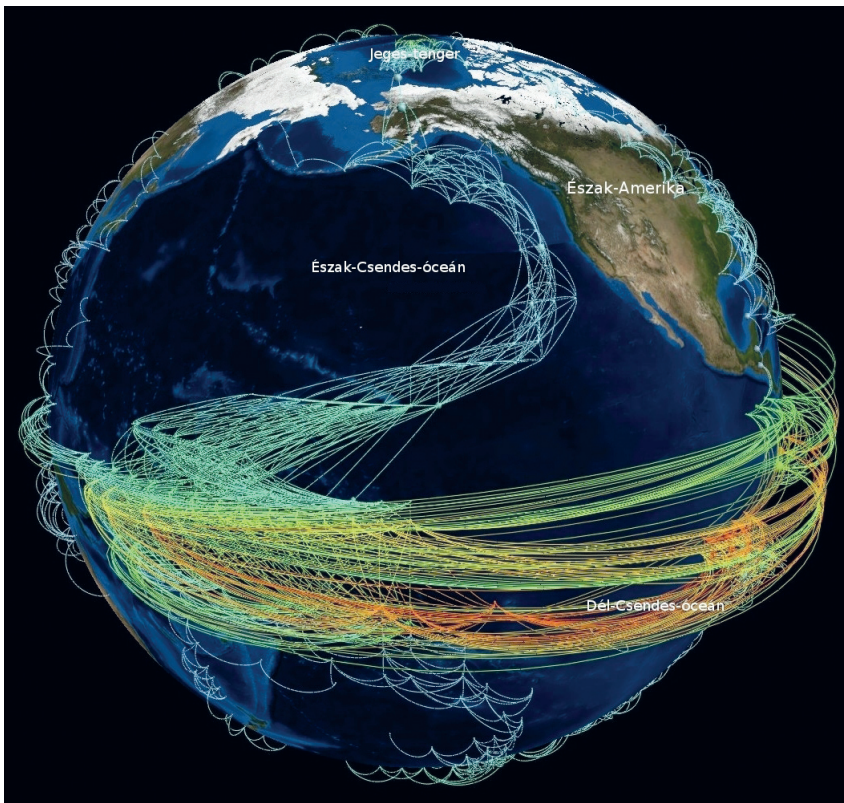
A légköri cirkulációban jelentős változékonyságok mutatkoznak több időskálán is: a pár napostól (egy vihar vagy frontátvonulás) a pár hetes-hónapos skálán át (az évszakhoz képest hűvösebb vagy éppen melegebb idő) a több éves, akár évszázados változásokig (klíma-változás). A pár napos változékonyságot a mai modern numerikus időjárási előrejelző modellek többsége jól visszaadja, azonban már több hétre előre e modellek segítségével sem vagyunk képesek előrejelezni az időjárást a légkör kaotikussága miatt. Az évszázadok alatt zajló változások paleoklimatológiai vizsgálatokból ismertek, manapság pedig klímamodellek segítségével próbáljuk előrevetíteni, vajon milyen lesz a századvég időjárása [1].

A köztes, pár hetes-hónapos változékonyság vizsgálatára alkalmas módszer a távkapcsolatok felderítése. A távkapcsolatok általában nagy földrajzi területeket fednek le, amelyek gyakran távol esnek egymástól. Ezek többnyire hosszú időskálájú változékonyságok, tipikusan néhány hétig-hónapig tartanak, vagy esetleg pár egymást követő évet ölelnek fel. A távkapcsolatok általában (légköri) oszcillációk formájában figyelhetők meg, vagyis két távoli terület bizonyos meteorológiai jellemzőjében (gyakran a tengerszinti nyomásban vagy a felszíni hőmérsékletben) periodikus váltakozás mutatkozik. A földi távkapcsolati rendszert jól érzékelteti az **1. ábra**, ahol a Földön a távol eső, mégis összefüggő pontokat kötik össze a színes vonalak [2, 3]. Jól

látható például, ahogy az egész Dél-Csendes-óceánt átívelik a vonalak, csak az óceán két partja között mutatva erős kapcsolatot. Ugyanennek a távkapcsolati hálózatnak másik nézete látható a címlapunkon.

Általában az oszcillációkat valami egyszerű index bevezetésével mérik, amely a pozitív és negatív fázis megkülönböztetését teszi lehetővé. Az index nagysága az ingadozás erősségére utal, de elsősorban csak a két fő ellentétes fázis kielemezése történik, ebből próbálunk becslést adni, hogy akár az oszcillációtól távoli helyen milyen időre lehet számítani. A módszertől nem várhatunk pontos értékeket, többnyire évszakokat jellemezhetünk így, amiben természetesen szélsőséges események is előfordulhatnak. Egy-egy terület időjárását több légköri oszcilláció is befolyásolja, amelyek együttes eredője nyilvánul végül meg az időjárásban, így általában a távkapcsolatokon alapuló előrejelzések beválása kicsi valószínűségű. Nem csoda tehát, hogy az évszakokra készített előrejelzések még mindig csak gyerekcipőben járnak. Próbálkozások persze mindig akadnak, emlékezzünk csak vissza, milyen jóslatok érkeztek az idei nyárra: voltak, akik úgy gondolták, hogy az idei lesz az elmúlt 200 év leghűvösebb nyara.

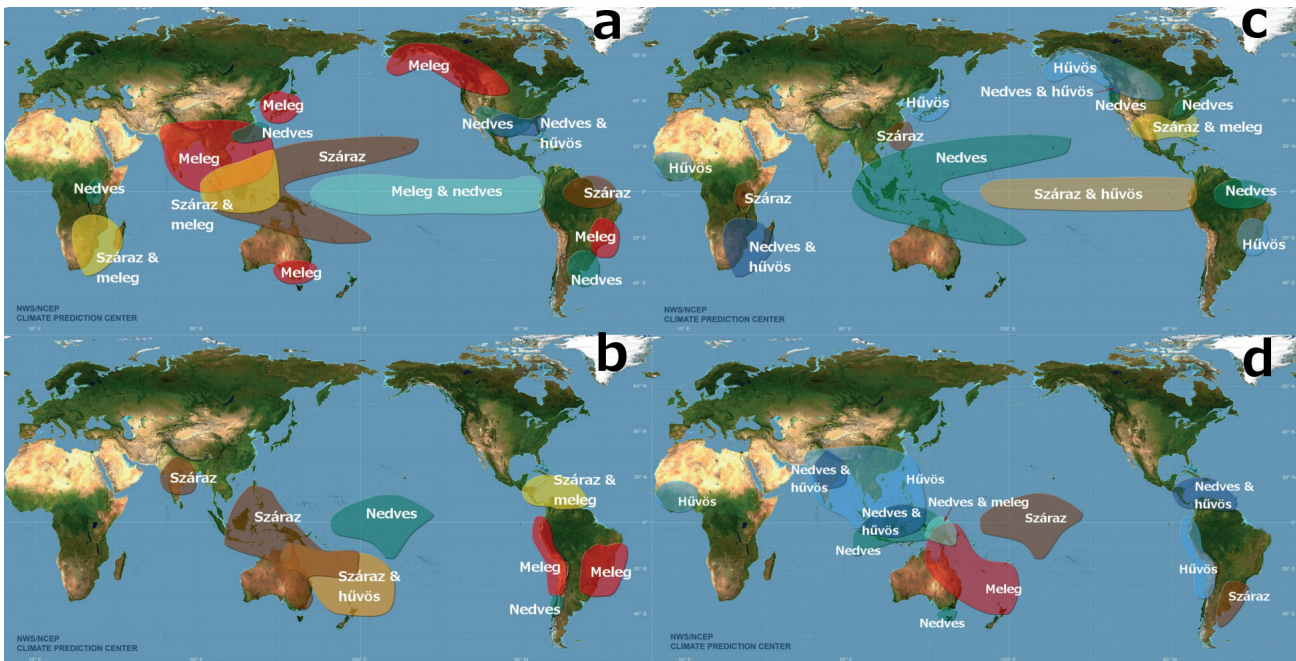
1. ábra. A földi távkapcsolati hálózat az elmúlt 60 év hőmérsékletváltozásai alapján [2,3]



El Niño-Déli oszcilláció

Az első távkapcsolat felfedezése Sir Gilbert Walker nevéhez fűződik, aki a XIX. század végén fedezte fel a Déli oszcilláció jelenségét. Észrevette, hogy a Csendes-óceán és az Indiai-óceán felett uralkodó légnyomások mindig ellentétesek: amikor az egyik medencében alacsony légnyomás uralkodik, akkor a másikban magas, és fordítva. Ezt a légnyomáskülönbséget hívatott mérni a Déli Oszcillációs Index (SOI), amelyet az észak-ausztráliai Darwin és Tahiti tengerszinti légnyomásának különbségéből számítunk. Ez az oszcilláció tulajdonképpen az El Niño és La Niña események légköri indikátora, ezért ezt a jelenséget együtt gyakran ENSO (El Niño-Southern Oscillation) rövidítéssel emlegetik.

Az El Niñónak, mint a legjelentősebb távkapcsolati rendszernek, meglehetősen nagy irodalmi háttere van, és sok helyen olvas-



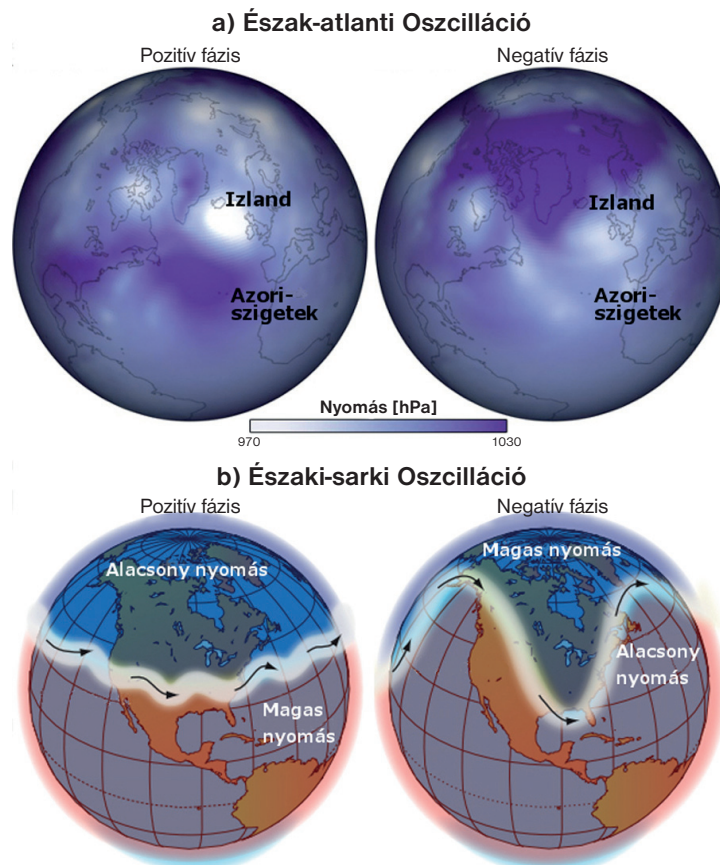
2. ábra. Az El Niño (a és b) és La Niña (c és d) események hatása a Föld különböző területein a téli (a és c) és nyári (b és d) időszakban. A színek jelentése: sötétbarna – száraz, türkiz – nedves, vörös – meleg, világoskék – hűvös, sárga – száraz & meleg, világosbarna – száraz & hűvös, világos türkiz – nedves & meleg, sötétkék – nedves & hűvös [7]

hatunk erről a jelenségről [pl. 4,5,6]. Ezért itt csak dióhéjban szeretném jellemezni ezt a jelenséget. El Niño évben Dél-Amerika északnyugati partjainál a tengerfelszín hőmérséklete megemelkedik, a tápanyagban gazdag hideg víz feláramlása elmarad, ezzel nagy károkat okozva a halászatban. Általában karácsony környékén kezdődik, nevét a „gyermek Jézusra” utalva kapta, 9–12 hónapig tart, és kb. 2–7 évente fordul elő. A La Niña ezzel szemben a szokásosnál hidegebb tengerfelszín-hőmérsékletet eredményez. A meleg (El Niño) fázisban a Csendes-óceán nyugati medencéjében magas légnyomás, míg a hideg (La Niña) fázisban alacsony légnyomás uralkodik, így lesz El Niño években a SOI negatív, La Niña években pozitív.

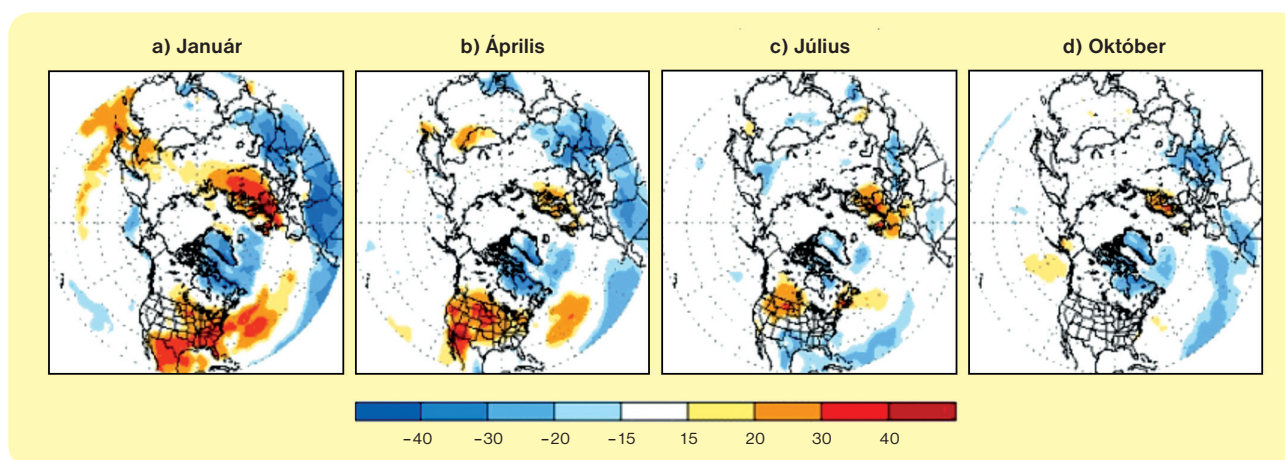
A 2. ábrán látható a meleg és hideg epizódok globális hatása télen és nyáron. El Niño évben a meleg tengerfelszínét követve a trópusi zivatar aktivitás keletre húzódik Indonéziából a Közép-Csendes-óceáni régióba, ezzel Észak-Ausztráliában szárazságot okozva télen és nyáron is. Az indiai monszun kevesebb csapadékot ad az északi félteke nyarán, különösen Északnyugat-Indiában, ahol a növénytermesztésre ez rendkívül kedvezőtlenül hat. Eközben a meleg fázisban az átlagosnál csapadékosabb időjárást figyeltek meg Észak- és Dél-Amerika bizonyos részein. La Niña években a globális hatások nagyjából ellentétesek a meleg epizódokban tapasztaltaktól.

Az ENSO kialakulásával kapcsolatban több elmélet is létezik. Az egyik szerint az

3. ábra. a) Az Észak-atlanti Oszcilláció és b) az Északi-sarki Oszcilláció pozitív és negatív fázisa. Pozitív fázisban mindkét oszcillációnál az erős nyugati áramlás a jellemző, míg negatív fázisban a meridiális (észak-dél irányú) áramlás a domináns [11,12]



A légköri hosszú távú kapcsolatok titka



4. ábra. Az Észak-atlanti oszcilláció pozitív fázisának hőmérsékleti távkapcsolatai a Föld északi féltekén a) januárban, b) áprilisban, c) júliusban, d) októberben. A piros színek jelzik az erős kapcsolatot, míg a kékek az antikorrrelációt, azaz ellentétes irányú kapcsolatot [13]

El Niño jelenség a légkör-óceán rendszer természetes oszcillációja, amelyet kaotikus folyamatok (időjárás) tesznek szabálytalan-ná. A másik elmélet a rendszert stabilnak feltételezi, ahol az El Niño eseményeket egy külső kényszer váltja ki, például a Madden-Julian-Oszcilláció (lásd később). Annak ellenére, hogy azt nem tudjuk megmondani, mikor kezdődik a következő El Niño esemény, ha már tapasztaljuk az első jeleit, a fejlődését már előre tudjuk jelezni [8].

A légköri-óceáni oszcillációk közül kétségtelenül a fent említett ENSO jelenségkör a legismertebb, de természetesen nem az egyetlen. A meteorológusok körében jól ismertek más óceáni vízfelszín-hőmérsékletekben bekövetkező változások, mint például az Atlanti-óceán Többévtizedes Oszcillációja (AMO) vagy az Évtizedes Csendes-óceáni Oszcilláció (PDO), illetve az óceáni medencék feletti nyomási ingadozások, mint például az Észak-atlanti Oszcilláció (NAO) és a Csendes-óceáni-Észak-amerikai Rendszer (PNA). A továbbiakban részletesebben csak az utóbbi két jelenségről lesz szó.

Észak-atlanti Oszcilláció

Európa időjárását elsősorban az Észak-atlanti Oszcilláció (North Atlantic Oscillation, NAO) határozza meg. Az Atlanti-óceán északi medencéjében Izland környékén többnyire alacsony nyomás uralkodik, míg az Azori-szigetek felett magas a légnyomás. Így e két terület nyomáskülönbségével jól jellemezhető ez az oszcilláció [9]. Mivel a magas és alacsony nyomás helyzete közel változatlan, így egy átlagos értéktől vett eltéréssel definiálhatunk pozitív és negatív fázist. A két szélsőséges állapot látható a 3a. ábrán. Pozitív NAO-index esetén nagyon erős a két gócpont, Izland térségében egy nagyon mély ciklon, míg az

azori térségben egy nagyon magas nyomású anticiklon található. Ilyenkor a nyugati szelek meglehetősen erősek, sok viharral hűvös nyarat és enyhe, nedves telet hoznak Közép-Európába. A negatív fázisban a nyomáskülönbség kicsi, a nyugati szelek gyengébbek, ilyenkor a viharok nyomvonala délebbre húzódik a Földközi-tenger felé, Európában ilyenkor köszöntenek be a zord telek (mint például 2009–2010-es tél). Meglepő módon azonban nem csak Európa időjárását határozza meg ez az oszcilláció. Erős kapcsolat mutatható ki Észak-Amerika és Ázsia északi területeinek hőmérséklet- és csapadékviszonyaiban (4. ábra), sőt még Japánban is.

A NAO ingadozásának alapvető mechanizmusa több mint 100 éves kutatás után még mindig nem teljesen tisztázott, annyit tudunk, hogy a légkör tranzienis örvényei fontos szerepet játszanak az oszcilláció vezérlésében, emellett úgy tűnik, hogy az El Niño és a NAO közötti kapcsolat meglehetősen gyenge [10].

Északi-sarki Oszcilláció

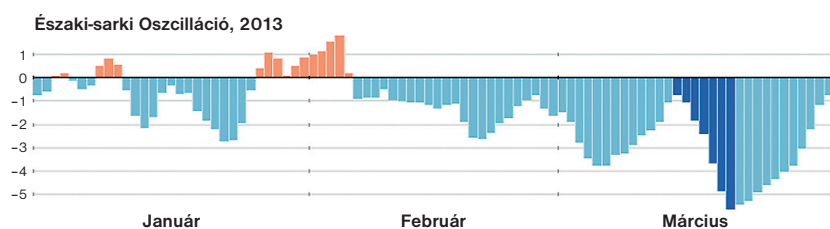
Rendkívülinek mondható téltre az idén is volt példa, mindannyiunkban még élénken él a márciusi hófúvással tarkított havazás, amely csaknem az egész országot megbé-

nította. Ugyan ki gondolná, hogy ezért a szélsőséges időjárásért a Föld egy másik pontjának időjárása a felelős? Pedig így van, méghozzá az Északi-sarki Oszcillációban (Arctic Oscillation, AO) mutatkozó erős anomália okozott zord időjárást egész Európában és Észak-Amerikában. Ez az index a sarki területek és a közepes szélességek közötti nyomáskülönbséget jellemzi (3b. ábra) [14]. Az 5. ábrán látható, hogyan alakult ennek az indexnek az értéke 2013. év elején. Csaknem végig negatív fázisban volt, márciusban különösen magas értékekkel [15]. Ez azt jelenti, hogy ilyenkor az áramlás meridionális (észak-dél irányú), így a meleg levegő egészen magas szélességeig jut, miközben a sarki hideg a mérsékelt övbe áramlik. Pozitív fázisban ezzel szemben a zonális (nyugat-kelet) irányú áramlás olyan erős, hogy a sarki hideg levegő az Északi-sark felett marad, nem tud a közepes szélességekre jutni.

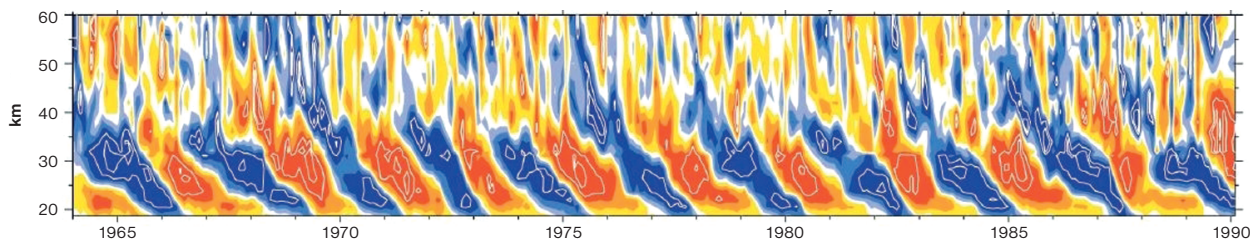
Az Észak-atlanti Oszcilláció erőteljesen kapcsolódik az Északi-sarki Oszcillációhoz, ezért a két változást együttesen figyelik. Mindkét oszcilláció elsősorban a téli időszakban jelentős.

A légköri áramlások szempontjából erős határ az Egyenlítő, emiatt az északi féltekén megjelenő oszcillációk hatása a déli félteke területére kicsi. Viszont a Föld déli féltekéjén szintén megfigyelhetők ingado-

5. ábra. Az Északi-sarki oszcillációs index értéke 2013. január 1-től április 1-ig. A sötétebb kék színekkel jelölt héten többek között Magyarországon is rendkívüli időjárás uralkodott [15]



Homonnai Viktória: A légköri hosszú távú kapcsolatok titka



6. ábra. A Kvázi-kétéves Oszcilláció (QBO) megjelenése az egyenlítői szélben. Vörös színek jelölik a nyugati szeleket, míg a kékek a keleti szeleket. [17]

zások, így például az Északi-sarki Oszcillációhoz hasonlóan megfigyelhető egy Antarktisz Oszcilláció (Antarctic Oscillation, AAO) is a Déli-sark környékén, amely az Antarktisz és a déli közepes szélességek időjárását befolyásolja.

Csendes-óceáni–Észak-amerikai Rendszer

A Csendes-óceáni–Észak-amerikai Rendszer (Pacific-North American Pattern, PNA) az Észak-atlanti Oszcillációhoz hasonlóan működik: a Csendes-óceán két pontja közötti nyomáskülönbséget méri, nevezetesen Hawaii és az alaskai Aleut-szigetek között. A NAO-hoz hasonlóan jellemezhető a két fázis, amely elsősorban Észak-Amerika időjárását befolyásolja a téli hónapokban, de hatása Kínában is érezhető. A PNA pozitív fázisában az átlagnál melegebb idő tapasztalható Nyugat-Kanadában, és az USA nyugati felén, míg délen, délkeleten átlag alatti hőmérsékletekkel találkozhatunk. Az Alaskai-öbölben és északnyugaton növekszik a csapadék mennyisége, míg az USA délnyugati részén ilyenkor a szárazság okoz problémákat. Ha a PNA-index negatív, a hatások megfordulnak.

Mai ismereteink szerint a PNA-mintázat kialakulásában a belső légköri folyamatok mellett az ENSO-nak is jelentős szerep jut, ugyanis a trópusi Csendes-óceán tengerfelszín-hőmérsékletének változására nagyon érzékenyen reagál ez a légköri oszcilláció [10].

Madden–Julian-Oszcilláció

A lista még közel sem teljes, említést kell tenni a trópusok jellegzetes oszcillációról is. Természetesen a legjelentősebb az ENSO, míg a trópusi vihartervékenységet a Madden–Julian-Oszcilláció (MJO) befolyásolja erősen. Ez egy 30–60 napos oszcilláció, amely során a trópusi csapadékos nem csapadékos régiók váltakoznak, miközben ez az ingadozás terjed kelet felé [16]. Periódusa olyan, hogy a Csendes-óceáni és az Atlanti-óceáni medence mindig ellentétes fázisban van, így a trópusi ciklonok aktivitása mindig csak az egyik óceáni medencében nagy. Mivel a trópusi

ciklonokat a Csendes-óceáni medencében tájfunoknak, míg az Atlanti-óceáni medencében hurrikánoknak nevezzük, ezért ez gyakorlatilag azt jelenti, hogy nem kell egyszerre pusztító hurrikántól és tájfunról is tartanunk. Az MJO hatása az ENSO-ra az El Niño és La Niña epizódok fejlődésében és intenzitásában mutatkozik meg, emellett az ázsiai és ausztrál monszuntevékenységre is hatást gyakorol ez az oszcilláció. Észak-Amerika Csendes-óceáni partvidékén előforduló özönvízszzerű esőzések (ún. Pineapple Express) kialakulásában is szerepet játszik ez az ingadozás.

Kvázi-kétéves Oszcilláció

A troposzféra a Föld légkörének alsó kb. 10 km-es rétege, ahol az időjárási folyamatok zajlanak. E fölött helyezkedik el a sztratoszféra, amelynek magas ózontartalma megvédi minket a káros ultrabolya sugaraktól. Ebben a rétegben is megfigyelhetők olyan jelenségek, amelyek hatnak a troposzféra áramlataira. Ilyen a trópusok időjárását befolyásoló kiemelt ingadozás a Kvázi-kétéves Oszcilláció (Quasi-biennial Oscillation, QBO), amely egy kvázi-periodikus oszcilláció a sztratoszferikus egyenlítői zónális szélben. A nyugati és keleti szelek váltakozásának periódusa 20 és 36 hónap között változik, az átlagérték kb. 28 hónap, vagyis valamivel több, mint két év, nevét is innen kapta ez a jelenség [17]. Kb. 30 km-es magasságból indul a jel, amely havonta 1 km-t halad lefelé a sztratoszféra aljáig. A keleti szelek általában erősebbek és rendszertelenebbek, mint a nyugatiak (6. ábra).

A szélben jelentkező oszcilláció a légkör kémiai összetevőinek eloszlására is hatással van, ezért ez az ingadozás a hőmérséklet mellett például a teljes ózonszint időszoriban is jelentkezik. Hatása a sztratoszféra cirkulációjában is tetten érhető erősítve, vagy éppen gyengítve azt. Ez az oka annak, hogy a QBO jele a trópusok mellett a sarkvidéki területeken mutatható ki elsősorban.

Összefoglalás

A távkapcsolatok feltérképezése, és még inkább a mögöttük rejtőző dinamika meg-

értése izgalmas témakör. Mivel ezek a jelenségek pár hetes, hónapos időskálán mozognak, ahol lehetetlen küldetés előre jelezni az időjárást, ezért különösen fontos, hogy e folyamatok mozgatórugóit megértsük, és így esetleg hatékony előrejelzési eszközökké váljanak. Hiszen az évszakos előrejelzésekre az igény a társadalom és ipar felől meglehetősen nagy, gondoljunk csak a nyaralás megtervezésére vagy például egy energiaszolgáltató esetében jó előre tudni, mennyivel növekszik meg az energiafelhasználás egy szokásosnál hidegebb tél során. ❄

A munka az OTKA NK100296 pályázat támogatásával készült.

IRODALOM

- [1] <http://www.cpc.ncep.noaa.gov/data/teledoc/telecon- tents.shtml>.
- [2] Juergen Kurths: *Networks with Complex Topology* című előadása, SIAM Conference on Applications of Dynamical Systems, Snowbird, Utah, USA, 2013. Az ábrát Thomas Nocke, a Potsdam Institute for Climate Impact Research munkatársa készítette.
- [3] J. Donges et al., 2009: *The backbone of the climate network*. *Europhys. Lett.* 87, 48007.
- [4] http://www.atmosphere.mpg.de/enid/2_raml_si_rendszerek/_El_Ni_o_27k.html.
- [5] http://en.wikipedia.org/wiki/El_Ni%C3%B1o-Southern_Oscillation.
- [6] Trenberth, Kevin E., 1997: *The Definition of El Niño*. *Bull. Amer. Meteor. Soc.*, 78, 2771–2777.
- [7] http://www.nws.noaa.gov/climate/local_data.php?wfo=pajk.
- [8] Wang, C., C. Deser, J.-Y. Yu, P. DiNezio, and A. Clement, 2012: *El Niño-Southern Oscillation (ENSO): A review*. In *Coral Reefs of the Eastern Pacific*. P. Glynn, D. Manzello, and I. Enochs, Eds., Springer Science Publisher, in press.
- [9] Marshall, J., Kushnir, Y., Battisti, D., Chang, P., Czaja, A., Dickson, R., Hurrell, J., McCartney, M., Saravanan, R. and Visbeck, M., 2001: *North Atlantic climate variability: phenomena, impacts and mechanisms*. *Int. J. Climatol.*, 21: 1863–1898.
- [10] Liu, Z., and M. Alexander, 2007: *Atmospheric bridge, oceanic tunnel, and global climatic teleconnections*. *Rev. Geophys.*, 45, RG2005.
- [11] <http://www.climatewatch.noaa.gov/article/2011/long-distance-relationships-the-arctic-and-north-atlantic-oscillations>.
- [12] <http://www1.ncep.noaa.gov/pub/data/cmb/bams-sotc/2010/bams-sotc-2010-brochure-hi-rez.pdf>.
- [13] http://www.cpc.ncep.noaa.gov/data/teledoc/nao_tmap.shtml.
- [14] http://en.wikipedia.org/wiki/Arctic_oscillation
- [15] <http://earthobservatory.nasa.gov/IOTD/view.php?id=80804&src=eoaiotd>.
- [16] http://www.cpc.ncep.noaa.gov/products/precip/CWlink/MJO/MJO_summary.pdf.
- [17] Baldwin, M.P. et al., 2001: *The Quasi-Biennial Oscillation*. *Rev. Geophys.*, 39, 179–229.

MÁRFY JÁNOS–RÁCZ ZOLTÁN

A jégkorszakok rejtélyei

Az utóbbi évtizedekben jelentős fejlődésen mentek keresztül a Föld múltbeli klímájának rekonstrukcióját végző tudományágak. Így viszonylag pontos képünk van bolygónk mintegy 5 millió évet átölő éghajlatáról, illetve az ezen időskálán domináló eljegesedésekről, az ún. glaciális-interglaciális oszcillációkról. Cikkünkben áttekintjük a klímát nagyban meghatározó oszcillációk egyszerűbb mechanizmusait, s rávilágítunk a magyarázatul szolgáló elméletek erőnyeire és problémáira.

A mindennapi életben komplexnek az átláthatatlant, a nehezen érthetőt nevezzük, s nincs ez másképp a tudományban sem. Az éjszakai égen vándorló bolygók mozgását bonyolultnak, misztikusnak tekintették évezredekken keresztül (olyannyira, hogy a mozgás megértésére tett erőfeszítések köré felépült egy, az emberek félelmére és hisztikusára alapított szakma, az asztrológia). A jelenségek azonban egyszerűvé, érthetővé váltak, miután Kepler észrevette, hogy a bolygók pályái ellipszisek a Nap-centrikus leírásban, Newton pedig megmutatta, hogy az ellipszisek a gravitációs kölcsönhatásból következnek.

A klímatológia tudománya is komplex, misztikusnak tűnő dologgal foglalkozik; a Föld időjárásának több évtizedre (esetleg évszázadokra, évmilliókra) vonatkozó trendjeit próbálja felderíteni. Mivel az időjárás az embereket közvetlenül érinti, a hiányos megértés társadalmi vonzatai itt még súlyosabbak. Ismert például a középkorban egy több évszázados lehülési periódus (a kis jégkorszak, XIV–XIX. század), amelynek során a 0,5–1°C-os átlagos hőmérsékletcsökkenés erősen visszavetette a mezőgazdasági termelést, így a mai Németország és Svájc bizonyos területein egy évtized alatt néha nyolcszor nem volt termés, s ez éhínségekhez vezetett [1]. A nehézségekre az emberek bűnbakot kerestek és találtak is: mindennaposá vált az időjárásért felelőssé tett boszorkányok elégetése.

Jelenleg a Föld felszínének átlaghőmérséklete mintegy 0,5–1°C-kal magasabb a 100–150 évvel ezelőttihez képest. A köznapi szóhasználatban ezt nevezzük globális felmelegedésnek. Kiváltó okként az üvegházhatású gázok, főként a szén-dioxid légköri koncentrációjának növekedését jelölik meg a klíma modellek nagyszkalájú szimulációival foglalkozó kutatók. Az ipari forradalom óta az atmoszféra szén-dioxid-tartalma valóban mintegy 30%-kal emelkedett, s a növekedés egyik forrása az emberi tevékenység, ami persze felveti az emberek felelőségének (esetleg bűnösségének) kérdését. A teendők szükségessége, fontossági sorrendje és a terheket viselő körének meghatározása óhatatlanul politikai elemeket visz a kutatásokba, illetve az eredmények interpretációjába. Láthatóan itt is felbukkannak egy gyorsan fejlődő, részben még kiforratlan tudomány nemkívánatos melléktermékei.

Az asztrológia és a boszorkányok példái óvatosságra intenek a klimatikus változások súlyosságának és okainak megítélésében. A Föld klímája visszacsatolásokból álló komplex rendszer, amelyben különböző időskálán ugyan, de minden összetevőnek (légkör, óceánok, kontinensek, jégmezők, bioszféra stb.) lényeges szerepe van. Azonban a komponensek alapvető fontosságú tulajdonságai, mint például a légkör esetén a felhőképződés, valamint a komponensek kölcsönhatásai a mai napig nem teljesen tisztázottak. Így a tudományos eredményekkel szembeni kötelező kétkedéssel kell kezelnünk a modellek szimulációjából következő eredményeket, így azt is, hogy az utóbbi évszázad változásait gyakorlatilag a klímarendszer egyetlen eleme okozta.

A következőkben az elméletek problémáit szeretnénk érzékeltetni a hosszabb időskálán végbemenő klímaváltozások, a jégkorszakok modelljeinek vizsgálatával. Megmutatjuk, hogy a klíma-

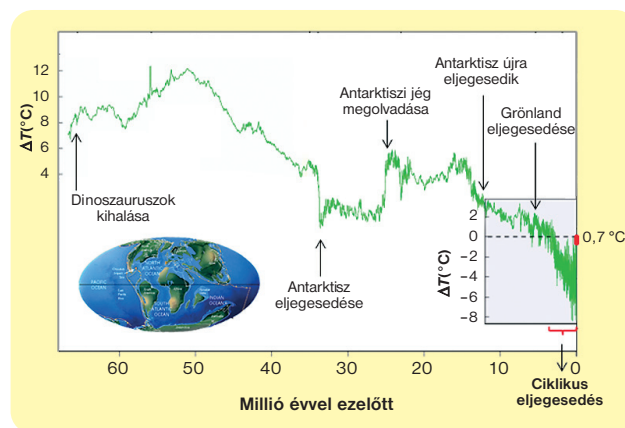
változások e dramatikus aspektusában, a ciklikusan megjelenő eljegesedésekben nem értjük sem a hőmérsékletváltozás nagyságát (8–10 °C), sem pedig a változások 100 ezer éves periódusát. Bár ezek a problémák nem tűnnek relevánsnak mindennapi időjárásunkban, a magyarázat hiánya azt jelenti, hogy a klímadinamika lényeges elemei hiányoznak az elméleti megfontolásokból, s ez megkérdőjelezi az éghajlati előrejelzések megalapozottságát.

Mit nevezünk jégkorszaknak?

A múltbeli hőmérséklet rekonstrukciója külön tudomány, amelynek részleteivel itt nem foglalkozunk. Fogadjuk el, hogy az antarktisi és grönlandi jégfuratokban, a mélytengeri üledékekben, valamint a geológiai kőzetformációkban mért izotópkoncentrációarányok [2] információt szolgáltatnak a Föld atmoszférájának hőmérséklet-változásairól, s e változásokat szeretnénk megérteni.

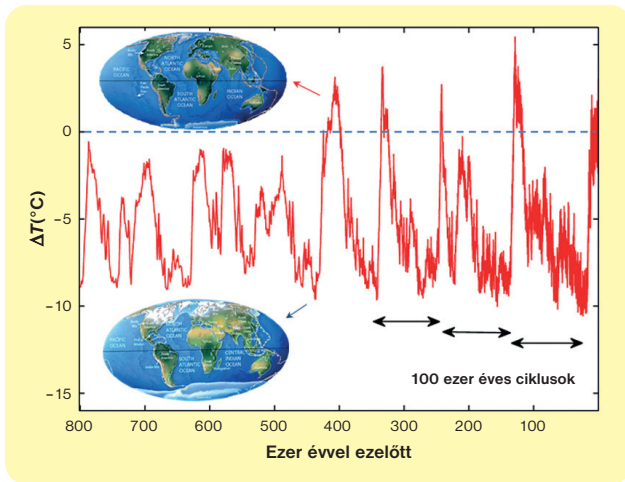
A jégkorszakot mint terminológiát a Föld történetének arra az időszakára használják, amikor valamelyik sarki területet jégtakaró fedi. Ebben az értelemben ma is jégkorszakban élünk, mely mintegy 35 millió éve kezdődött az Antarktisz eljegesedésével (1.

1. ábra. A Föld felszíni hőmérsékletének eltérése (ΔT) a szaggatott vonallal jelölt jelenkori értékhez képest, valamint a főbb klimatikus események az elmúlt 65 millió évben. A bal alsó sarokban a kontinensek 65 millió évvel ezelőtti elhelyezkedése látható. Fontos különbségek a jelenlegi helyzethez képest: Ausztrália még nem vált el az Antarktistól; India még útban van Ázsia felé, s ennek megfelelően még nincs Himalája; a közép-amerikai csatorna igen széles, s Európa nagy részét tenger fedi. A mélytengeri- és a jégfuratokból rekonstruált hőmérsékletek a jobb alsó sarokban kiemelt 12 millió éves időtartamra vonatkoznak. Az utóbbi 150 év hőmérséklet-emelkedésének nagyságát pedig a függőleges piros vonal mutatja



ábra). Bár a jégtakaró átmenetileg megolvadt, mégis a fenti időpontot tekintjük a „modernkori” jégkorszak kezdetének. Ekkor ugyanis a kontinensek részben átrendeződtek (az Antarktisz elvált a dél-amerikai és az ausztrál kontinensről), megteremtve a lehetőségét az Antarktisz körülölelő hideg tengeráramlat kialakulásához, ami jelentősen csökkentette az Egyenlítő és a sarkvidék közötti hőáramlást [3,4] és idővel (az Antarktisz második eljegesedésekor) stabilizálta a jégmezőket. Az 1. ábrán szintén megfigyelhető, 15 millió éve tartó általános lehülési folyamatra nem ismert meggyőző magyarázat, minden valószínűség szerint a tengeráramlatok folyamatos átrendeződése okozza.

Hétköznapi értelemben, bár megtévesztő, de jégkorszaknak nevezzük az utóbbi néhány millió évben *ciklikusan* bekövetkezett



2. ábra. A hőmérséklet változása az utolsó 800 ezer évben egy antarktisi jégfúratból rekonstruálva [5]. A nulla szintet itt az elmúlt 50 év átlaga adja (szaggatott kék vonal). Az ábra bal felső, ill. alsó sarkában a jégtakaró jellegzetes eloszlása látható a meleg (interglaciális), ill. a hideg (glaciális) időszakokban

erőteljes eljegesedéseket is. A jégkorszakon belül ciklikusan ismétlődő hidegebb (8–10°C-kal alacsonyabb hőmérséklet, kiterjedt jégtakaró) és melegebb (jelenlegi hőmérséklet, kisebb jégtakaró) időszakokat találunk (**2. ábra**), amiket glaciális-interglaciális oszcillációknak neveznek.

Eljegesedések az elmúlt 800 ezer évben

Az eljegesedések történetéről szóló legmegbízhatóbb információink a sarki jégmezőkből származnak [5]. Ahogy megyünk visszafelé az időben, az időmeghatározás egyre bizonytalanabbá válik, azonban a jégfúratok rétegeinek kormeghatározása az utolsó 800 ezer évre jónak mondható. Mivel a jégmagokból számolt hőmérséklet megegyezik más mérésekből, elsősorban a tengeri üledékekből kapott eredményekkel, a 2. ábra hőmérsékleti idősorára, mint magyarázatra váró kísérleti tényre tekinthetünk.

A 2. ábra jól mutatja, hogy az eljegesedéseket 100 ezer éves periódussal szakítják meg 10–20 ezer éves meleg időszakok. A gyors felmelegedéseket sokkal lassabb lehülési periódusok követik, általában a hőmérséklet menete fűrészfogszerű. A lehülésből következő drasztikus eljegesedés mértékét jól jellemzi, hogy az utolsó eljegesedés mélypontján (nagyjából 22 ezer évvel ezelőtt), az óceánok szintje 120–130 méterrel alacsonyabb volt a mainál, s az elpárolgott víz mintegy 50 millió km³-nyi szárazföldi jégként volt jelen.

1. A földi klíma belső időskálái

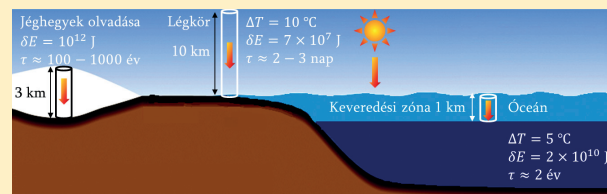
A **3. ábrán** 1 m² keresztmetszetű, tipikus magasságú légköri, óceáni és jégmező állapotváltozásokra számoltuk ki a fenti megfontolásból következő időskálákat. A légkör esetén 10 km az a karakterisztikus magasság, ahol a meteorológiai folyamatok lejátszódnak, s a hőmérséklet jellegzetes változása egy front esetén $\Delta T \approx 10^\circ\text{C}$. Mivel ismerjük a levegőoszlop átlagos sűrűségét ($\rho \approx 0,7 \text{ kg/m}^3$) és fajhőjét ($C \approx 1000 \text{ J/kg}^\circ\text{C}$), ezért az 1 m² keresztmetszetű oszlop energiaváltozását következőképp számíthatjuk:

$$\delta E = C \cdot \rho \cdot \Delta T \cdot 1 \text{ m}^2 \cdot 10 \text{ km} \approx 7 \cdot 10^7 \text{ J.}$$

Mivel az 1 m²-re eső energiaáram $J_E = 342 \text{ J/s}$, a változások karakterisztikus idejére a következő becslést kapjuk:

$$\tau = \delta E / J_E \approx 2 \cdot 10^5 \text{ s} \approx 2\text{--}3 \text{ nap,}$$

ami megfelel a szokásos meteorológiai változások időtartamának.



3. ábra. Jégmezők olvadásának, valamint jellegzetes légköri ($\Delta T \approx 10^\circ\text{C}$) és óceáni ($\Delta T \approx 5^\circ\text{C}$) állapotváltozások élettartama

A számolás hasonló az óceánra, csak a víz fajhőjét és sűrűségét kell behelyettesíteni a fenti formulákba, s a vízoszlop magasságaként a keveredési zóna magasságát kell használnunk. A jég esetén az olvadás igényli a legtöbb energiát, s az egységnyi tömegre vonatkozó olvadáshő ($Q \approx 3 \cdot 10^5 \text{ J/kg}$) határozza meg a szükséges energiát ($\delta E = Q \cdot \rho \cdot 1 \text{ m}^2 \cdot 3 \text{ km} \approx 10^{12} \text{ J}$). Végül is a légkörre a ciklonok időskáláját, az óceánokra pedig a nagyobb örvények élettartamát kapjuk meg, ezek azonban nem összemérhetők a 100 ezer éves periódussal. A nagy jégmezők kialakulása és olvadásra már hosszabb időszakot vesz igénybe ($\tau \approx 100$ év), különösen ha figyelembe vesszük a jégmezők nagyobb fényvisszaverő képességét (albedóját) és a napsugárzás kisebb beesési szögét magasabb szélességeken. De így is csak egy nagyságrendet tudunk növelni az időskálán, s a $\tau \approx 1000$ év a kívánt 100 ezer éves periódusnál még mindig két nagyságrenddel kisebb!

A 2. ábra magyarázatát keresve kézenfekvő a kérdés, hogy következhet-e a 100 ezer éves periódus a légkör, az óceánok, vagy a jégmezők dinamikájából. Ennek megválaszolásához vizsgáljuk meg, hogy a klíma említett összetevőinek természetes változásai milyen időskálán mennek végbe (l. részletesebben az **1. blokkban**). Induljunk ki abból, hogy minden változáshoz energiára (δE) van szükség, s a változás gyorsaságát, az energiát szállító energiaáram (J_E) nagysága határozza meg. A klíma esetében az energiaáram elsősorban a Napból jön, s a Föld felszínére vett átlagban $J_E = 342 \text{ W/m}^2$. Tehát, ha a változáshoz (pl. jellegzetes hőmérsékletváltozáshoz) szükséges energiát egységnyi keresztmetszetű, tipikus magasságú lég-, víz-, vagy jégoszlopokra számoljuk, akkor az energia-megmaradásból következik, hogy $\delta E = J_E \tau$, azaz a folyamat jellegzetes időskálája $\tau = \delta E / J_E$.

Az 1. blokkban kapott becslések arra utalnak, hogy az eljegesedések periódusát, legalábbis első közelítésben, nem a Föld éghajlati rendszerének belső oszcillációiban kell keresni. Természete-

A jégkorszakok rejtélyei

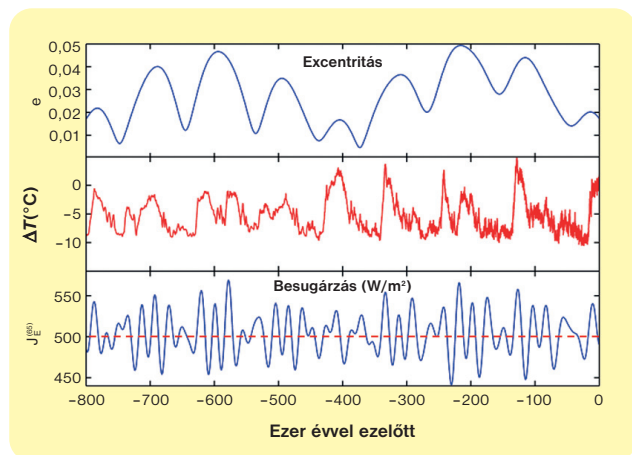
sen nem kizárt, hogy visszacsatolások sorozata révén kaphatunk 100 ezer éves periodicitású folyamatot, mint ahogy erre példát is mutatunk a későbbiekben, de egy ilyen időskálájú folyamat inkább külső okot sugall. A becslésekből az is látható, hogy a jégmezők dinamikája lassú, tehát bármi is legyen a külső ok, az valószínűleg a jégmezőkhöz csatolva fejt ki hatását.

A Milanković-elmélet

Külső ok sok minden lehet (a Nap tulajdonságainak változása, a Naprendszer mozgása a Tejút porfelhőin keresztül stb.), a 100 ezer éves periódus mindenestre csillagászati hátteret sejtet. Az első jelentős elmélet e területen Milanković [6] nevéhez fűződik, aki észrevette, hogy a Föld pályájának változásai befolyásolják a Naptól jövő energiaáramot. Mivel főképp ez az áram hajtja a klímát, javaslata szerint ez okozza annak változásait is. A **4. ábra** legfelső görbéje a Föld pályájának excentricitását (a kör alaktól való eltérés jellemzőjét) mutatja. Feltűnő, hogy az excentricitás 100 ezer éves periódussal változik, s maximumai jó közelítéssel egybeesnek az antarktisi jégfuratok hőmérsékleti maximumaival (*4. ábra* középső görbe). Ezt az egybeesést kombinálva a mechanikából ismert ténnyel, hogy a Naptól érkező átlagos energiaáram nő a pálya excentricitásával, egy egyszerű elmülethez jutunk: a Földnek két jellegzetes klímaállapota van, s az eljegesedés állapotából a 100 ezer évente, az excentricitással együtt megnövekvő energiaáram visz át a magasabb hőmérsékletű állapotba.

Az excentricitáson alapuló elméletnek a **2. blokkban** tárgyalt problémáin túl vannak egyéb nehézségei is (pl. a 100 ezer éves periódus mellett a *4. ábrán* látható egy igen nagy amplitúdójú 400 ezer éves periódus is, aminek semmi jele sincs a hőmérsékleti adatokban), ezért Milanković is elvetette az eljegesedések ezt a magyarázatát. Kifinomultabb elmélete is csillagászati okokra, de nem a földpálya excentricitására vezet vissza a 100 ezer éves periódust. Elméletének két alapfogolata közül az első könnyen érthető. Az eljegesedéseket a jégmezők mozgása határozza meg, s ebből következően a besugárzás J_E változásai a jégmezők határán lényegesek. A második gondolat is kézenfekvőnek tekinthető. Mivel a szárazföldek léte az északi féltekén megkönnyíti a jégmezők kiterjedését, ezért a ciklusok eljegesedést meghatározó folyamatok itt mennek végbe. E két feltételezésből következik, hogy a klímát meghatározó faktor a jéghatárok mozgási tartományában, a 60–70. északi szélességi fokon bejövő energiaáram, tehát ennek változá-

4. ábra. Az excentricitás (felső ábra) és az antarktisi jégfuratból rekonstruált hőmérséklet (középső ábra), valamint a Föld pályaelemeiből közelítően a jégmezők határán, a 65. északi szélességi fokon számított napi átlagos besugárzás június közepén (alsó ábra)



sait kell vizsgálnunk. Itt már lényeges szerepet játszanak olyan csillagászati faktorok is, mint a Föld forgástengelyének a pálya normálisától mért hajlásszöge (jelenleg $23,5^\circ$, egyébként 41 ezer éves periódussal billeg $22,1^\circ$ és $24,5^\circ$ között), valamint a forgástengely irányítottágának változása (a tengely 23 ezer év alatt ír le egy teljes fordulatot).

Vizsgáljuk tehát a 65. északi szélességi fokon érkező sugárzást (*4. ábra* alsó része), ami első ránézésre igen bonyolultnak tűnik. Az átlag körül nagyjából 10%-os fluktuációkat látunk a bejövő energiaáramban, ami előző számításaink alapján már tud $8\text{--}10^\circ\text{C}$ -os hőmérsékletváltozást okozni. A besugárzás jelében azonban csak 41 és 23 ezer év periódusidejű komponensek láthatók (ami érthető, hiszen mind a forgástengely dőlésszöge, mind pedig a forgástengely irányítottága erősen befolyásolja a magas szélességi fokon beérkező sugárzást). A jelben ott van a 100 ezer éves periódus is, de szabad szemmel nem látható, mivel az amplitúdója két nagyságrenddel kisebb, mint a másik két komponensé.

Most tehát a kérdés az, hogy a besugárzási jelből (*4. ábra* alsó része) hogyan lesz a megfigyelt hőmérsékleti jel (*4. ábra*, középső része). Természetes feltételezni, hogy erős besugárzás esetén a jégmező kiterjedése csökken, s a jégmezők visszavonulásával a hőmérséklet nő, s fordítva, kis besugárzás a jégmezők növekedésére és egyben a hőmérséklet csökkenésére vezet. A kétfajta visel-

2. Az excentricitáson alapuló elmélet kritikája

Sajnos a fenti elmélet magában hordja a komplex rendszerek magyarázatának egy gyakran előforduló problémáját. Hogy ezt megértsük, becsljük meg elméletünkben az excentricitás okozta hőmérsékletnövekedést, ami a megfigyelések szerint $8\text{--}10^\circ\text{C}$ (lásd a *2. ábrán*). A Föld pályája közel kör, s az excentricitás maximális értéke kicsi $e \approx 0,05$ (lásd *4. ábra*). A számítások szerint a Naptól jövő energiaáram J_E változása, ΔJ_E az excentricitás négyzetével arányos ($\Delta J_E \approx e^2 J_E$), s a relatív változás nagyságrendje $\Delta J_E / J_E \approx e^2 \approx 10^{-3}$. Az okozott hőmérsékletváltozást megbecsülhetjük abból a feltételezésből, hogy a Föld abszolút hőmérsékletét (T_F) a bejövő és kimenő energiaáramok egyensúlya $J_E = J_E^k$ határozza meg, s a Föld fekete testként sugároz, azaz (a *Stefan-Boltzmann-törvény* szerint) a kimenő áram a hőmérséklet negyedik hatványával arányos, $J_E^k = aT_F^4$. Kis változások esetén a hőmérséklet megváltozását (ΔT_F) következőképpen írhatjuk:

$$\Delta J_E^k = a(T_F + \Delta T_F)^4 - aT_F^4 \approx 4aT_F^3 \Delta T_F = 4J_E^k \frac{\Delta T_F}{T_F},$$

s mivel $J_E = J_E^k$ és $\Delta J_E = \Delta J_E^k$, fentiekből következik, hogy az excentricitás okozta direkt hőmérsékletváltozás nagyságrendje:

$$\Delta T_F = \frac{T_F}{4} \frac{\Delta J_E}{J_E} = \frac{T_F}{4} \frac{\Delta J_E}{J_E} \approx 0,07^\circ\text{C},$$

ahol a Föld jelenlegi hőmérsékletét $T_F = 15^\circ\text{C} = 288^\circ\text{K}$ -nek vettük. Látható, hogy a fenti becslés két nagyságrenddel kisebb a hideg és meleg időszakok közötti $8\text{--}10^\circ\text{C}$ -os hőmérsékletkülönbségénél, tehát az elmélet nem állja ki a tesztet. Ilyenkor szoktak arra hivatkozni, hogy a Föld egy komplex, nemlineáris rendszer, s a nemlineáris effektusok felerősíthetik az excentricitás oszcillációit két nagyságrenddel. Itt azonban látnunk kell azt a csapdát, amibe a lelkes kutatók gyakran beleesnek. Az eljegesedések és az excentricitás változásai között észrevett korreláció nagyon szép, de egy elmélet alapjává tenni, majd a problémákat a szőnyeg alá söpörni az ismeretlen nemlineárisokra hivatkozással veszélyes. Kívülállók számára ez ugyanis elfogadottá tehet olyan összefüggéseket, amelynek esetleg nincs semmi alapja.

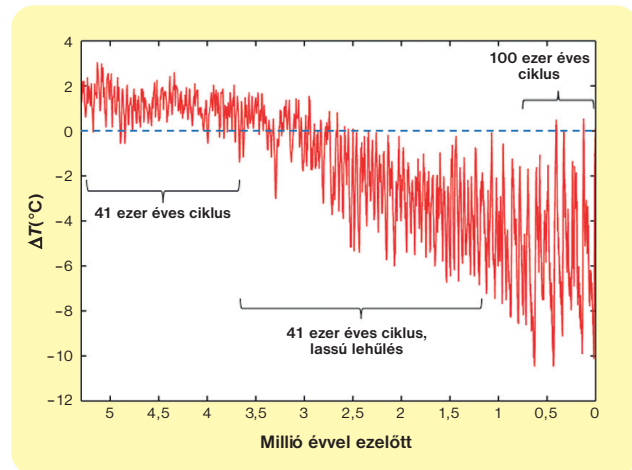
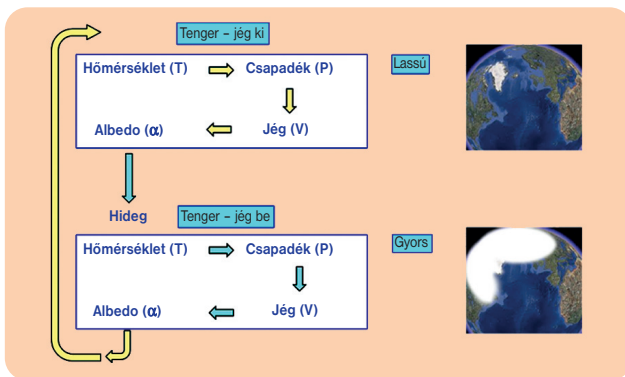
kedés a besugárzás kritikus értékénél, mondjuk 500 W/m^2 -nél változik (vízszintes szaggatott vonal a 4. ábra alsó részén). Ha még azt is feltételezzük, hogy a jégmezők növekedése, illetve csökkenése a besugárzás kritikus értékétől való eltéréseivel arányos, akkor el is érkezünk az eljegesedési elméletek következő nagy csoportjához [7].

A fenti gondolatokat egyenletekben megfogalmazva, s a kritikus $J_E^{(cs)}$ értékét megfelelően választva, megjelenik a megfigyelt hőmérsékleti jel 100 ezer éves periódusa. Sajnos azonban nagy amplitúdóval megmaradnak a 41 és 23 ezer éves periódusok is, s ettől a ponttól kezdve a versengő elméletek abban különböznek, hogyan szűrjük ki az utóbbi periódusokat. Egy könnyen érthető megoldás arra épül [2], hogy a jégmező változásait nem a jég térfogatának pillanatnyi nagysága, hanem egy t_m időre visszamenőlegesen átlagolt értéke határozza meg. Az átlagolás a jégmezők t_m idejű memóriájának tekinthető, s e memória bevezetését motiválhatja a jégmezők lassú változása. Ha $t_m \approx 50\,000$ évet választunk (s az egyenletek néhány egyéb paramétereit is hangoljuk), akkor a 4. ábra középső részén az eljegesedések időskáláján túlmenően a felmelegedések ideje és nagysága is jó közelítéssel megkapható. Így az eredeti problémánk magyarázatát visszavezettük egy 50 ezer éves memória problémájának megfejtésére. Sajnos a bevezetett memória konkrét mechanizmusáról nincs elképzelés, s az 50 ezer éves érték sem értelmezhető fizikailag. Hasonló nehézségek merülnek fel minden olyan elméletben, amik főképp csillagászati okokra vezetnek vissza az eljegesedések menetét, tehát megoldatlan maradt a jégkorszakok rejtélye.

Alternatív elmélet: belső oszcillációk

A csillagászati magyarázatok nehézségei elvezettek a Föld klíma-rendszerének belső dinamikáján alapuló eljegesedési elméletek kidolgozásához. Emlékezzünk, a jégmezők belső dinamikájára az 1000 éves időskálát úgy kaptuk, hogy kiszámítottuk, mennyi idő alatt olvastja el a Nap a 3 km-es jégoszlopot. Ez az időskála nagyjából meg is felel a hidegből a melegbe átmenet időtartamának. Nem vizsgáltuk azonban a jégmezők növekedését, amit az atmoszféra és az óceánok globális áramlása határoz meg, ugyanis ez a cirkuláció szállítja a nedvességet a magasabb szélességi fokokra. A jégmezők kialakulásához szükséges időt megbecsülhetjük ab-

5. ábra. A ciklikus eljegesedések magyarázatául szolgáló alternatív mechanizmus vázlata. A nyílak az okból az okozat felé mutatnak, s a sárga nyílak növekedésre, a kékkek pedig csökkenésre utalnak. Például a felső dobozban a megemelkedő hőmérséklet a csapadék növekedésére, az pedig a jég mennyiségének növekedésére vezet, s ebből következően az albedó nő (mindenütt sárga nyílak). Az albedó növekedése viszont csökkenti a hőmérsékletet (kék nyíl) és átjutunk az eljegesedő állapotba (alsó doboz). A klímarendszernek a jobb oldali képeken ábrázolt két állapotát („lassú” lehülés, ill. „gyors” felmelegedés) a tengeri jég léte vagy nemléte választja el



6. ábra. Mélytengeri üledékekből rekonstruált hőmérséklet-eltérési idősor az elmúlt öt millió évre. A nulla szintet itt az utolsó 100 év átlaga adja (szaggatott kék vonal). Látható, hogy az eljegesedések 41 ezer éves ciklusára egy lassú hűlési trend rakódik, s a 100 ezer éves ciklus 800 ezer évvel ezelőtt vált dominánssá

ból, hogy pl. Grönland északi részén 15 cm/év a csapadékmenyiség. Ebből következik, hogy egy 3 km magasságú jégtakaró felépítéséhez $20\text{--}30$ ezer év kell, s ha számolunk a jégmezők olvadásával is, akkor már közel járunk a gleccserek növekedési szakaszában mutatkozó $80\text{--}90$ ezer éves időtartamhoz.

Most tehát van két különböző időskálánk (1000 év és a $80\text{--}90$ ezer év), mely az éghajlati rendszer két fázisát, a hosszú ideig tartó lehülést és a gyors felmelegedést jellemzi. A kérdés az, hogy mi kapcsolja át a klímát az egyik állapotból a másikba, honnan ered az oszcilláció? Az alábbiakban vázlatosan ismertetjük a legmeggyőzőbbnek tekinthető elméletet, az ún. tenger-jég kapcsoló elméletét [8].

Induljunk ki az éghajlati rendszer azon állapotából, amikor az északi féltekén a szárazföldi jégtakaró területe minimális, a tengerjég elolvadt (a napi sajtó ezt vizionálja a közeljövőre, az elméletben pedig ezt nevezik a *tenger-jég ki* állapotnak). Ekkor az albedó értéke minimális, a hőmérséklet maximális (5. ábra jobb felső része). A nyílt tengerek és a magasabb hőmérséklet északon intenzív párolgásra vezet, s jelentős mennyiségű csapadék jut a sarki területekre, s ennek nagy része hó formájában hullik a szárazföldekre. A gleccserek lassan gyarapodni kezdenek, a jégmezők megindulnak délre, s ennek fontos következménye van. A kiterjedő jégmezők ugyanis növelik az albedót, ami csökkenti a sugárzásbevételt, ami pedig lehülést indít el. A fent leírt *hőmérséklet magas* \rightarrow *csapadék nő* \rightarrow *jégmező nő* \rightarrow *albedó nő* \rightarrow *hőmérséklet csökken* folyamatot ábrázoltuk az 5. ábra felső, bekeretezett részében.

A jégmezők lassan terjednek ki, s ebből következően az albedó lassan nő, és az egész lehülési folyamat nagyon lassú. A folyamat addig tart, amíg az északi tengerek felszíni hőmérséklete el nem eléri a fagyáspontot. Ekkor a lassú fázishoz képest igen gyorsan kiterjed a tengerek jégtakarója egészen az $50\text{--}60$. szélességi fokig, s létrejön a *tenger-jég be* állapot, ami indítja a felmelegedési folyamatot. Mivel hideg van és az északi tengerek befagytak, a jégmezők csapadék-utánpótlása elapad. A besugárzás miatt a jégmezők továbbra is olvadnak, de a csökkenő csapadék ezt nem tudja kompenzálni, s a jéghatár elindul északra. Ez viszont csökkenő albedóhoz, s a növekvő sugárzásbevétel miatt emelkedő hőmérséklethez vezet (5. ábra alsó bekeretezett része a fenti *hőmérséklet alacsony* \rightarrow *csapadék csökken* \rightarrow *jégmező fogy* \rightarrow *albedó csökken* \rightarrow *hőmérséklet nő* folyamatot ábrázolja).

A jégkorszakok rejtélyei

Amikor a tengerfelszín hőmérséklete eléri az olvadáspontot, a tengerek jégtakarója rövid időn belül visszahúzódik a Jeges-tengerre, a száraz, hideg, eljegesedési időszak lezárul, majd további olvadással elérünk a *tenger-jég ki* kezdőállapotunkhoz, s a ciklus kezdődik előlről. Úgy tűnik, rátaláltunk a keresett mechanizmusra, mely nemcsak az eljegesedések oszcillációját, hanem a jelek fűrészfogszerű alakját is magyarázza a lehülés lassú és a felmelegedés gyors dinamikájával. Kérdés, hogy az elmélet megadja-e a ciklus 100 ezer éves hosszát. Itt jelentkezik a komplex rendszerek megértésének újabb jellegzetes problémája. Ha az említett mechanizmust egyenletek formájában írjuk le, akkor ezekben az egyenletekben annyi az ismeretlen, vagy a nagyon nehezen megbecsülhető paraméter, hogy bár a paraméterek hangolásával eljutunk a 100 ezer éves ciklushoz, azonban nem érezzük azt, hogy megoldottuk a jégkorszakok rejtélyét.

Maradtak persze egyéb megoldatlan problémák is a jégkorszakok időskálájával kapcsolatban. Például ismert, hogy a ciklikus eljegesedések jellegzetességei változtak az elmúlt ötmillió évben. A **6. ábrán** látható hőmérsékleti jel [9] az eddig tárgyalt 100 ezer éves periódust csak az utolsó 800 ezer évben mutatja. Azelőtt lehülési trend és egy gyorsabb, 41 ezer éves periódusú oszcilláció figyelhető meg (emlékezzünk, a Föld forgástengelyének billegése hasonló periodicitást mutat). Elmélet nincs, de világos, hogy a 41 és 100 ezer éves periódusú állapotok közti átmenet megértése fon-

tos lenne a jelenlegi klimatológiai dialógusokban. Egyben ez a feladat a klimatológia legszebb problémái közé tartozik. Megoldásához azonban valószínűleg túl kell lépünk a komplex rendszerek leírásának jelenlegi technikáin. ❄

A munka az OTKA NK100296 pályázat támogatásával készült.

IRODALOM

- [1] Behringer W., Witchcraft Persecutions in Bavaria (Cambridge University Press, 1997); *A klíma kultúrtörténete* (Corvina, 2010).
- [2] Berger, A. and Loutre, M. F., *Insolation values for the climate of the last 10 million years*. Quaternary Sci. Rev. **10**, 297–317, (1991); Berger, W. H., *The 100 kyr ice-age cycle: internal oscillations or inclinational forcing?* Int. J. Earth Sciences **88**, 305–316 (1999).
- [3] Diester-Haass, L. and Zahn, R., *Eocene-Oligocene transition in the Southern Ocean: History of water mass circulation and biological productivity*. Geology, **24**, 163–166 (1996).
- [4] Crowley, T. J. and North, G. R., *Paleoclimatology* (Oxford University Press, New York, 1991).
- [5] EPICA community members: *Eight glacial cycles from an Antarctic ice core*. Nature **429**, 623–628 (2004).
- [6] Milankovich, M., *Mathematische Klimalehre und Astronomische Theorie der Klimaschwanken*, Handbuch der Klimatologie (1930), Bd.1 Teil A.
- [7] Imbrie, J. et al., *On the structure and origin of major glaciation cycles: the 100 000 year cycle*, Paleoclimatology **8**, 699–735 (1993).
- [8] Tziperman E. and Gildor H., *A sea-ice climate-switch mechanism for the 100 kyr glacial cycles*. J. Geophys. Res. **106**, 9117–9133 (2001).
- [9] Lisiecki, L. E. and Raymo, M. E., *A Pliocene-Pleistocene stack of globally distributed benthic records*. Paleoclimatology **20**, PA1003 (2005).

A Kenai-félsziget jégmezője 300 négyzetmérföldnyi területet borít be Alaszkában

(Forrás: <http://www.alaskatourjobs.com/wp-content/uploads/2011/07/kenai-fjord-harding-icefield.jpg>)



JÁNOSI IMRE

Energia és társadalom

Drasztikus fázisátalakulás küszöbén állunk?

Nem sok dolog változott a gépkocsik felépítésében 1900 óta – jelenti ki Lawrence Burns, a General Motors kutatásért felelős egykori alelnöke, jelenleg egyetemi tanár [1]. A világon ma található, kb. 1 milliárd személy- és tehergépkocsi 95 százaléka benzint, illetve dízel üzemanyagot használ, önmagában a globális széndioxid-kibocsátás egyötödét adva. Eltekintve azon figyelemre méltó adatoktól, melyek szerint pl. az összes, szorosan egymás mögött parkoló gépjárműből álló kocsisor mintegy százszor érné körül a Földet, vagy azt, hogy a közlekedési balesetek évente több mint 1,24 millió halálos áldozatot szedve az egyik legveszélyesebb globális „járvány” forrásai [2], könnyen látható, hogy a jelenlegi közlekedési rendszer is a fenntarthatatlan kategóriába esik. Az állandó zsúfoltság, forgalmi dugók, parkolási gondok és hasonlóak mellett természetesen a tankolásnál fizetett számlák jelzik talán legjobban a globális problémákra kevésbé fogékony átlagembereknek is, hogy valamiféle komoly változás előtt állunk.

„Peak Oil”

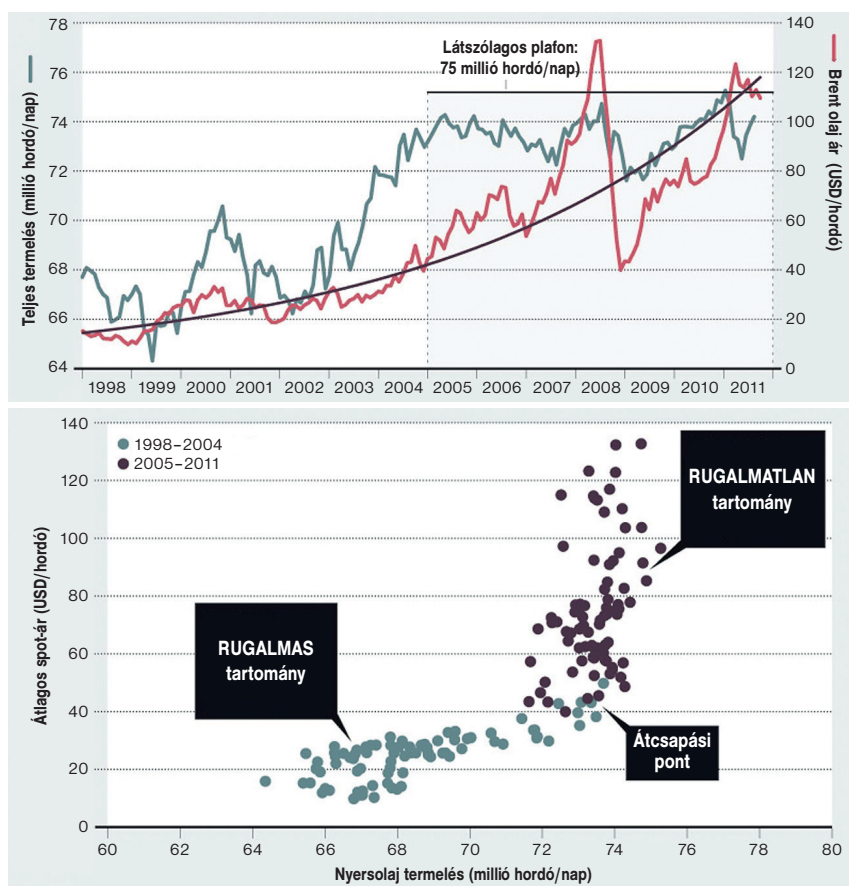
Az egyre több nyelven hétköznapi használatba jött angol kifejezés arra az (egyelőre ismeretlen) időpontra utal, mikor a kőolaj kitermelése elér egy maximális értéket, és ezután egyre csökkenő mennyiség kerül a felszínre, illetve a piacra. *M. King Hubbert* 1956-ban publikált elmélete számít a hasonló elképzelések „ősforrásának”, amely valójában minden véges mennyiségben rendelkezésre álló alapanyag felhasználására hasonló időfejlődést jósol [3]. A szimmetrikus, nagyjából haranggörbe alakú Hubbert-csúcs maximumhelye a „peak oil” pont, egyes földrajzi területekre illetve anyagfajtákra csak mérésekkel illeszthető paraméterekkel. Mindenesetre pár évtizednyi adatsor elég volt ahhoz, hogy a nem is túl bonyolult teória viszonylag pontos jósolatot adjon az USA kőolaj kitermelésének 1971 után valóban bekövetkezett visszaesésére. A különböző földrajzi területek „Hubbert-csúcsai” időben eltolva jelentkeznek, pl. az Északi-tenger olajmezőin az USA-hoz ké-

pest évtizedekkel később kezdődött és tetőzött a kitermelés.

Bizonyos frissen megjelent tanulmányok szerint a globális kitermelés is elérkezett már a „peak oil” ponthoz. Egy 2012 elején publikált cikk [4] legfontosabb eredményeit mutatja be az **1. ábra**. A felső panel idősorai szerint a globálisan kitermelt mennyiség 2005-től lényegében elért egy kb. napi 75 millió hordós felső küszöbértéket, bár a 2012-es adatok ismételt emelkedésre utalnak (napi 75,6 millió hordó [5]), köszönhetően a meglódult palaolaj kiaknázásnak. Ugyanezen időszakban az árak trendje töretlenül emelkedő maradt, hatalmas kilengések mellett (*1. ábra*, felül, piros görbe). Ennél is beszédesebb az *1. ábra* alsó paneljén ábrázolt összefüggés, amely a kitermelt mennyiség és az azonnali piaci árak között fennálló korrelációkat ábrázolja kétszer hétéves időszakban. A közgazdaságtan nyelvezetét használva az első periódus (1998–2004) „ár-rugalmasnak” tekinthető, melyben a növekvő igényeket minden további nélkül képes volt kielégíteni a növekvő kitermelés, vi-

Dél-keleti csúcsforgalom egy távol-keleti nagyvárosban (Forrás: <http://blogs.rftdata.co.uk/photo-diary/files/2013/07/RUSH-HOUR.jpg>)





1. ábra. Fenn: globálisan kitermelt kőolaj mennyiség (bal oldali tengely, zöldes görbe), illetve a Brent típusú olaj hordónkénti ára (jobb oldali skála, piros görbe) az idő függvényében. Lenn: korreláció a kitermelt mennyiség és az azonnali (spot) piaci ár között két, egymást követő hétéves időszakra. (Forrás: [4])

szonylag enyhe általános áremelkedés mellett. A szerzők szerint 2005-ben bekövetkezett „átcsapás” után egy „ár-rugalmatlan” periódusba kerültünk, mikor is a tovább növekvő igényeket nem képes már kielégíteni semmilyen létező tartalékkapacitás sem, erre a piaci árak kiszámíthatatlan ingadozásokkal reagáltak.

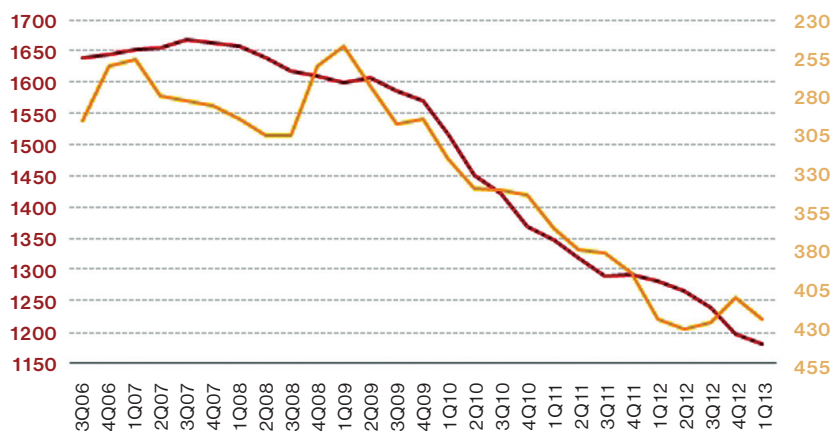
Az itt felvázolt képet persze árnyalhatja a már említett palaolaj növekvő mértékű kiaknázása. Érdekes lehet talán, hogy ezt a fajta, magas szerves anyag tartalmú üledékes kőzetekből különböző eljárásokkal kinyerhető nyersolajfajtát Svájcban és Ausztriában már a XIV. századtól használják [6]. Az ipari méretű előállítás a hagyományos kútból történő kitermeléshez képest ez idáig gazdaságtalanul drága volt, a jelenlegi árszinten azonban láthatóan már sok társaságnak igencsak megéri.

Az üzemanyagválság képzete

Kétségtelen tény, hogy emelkedő üzemanyagárak esetén a fogyasztás általában csökkenő tendenciákat mutat. A 2. ábrán lát-

ható, hogy pl. hazánkban, a 2007-es csúcs-évhez képest majdnem 30 százalékkal csökkent a benzin éves fogyasztása az idei év elejére. A kissé trükkös grafikon jelzi, hogy ebben fontos szerepet játszik a benzinku-

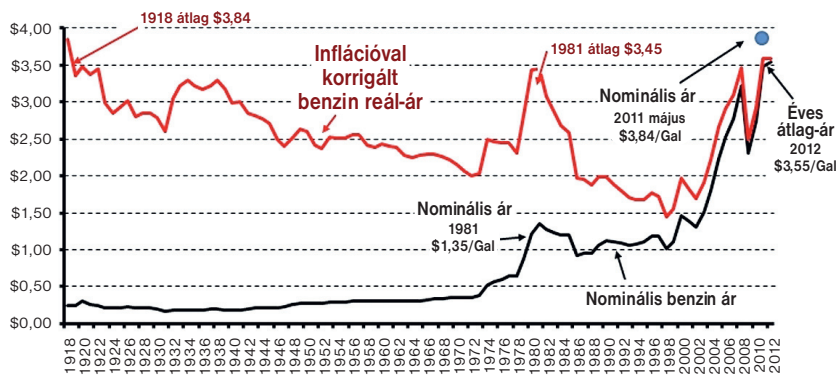
2. ábra. A hazai benzinfogyasztás (piros görbe, millió liter, 12 havi gördülő) és a benzin negyedéves átlagos kiskereskedelmi ára (narancssárga, Ft/liter, jobb oldali fordított skála). (Forrás: MÁSZ, <http://www.portfolio.hu>)



taknál fizetendő ár (a vizuális trükk a jobb oldali fordított skála megfelelő tartományának kiválasztása). Ám biztosan kijelenthetjük, hogy az ár mellett sok más tényező is közrejátszik, melyek közül kiemelkedő a jövedelmi helyzet: ki foglalkozna az árral, ha otthon vasvillával kellene szellőztetni a papírpénzt, hogy be ne penészedjen! Az autózás járulékos költségei sem játszanak elhanyagolható szerepet, ha a biztosítási díjak vagy javítási költségek emelkednek, a fogyasztás még változatlan benzinárak mellett is csökken. A dízelolaj esetén enyhébb, de hasonló a trend, bár ott nem az ár a legfontosabb tényező, hanem a gazdaság általános állapota (ha nincs mit fuvarozni, a teherautók nem közlekednek).

Tulajdonképpen igaz a kijelentés, mely szerint a benzin ára még soha nem volt ilyen magas, érdemes azért megnézni a 3. ábra piros görbét, amely az USA átlagos reálárának időbeli alakulását mutatja be (ez a 2013. februári dollár árfolyamon alapuló infláció korrekcióval készült). A mai reálárszint a közelmúltban is legalább kétszer előfordult már, 1979-ben és 2008-ban. Az előbbit az iráni forradalomnak majd az ezt követő iráni-iraki háborúnak, az utóbbit a 2004-től fellépő folyamatos piaci zavaroknak és polgárháborús illetve nemzetközi konfliktusoknak tulajdonítják. Figyelemre méltó, hogy különösen a 2008. július eleji csúcsot követően milyen gyorsan visszazuhan az ár, de az elemzések szerint az 1979-es tetőzés is inkább pániknak, mintsem valós ellátási gondoknak volt köszönhető (a kiesett iráni és iraki mennyiséget gyorsan pótolta Szaúd-Arábia és a többi OPEC ország). Az is tény, hogy a közhangulat is erősen hasonló volt a maihoz, az USA gazdaságának teljes összeomlásáról szóló víziók napi gyakorisággal tüntek fel a különböző sajtótermékek vezető hírei között.

Az első komolyabb olajválság valójában



3. ábra. A benzin nominális és 2013. februári \$ árfolyamon számított realára, \$/gallon (1 gallon = 3,785 liter) egységekben. (Forrás: Timothy McMahon, InflationData.com)

már 1973-ban lezajlott, annak ellenére, hogy sem a nominális, sem a realárak nem tűnnek túl magasnak. (A 3. ábra éves átlagértékeket illusztrál, ami a rövid idejű, nagy amplitúdójú csúcsokat erősen kisímitja, pl. a válság elején az árak megnégyszereződtek.) Az olajexportáló országok szervezete, az OPEC, 1973 októberében szállítási embargót hirdetett az Izrael az ún. Jóm kippúri háború során támogató országok irányába. Piaci értelemben az embargó valójában soha nem valósult meg, de messzemenő, máig ható politikai és gazdasági következményekkel járt. NATO tagországok sora és Japán például mindenféle háttéralkukkal próbálta meg a szállítások valódi leállításának megakadályozását, eközben szárnyaló infláció, tőzsdei válságok és mérhető gazdasági visszaesés indult el számos országban (a Bretton Woods-i pénzrendszer ekkorra már szétesett). A közgazdászok egészen új jelenségekkel találkoztak, pl. ekkor jelentkezett először a „stagfláció”, azaz a magas munkanélküliség és csökkenő termelés együtt a magas inflációval. Ma már viccesnek tűnő események, mint pl. a Japánban és az USA-ban lezajlott vécépapír pánik (teljesen alaptalanul elterjedt, hogy olajellátási problémák miatt leáll a gyártás) mellett többek között a nyári óraátállítást is ennek a pánikidőszaknak köszönhetjük.

Hosszabb távon is figyelemreméltó változások indultak: a fűtési rendszereket ekkortól állították át tömegesen földgázra, megjelentek és sikeressé váltak az alacsony fogyasztású kisautók, elindult az alternatív energiaforrások keresése (kezdetben a nukleáris technológia dominált, később a valódi megújuló források is kezdtek fejlődni), óriási beruházások történtek a kevésbé olajfüggő, pl. elektronikai iparban stb.

Mi zajlik napjainkban? Mint látható, a korábbi üzemyag „válságok” sora valódi alapanyagihiány nélkül következett be, gyakran túlreagálva csekély mértékű eseményeket. (Ez év elején néhány híradás olajfinomítók bezárásáról elég volt egy 14%-os áremelkedés kiváltásához az USA-ban.)

Az teljesen valószínűtlen, hogy az olaj egyik napról a másikra elfogyva, a Hubbert-elmélet is egy sima, fokozatos csökkenést feltételez. Az viszont tény, hogy a hetvenes évekkel összehasonlítva egy újabb globális kihívással nézünk szembe: a Föld népessége töretlenül növekszik.

Egy kis demográfia

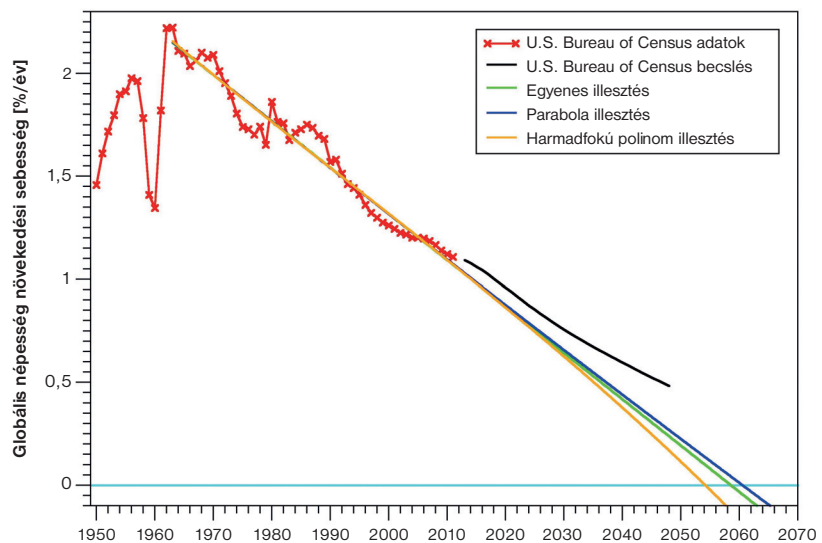
1974 volt az az év, mikor a globális népesség átlépte a 4 milliárdos értéket, 2011 végén pedig elérte a 7 milliárdot. Az ENSZ legfrissebb prognózisai szerint 2025-re 8,1 milliárd fő lehet a Föld népessége, 2050-re 9,6 milliárd, 2100-ra pedig 10,9 milliárd [7]. A legnagyobb növekedés Afrikában várható, 2,4 milliárddal lesznek többen 2050-re. A világ második legnépesebb országának, Indiának a lakossága 1,6 milliárdra nőhet a jelenlegi 1,3 milliárdról, átvéve ezzel a világ legnépesebb országa címet Kínától (ahol a

tetözést 2030 körül várják). Eközben a fejlődő országokból induló migráció nélkül a fejlett országok népessége csökkenne. Mindemellett a globális népesség növekedési rátájának adatai (4. ábra) azt mutatják, hogy valami történt a hatvanas évek elején.

Az 1963-as évben tetőzött 2,2%/év növekedési sebesség 2012-re megfeleződött (1,1%/év) némi ingadozással ugyan, de folyamatosan csökkenő értékekkel az utóbbi 50 évben (4. ábra). Mi történhetett 1963-ban? Az egyik közkeletű vélekedés szerint az egy családban születő gyermekek számának csökkenése az oktatás és a modern fogamzásgátlási eljárások elterjedésének az eredménye. Statisztikai adatok bizonyítják, hogy az egyetemi végzettséggel rendelkező nők első gyermekvállalása egyre idősebb korra tolódik ki, illetve a főállású anyaság egyre jobban visszaszorul a világ sok országában. A probléma komplexitása abban rejlik, hogy egy ifjú pár gyermekvállalási kedve mindig lokális körülmények mérlegelésén alapul (lakás, munkahelyi kilátások, anyagi biztonság stb.), nehezen elképzelhető, hogy tömegeket befolyásolna a globális népességi, gazdasági és környezeti adatok ismerete (klímaváltozás, élelmiszerellátás, vagy energiaforrások). Kivételként meg lehet említeni a kínai „egyke” programot (1979-től a városban élő pároknak csak egy gyermek születését engedélyezik, megakadályozva ezzel állításuk szerint 650 millió szülést 2011-ig), de ez inkább egyedi kielégésnek tekinthető, mint általános gyakorlatnak. Hasonló okok miatt még a legbonyolultabb előrejelzési módszerek sem adnak megbízhatóbb becslést, mint az egyszerű interpolációs eljárások (4. ábra), senki nem láthatja előre a jövőt.

Az viszont tény, hogy a várakozások szerint hamarosan tízmilliárdos globális né-

4. ábra. A Föld népesség növekedési sebessége az idő függvényében. A folytonos vonalak különböző „előrejelzéseket” jelölnek



Energia és társadalom

pesség számára az „amerikai álom” élet-színvonal elérése teljesen lehetetlen a jelenlegi erőforrások végessége miatt. Az energiaforrások felhasználása mindig gyorsabban nő, mint a népesség, mert a növekedéssel együtt járó általános életszínvonal emelkedése egyre több fogyasztást jelent, ami folyamatosan növekvő mértékben terheli a rendelkezésre álló forrásokat. A termelés, szállítás, elektromos hálózat, fűtés-hűtés, nagyobb lakás, autó, háztartási berendezések stb., energiaellátása napi szintű kihívás, amivel egyre nehezebben tartanak lépést a rendelkezésre álló elosztó rendszerek. Az is világossá vált, hogy a jelenleg ismert és alkalmazott alternatív (megújuló) energiaforrások nem képesek kiváltani a fosszilis tüzelőanyagokat az energiasűrűség rendkívül alacsony értéke és a kiszámíthatatlan hozzáférés miatt [8], legalábbis amíg nem sikerül megtalálni az olcsó, ipari méretű energiatárolás technológiáját.

Mi várható?

Innentől kezdve ezen írás záró része nyugodtan sorolható a tudományos-fantasztikus irodalomba, hiszen az egyetlen tanulság az emberiség történelméből az, hogy soha, senki nem tudta megmondani a jövőben várható történelem fő vonalait sem. (Ne téveszsen meg minket az elismert jövőlátó tudósok dokumentált teljesítménye, mert ez inkább a lottóhúzás eredményéhez hasonlítható. Jó példa erre a tőzsdeguruk tevékenysége, akik a pusztaszerencse alapján lettek váteszek: az árfolyamok változásának minden lehetséges kimenetére születnek tömeges „előrejelzések”, akinek bejön, megdicsőül, a milliányi veszteségről szó sem esik.) Ettől függetlenül még eljátszhatunk a gondolattal, hogy mi történhet, ha a jelenlegi tendenciák tovább folytatódnak, pl. az energiaellátás terén.

Az emberi társadalomhoz hasonló komplex rendszerek sajátossága, hogy a belső viszonyok, kölcsönhatások apró lépésekkel történő fokozatos változása is vezethet drasztikus, fázisátalakuláshoz hasonló összeomlásokhoz (lásd *Kondor Imre* írását ebben a számban). Gondoljunk csak a Brazíliában történt júniusi eseményekre: a tömegközlekedés árainak pár százalékos emelése milliós tüntetésekhez, utcai lázongásokhoz vezetett. Természetesen a háttérben sokkal több történés vezetett az ingerküszöb átlépéséhez, ám tény, hogy a döntéshozók nem sejtették, hogy ez a látszólag kicsiny lépés mekkora zavargásokat okoz. A többször említett Hubbert-görbe szerinti ellátási időfejlődés hasonló fejleményekhez vezethet. A tovább emelkedő üzemanyagárak még jobban széthúzhatják a jövedelmi különbségeket, a vesztes oldalról azok állnak majd, akik kénytelenek lesznek elszervezni az alapvető ellátási javak árába

épített magas szállítási költségeket, és az elszálló közüzemi és közlekedési számlákat. Végtelen esetben az is elképzelhető, hogy a tömeges energiaellátás forrása a megújuló technológiákra szorítkozik majd, azaz ha fúj a szél, vagy süt a nap, akkor van elektromos áram a konnektorokban, szélcsendes éjszakákban meg nincs.

Az idősebb generáció jól emlékszik még arra a pár évtizeddel korábbi helyzetre, amikor sem a háztartási fagyaszto, sem a fagyasztott élelmiszerek nem voltak ennyire elterjedtek. A borsó érésékor borsót ettünk borsóval, a kelkáposzta szezonban egymást váltotta az asztalon a főzelék és a rakott kel, ha meg bedurrant a paradicsom, akkor csak piros színű ételek kerültek tálalásra (mondjuk, ezt a befőzés el tudta kissé húzni).

Az energia-hozzáférés azonban teljesen más módon alakult az iparilag fejlett országokban. A fogyasztást a pillanatnyi igény szabja meg, az elektromosáram-ellátás jobbra zavartalan, akkor tankolunk, amikor csak óhajtunk (és ki tudjuk fizetni). Csak egy hosszabb áramszünet vagy üzemanyaghiány döbrent rá minket, mennyire ki vagyunk szolgáltatva e forrásoknak. A leolvadó hűtőgép, az Internet hiánya, pislákoló gyertyafény, vagy kényszerű gyalogolás többségünket meglehetősen felidegesíti (a mosógép kiesése nem zavaró annyira, mert egy-két óra áramszünet után úgyszincs vezeték víz). Ritkán jut eszünkbe, hogy ennél nagyságrendekkel komolyabb gondok jelentkezhetnek, ha egy kórház, termelőüzem, vagy az áruszállítás marad tartósan energia nélkül. Nemigen tudjuk elképzelni, miféle társadalmi átalakulással járna, ha az energiafogyasztást nem a pillanatnyi igény, hanem a pillanatnyi hozzáférhetőség szabná meg. Jelenleg a szabályozás a háztartásokat kezeli kiemelt figyelemmel, azaz pl. gázhiány esetén először a nagy ipari fagyaszto szervenvednek el korlátozásokat, és csak a legvégső esetben jönne a lakossági ellátás kikapcsolása. Nyilvánvaló módon ez a prioritási sorrend kényszerűen megfordulna, ha az ellátási zavarok tartósan fennállnának. Még egy fűtetlen és sötét lakásban is ennünk kell, szükségünk lehet gyógyszerre, ruhára, valahogy el kell jutnunk az iskolába, munkahelyre, orvoshoz. A modern társadalmakban még a mainál is jobban felértékelődne a távolsági kommunikáció szerepe, azaz pl. a mobiltelefon-hálózat központi energiaellátása lenne a legutolsó, amiben korlátozásokat lehetne bevezetni. Ugyanígy az időjárás-előrejelzés még fontosabb szerepet kapna, ehhez a jelenlegi észlelő hálózat óriási fejlesztésére lenne szükség (a szél pontos előrejelzése az egyik legnehezebb feladat a meteorológusoknak). Azt is nehezen tudnánk elviselni, ha egy szeles hétvégét munkával kellene tölteni, csak mert éppen akkor van áram, utána meg egy hétig kényszerből (fizetés nélküli) szabadságot rendelnének el. Az isko-

laév közepe a nyári hónapokra esne, a hosszú (sötét) téli szünet viszont hó hiányában minden bizonnyal kissé unalmas lenne. Sokan észrevennék a pultokról eltűnő banán és narancs hiányát, vagy azt, hogy a boltban csak egyféle sör kapható (ha a sörgyár lovas kocsival elérhető távolságban van).

Szerencsére a felvillantott pesszimista forgatókönyv nem látszik rövidtávon reálisnak. Ám a modern parlamenti demokrácia egyik nemkívánatos következménye, hogy a kormányok túlnyomó többségének tevékenysége egy választási ciklusra (általában 4 évre) koncentrál, nagyon kevés helyen valósulnak meg hosszú távú fejlesztések. Ez az oka annak is, hogy egy pár évvel ezelőtt, ugyanebben a témában megjelent írásom [8] nemigen tűnik még elavultnak. Időközben néhány apróbb változás történt azért, a legfontosabbakat röviden megemlíteném. Először is az eltelt hat évben a globális népesség legalább fél milliárd fővel gyarapodott, akiknek a fele már sóvárogva nézegeti a babakocsi mellett elszáguldó csillogó autót-motorokat. Az üzemanyagárak tisztán jelzik az ellátási gondokat, komoly és tartós jövőbeli áresésről nemigen érdemes álmodozni (a manapság sokak által panaceaának tekintett palaolaj és palagáz nem termelhető ki ennél lényegesen alacsonyabb árak mellett). Töretlenül nő a szél- és napenergia-kapacitások telepített nagysága, de globális szinten ez még mindig nem számít jelentősnek a maga 3% körüli részarányával. Ezzel szemben kormányok sora szakította félbe a korábban elkezdett bioüzemanyag-támogatási programokat, felismerve a csekély hozadékot és komoly élelmiszer-ellátási kockázatokat. A nagy nemzetközi fúziós és új generációs nukleáris technológiai projektek egyre csak késnek, csökkenő anyagi és erkölcsi támogatás mellett akadozva haladgatnak. Márpedig a környaalazódó energiaproblémák megoldása talán évtizedes futamidőt sem enged, különösen a régiókban jelenleg elhúzódó gazdasági-társadalmi válság, tetézve egy energiakrizissal könnyen vezethet robbanáshoz. Jobb lenne időben felkészülni, hazafelé menet veszek is egy új gumit a bringára – meg néhány gyertyát. ❁

Ezen munka az OTKA NK100296 sz. pályázatának támogatásával készült.

IRODALOM

- [1] Burns, L. D. (2013) *Sustainable mobility: A vision of our transport future*. Nature, **497**, 181–182.
- [2] <http://www.who.int/mediacentre/factsheets/en/>
- [3] Markó, L. (2010) *Meddig tartanak ki a Föld kőolaj-készletei?* Természet Világa, **141**, 194–198; **141**, 247–250.
- [4] Murray, J., King, D. (2012) *Climate policy: Oil's tipping point has passed*. Nature, **481**, 433–435.
- [5] <http://www.eia.gov/cfapps/ipdbproject/iedindex3.cfm>.
- [6] <http://hu.wikipedia.org/wiki/Palaolaj>.
- [7] *World Population Prospect: The 2012 Revision*. (2013) New York: United Nations. <http://esa.un.org/wpp/>.
- [8] János, I. (2007) *Krisis vagy hisztéria? Emberiség és energia*. Természet Világa, **138**, 150–153.

KONDOR IMRE

A komplexitás kihívása

Az utóbbi néhány évtizedben a tudományos életben és irodalomban egyre gyakrabban találkozhatunk olyanféle kifejezésekkel, mint komplexitás, komplex rendszerek, a komplexitás tudománya. Számos tudományos cikk, könyv és konferencia foglalkozik komplex rendszerekkel, és egyre több egyetemi tanszék, kutatóintézet nevében fordul elő a komplexitás szó – a Google a „complex systems” keresőszóra 145 millió találatot dob ki. A komplexitás tehát divatba jött, sokan úgy tekintenek rá, mint a XXI. század tudományának egyik meghatározó irányzatára. Másoktól, köztük komoly tudósoktól is hallani olyan véleményeket, mely szerint a komplex rendszerek tudománya nem is létezik, az így megjelölt rendszereket egyenként, nagyon eltérő tulajdonságaikat külön-külön vizsgálva kell tanulmányozni, és a köztük felfedezni vélt hasonlóságok mind felületesek.

Mi volna hát ez a vitatott komplexitás, milyen definíciót tudunk adni rá? A Google találatok között az első két helyen Wikipedia szócikkek jönnek fel, melyek bevezetésükben elég jól érzékeltetik, miről is van szó, de rögtön leszögezik azt is, hogy kielégítő, minden aspektust lefedő és általánosan elfogadott definíció nem létezik. Jellemző, hogy amikor az EU Bizottság által a gombamód szaporodó európai komplexitás-kezdemenyzések koordinációjára létrehozott Complexity-Net munkacsoport megkísérelte a komplexitás fogalmát meghatározni, a próbálkozást háromszori, többszörös nekifutás után mint reménytelen időpocsékolást feladta.

Ez a cikk nem is kísérli meg, hogy logikailag kielégítő definíciót adjon, megelégszik azzal, hogy megemlíti a komplex rendszerek néhány olyan tulajdonságát, amelyek az érintett kutatók, vagy legalábbis nagy többségük a komplexitás lényegi ismérveinek tekint, továbbá felsorol néhány olyan ténylegesen létező rendszert, amelyeket talán mindenki elfogadna komplexnek. Mint-hogy a komplexitás problémája a matematikától kezdve a társadalomtudományokig minden területen fellép, a lényeges ismérvek közül az egyes tudományok más és más vonást emelnek ki. A Természet Világa jelen különszáma erre számos példát szolgáltat. Itt a komplexitás két olyan ismérvét veszem kicsit közelebbről szemügyre, amelyeket magam meghatározóan fontosnak érzek, de amelyek sokszor nem kapnak kellő hangsúlyt az irodalomban: az irreducibili-

tást és a meghatározó távolság-, idő- és energiaskálák összeolvadását.

A komplexitás tudományát úgy is jellemezhetjük, ha szembeállítjuk a tudomány megelőzőleg általánosan elfogadott karakterével. Beszéljünk tehát előbb egy kicsit az egyszerűségről, a tudománytól elvárt lehető legegyszerűbb, takarékos leírás követelményéről.

Az egyszerűség dicsérete

Az elméletalkotásban mindig is érénynek tekintették az egyszerűséget, a logikai takarékoságot. Ennek illusztrációjára álljon itt néhány idézet, melyeket 5 perc alatt összegyűjthetünk az internetről:

Einstein: *A [tudományban a] dolgokat olyan egyszerűvé kell tenni, amennyire lehetséges, de nem egyszerűbbé.* – Ez kétségtelenül szellemes, de nem nagyon igazít el az egyszerűség vagy bonyolultság kívánatos mértéke tekintetében.

Richard Feynman: *Az igazság mindig felismerhető a szépségről és egyszerűségről.* – Itt megjelenik a szépség motívuma is, amely számos más alkotónál is viszsztatér.

Például Paul Dirac, a XX. századi fizika egyik óriása, közismerten hatalmas jelentőséget tulajdonított a tudományos elméletek szépségének. A róla elnevezett egyenlet vitathatatlanul elegáns és szép:

$$-i\hbar\gamma^\mu\partial_\mu\psi + mc\psi = 0$$

olyannyira, hogy Ericeben (Szicília) az Ettore Majorana konferenciaközpont előadótermében lendületes kézírással, felnagyítva és kovácsolt vasból kikalapálva ez az előadói emelvény dísz.

Hasonlóképpen jó véleménye volt az egyszerűségről Leonardo da Vincinek is, aki szerint: *Az egyszerűség a végső kifinomultság.*

Az egyszerűségnek, az elméletalkotásban felhasznált posztulátumok minél kisebb számának elvét a középkori angol szerzetes-filozófus után *Occam borotvajaként* szokás emlegetni, azonban ennek az elvnek a gyökerei a középkornál sokkal meszebbre nyúlnak vissza. Már Arisztotelésznél is ezt olvashatjuk: *Feltehetjük, hogy az a bizonyítás, amely kevesebb feltevést vagy hipotézist használ, ceteris paribus, magasabb rendű.*

Newton, talán minden idők legnagyobb természettudósa, ekként vélekedik a Ter-

mészet ízléséről: *A Természet az egyszerűségben leli örömét, nem pedig a fölösleges okok pompázatoságában.*

Ezek után a vonzó, ámde mégis csak szubjektív megállapítások után idézzük most Lavoisier-t: *Ha az egész kémiát kielégítően meg tudjuk magyarázni a flogiszton nélkül, ez már elegendő ahhoz, hogy végtelenül valószínűnek tartsuk, hogy ez az anyag nem létezik, hogy merőben hipotetikus, csupán szükségtelen feltevés.* Vegyük észre, hogy itt az egyszerűség, a fölösleges feltevések elhagyásának követelménye már kiválasztó elvként működik. Több mint száz évvel később hasonló megfontolások ítélték a flogiszton sorsára az étert is.

Ha meggondoljuk, a minél takarékosabb, minél szebb, tömörebb és egyszerűbb elméletekre való törekvés nyilvánul meg a fizika mint *par excellence* redukcionista tudomány széles fejezeteinek extrémum elvekbe történő sűritésében, a kölcsönhatások egyesítésében, vagy a szimmetria-elveknek a fizikában játszott kiemelkedő szerepében is. Lássuk megint Einstein: *Minden tudomány magasrendű célja... hogy a lehető legnagyobb számú empirikus tényt a lehető legkisebb számú feltevésből vagy axiómából vezesse le.*

Nem tudhatjuk, hogy amikor Einstein itt *minden* tudományról beszélt, az olyan kevésbé axiomatizálható tudományokra is gondolt-e, mint mondjuk a szociológia vagy a művészettörténet. Valószínűleg nem. Az egyszerűség dicséretét zengő vélemények után álljon most itt Jakob Burckhardt figyelmeztetése: *A komplexitás tagadása a zsarnokság lényege.*

Az említett idézetek mind a redukció, ill. információ tömörítés körül forognak. A világból ránk zúduló benyomások, ingerek áradatában tökéletesen lehetetlen volna eligazodnunk nagymértékű tömörítés nélkül. Redukció nélkül nem lennének fogalmaink, beszélni sem tudnánk. Amikor a selyempincsit és a dán dogot egyaránt kutyának nevezük, számos, nem is csekély különbségtől tekintünk el. A kutya fogalma, mint minden fogalom, igen erős redukció eredménye.

Nem világos azonban, hogy minden egyes tudományos problémát redukálni lehet olyan mértékig, hogy egyszerűvé és esztétikailag vonzóvá váljék. A tömörítés adódhat tudatlanságból, érdekből, vagy szándékos, értékvezérelt választásból.

Mindemellett az egyszerűség és szépség szubjektívek. Ami az egyik embernek egy-

szerű, a másíknak tökéletesen felfoghatatlan lehet, a szépségről meg vitatkozni sem érdemes.

Lehet, hogy az egyszerűség iránti obszesszív vágy az emberi intelligencia szerkezetéből adódik (pl. a limitált rövid távú memóriánkból)? Lehet, hogy ekként evolúciós háttér van? Hogy a túlélésért folytatott küzdelemben a gyors és takarékos reakción múlhott a siker (ha három baljós folt emlékeztet egy leopárdra, akkor ugrani kell), ami kifejlesztette bennünk azt az előre drótozott képességet, hogy „összeköszük a pontokat”, valamiféle értelmet vetítünk bele a lét káoszába, hogy szabályosságokat fedezzünk fel (ott is, ahol esetleg nincsenek), hogy elméleteket alkossunk, és hogy „Occam seprűjével” (copyright by Sidney Brenner) a szőnyeg alá söpörjük a kényelmetlen részleteket?

A redukcióra, az axiomatizálásra való törekvés még az olyan nehezen (vagy egyáltalán nem?) redukálható tárgyak leírásában is megjelenik, mint a gazdaság vagy a pénzügyek. Ez teszi érthetővé, hogy a közgazdaságtan főárama néhány egyszerű, nyilvánvalóan igaznak tétélezett állításra építette fel a neoklasszikus egyensúlyelmélet impozáns épületét. Igazán kár, hogy a nyilvánvalónak tétélezett állítások nyilvánvalóan nem igazak, így az olyan apró kis zavar, mint a 2007 óta tartó világválság, még elvont lehetőségként sem szerepel ebben az elméletben. A társadalom vagy a gazdaság mintha komplexebbek lennének annál, hogysem néhány axiómában megragadhatóak legyenek.

A komplex rendszerek elméletében kétségtelenül létezik törekvés arra, hogy megtaláljuk a kulcsot általában a komplex rendszerek, vagy speciálisan valamely konkrét komplex rendszer egyszerűsítésére. Ha egy-egy konkrét esetben sikerülne valami ilyesféle redukciót elérni, az csak azt jelentené, hogy a kiinduló rendszer maga nem volt komplex, legfeljebb csak nagyon bonyolult. Az irreducibilitás a komplexitás egyik meghatározó tulajdonsága.

Amikor ilyen címetek olvasok, hogy „Complex systems made simple”, a Monty Python: The All-England Summarize Proust Competition című jelenete jut eszembe, ahol is a versenyzőknek 15 másodpercük van, hogy Proust hétkötetes, 3200 oldalas regényfolyamát összefoglalják. A verseny eredményét nem árurom el, de javasolom, hogy aki még nem látta, nézze meg az interneten.

Van, ami nem tömöríthető, nem redukálható. Van, ami komplex.

A komplex rendszerek néhány jellegzetes tulajdonsága

A következőkben felsoroljuk a komplex rendszerek néhány jellegzetesnek tekinthető vonását.

- A komplex rendszerek (mindig? tipikusan?) nagyszámú alkotóelemből állnak.
 - Ezek az alkotóelemek és a köztük működő kölcsönhatások tipikusan nem egyformák, a rendszer heterogén, általában jellemző a szimmetriák hiánya.
 - Erős kölcsönhatás működik az alkotórészek között, kollektív effektusok lépnek fel.
 - Az egész olyan váratlan tulajdonságokat mutat, amelyekre nem következtethetünk a *részeinek* az ismerete alapján (emergencia).
 - A jellemző (távolság-, idő-, energia-) skálák nem különülnek el, összefolynak.
 - A rendszer nem lineáris.
 - A rendszer dinamikája multiattraktor szerkezetű, komplikált attraktorok léphetnek fel, nagyszámú vonzási medencét találunk.
 - A rendszer érzékeny a kontrollparaméterekre, kezdő- és peremfeltételekre. Hosszú távú korrelációk jellemzik, nem vágható részekre lényeges tulajdonságainak sérülése nélkül („több, mint a részeinek összege”).
 - Apró szerkezeti részletek is fontosak, a rendszer leírásához nagyszámú változó szükséges, a rendszer irreducibilis, azaz nem redukálható lényeges vonásainak elvesztése nélkül, viselkedése sokszor annyira bonyolult, hogy akár véletlenszerűnek is tűnhet.
 - Fejlődését történetiség, hosszú távú memória jellemezheti, sokszor csak egyetlen realizációja figyelhető meg, a kísérleteket nem lehet megismételni.
 - A fejlődést az előzmények kondicionálják – ezek határt szabnak a racionális választásnak, döntésnek vagy kontrollnak.
 - A rendszer viselkedésének lassú, hosszú távú módusai kondicionálják a gyorsabb, rövidebb skálájú módusokat (mode slaving).
 - Rendelkezhet az adaptáció, tanulás, önszerveződés, reprodukció képességével.
 - A fejlődést magáról a rendszerről szerzett ismeretek is kondicionálhatják, a tudás, az önmagára való reflexió befolyásolhatja a viselkedését (öntudat, önbejelítő próféciák, kollektív mítoszok).
- Az említett tulajdonságok logikailag nem mind függetlenek egymástól, másrészt a lista biztosan nem is teljes, de már így is sokkal bővebb annál, mint amit ennek a cikknek a korlátozott keretei között megbeszélhetünk. Ezért csak egy-két jellemző tulajdonság tárgyalására szorítkozunk, azt remélve, hogy néhány másikat e különszám többi cikke fog megvilágítani.

Ezek után adjunk néhány példát komplex rendszerekre:

- a sejt,
- az agy,
- az élőlény,

- egy ökoszisztéma,
- a pénzügyi rendszer,
- a gazdaság,
- a társadalom.

A fogalmat néhány ellenpélda is segíthet megvilágítani:

- A másodfokú egyenlet (dacára annak, hogy kellően ravaszul feltett kérdésre olyan elképesztően bonyolult válaszokat generál, mint a Mandelbrot-halmaz).
- A hidrogénatom (noha „több – tömegére nézve aktuálisan kevesebb –, mint a részeinek összege”)
- Az ideális gáz (pedig a hőmérséklet emergens fogalom).

Ezek a triviális ellenpéldák azt mutatják, hogy egészen egyszerű rendszerekben is megjelenhet a komplexitás egyik vagy másik jellemzőnek tartott aspektusa, és amikor a komplexitás definíciójáról vitázunk, ezeket a vonásokat minden kutató a saját ízlésének vagy érdeklődésének megfelelő módon hajlamos súlyozni. Hogy a magam elfogultságai felől ne hagyjak kétséget, már itt bevallom, hogy számomra az irreducibilitás tűnik központi jelentőségű ismérveknek, amely a valamiképpen legtöbb felsorolt tulajdonsággal kapcsolatban áll.

Információtömörítés

Kezdjük a komplexitás különböző aspektusainak a tárgyalását a matematikában való megjelenésével. Mindenki megtanulja az iskolában, hogy az egészek után a legegyszerűbb fajta számok a racionális számok, amelyek két egész szám hányadosaként, vagy ami ugyanaz, véges, vagy legfeljebb ismétlődő szakaszos tizedes törtekként állíthatók elő. Ennek megfelelően a racionális számok egyszerűen kódolhatók, két egész szám hányadosát kiszámító program egészen rövid.

A racionális számok korántsem merítik ki a valós számok teljes halmazát. Ellenpéldaként azt is megtanuljuk, hogy pl. a $\sqrt{2}$ nem racionális. Noha a $\sqrt{2}$ tizedestört alakjában semmiféle ismétlődés vagy egyéb szabályosság nincs, egy egyszerű kód mégis könnyen meghatározza az egymást követő jegyeket. Például a következő iteráció $X_{n+1} = 1/(2 + X_n)$ nagy n -re a $(\sqrt{2} - 1)$ -hez konvergál, valahányszor a -2 fölött indítjuk az iterációt, és gyors egymásutánban produkálja a $\sqrt{2}$ -nek a tizedesvessző utáni jegyeit.

Itt azt látjuk, hogy egy rendkívül egyszerű előírás látszólag nagyon bonyolult, soha sem ismétlődő, véletlennak látszó sorozatot generál. Megfordítva: egy látszólag véletlen számsor egyszerű receptbe kódolható, rövid algoritmusba tömöríthető, a végtelen számsorban található információ nagyon erősen redukálható.

A komplexitással összefüggésben számunkra most az a vonás lesz a meghatározó, hogy egy szám a fentiekhez hasonló mó-

don előállítható-e valamilyen meghatározott algoritmussal, azaz létezik-e olyan eljárás, amely lehetővé teszi a szám egymást követő jegyeinek a meghatározását, vagy sem. Ha létezik, akkor a számot *kiszámíthatónak* nevezzük. Láttuk, hogy a racionális számok mind kiszámíthatóak, de kiszámítható az irracionális $\sqrt{2}$, sőt számos ún. transzcendens szám, mint pl. a π is. A kiszámítható számok mindenütt sűrűn helyezkednek el a számegyenesen, ami azt jelenti, hogy bármely két kiszámítható szám között talá-lunk kiszámítható számot, pl. a kettő számtani közepét. Felmerül a kérdés, hogy vannak-e egyáltalán nem kiszámítható számok. Hogy válaszolhassunk, a kiszámítható számok számosságát kell meggondolnunk. A kiszámítható számokat az őket előállító algoritmusok definiálják, számosságuk tehát megegyezik ezeknek az algoritmusoknak a számosságával. Az algoritmusok azonban diszkrét szimbólumokkal leírható utasítások, ezért sorba állíthatók (pl. abcéssorrendben, vagy hosszúság szerint, vagy akármi egyéb módon). Ekként az algoritmusok sorszámmal láthatók el, vagyis annyi van belőlük, mint a természetes számokból: megszámlálhatóan végtelen sok. Ebből következőleg a kiszámítható számokból is „csak” megszámlálhatóan végtelen sok van.

Gondoljuk most el az összes, mondjuk a 0 és 1 közé eső kiszámítható számot. A fentiek szerint ezek mind sorba állíthatók, tehát ugyanannyian vannak, mint az egész számok. Cantor nevezetes diagonális érve szerint azonban az összes 0 és 1 közé eső *valós* számok *nem* megszámlálhatóak, többen vannak, mint az egész számok, végtelenségük minősége más, magasabb az egész számok végtelenségének minőségénél. A (0,1) intervallumba eső valós számok többsége tehát nem kiszámítható, ezek a számok nem állíthatók elő semmiféle egyszerűbb algoritmussal, nem tömöríthetők, nem redukálhatók. Azt látjuk tehát, hogy a 0 és 1 közé eső valós számok döntő többsége irreducibilis, semmi más módon nem ragadható meg, mint végtelen sok jegyük tényleges felsorolásával.

A fenti meggondolás Turing nevezetes 1936-os gondolatmenetének („*On computable numbers, with an application to the Entscheidungsproblem*”) vulgarizált változata, eredeti, szigorúan megfogalmazott alakjában a matematika nem-teljességi tételei közé tartozik.

Kolmogorov és Chaitin javaslatára az algoritmikus információelméletben egy szám-sorozat, vagy bármely karakterfüzér (string) komplexitását annak a legegyszerűbb algoritmusnak a hosszával mérik, amely a fű-zért előállítja. E mérték szerint pl. a $\sqrt{2}$ -t megadó végtelen tizedestört nagyon egyszerű, míg a nem kiszámítható számok maximálisan bonyolultak. Vegyük észre, hogy ez a mérték a teljesen véletlen sorozathoz rendeli a legnagyobb komplexitást, ami nincs

igazán összhangban a komplexitás intuitív fogalmával. A mérték további kifogásolható vonása, hogy maga a legrövidebb algoritmus nem meghatározható, így az sem dönthető el, hogy egy szám nem kiszámítható, vagy csak egyelőre nem jöttünk rá az őt előállító algoritmusra.

A matematikában számos egyéb komplexitás fogalom létezik, de a fenti kis gondolatmenet talán érzékelteti a komplexitás egyik fontos aspektusát, az irreducibilitást.

Nem-teljesség

Az axiomaticus módszert Eukleidész óta a tudományos információ legmagasabb rendű szervezési formájának tekintjük. Végső alakját Hilbert programjában nyerte el. Az általa elképzelt formális axiomaticus rendszer elvileg lehetővé teszi az axiómákból következő igaz tételek mechanikus előállítását. Egy ilyen rendszerrel szemben támasztott nyilvánvaló követelmény a konzisztencia (vagyis hogy ne lehessen az axiómákból egymásnak ellentmondó tételeket levezetni), valamint a teljesség (vagyis hogy az axiómák hatálya alá eső területen minden igaz állítás levezethető legyen belőlük). A matematikában hatalmas megrázkódtatást okozott Gödel 1931-es eredménye, mely szerint ha a számelmélet (pozitív egészek plusz összeadás és szorzás) konzisztens, akkor nem teljes. A bizonyítás önreferencián alapul: Gödel levezetett egy olyan állítást az egészekről, amely azt állította önmagáról, hogy nem bizonyítható.

A nem-teljességi eredmények azt mutatják, hogy a matematika hatalmas része irreducibilis, nem tömöríthető, nem vezethető vissza néhány egyszerű állításra, és a bizonyítások nem gépesíthetők, nem tudunk olyan algoritmust szerkeszteni, amely a következtetési szabályok betartása mellett mechanikusan előállítaná a kérdéses matematikai terület összes igaz tételét.

Utólag megítélve ez inkább megkönnyebbülés, mint tragédia. Érdekes eltűnőnünk azonban, hogy ha egy ilyen egyszerű, ráadásul általunk alkotott rendszer, mint az egész számok matematikája, ennyire tele van titokzatos, igaz, de nem bizonyítható állításokkal, akkor milyen esélyei vannak annak a törekvésnek, hogy a milliárdnyi, saját akarattal és érdekekkel rendelkező ember által alkotott társadalom vagy gazdaság törvényeit néhány egyszerű állításra vezethessük vissza.

Nehéz számítási problémák

Minden nem-teljességi eredmény aszimptotikus, lényegesen kihasználja, hogy a szóban forgó halmazok végtelen nagyok. A gyakorlatban soha nem használunk végte-

len halmazokat, valós számokat, csak racionálisakat sőt, még azok közül is csak az egyszerűbbeket, a véges tizedes törteket, hiszen valószínűleg semmilyen mennyiséget nem ismerünk 20 jegynél pontosabban. Lehet, hogy ez az egész irreducibilitás pusztán elméleti kérdés, és valódi, véges problémákban minden eldönthető véges idő alatt, legrosszabb esetben kimerítő kereséssel?

Nem ilyen egyszerű a helyzet. A számítási problémákat osztályozhatjuk aszerint, hogy milyen gyorsan nő a bevitt adatok N méretével annak az algoritmusnak a futási ideje, amelyek megoldja őket. A kezelhető problémák lineáris ($\sim N$), kvadratikussá ($\sim N^2$), vagy más alacsony fokszámú polinomiális növekedést produkálnak. A kemény problémák exponenciálisan növekvő hosszúságú algoritmusokra vezetnek, ezért praktikus-san megoldhatatlanok. A gyakorlatban is fontos, komplex számítási problémáknak van egy széles osztálya, az ún. NP-teljes osztály¹, melyről azt sejtjük, hogy az ebbe az osztályba tartozó problémák csak exponenciálisan növekvő idő alatt oldhatók meg. A számítási komplexitás problémája nem csupán gyakorlati jellegű, nem technológia kérdése: ha egy számítás hosszabb időt vesz igénybe, mint az Univerzum hátralévő életkora, akkor az bármilyen értelemben véve is kezelhetetlen.

Skálák összeolvadása

Különböző elméleti megfontolások arra utalnak, hogy a Természetben létezik egy 10^{35} m nagyságrendű jellegzetes hosszúság, a Planck-hossz, amely valamilyen értelemben a legkisebb elérhető vagy felbontható távolság. Másrészt a megfigyelhető Univerzum mérete nagyságrendben 10^{26} m. Ezen a szédítő, 61 nagyságrendet átfogó skálán igyekszünk megérteni a jelenségeket. A feladatot lényegesen megkönnyíti, hogy a Természet valamiféle hierarchikus szerkezetet mutat, a különböző jelenségek gyakran jól szeparált távolság (valamint idő, energia stb.) skálákon zajlanak, és a hierarchia egy adott szintjén lezajló folyamatokat úgy lehet leírni, mintha függetlenek lennének az alacsonyabb (vagy magasabb) szintektől. Például: az atomi spektrumok első közelítésben érzéketlenek arra, mi megy végbe a magban. Csak a mag tömege és töltése számít igazán. Ha hidat akarunk építeni a folyón, nem szükséges a felhasznált szerkezeti anyagok atomi szerkezetéről elmélkedni, elég, ha néhány makroszkopikus jellemzőt (sűrűség, rugalmassági modulusok, hőtágulási együttható, stb.) tekintetbe vesszünk. A skálák elkülönülése hatalmas mértékben leegyszerűsíti a feladatunkat. Ugyan-

¹ Megjegyzés: ezzel a témával foglalkozik *Ercsey-Ravasz* és *Toroczkai* cikke.

A komplexitás kihívása

ez teszi lehetővé a különböző tudományok illetékességi körének a kijelölését, és ez kelti a különböző tudományok egymástól való függetlenségének az illúzióját is.

Amikor a skálák jól szeparálódnak, a redukció működik. A skálák azonban számos fontos esetben nem különülnek el; a jellegzetes skálák hiánya magyarázza meg a hatványfüggvény-szerű eloszlások feltűnő gyakoriságú fellépését a komplex rendszerek elméletében. A gazdasági szereplők között óriási különbségek vannak, a vállalati méretek és jövedelmek hatványfüggvény-szerű eloszlást mutatnak, a leghatalmasabb játékosok gazdasági ereje azonos nagyságrendbe eshet, mint az egész gazdaságé. Az óriásbankok mérlegfőösszege nem egy országban meghaladja az egész ország bruttó hazai termékének (GDP) az értékét, ami még azt is kérdésessé teszi, hogy mi a része minnek: a bank az ország gazdaságának, vagy a gazdaság a banknak. Egy ennyire heterogén, erősen kölcsönható sokaságra értelmetlen átlagolni, a makroökonómia levezetése a mikroökonómiából valószínűleg lehetetlen feladat. A nagy számok törvényeire, a határeloszlásokra való hivatkozás súlyosan félrevezető, az ingadozások sokkal erősebbek, mint amit a tankönyvek kedvenc haranggörbéi megengednének. Ezek az eltérések egyáltalán nem is ártatlanok: a nagy események valószínűsége összemérhetetlenül nagyobb annál, mint amit a standard elméletek jósolnának. Az 1987. október 19-i Fekete Hétfőn a világ tőzsdéin bekövetkezett zuhanás „25 szigmányi” esemény (vagyis a napi árfolyam-ingadozások szórásánál 25-ször nagyobb kilengés) volt; a normális (Gauss-) eloszlás szerint ekkora ingadozásnak a Világegyetem egész története során nem lett volna szabad bekövetkeznie.

A skálák összeolvadása általában jellemző a komplex rendszerekre. A biokémiai folyamatokban, az idegrendszerben, a társadalomban vagy a gazdaságban nincsenek jól elkülönülő (energia-, idő-, vagy térbeli) skálák.² Ezért nem működnek a határeloszlás tételek, nincs egyszerűsödés a nagy számok limeszében; a rendszer irreducibilis, nagyon sok részlet pontos ismerete kell ahhoz, hogy megbízható leírást adhassunk róla.

A közgazdaság-tudományban újabban számos kísérlet történt az egyensúly-elmélet irreális egyszerűsítő feltevéseinek meghaladására. Terjedőben vannak az ún. heterogén ágens modellek, amelyek a gazdaság szereplőinek részletesebb, gazdagabb leírását teszik lehetővé, és megengedik a szereplők alkalmazkodását, tanulását is. Miközben ettől a leírás kétségtelenül valószínűbb lesz, a modellt meghatározó paraméte-

rek száma hatalmasan megnő, s ettől ezeket a modelleket nagyon nehéz a megfigyelésekhez kalibrálni és a tényleges gazdasági folyamatokkal összevetve igazolni őket.

Mindemellett a leírás bonyolódásával egyre távolabb kerülünk attól, amit szokásosan elméletnek tekintünk. Egy olyan modellt, ami 1000 paramétértől függ lényegesen, lehetetlen fogalmilag megragadni, nem tudjuk vele kapcsolatban elérni azt a belső meggyőződést, hogy „értjük”, mi történik a modellben, legfeljebb tudomásul vesszük azokat az eredményeket, amelyeket a számítógép kidob.

Van egy további nehézség is. A komplex rendszerek fejlődésük során gyakran érkeznek el olyan választópontokhoz, ahonnan a további fejlődés több, különböző irányt vehet, és a választást véletlenek, lényegtelennek tűnő előzmények határozhatják meg, de miután a választás megtörtént, a korábbi állapothoz már lehetetlen visszatérni. Nyilvánvaló, hogy minden lehetséges ilyen véletlen körülményt soha nem leszünk képesek beleépíteni a mégoly realiztikus modelljeinkbe sem.

Konkrét példa gyanánt gondoljunk a jelenlegi válság kibontakozására. A Lehman bankház összeomlását követő napokban, 2008 szeptemberében az Egyesült Államok politikai és pénzügyi vezetői drámai döntés elé kerültek: vagy sorsukra hagyják a megrendült bankokat, és ezzel kirobbantanak egy világméretű gazdasági összeomlást, vagy minden saját konzervatív, az állami beavatkozást elutasító politikai meggyőződésükkel ellentétben (2008-ban még G. W. Bush volt az elnök!) keresztülnyomnak a hevesen ellenálló Kongresszuson egy példátlan méretű mentőcsomagot, ami ráadásul a saját republikánus szavazók dühödt ellenkezésébe is ütközik. Több független kommentátor szerint abban, hogy a bankmentő kampány beindult, döntő szerepe volt az amerikai központi bank, a Federal Reserve elnökévé kevéssel előbb kinevezett Ben Bernanke-nek, aki megelőzőleg, princeptoni professzorként, az 1929-es világválság kialakulásával, és annak egészen a II. világháborúig vezető következményeivel foglalkozott. A válságot így sem lehetett teljesen elkerülni, de a mentőcsomagnak és az azt követő hatalmas méretű pénzkibocsátásnak a következtében az Egyesült Államok kevésbé szenvedte meg, s mintha hamarabb épülne fel belőle, mint a különböző megszorításokkal operáló Európa, amelynek a pénzhígitást sokáig elutasító magatartását viszont a vezető európai gazdasági hatalomnak, Németországnak megint csak az 1920-as évekig visszanyúló inflációs főbiái határozták meg.

Az ilyen példák láttán komoly kételyek támadhatnak bennünk a heterogén ágens modellek teljesítőképességével kapcsolatban. Elképzelhető olyan tudományos modell, amelyben a FED elnökének a személyisé-

ge és szakmai előzményei, vagy a német társadalom kollektív inflációs emlékei is szerepelnek? Bármely létező, vagy elképzelt modell képes lehet-e ezeknek a politikai döntéseknek a következményeit megjósolni? Mondjuk az Európai Unió esetleges politikai megroppanását, vagy a dollár hosszabb távon bekövetkező esetleges értékvesztésének a világpolitikai következményeit?

Az ilyesfajta nehézségek egyáltalán nem korlátozódnak a gazdasági folyamatok leírására. Az élővilág evolúciója ugyancsak tele van ilyen feltételes, véletlenszerű elágazási pontokkal, amik lehetetlenné teszik pl. az evolúció további menetének megjóslását, legfeljebb annyit engednek meg, hogy a bekövetkezett fejleményeket utólag megpróbáljuk értelmezni.

Úgy tűnik, a nagy szabadsági fokú, heterogén, erősen kölcsönható rendszerek leírásában, megértésében, kontrolljában és viselkedésük előrejelzésében komoly, talán áthághatatlan akadályokba ütközünk.

Irreducibilitás

Korunkat gyakran jellemzik az információ, az adatok túlzott bőségével, számos tudományterületen az áradó adattömeg kezelése, értelmezése okoz gondokat. Itt most arra szeretnék rámutatni, hogy ezzel ellentétben egy komplex rendszerről valójában soha nem lehet túl sok, vagy akár csak elegendő mennyiségű adatunk.

Képzeld el, hogy egy bonyolult jelenség magyarázatát keressük. Tételezzük fel, hogy a jelenség egy mérőszámmal jellemezhető, jelöljük ezt a mérőszámot y -nal. Az y lehet egy kardiológiai osztályra infarktussal bekerülő betegek túlélési esélye, egy vállalat éves profitja, az ország GDP-jének éves változása, vagy bármi hasonló jellemző. Ez az y egy egész sor tényezőtől függhet. A kardiológiai példában ilyenek lehetnek a kórház felszereltsége, a személyzet felkészültsége, a kórház vonzáskörzetében élők életkor, jövedelem, műveltségi szint szerinti eloszlása, a lakosság táplálkozási és dohányzási szokásai, a közlekedési viszonyok, amelyek meghatározzák a kórházba jutás idejét, és egy sereg további faktor – egy kanadai epidemiológiai cikkben több mint 40 ilyen tényezőt soroltak fel. A gazdasági példákban könnyedén össze tudnánk szedni hasonló számú magyarázó változót. Jelöljük ezeket x_i -vel. Első közelítésben sokszor feltételezik, hogy az y lineáris függvénye a magyarázó változóknak, de a valóságban ez a legtöbb esetben nincs így. Tegyük fel, hogy rendelkezésünkre állnak a $t = 1, 2, 3, \dots, T$ időpontokban mért összetartozó $y(t)$ és $x_i(t)$ értékek. Ha a magyarázó változók N száma nem túl nagy, a rendelkezésre álló idősorok viszont hosszúak, vagyis $T \gg N$, és az összefüg-

² Megjegyzés: erre utalnak a faj-, illetve stratégiatársulások kialakulási is Boza és Scheuring, illetve Szolnoki cikkében.

Kondor Imre: A komplexitás kihívása

gés jó közelítéssel lineáris, akkor a feladat a matematikai statisztika módszereivel értelmesen megoldható, a megoldás meg fogja mutatni, hogy az egyes x_i változók mekkora súllyal szerepelnek az y meghatározásában, de azt is eldönti, hogy elegendő számú változót vettünk-e fel, ill. hogy az így megalkotott modell mennyire megbízható, mennyire használható a következő mérések eredményének előrejelzésére.

A nehézségek akkor kezdődnek, amikor a magyarító változók száma, a modell dimenziója, N nagyon nagy. Ilyenkor gyakran előfordul, hogy nincs elegendően nagyszámú mérési adatunk, pl. azért, mert túl költséges lenne beszerezni őket, vagy azért mert, nem is léteznek (pl. egy sokdimenziós GDP előrejelző modellhez nem találunk elegendő számú adatot, mert a GDP-t csak a II. világháború vége óta mérik, évi gyakorisággal). Az ilyen helyzet már a legegyszerűbb, lineáris esetben is a becslési hiba megnövekedéséhez vezet, a modell előrejelző képessége elvész. Amikor $T=N$, a becslési hiba minden határon túl nő, a $T < N$ esetben pedig a feladat értelmetlenné válik, kevés számú mérési adat nem határoz meg egy magas dimenziós számú modellt (egy pont nem határoz meg egy egyenest, két pont nem határoz meg egy síkot).

Ha belegondolunk, a mindennapi élet tele van ilyen magas dimenziós problémákkal: a közlekedésben nagyszámú másik jármű mozgásának rövid idejű megfigyelése alapján döntünk arról, hogy gyorsítsunk, vagy fékezzünk, iskolai, munkahelyi, társadalmi érvényesülésünket barátok támogatják, és versenytársak akciói keresztezik, az üzleti életben, a családban, az élet minden területén igen magas dimenziós problémákkal küzdünk. Valami módon legtöbbször mégis sikerül túlélünk a közlekedést és az étellel járó versenyeket, konfliktusokat. Hogyan csináljuk? Az biztos, hogy nem készítenünk optimálisan illesztett statisztikai modelleket. Ehelyett a felhalmozódó tapasztalatokból egyszerű hüvelykujj-szabályokat vonunk le, egyszerű heurisztikákat alkalmazunk, amelyek persze néha kudarcra vezetnek, de szerencsés esetben ebből tanulva javítgatjuk a heurisztikus eljárásainkat.

Valami ennek megfelelően csinál a modern statisztika, a statisztikus tanuláselmélet is. Miután felismeri, hogy a kevés adatra illesztett modell lényegében a zaj alapján jósol, tudatosan torzít a becslésen, az adatoktól függetlenül rendelkezésre álló ismereteket vagy elvárásokat vezet be a magyarító változók értékének a meghatározásába. Formálisan ez a modellhez adott regularizáló tagok segítségével történik.

A legegyszerűbb ilyen háttérismeret az lehet, ha valahonnan eleve tudjuk, hogy a probléma valójában alacsony dimenziós, vagyis csak néhány magyarító változóval kell számolnunk, de nem tudjuk, hogy melyek

ezek. Létezik olyan regularizációs eljárás, ami elvégzi a jelentős változók kiválasztásának a feladatát akár igen zajos adatokból is. Máskor azt gondoljuk, hogy a komplikált környezetben, pl. a tőzsdén, a legbiztonságosabb stratégia, ha az ismert magyarító változók között nem teszünk lényeges különbséget, nagyjából egyenletesen osztjuk el a befektetéseinket az értékpapírok egy bizonyos készletén. Olyan regularizációs eljárás is létezik, amely az egyenletes eloszlás irányába húzza el az empirikus adatokból leszűrhető következtetéseinket.

A regularizáció, a gépi tanulás eszközei igen hatékonyak a látszólag komplex feladatok megoldásában, pl. a kézírás felismerésében, a beszédfelismerésben vagy a zajos képek szűrésében, visszaállításában. Ezekben a feladatokban viszonylag csekély számú magyarító változó szerepel, a többiek súlya egy bizonyos indextől kezdve nagyon kicsi, van egy karakterisztikus dimenziószám, ami fölül nem kell mennünk. Az ilyen természetű problémák látszólag nagyon bonyolultak, de valójában jól tömöríthetők, redukálhatók, nem igazán érdemlik meg a komplex minősítést.

Az igazán komplex problémákat viszont éppen az jellemzi, hogy nincs olyan karakterisztikus dimenziószám, amely fölött a magyarító változók lényegében eltűnnének. Ez már a lineáris esetben is előfordulhat, ha az x_i változók együtthatói az i függvényében hatványszerűen csökkennek. Erre a gazdaság adhat példát: minthogy a gazdasági szereplők tökéreje hatványfüggvény-szerű eloszlást mutat, első közelítésben a piacra tett hatások is ilyen. A legnagyobb játékosok aránytalanul nagy járulékot adnak a piaci ingadozásokhoz, viszont a kicsikből nagyon sok van, ezért ők sem felejthetők ki, amikor pl. a gazdaság következő évi teljesítményét akarjuk előre jelezni. Nem ismeretes olyan regularizációs eljárás, amely egy ilyen komplex statisztikai feladatot megoldna oldani. A hatványszerűen eső együtthatókkal bíró modell a komplexitásra jellemző irreducibilitást mutatja: nem strukturálatlan, nem minden együttható egyenlő, de olykor az igen kis súlyú változók is jelentős kilengést okoznak, és a modell kalibrációjához tetszőlegesen sok mérési adat sem igazán elegendő. A feladatot óriási mértékben nehezíti, ha a magyarító változók között kölcsönhatások vannak, a modell nem lineáris. Ha egy ilyen természetű folyamat véges számú magyarító változóval rendelkező modellel írunk le, akkor a modell az esetek nagy részében értelmes előrejelzéseket adhat, de a folyamat ritkán előforduló nagy kilengéseit nem fogja tudni megjósolni.

Pontosan így viselkedtek a központi bankok és más elemző szervezetek kedvenc DSGE (Dynamic Stochastic General Equilibrium) modelljei a 2007–08-as válság esetében: 2008 májusában még semmit nem

jeleztek előre abból, hogy az amerikai gazdaság alig 2%-át kitevő „subprime” jelzálogpiac válsága pár hónapon belül az egész amerikai gazdaságot, majd közvetlenül utána az egész világ hitelpiacát megbénítja. A gazdasági folyamatok modellezését, számos egyéb probléma mellett, ilyen természetű okok is nehezíthetik: a nagyszámú, heterogén szereplő összjátéka kiszámíthatatlan fluktuációkat eredményezhet, amelyeket a nemlineáris kölcsönhatások váratlan mértékben felnagyítanak, és lényegében előreláthatatlanná tesznek.

Mit várhatunk ezek után a komplex rendszerek megértését, modellezését, előrejelzését vagy kontrollját illetően? Alacsony dimenziós, néhány paraméterrel jellemzett, fogalmilag könnyen megragadható elméletekre számítani nyilvánvalóan alaptalan illúzió. Lehetséges, hogy bizonyos problémákban nem marad más választásunk, mint óriási méretű számítógépes szimulációkkal „eljátszani” a valóságos modellt, abban reménykedve, hogy eközben sikerül kifejlesztenünk valamiféle intuíciót a főbb változók szerepét illetően. Ez a tudományt egészen új irányokba terelné, és felidézti azt a veszélyt, hogy a civilizáció meghatározó alrendszereit (közlekedés, élelmiszerellátás, energia-, víz- és információs szolgáltatás, közegészségügy, biztonságpolitika, választások stb.) értelmünk számára felfoghatatlan bonyolultságú algoritmusok uralmának engedjük át.

Valamivel biztatóbb példát a biológia szolgáltatott a múlt évben: a Stanford Egyetem és a J. Craig Venter Intézet kutatói először voltak képesek egy valódi (igaz, hogy a legkisebb számú génnel rendelkező) élőlény, a *Mycoplasma genitalium* nevű baktérium teljes életciklusát számítógépes szimulációval utánózni, beleértve a sejtosztódást is (Cell, Volume 150, Issue 2, 389–401, 20 July 2012). A kutatók a baktérium molekuláris szerkezetének minden elemét (DNS, RNS, fehérjék stb.) és kölcsönhatásaikat figyelembe vették. A nagyon magas dimenziós problémát számos kisebb dimenziós alprobléma moduljaiból építették fel, amelyeket részben saját, részben mások kísérleti eredményei alapján külön-külön ellenőriztek, majd integráltak. Az így kapott modell nem elegáns és nem „szép”, viszont az első lényegében hű leírását adja egy ténylegesen létező élőlénynek.

Nem láthatjuk előre, hogy a komplex rendszerek kutatása milyen irányt vesz a jövőben, abban azonban biztosak lehetünk, hogy ha ezeket a problémákat egyáltalán képesek leszünk leküzdeni, az többet jelent majd a tudomány ismétlődő forradalmainál, „paradigmaváltásainál”: az irreducibilis problémákkal való birkózás alapvetően fogja megváltoztatni annak a tevékenységnek az egész karakterét, egész stílusát, amit tudománynak, tudományos kutatásnak tekintünk. ❀

BOZA GERGELY–SCHEURING ISTVÁN

Diverzitás és komplexitás a mutualista kapcsolatokban

Ki ne heverészett volna már egy nap-sütéses erdei tisztáson kora nyári délután? Már nyílik ilyenkor a harangvirág és a százsorszép, távolabb egy csoportban talán boglárka is sárgállik, a rét közepén pedig néhány öreg tölgyfa terpeszkedik. (Nem beszélve arról a több tucatnyi lágyszárú növényfajról, amelynek a legtöbb turista még a nevét sem ismeri.) Kicsit távolabb, talán épp egy másik rét fölött, olyvek köröznék. Néha vijjogó hangjuk is ide hallik, bár a sáskák, méhek és az énekesmadarak elég nagy „zajt” csapnak a közelben.

Egy erdei rét, valamennyi élőközösséghez hasonlóan, nem véletlenszerűen egymás mellé rakogatott élőlények összessége, sokkal inkább rendkívül szövevényes kapcsolatban lévő fajok együttese. Sok esetben az egyedek (akár különböző, akár azonos fajba tartoznak) versengenek valamilyen forrásért (pl. a vízéért, a fényért vagy valamilyen táplálékért), ilyenkor egymás túlélési esélyét és vagy szaporodási sikerét (azaz rátermettségét) kölcsönösen csökkentik. Például egy napfényért versengő fa kénytelen a szár növekedésére fordítani a fotoszintézis során megszerzett energiát, különben a szomszéd fa magasabbra nő, és leárnyékolja. Ez költséges a fának, hisz annyival is kevesebb energia jut a szaporodásra, és hátrányos az alacsonyabb szomszédnak, hisz kevesebb napfény jut el hozzá. A növényeket rágó sáskák vagy a pockokra vadászó olyvek számára a másik faj táplálék, a kapcsolat a táplálkozó félnek előnyös, a tápláléknak viszonyt nyilván hátrányos. De van a kölcsönhatásoknak egy olyan népes csoportja is, mely a résztvevő felek számára kölcsönösen előnyös. Ott vannak például a réten a méhek, melyek a nektárért cserébe beporozzák a virágokat. E kapcsolat vélhetően mind a két fél rátermettségét növeli. Aztán – bár szabad szemmel nem látható – legalább ilyen jelentős, kölcsönösen előnyös kapcsolatrendszer alakul ki a talajlakó mikroorganizmusok és a növények között is, és a felsorolást folytathatnánk tovább.

A különböző fajokhoz tartozó egyedek között meglévő *kölcsönösen előnyös kapcsolatokat* összefoglalva *mutualizmusnak* nevezzük [1]. Írásunkban a mutualista kapcsolatok sokféleségéről, azok evolúciós változásait érintő kérdésekről, valamint az

egész bioszféra működésében betöltött kulcszerepéről szólnak. A mutualizmus helyett gyakran használják a *szimbiózis* kifejezést, mely sokaknak talán ismerősebben cseng. A két fogalom valóban nagyon hasonló dolgot takar, esetenként felcserélve is használható. A szimbiózis azonban a mutualizmus olyan alosaja, amikor szoros, tartós fizikai kapcsolat van a kölcsönhatásban résztvevők között. A szimbiózis egyik nagyon speciális formája az *endoszimbiózis*, ebben az esetben a szimbionta a gazda szervezetben, sejtjeiben él, s adódik át generációról generációra. A bioszférában fellelhető kölcsönösen előnyös kapcsolatok közül a legrégebbiek éppen az ősi endoszimbiotikus kapcsolatokról kialakult sejtszervecskék, sejtorganelumok, mint a kloroplasztisz vagy a mitokondrium.

A mutualista kapcsolatok diverzitása

Amikor a biológiában sokféleségről (diverzitásról) beszélünk, minden esetben pontosan meg kell határozzuk, hogy milyen szempontból vizsgáljuk a diverzitást. *A biológiai sokféleség maga is sokféle.* Vizsgálhatjuk egy populáción belül a genetikai sokféleséget, vagy az egyedek viselkedésbeli diverzitását, rákérdezhetünk egy élőközösségben élő fajok sokféleségére, de akár a köztük lévő kölcsönhatások diverzitását is mérhetjük. Tehát, amikor biológiai sokféleségről beszélünk, akkor minden esetben meg kell, hogy mondjuk, hogyan is értjük azt. Cikkünkben elsősorban a mutualista kapcsolatokba, és azok sokféleségébe szeretnénk bepillantást nyújtani. Ahogy a korábban bemutatott néhány példa is érzékelteti, a mutualizmus egymástól sokféle szempont alapján jól elkülöníthető kapcsolatrendszereket fog egybe, egy nagyon diverz jelenségcsoport gyűjtőnéve.

A mutualista kölcsönhatásokat leegyszerűbben úgy tudjuk elképzelni, mint különböző fajtól származó egyedek között kialakuló cserekereskedelmet, amiben a felek változatos „árucikkeket”, „szolgáltatásokat” adhatnak partnerüknek [1, 2]. A már említett beporzás esetén a beporzást végző fél (rovar vagy gerinces) egy *szolgáltatásért cse-*

rébe tápanyaghoz jut. Ilyen jellegű kapcsolat még a magterjesztés, a paraziták eltávolítása a gazdaállatból és a gazdanövény védelme a növényevők ellen. Közös jellemzője ezeknek a kapcsolatoknak, hogy valamilyen energiagazdag táplálékért cserébe a gazda védelmét vagy szaporodását elősegítő segítséget nyújt a mutualista partner, mely utóbbi esetben a szolgáltatás általában valamilyen szállítással függ össze (pollen, mag szállítása).

A bevezetőben már említettük, hogy a talajban élő mikroorganizmusok (gombák, baktériumok) és a növények között is igen gyakran kialakulnak mutualista kapcsolatok. Ilyenkor a növény valamilyen energiagazdag szerves anyagért (cukor, keményítő) számára nehezen elérhető szerves anyagot kap, azaz *tápanyagért tápanyag* jár cserébe. Az olyan esetek, amikor a *szolgáltatásért szolgáltatás* jár viszonzásul, sokkal ritkábbak, és sokszor valamilyen tápanyag is szerepet játszik a kölcsönhatásban, de például a hangya-növény kapcsolatokban a növény általában nemcsak táplálékkal látja el a hangyát, hanem szolgáltatásként speciális, a növény számára sokszor igen költséges lakóhelyet is biztosít a kolónia számára. A hangyák cserébe megvédik a növényt egyes növényevőktől, illetve a gazdanövény közelében növekvő más növények káros hatásától. Ez utóbbi olyan hatásos lehet – mint azt kimutatták a dél-amerikai trópusi esőerdőkben élő citromhangyák (*Myrmelachista schumanni*) és gazdanövényük (*Duroia hirsuta*) kapcsolatának vizsgálata során –, hogy az egyébként sokszínű, buja erdőben, kizárólag a gazdanövény által elfoglalt helyen, kísérletesen üres foltok alakulnak ki, melyeket a helyiek az „Ördög kertjének” neveznek. Természetesen nem természetfeletti erők kertészkedése, hanem a hangyák által a nemkívánatos növényekbe fecskendezett, mérgező hangyasav révén alakulnak ki ezek az üres foltok, amit aztán a gazdanövény akadály nélkül elfoglalhat. Egyes tengeri rákok, például a remeterákok, hogy védjék magukat a ragadozók ellen, tengeri rózsákat vagy egyéb mérgező élőlényeket pakolnak a hátukra, melyek cserébe a rák által elhullajtott táplálékból fogyasztanak, illetve a rák hátán utazva helyzetet változ-

tathatnak, terjedhetnek. A gazdaszervezet és a benne (rajta) élő mikrobaközösségek kapcsolata még összetettebb: a gazda a tápanyag mellett speciális életfeltételeket is biztosít a mikrobák számára, és a mutualista mikrobák nemcsak tápanyaggal fizetnek érte, hanem védenek a kórokozók ellen is. Éppen ilyen kapcsolatban vannak a gerincesek, így az ember bélrendszerében élő baktériumközösségek a gazdaállattal (lásd pl. [3]). Azaz a kölcsönhatás sokrétű, a felek szolgáltatásokat és tápanyagokat is juttatnak a másik félnek (**1. táblázat**) [1]. A legtöbbször számos faj dinamikusan változó laza kapcsolati hálózatban alakítja ki a „cserekereskedelmet”, és egy cserepartner időleges kiesése sem okoz különösebb zavart (azaz a kapcsolat nem specifikus), és a legtöbb faj a partner hiányában is életképes marad (azaz a kapcsolat fakultatív). De vannak olyan esetek is, amikor két faj között nagyon szoros mutualista kapcsolat van, s az egyik fél tartós hiánya a másik fél populációjának az eltűnéséhez vezet, azaz a kapcsolat obligát (**1. táblázat**).

A mutualista kapcsolatok evolúciós stabilitása

A fizikai vagy kémiai rendszerek vizsgálatakor talán a legfontosabb kérdés a rendszer dinamikai viselkedése, azaz, hogy milyen időbeli és térbeli mintázatokat tapasztalunk benne, és ennek mi az oka, valamint ezen mintázatok hogyan függenek a rendszer tulajdonságaitól. A biológiai rendszerekben e kérdés mellett minden esetben legalább olyan fontos kulcskérdés a jelenség evolúciós háttere is. A mutualista kapcsolatrendszerre gondolva, két izgalmas evolúciós kérdést fogalmazhatunk meg: 1. Hogyan alakultak ki az evolúció során a fajok közötti kölcsönösen előnyös kapcsolatok? 2. Milyen mechanizmusok akadályozzák meg, hogy a család mutánsok kihasználják a mutualistákat, s így szétrombolják a mutualista kapcsolatokat?

Az előző fejezetben hangsúlyoztuk, hogy a mutualizmus nagyon sokrétű jelenségcsoport gyűjtőneve, ezért nem várhatjuk, hogy e jelenségcsoport evolúciós eredetét és stabilitását egyetlen mechanizmus képes lesz megmagyarázni. Mint oly gyakran a biológiában: nemcsak a jelenségek rendkívül összetettek, hanem azok magyarázatai is igen sokfélék. Ráadásul az 1. kérdésre megbízható választ csak a régmúlt (60 vagy 400 millió évvel ezelőtti) élővilágának pontos ismeretében adhatunk. A biológiai régmúlt feltárása azonban komoly kihívás, hiszen az evolúcióból részben történeti tudomány, ahol a leletek megléte vagy hiánya döntő fontosságú lehet a történet feltárásában. A 2. kérdésre viszont a ma élő mutualista kapcsolatok vizsgálatával is nagy eséllyel tudunk pontos válaszo-

kat adni. Természetesen a két kérdés összefügg, hiszen ha tudjuk, hogy a ma élő mutualisták hogyan védekeznek a család ellen, akkor az sok esetben arról is informál, hogy hogyan jelentek meg ezek a kapcsolatok az evolúció során, és fordítva. A cikkben főként a 2. kérdésben megfogalmazott evolúciós stabilitás problémájára fogunk koncentrálni.

A mutualista kapcsolat lényege, hogy az együttműködő felek számukra költséges javakkal (nektár, szerves molekulák, szolgáltatások stb.) látják el a partnerüket [1, 2]. Képzeljük el, hogy valamelyik mutualista fél egyedei között megjelenik egy olyan

mutáns, aki kevesebbet (vagy egyáltalán nem) szolgált partnerének (alig vagy egyáltalán nem termel nektárt, nem ad nitrogént a szénhidrátért cserébe stb.). Egy-két gén véletlenszerű megváltozása is elegendő az ilyen jellegű változatok megjelenéséhez. Ha az így megjelent kevésbé mutualista egyedek (szélsőséges esetben családok) a mutualista kapcsolatból ugyanakkora nyereségre tesznek szert, mint a költségesebb viselkedést követő együttműködők, akkor több energiájuk marad, melyet túlélésükre, szaporodásukra tudnak fordítani, tehát nagyobb eséllyel fognak elterjedni a populációban. Ennél fogva újabb és újabb, egyre

1. táblázat. A mutualista kapcsolatok osztályozása a kapcsolat élettani fontossága, a kapcsolat fizikai szorossága és a partnerek egymásnak nyújtott szolgáltatása alapján

Csoportosítási szempont	Csoport	Biológiai példa	Rövid leírás
Kapcsolat fontossága alapján	Obligát	<i>Füge –fügedarázs</i>	A fügedarázs kizárólag a fügenövény virágzatába helyezett, és abban kifejlődő petékkel képes szaporodni. A nősténydarazsak megtermékenyítése is a termésben zajlik. A fügevirág megporzását kizárólag a fügedarazsak végzik.
	Fakultatív	<i>Tisztogató-halak és klienseik</i>	A korallzátanyoknál gyakran megfigyelhető kölcsönhatás, melynek során a nagyobb halak külsejéről, szájából távolítják el a parazitákat, beteg szöveteket a tisztogató halak. Bár mindkét fél számára előnyös a kapcsolat, mindegyikük képes a kölcsönhatás létrejötte nélkül is túlélni, szaporodni.
Fizikai kapcsolat szorossága alapján	Szoros, folytonos fizikai kapcsolat: szimbiózis	<i>Gazdanövény nitrogénfixáló baktériumok</i>	A gazdanövény speciálisan erre kialakított képletében, a gyökérgümőben lát „vendégül” baktériumokat, melyek szénhidrátokért cserébe a levegőből megkötött szerves nitrogént juttatnak a gazdának. A gyökérgümő napokig, esetleg hetekig fennmarad, miközben folytonos az anyagcsere a partnerek között.
	Nincs folytonos fizikai kapcsolat	<i>Beporzás</i>	A partnerek csak igen rövid ideig kerülnek kapcsolatba beporzáskor; a beporzást végző állat esetenként csak néhány másodpercig marad egy partneren.
Szolgáltatás típusa alapján	Tápanyagért – tápanyag	<i>Növény –mikorrhiza</i>	A cserekereskedelem alapja valamely a növény számára nehezen elérhető szervetlen vegyület, mint víz vagy foszfor, melyért cserébe a fotoszintézis során előállított szénhidrátokat, cukrokat ad a gazdanövény a gombának.
	Tápanyagért – szolgáltatás	<i>Hangyák és a <i>Lycaenid</i> lepke hernyói</i>	A hangyák a védelemért cserébe a hernyók egy speciálisan erre a célra kifejlődött szervéből kinyert nektárral táplálkoznak.
	Szolgáltatásért – szolgáltatás	<i>Gazdanövény –hangyák</i>	Számos hangya alakít ki kolóniát gazdanövényeken, melyek akár a növény szaporodását is hátráltathatják. A hangya azonban a menedékért cserébe távol tartja a növényt veszélyeztető növényevőket, szomszédos növényeket.

Diverzitás és komplexitás a mutualista kapcsolatokban

kevésbé mutualista mutánsok fogják kiszorítani a korábbi típusokat, és ezért egy idő után teljesen megszűnik a javak szolgáltatása. Tehát, amennyiben az előbbi feltevések igazak, a mutualista kapcsolatok evolúciósan instabilak [2, 4]. Azaz a természetben ritkán, s csakis időlegesen látnánk ilyen kölcsönhatásokat. Azonban, ahogy azt a bevezetőben is említettük, a tapasztalatunk merőben eltér ettől, hiszen a mutualista kapcsolatok behálózják az élővilágot. A fajok döntő többsége hosszú ideje tagja legalább egy, de gyakran több kölcsönösen előnyös kapcsolatrendszernek!

Vajon mi lehet ennek az ellentmondásnak az oka? Célszerű ilyen esetekben megvizsgálni az érvelésben használt feltevéseket. Előzőleg azzal érveltünk, hogy olyan csaló mutáns tud elterjedni, és tudja lerombolni a mutualizmust, mely kevesebbet fektet be, de a jó mutualistához hasonló nagyságú nyereségre tesz szert. De valóban ugyanakkora lesz a nyereségük a csalóknak, mint a mutualistáknak? Rövid válaszuk, hogy nem! Számos, ma már kísérletekkel is igazolt mechanizmus akadályozza meg, hogy a csalók a mutualista partnerekkel minden szempontból azonos kapcsolatokat alakítsanak ki, mint mutualista fajtársaik (**2. táblázat**) [1, 2, 5].

E gondolatmenet alapján talán a legkönnyebben a növények gyökérzetéhez kapcsolódó mikroorganizmusok (mikorrhizák, nitrogénfixáló baktériumok) közötti kölcsönösen előnyös kapcsolatrendszer evolúciós stabilitása érthető meg [6]. Ugyanis, ha a gyökérzet valamely részén megjelenik és elkezd szaporodni egy önző mikrobatorzs, akkor a gyökérzet azon része valamilyen hátrányba kerül, például kevesebb nitrogénhez vagy foszforhoz jut. Ennek hatására a gyökér rész növekedése is lassabb lesz vagy akár teljesen le is állhat. Természetesen a sanyarúbb körülmények között élő gyökér rész a mikrobáknak is kevesebb szerves molekulát fog szállítani. A csalás visszahat a csalókra, azok nyereségét közvetlenül csökkenti, s így terjedésüket is megakadályozza [1, 2].

A szakértők azonban vitatkoznak azon, hogy a növény reakciója csupán egy automatikus élettani válasz, vagy a növény aktívan bünteti a csalókat azzal, hogy az életlenül indokoltnál rosszabb körülményeket teremt a csaló által elfoglalt gyökérrészekben.

A biológia komplexitást ismerve nem nehéz megjósolni, hogy mind a két mechanizmusra fogunk egyértelmű példát találni a közeljövőben. Másrészt a gazdanövény sem tudja becsapni a mikroorganizmusait. Ugyanis az a növény, mely nem, vagy alig szolgáltat tápanyagokat a mikroorganizmusoknak, az a saját „jötevőit” sanyargatja, így azok rosszabbul teljesítenek, végül maga is rosszabbul fejlődik, mint a partnereit jól ellátó növények. Ennek következtében a

A mutualizmust fenntartó mechanizmusok	A mechanizmus rövid leírása	Biológiai példák
Partnerek csatolt érdeke	A gazda sikere növeli a szimbióta túlélési és szaporodási esélyeit, és viszont.	A gazda szervezetében, sejtjeiben élő partnerek, az endoszimbióták hosszú távú, szoros kapcsolata.
Partnerválasztás, előszűrés	A felek a kapcsolat kialakulásának kezdeti fázisában jelek alapján igyekeznek a megfelelő partnert kiválasztani, és a csalókat kizárni.	A talaj mikrobiális közösségéből a megfelelő nitrogénfixáló <i>Rhizobium</i> baktériumok kiválogatása különféle kommunikációs mechanizmusokon keresztül.
Partnerelutasítás, szankciók	A kapcsolat során a felek folyamatosan vizsgálják partnereik befektetéseit. Amennyiben valamelyik fél csalóként viselkedik, partnere csökkenti vagy megszünteti a befektetést.	A nitrogént nem, vagy csak kevésbé fixáló <i>Rhizobium</i> baktériumok leépítése, szelektív leválasztása; függő-függedarázs kapcsolatban a függenővíny szelektíven leválasztja a túl sok petét tartalmazó virágzatokat.
Melléktermék mutualizmus	Az egyik fél nyújtotta nyereség ugyan nem a partner megsegítésére fejlődött ki, ám bizonyos esetekben a partner mégis élvezheti ennek előnyét.	A remeterák és tengeri rózsza kapcsolatában a tengeri rózsza védekezési mechanizmusa a saját védelme érdekében alakult ki, ám a rák hátán található egyed, mellékesen, a rákot is megvédi. A tengeri rózsza viszont olyan táplálékot kap a ráktól, mely a rák számára hulladék.

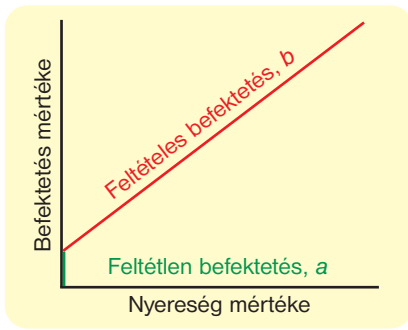
2. táblázat. A mutualista kapcsolatok evolúciós stabilitását fenntartó mechanizmusok osztályozása

csaló növény idővel kiszektálódik a populációból.

A bemutatott esetekben fontos tényező, hogy a növény, az állatokkal ellentétben, könnyen fölláldozhatja egy részét, hogy szabaduljon a veszélyes lakótársaktól, hiszen ezzel az egész élőlény működőképessége nem sérül. Egy állat azonban nem szabadulhat meg következmények nélkül az egyik lábától vagy mondjuk a fél veséjétől. Ezzel szemben az állatok általában aktív mozgásra képesek, ezért válogathatnak a partnerek között, s így tudják elérni, hogy a csalók ne jussanak akkora nyereséghez, mint a mutualisták [1, 2, 5]. Például a rovarok által beporzott növények a nektártermelés visszafogásával képesek lennének költségeiket csökkenteni, így tudnák becsapni a rovarokat. Azonban a rovarok egy-egy rossz tapasztalat után kerülnek azokat a növényeket, ahol nem kaptak elegendő nektárt. Olyan ez, mint egy piac [1]: „nézelődnek, kóstolgatnak”, s onnan vásárolnak, ahol az adott választékot felmérve ez leginkább megéri. Éppen ezért a virágok a legtöbb esetben erős versenyben vannak a beporzók kegyeiért, ezért ontják is a nektárt bőségesen. Azonban a rovar, ha egyszer bemászott a nektárért, már nem tud csalni, a virágpár a testére tapad, s a beporzást ezek után akarva, akaratlan elvégzi. Ezekben a mutualista kölcsönhatási hálózatokban olyan állatok (pl. poszméhek vagy kolibrók) lehetnek a kölcsönhatás parazitái, melyek alulról megfűrik a virágot, s úgy szívják ki a nektárt.

A mérések alapján azonban ezen fajok negatív hatása a növényekre elenyésző, sőt esetenként pozitív is lehet amiatt, hogy ez a viselkedés serkenti a nektártermelést.

A beporzók tehát mintát vesznek, és így válogatnak a virágok között. De vajon mit tudnak a növények és az állatok válogatni a mikrobiális partnerek között? Hiszen jó megoldás az, ha nem engedi a gazda a csalók elterjedését, de még hatékonyabb lenne, ha idejekorán ki tudná szűrni őket. Ez a módszer az állatok körében különösen fontos, hiszen ahogy korábban hangsúlyoztuk, utólagos szankciókat az állatok nem tudnak olyan egyszerűen alkalmazni, mint a növények. Az előzetes szűrés (**2. táblázat**) szerepét tölti be többek között a gerincesek komplex immunrendszere, mely felismeri és elpusztítja a gazdaszervezetbe jutott idegen sejteket. Az állat bőrfelületén vagy például a vastagbél hámszövetjében (emésztőrendszer is külső felület, csak be van gyűrődve) a veszélytelen vagy még gyakrabban a gazda számára előnyös baktériumközösségek élnek. Ennek oka, hogy egyrészt a gazda olyan körülményeket teremt, melyek kedveznek a mutualista mikrobáknak, másrészt egy fajgazdag, stabil mikrobaközösség nem teszi lehetővé egyetlen törzs számára sem, hogy túlzottan elszaporodjon. Ha valahol mégis a hámszövetet támadó mikrobák jelennének meg nagyobb számmal, akkor az immunrendszer azonnal közbelép, s megakadályozza azok terjedését. Annak ellenére, hogy a növények könnyeb-



1. ábra. A befektetési stratégia sémája. Az a paraméter a partner választától független befektetés mértékét adja meg, a b paraméter (a piros egyenes meredeksége) pedig azt határozza meg, hogy a partnertől kapott nyereség mennyire növeli a befektetést a következő körben

ben meg tudnak szabadulni a káros mikrobáktól, egyfajta előszűrést ők is alkalmaznak. Példaként a pillangósvirágúak és a nitrogénfixáló baktériumok együttműködésének kiépülését mutatjuk be kicsit részletesebben [6]. A pillangósok gyökerén gümők találhatóak. Ezekben a gümőkben a talajból felvett speciális *Rhizobium* baktériumok élnek, melyek a levegő nitrogénjét megkötik, és szerves szénvegyületekért cserébe a növény számára felvehető nitrogén formájában szállítják. A talajban nagyon sokféle mikroba él, ezek közül kell a növénynek a legjobb partnereket kiválasztania. A kapcsolatot megelőzően a gyökérszövet és a *Rhizobium* baktériumok jeleket adnak egymásnak. A növényi jelek (flavonoidok) hatására a baktériumok a jelforrás irányába mozognak, a baktériumok által kibocsátott jelmolekulák (Nod-faktorok) hatására a gyökérszövet meggyöngyösödik, és így a baktériumokat csapdába zárják, mely folytatá-

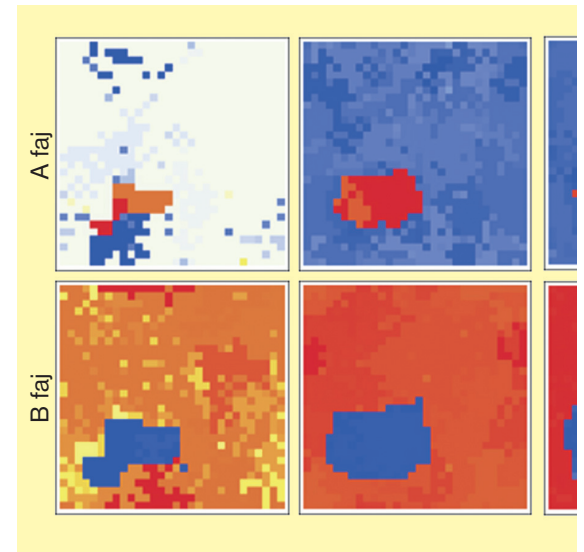
saként aztán bejutnak a növényi szövetekbe. Itt alakul ki a számukra speciálisan kialakított lakótér, a gümő, mely táplálja és menedéket ad a baktériumoknak, de el is zárja a gyökér további részeitől. A részletek ismertetése nélkül csupán megjegyezzük, hogy a növény-mikorrhiza kapcsolat kommunikációja a jelek szerint nagyon hasonlít a növény-rhizobium kommunikációs rendszerre, mely valamilyen evolúciós rokonságra utal.

A kommunikációs kapcsolat nyilván nem zárja ki teljesen a csalókat, hiszen nem elképzelhetetlenek olyan egyedek, amelyek a kommunikációban aktívan részt vesznek, de magában a mutualizmusban már nem. Ennek ellenére a kommunikáció előzetes szűrésként jól működhet, különösen, ha ez a kommunikáció költséges és (vagy) a kommunikációért és a mutualista viselkedésért felelős gének részben megegyeznek vagy kapcsolatosan öröklődnek (a részletes magyarázatért ajánljuk Számadó 1998-ban a *Termetzet Világában* megjelent írását [7]).

Ismerünk tehát számos mechanizmust (2. táblázat), mely a mutualista kapcsolatokat stabilizálja, a csalókat szűri és/vagy bünteti [1, 2, 5, 6]. Láttuk, hogy ha a partnerek tápanyagokat cserélnek, és az egyedek tartós kapcsolatban vannak, akkor a partnertől kapott tápanyag függvényében a kedvezményezett egyedek módosíthatják a befektetésüket. Kérdés, hogy ez a mechanizmus önmagában elegendő lehet-e a mutualista kapcsolat stabil fenntartásához? A következő fejezetben egy olyan modellt mutatunk be, mely erre a kérdésre keresi a választ.

A mutualizmus modellje

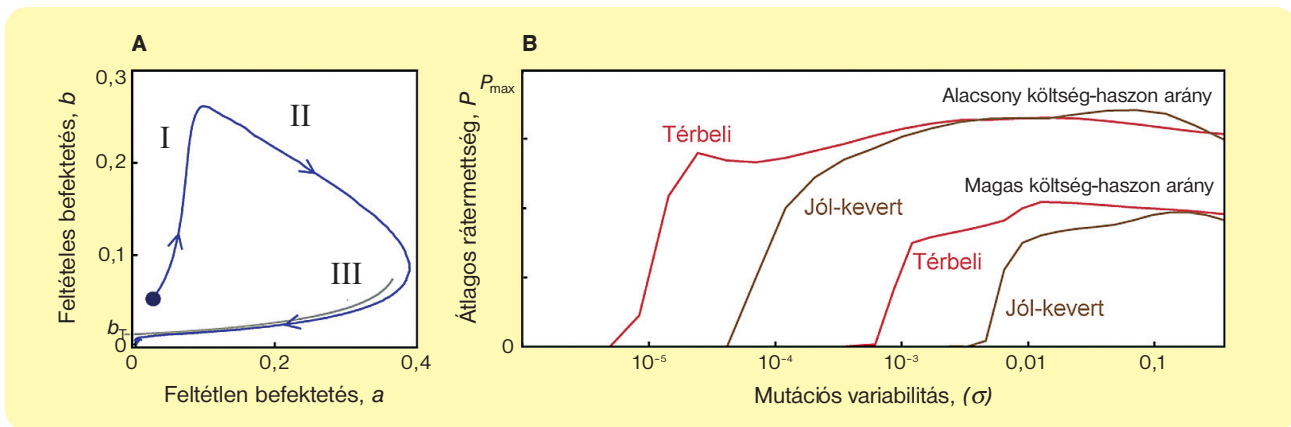
Modellünk két faj kapcsolatát írja le, melyek között hosszú távú (mondjuk tápanyagszállításra alapuló) mutualista kapcsolat



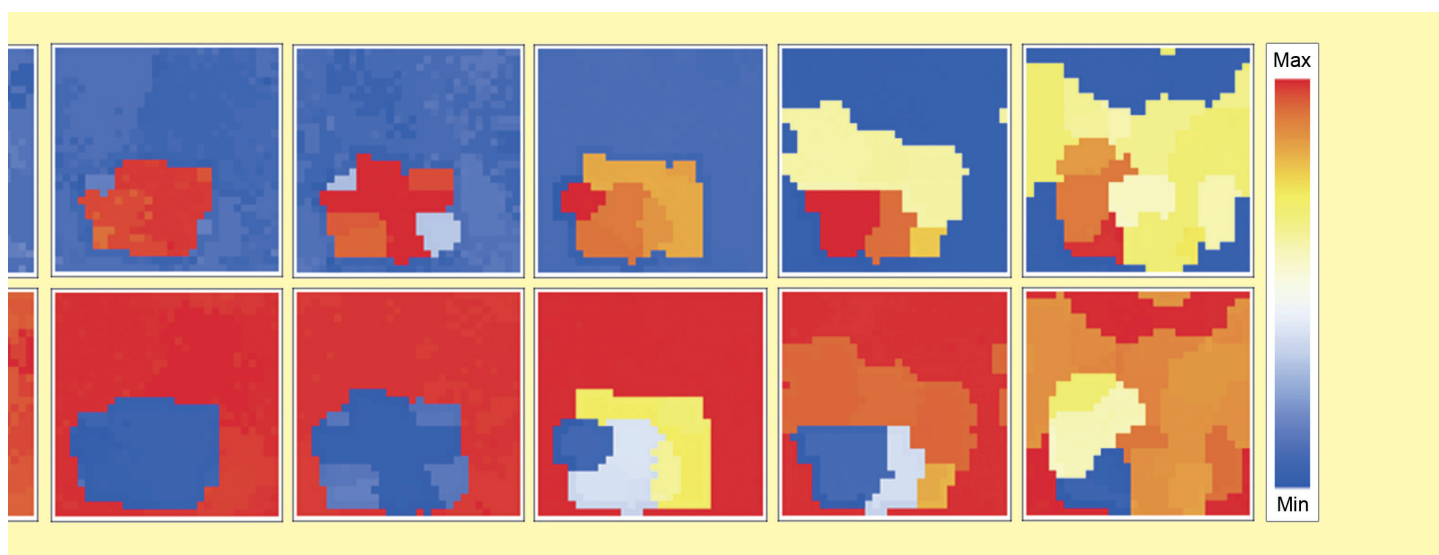
áll fenn [4]. A modell könnyebb vizsgálhatósága kedvéért tegyük fel, hogy a két faj minden tekintetben azonos. Ilyen tulajdonságú mutualista kapcsolatokat természetesen nincsenek, de a jelenségek megértése miatt érdemes a komplex rendszerek egyszerűsített modelljeit tekinteni. Mint látni fogjuk, ezen egyszerűsítések mellett is igen összetett eredményeket kaphatunk, melyek sokat segíthetnek a valós biológiai folyamatok jobb megértésében.

Feltesszük továbbá, hogy a fajnak azonos számú egyede van. A mutualista kapcsolatot a két faj egy-egy egyede között létrejövő, határozott ideig tartó kölcsönhatásként képzeljük el. Ezt úgy jelenítjük meg, hogy ez idő alatt a két egyed egymást követő diszkrét időpontokban (körökben) változtatható mértékű befektetést (azaz tápanyagot vagy szolgáltatást) ajánl fel partnerének. A befektetés természetesen költséggel jár a be-

2. ábra. (A) A befektetési hurok, mely három jól elkülöníthető fázisra osztható (I, II, III). A nyíl irányja az idő múlásának irányát mutatja a görbén, a szürke vonal a b_T küszöb feltételes befektetés mértékét az a feltétlen befektetés függvényében. (B) Az átlagos fitnessz (rátermettség) a mutációs variabilitás függvényében jól kevert, és térben explicit populációk és különböző költség-nyereség arányok esetén. Az ábrán jól látszik, hogy stabil mutualizmus csak úgy jön létre (azaz a rátermettség nagyobb, mint nulla), ha a mutáció keltette sokféleség elegendően nagy



Diverzitás és komplexitás a mutualista kapcsolatokban



3. ábra. Térbeli mozaik dinamika. A képen látható két mutualista populáció két egymással párhuzamosan elhelyezkedő négyzettrács négyzeteit foglalja el. (A faj a felső, míg a B faj az alsó sorban látható). Minden négyzetben egy-egy egyed található, melyeket a rátermettségük alapján színeztünk ki (az elérhető maximális és minimális rátermettség értékek között; Min, Max > 0). Látható, hogy a rátermettség alapján csoportokba sorolható egyedek a térben egymáshoz közel helyezkednek el, egy térben és időben változó mozaikos foltrendszert alkotnak

fektetést felajánló egyed számára, s hasznot hoz a fogadó félnek. Egy egyed nyeresége a kölcsönhatásban a realizált haszon és a költség különbsége lesz. Az azonos típusokhoz vagy az egész populációhoz rendelt rátermettséget (fitness) ezen nyereségek hosszú távú átlagaként számoljuk. A befektetések mennyiségét, a vizsgálatok tanúsága szerint, még a legegyszerűbb egysejtű *Rhizobium* baktériumok esetében is gének tucatjai befolyásolják. Ilyen bonyolultságú rendszer átható matematikai elemzése mai tudásunk szerint reménytelen, ezért a jelenséget a viselkedés szintjén modellezzük és az egyed mutualista befektetését két fő részre osztjuk: feltétlen és feltételes befektetésekre. Míg a feltétlen befektetés mértéke független a mutualista társ viselkedésétől, addig a feltételes befektetés a partner által szolgáltatott nyereség függvénye. Egy jó partner esetén a feltételes befektetési típus növekvő befektetést eredményez. Ezzel szemben, ha az egyed nem kap semmit a (csaló) partnerétől, akkor az egyed feltételes befektetése is nulla lesz. (Ez a felosztás nem összekeverendő az obligát/fakultatív felosztással, aminek a partnerek túlélése az alapja.) Tétélezzük fel, hogy ezen két befektetési típus mennyiségét két független, szabadon evolválódó tulajdonság határozza meg (1. ábra). A modellben tehát e két tulajdonsága definiálja az egyed befektetési stratégiáját, mely stratégia sikerét (túlélési esélyét és/vagy utódainak számát) a partnerével kialakított kölcsönhatás során felhalmozott nyeresége határozza meg. A sikerebb stratégiákat hordozó egyedek nagyobb eséllyel maradnak fenn, több utódot hoznak létre, mely utódok

öröklik a szülő stratégiáját is, ám apróbb, mutáció okozta eltérésekkel. Tétélezzük fel először, hogy ezek a mutációk nagyon ritkán, és nagyon kicsi változásokat okoznak a tulajdonságokban.

Alacsony, de egy adott küszöbérték feletti kezdeti feltételes befektetések (b_7) esetén a szelekció hatására növekedni fog a befektetési hajlam (azaz mind az a mind a b paraméter) (2A. ábra I.). Minél többet fektet be egy egyed, annál inkább növekszik nyeresége, mivel a feltételes befektetés a nyereség mértékében tovább növelheti a partner befektetését, mely aztán nagyobb nyereséggé fordítódik le az egyed számára. Így olyan stratégiák terjedhetnek el a populációban, melyek egyre többet és többet fektetnek be a kapcsolatba. A növekvő befektetésekkel egy időben megváltozik a befektetés minősége is. Kezdetben azok a stratégiák lehetnek sikeresek, amelyekre a feltételes befektetés a jellemzőbb, mivel ezek csak akkor növelik kis kezdeti befektetésüket, ha a partner is viszonozza azt. Mondhatjuk tehát, hogy ezek a stratégiák „óvatosak”. Azonban az óvatos stratégiák elterjedésével egyre nagyobb nyereségre tesznek szert az olyan stratégiák, melyek kevésbé óvatosak, azaz egyre nagyobb feltételes befektetéseket tesznek (2A. ábra II.). Ahogy a feltétlen stratégiák átveszik a feltételes stratégiák helyét, a csalók számára is szabaddá válik a pálya. A magas feltételes befektetések kedveznek a csalók elterjedésének, hisz nagy nyereséggel jár a csaló számára, ha partnere sokat fektet be, míg maga a csaló egyed szinte semmit, s így nem fizeti a befektetés költségét sem. A csalás mértéke itt fokozatos, egyre keveseb-

bet és kevesebbet befektető stratégiák terjednek el, míg végül senki nem fektet be semmit (2A. ábra III.). Bármilyen kezdeti állapotból ebbe a végős, befektetés nélküli állapotba juttatja el a populációt egy ún. befektetési hurkon keresztül. Hosszú távon a magas befektetések jellemezte mutualizmus tehát nem fenntartható a felvázolt modellben, a rendszer a leírt befektetési hurkot bejárva végül egy nem mutualista állapotba (attraktorba) jut (2A. ábra III. vége).

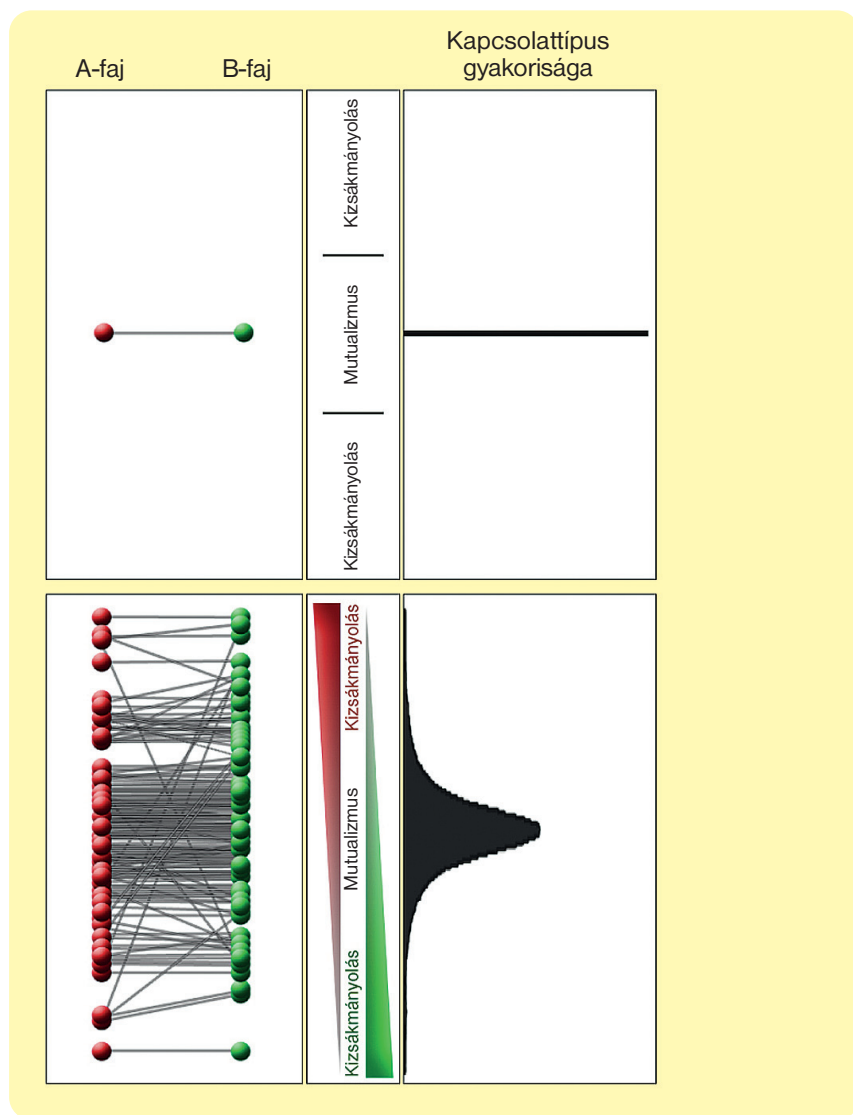
Egyes esetekben azonban megfigyelhető, hogy a befektetések mértéke, és a nyereség mennyisége a populációban végül egy stabil, állandó érték körül ingadozik. Ezt akkor észlelhetjük, ha növeljük a mutációs variabilitást, tehát annak a mértékét, hogy milyen gyakran, és a szülői (a , b) tulajdonságoktól mennyiben különböző mutáns utódokat hozunk létre. Az bizonyos, hogy a növekvő mutációs variabilitás eredményeképpen növekvő stratégia-diverzitás jön létre, azaz egyre több, különböző befektetési stratégia lesz jelen a populációban. De miért vezet ez a mutualista kapcsolat stabil megjelenéséhez?

Alacsony mutációs variabilitás esetén a populáció jellemzően először növekvő, majd csökkenő befektetéseken keresztül evolválódik, ahogy azt a 2A. ábra mutatja. Ezekben az esetekben a populációban egy időpillanatban az összes egyed közel azonos befektetési stratégiát követ. A mutációs ráta növelésével egyre többféle befektetési stratégia-kombináció jelenik meg, és egyre változatosabb stratégiákkal rendelkező mutualista párok kerülhetnek kölcsönhatásba egymással. Ennek eredményeképp a különféle

mutualista párok esetében a szelekció más-más stratégiákat részesít előnyben, így ez a stratégia-diverzitás fennmarad. Végeredményben, ha elegendően nagy mértékű a stratégia-sokféleség, akkor a befektetési hurok kis valószínűséggel evolválódik a nulla mértékű befektetések állapotába, hisz mindig lesznek olyan párok, melyek számára megéri növelni befektetéseiket. Ezek újraindíthatják a hurkot, legalábbis a populáció egy részében. A két folyamat, a befektetési hurok evolúciója, mely végül nulla mértékű befektetésekhez vezetne, és a mutáció, mely a létrehozott stratégia-diverzitás révén a befektetési hurok újraindulását eredményezheti, egyensúlyban, stabil szinten tarthatják a populációra jellemző átlagos befektetést (**2B. ábra**).

A bemutatott modellben a két mutualista populáció egyedei véletlenszerűen kerülnek kapcsolatba, majd a kölcsönhatás végzetével elválnak útjaik, s a következő kölcsönhatásban újabb véletlenszerűen kiválasztott egyedek alakítanak ki kapcsolatot. Úgy működik a modell, mintha a populációkat folyamatosan nagyon intenzíven kevernénk, azaz szakszóval élve, a rendszer jól kevert. A természetben megfigyelhető valós mutualista kapcsolatok esetén azonban ez ritkán van így. A növények és mikorrhizák vagy a baktériumok esetében például mindkét partnerre nagyon korlátozott helyváltoztató mozgás jellemző, vagyis ugyan azonos területet foglalják el hosszú időn keresztül. Az egyedek véletlenszerű mozgása tehát nem írja le megfelelően ezeket a kapcsolatrendszeret. A korlátozott mozgású populációk kölcsönhatását is modellezhetjük. Ilyenkor egy képzeletbeli rácsháló üres négyzeteit tekintjük elemi élőhelynek, és az egyedek ezekben a négyzetekben foglalhatnak helyet, egyesével. Így egy H hosszúságú és L szélességű rácsháló $N=H*L$ számú élőhelyet jelent, ahol maximum N számú egyed élhet. Minden egyednek n számú szomszédja van, mely jellemzően egy rácspont közvetlenül szomszédos 8 négyzetét jelenti.

Tapasztalataink szerint térben kötött populációk esetében jóval kisebb mutációs variabilitás elegendőnek bizonyul a mutualizmus fenntartásához (**2B. ábra**). A stratégiapárok ebben az esetben térben kapcsolatosan fognak elhelyezkedni, egy dinamikus változó térbeli mozaik struktúra alakul ki (**3. ábra**). A térbeli mozaik egy foltjában többé-kevésbé ugyanazt a befektetési stratégiát követik az egyedek egy populáción belül. Mivel az egyedek mozgása, elterjedése korlátozott, a stratégiák jól elkülönülten, tartósan fennmaradhatnak a négyzetrács különböző helyein. Ezenkívül, egyes esetekben létrejöhét egy olyan „szigetelő határteleg” az egyes foltok között, mely hosszú időn keresztül képes az egyébként sikeres stratégiák terjedését, invázióját gátolni a szomszédos foltba. A **3. ábrán** jól kivehe-



4. ábra. A mutualista kapcsolat komplexitása. Leegyszerűsített elgondolásban a mutualizmus a két résztvevő faj minden egyedének egyformán nyereséges kapcsolata (felső panelek). Valójában a mutualista kapcsolat a kölcsönösen előnyös, egyoldalúan előnyös (kizsákmányolás), esetenként kölcsönösen hátrányos kapcsolatok valamilyen együttese (alsó panelek). (A bal oldali paneleken a színes pontok a populációk egyedeit jelölik a nyereségük alapján függőlegesen elhelyezve, a szürke vonalak a közöttük lévő kapcsolatokat. A középső panel a nyereség változását illusztrálja az A- [piros] és a B- [zöld] faj egyedei esetén). Bár a populációkban leggyakoribb kölcsönhatástípusok a mutualisták (panel közepén lévő kapcsolatokat), de találunk aszimmetrikusan nyereséges (bal felső oldalt a jobb alsóval összekötő vonal), illetve az egyik vagy a másik partner javára egyoldalúan előnyös kölcsönhatásokat is (panelek felső és alsó harmadában lévő közel vízszintes vonalak) vagy esetenként kölcsönösen hátrányos kapcsolatokat is (bal alsó sarkot a jobb felsővel összekötő vonal). A jobb oldali panelek az A- és B-faj közötti egyes kapcsolattípusok gyakoriságát mutatja: míg felül csupán egyetlen mutualista típus van, alul a különböző kapcsolattípusok eloszlása látható

tő az is, hogy a hosszú ideig fennmaradó stabil foltpárokból gyakran nem igazán mutualista a kapcsolat. Míg az egyik fél magas, addig a másik alacsony rátermettség-értékhez jut (lásd pl. 2–4 paneleket, piros és kék foltok együtt járó elhelyezkedése a két populáció megfelelő térrészeiben). Hosszabb időre kiátlagolva azonban általában

mind a két fél jól jár a kapcsolatból, magas átlagos rátermettségre tesznek szert.

A szigetelő határtelegnek és a térben elkülönült mozaikos foltrendszernek köszönhetően a stratégia-sokféleség akkor is számottevő marad, ha az új változatokat előállító mutációs variabilitás jóval kisebb, mint a jól kevert esetben. Ennek oka egyszerű:

Diverzitás és komplexitás a mutualista kapcsolatokban

ha megjelenik egy sikeres csalo stratégia, nem fog tudni az egész populációban elterjedni, csupán abban a foltban, ahol megjelent, hisz a szigetelő határreteg meggátolja továbbterjedését. Így azok a kevésbé rátermett stratégiák, melyek egy jól kevert populációban hamar eltűnnek, az egymástól elválasztott foltokban hosszú időn át fennmaradhatnak.

A mutualisztikus kapcsolatok megértésének fontossága

Az előző fejezetben hangsúlyoztuk, hogy egyszerűsített modellünk a valós, komplex biológiai rendszer jellemzőinek csak egy szeletét volt hivatott megfogni. Így egy kezelhető modellhez jutottunk, viszont abban nem lehetünk biztosak, hogy a valós biológiai folyamatokról, azok komplexitásáról is megbízható információkat gyűjtöttünk ily módon (a probléma mélyebb megértéséhez ajánljuk Kondor Imre különszámunkban található „Komplexitás kihívásai” című esszéjét). Azt mondhatjuk tehát, hogy a modell alapján valószínűsíthető, hogy egyes mutualista kapcsolatokban (növény-mikroba, mikroba-mikroba) a mutációk fenntartotta viselkedésmód sokféleség minden más mechanizmus nélkül is fenn tudja tartani a mutualizmust. Ennek az esélye lényegesen nagyobb, ha az egyedek mozgása korlátozott. Megmutattuk, hogy az evolúciós egyensúlyban is viselkedésmód sokféleség van, a kapcsolat csak átlagban kölcsönösen előnyös, a kizsákmányolás (csalás) mindig jelen van, de nem válik domináns tényezővé a mutualista kapcsolatokkal szemben (**4. ábra**) [4].

Fontos tanulság továbbá, hogy a mutualizmusra ne úgy tekintünk, mint egy változatlan, meghatározott nyereséggel járó, kölcsönösen előnyös kapcsolatra két (vagy akár több) faj egyedei között (**4. ábra**, felső panelek), hanem sokkal inkább mint egy komplex kölcsönhatásrendszer működésének eredményeképpen kialakuló viszonyra, melyben a kizsákmányolás és a mutualizmus egyszerre van jelen (**4. ábra**, alsó panelek). Ha egy adott egyed egy jó mutualista partnerrel kerül kapcsolatba, akkor jellemzően magas befektetéseket tapasztalhatunk, mindkét fél nagy nyereségre tesz szert a mutualista kapcsolat eredményeképp. Ha viszont ugyanez az egyed egy másik, a nagy befektetést nem viszonozó, csalo típusú egyeddel kerül össze, akkor a csalo kizsákmányolja [4, 6, 8].

De nem csak a partner milyensége befolyásolhatja a kölcsönhatás kimenetelét. A környezet minősége, az elérhető tápanyagok, az elérhető víz mennyisége például igen erősen megváltoztathatja a mutualizmus jellegét [8]. Ismert jelenség például, hogy nitrogénben gazdag talajban, ahol a növény

maga is képes a számára szükséges nitrogént felvenni, a nitrogénfixáló baktériumokkal kialakított kapcsolatok ritkábbak, és a növény sokkal kevesebbet fektet be a baktériumok és az őket tartalmazó gyökérgümők fenntartásába [6]. Hisz mi értelme lenne nagy árat fizetni egy munkáért (a nitrogén fixálásért), ha az szinte ingyen elérhető a környezetből is. Igen hasonló változások figyelhetők meg a növény-mikorrhiza kapcsolatok minőségében is, ahol például a növény számára elérhető víz, vagy épp a foszfor mennyisége lehet a befolyásoló tényező [8]. Nem nehéz belátni ezek után, hogy egyes emberi hatások, mint például a műtrágyák alkalmazása, nemcsak a növények növekedésének és terméshozamának a mértékét, hanem a mutualista kapcsolatrendszerek minőségét is drasztikusan megváltoztathatják.

Hasonlóan komoly változásokat okozhat a föld alatti mikrobiális közösségekben a kiterjedt monokultúrák gazdálkodás, vagy épp a behurcolt, tájidegen fajok elterjedése is. Egyes fajok az egyik, míg mások más típusú mikrobiális partnerekkel alakíthatnak ki sikeresen kapcsolatot. Egy természetes közösségben kialakult egyensúlyt boríthat fel például egy behurcolt faj, nemcsak a bennszülött fajok kiszorításával, hanem a mikrobiális közösségben kialakult erőviszonyok megváltoztatásával. Mivel a növények nagy része több mutualista kapcsolatot is kialakít a talajban található mikrobiális közösségek tagjaival, így ezen közösség megváltoztatása, szétrombolása szinte minden, az adott helyen élő növényre hatással lehet.

Ha kicsit mélyebbre hatolunk a fajok közötti kölcsönhatások megismerésében, kiderül, hogy az emberi beavatkozásnak jóval komolyabb hatása lehet a természetes élőközösségekre, mint korábban gondoltuk. E kapcsolatrendszereket fenntartó mechanizmusok ismerete lehetővé ad arra, hogy megjósoljuk az egyes emberi cselekvések hatásait [6, 8]. Jó példa erre a globális klímaváltozás, és ami talán még fontosabb, a CO₂ változása a légkörben. Itt érdemes megjegyezni, hogy úgy 400 millió évvel ezelőtt az akkori légköri CO₂-szint a földi időskálán igen rövid idő alatt tízedére csökkent. Ez a hirtelen változás egybeesik a szárazföldi edényes növények megjelenésével. Az adatok alapján a CO₂ ilyen drasztikus csökkenését nagy valószínűséggel az edényes növényekhez kapcsolódó mikorrhizák okozták. A mikorrhiza gombák a talajban a mállást jelentősen megnövelik, és a szén-dioxidot különböző karbonát vegyületekké alakítják. Ezek a karbonátok az édesvizbe, majd az óceánokba jutva azok alján leülepednek, így folyamatosan kivonják a szén-dioxidot a légkörből.

Nem vitás, hogy a növekvő légköri CO₂ – többek között – befolyásolja a fotoszintézis hatékonyságát és a talajban található

különböző kötések formában fellelhető szén mennyiségét. Ez a változás viszont befolyásolja a növény-mikorrhiza kapcsolatokat is, melyek, mint láttuk, hatással vannak a globális CO₂-egyensúlyra. A hatás iránya azonban a kísérletek alapján nem egyértelmű. Míg egyes esetekben növekedett a föld feletti és a föld alatti biomassa (az élő anyag mennyisége), azaz a mutualizmus hatékonyabbá vált, más esetekben olyan átrendeződések voltak megfigyelhetők a természetes közösségekben, melyek a biodiverzitás csökkenéséhez, sok mutualisztikus kapcsolat visszaszorulásához vezettek. Fel kell készülnünk tehát a klímaváltozás okozta további változásokra, bár ezen változások nagyságáról és sebességéről vajmi keveset tudunk és értünk jelenleg [8].

Az elmondottak alapján kijelenthetjük, hogy nemcsak az evolúcióbiológusok belügye a mutualista kapcsolatokat fenntartó mechanizmusok megértése, hanem ezek az ismeretek közvetlenül hozzájárulhatnak környezetünk, élőhelyünk, a bioszféra egyensúlyának fenntartásának biztosításához is. Felelősségünk, hogy a kísérletes és elméleti kutatásokban eredményeink egyesített felhasználásával dolgozzunk a biológiai sokféleség megőrzésén, mely, mint cikkünk igyekezett megvilágítani, nemcsak a fajok számával vagy eloszlásával mérhető, hanem sokkal inkább a kapcsolatrendszerek sokféleségén, komplexitásán keresztül. Láthattuk, a feladat nem csekély, a jelenségek sokfélesége és a vizsgált rendszerek komplexitása is komoly kihívást jelent mind az elméleti mind a kísérletes kutatók számára. ❁

A cikkben ismertetett saját kutatásokat az OTKA K100296-os és K100299-es számú pályázata támogatta.

IRODALOM

- [1] Hammerstein, P. (ed.) 2003. *Genetic and Cultural Evolution of Cooperation: Report of the 90th Dahlem Workshop, Berlin, June 23–28, 2002*. Cambridge: MIT Press.
- [2] Sachs J. L., Mueller U. G., Wilcox T. P., Bull J. J. 2004. *The evolution of cooperation*. *The Quarterly Review of Biology* 79: 135–160.
- [3] Scheuring I. 2013. *Az ökoszisztéma védelmében: kulcsszereplők és segítőitársak*. *Természet Világa* 144 (2): 60–63.
- [4] Boza G., Kun Á., Scheuring I., and Dieckmann U. 2012. *Stabilizing mutualism through investment cycles, phase diffusion, and spatial bubbles*. *PLOS Computational Biology* 8 (11) e1002660.
- [5] Bull J. J., Rice W. R. 1991. *Distinguishing mechanisms for the evolution of cooperation*. *Journal of Theoretical Biology* 149: 63–74.
- [6] Kiers E. T. & Denison R. F. 2008. *Sanctions, Cooperation, and the Stability of Plant-Rhizosphere Mutualisms*. *Annual Review of Ecology, Evolution, and Systematics* 39: 215–236.
- [7] Számadó, Sz. 1998. *Az ösztimuláció ára van? Kommunikáció az állatvilágban*. *Természet Világa* 129 (10).
- [8] Johnson N. C., Angelard C., Sanders I. R., and Kiers E. T. 2013. *Predicting community and ecosystem outcomes of mycorrhizal responses to global change*. *Ecology Letters* 16: 140–153.

SZOLNOKI ATTILA

Komplex viselkedés társadalmi dilemmákban

A természettudományos módszerek térhódítása a hagyományosan a társadalomtudományok fennhatósága alá tartozó területekre látványosan felgyorsult az elmúlt évtizedben. Ennek egyik lényeges oka az, hogy a társadalom működését leíró adatok mennyisége drasztikusan növekszik. A *Google*-vezér, *Eric Schmidt* szerint manapság kétnaponta „termelünk” anynyi adatot, mint amennyi az emberiség hajnala óta 2003-ig keletkezett. Ugyanakkor az adatok elemzését a számítógépek hatékonyan segítik, és reális lehetőségünk van arra, hogy az alkalmas matematikai modellekre alapozva mélyebb összefüggéseket tárjunk fel a bonyolultságuknál és összetettségükönél fogva eredendően komplex társadalmi rendszerekben is.

A mindnyájunkat érintő társadalmi dilemmák egyike az együttműködés mechanizmusainak megértése, ami a legkülönbözőbb skálán szövi át mindennapjainkat [1, 2]. Az együttműködés fontosságát különösen annak hiányában érzékeljük. Példaként említhetjük a gazdasági és pénzügyi válságok jelenlegi sorozatát, de az erkölcs, a munkamorál és a szakértelem alacsony színvonalra is ide sorolható. A feladat súlyát jelzi, hogy a tekintélyes *Science* magazin néhány éve az emberiség legfontosabb 25 problémája közé sorolta annak megértését, hogy miként alakulhat ki az együttműködő magatartás olyan helyzetekben is, amikor a szereplők az egyéni érdekek mentén cselekednek [3].

A mai modern felfogás szerint a játékelmélet valóságos élethelyzeteket egyszerűsít le úgy, hogy a résztvevők számát néhány játékosra korlátozzuk, akik mindannyian csupán néhány lehetőség közül választhatnak. Ha az egyszerűség kedvéért az emberi magatartást két szélsőséges esetre korlátozzuk, akkor az ún. „együttműködő” játékos bizonyos áldozatot hoz a társai érdekében, míg az „önző” kihasználja az együttműködők hozzájárulását, és így azokkal szemben előnyt élvez. A dilemma helyzetet éppen az jelenti, hogy ha mindenki a vonzóbb élethelyzetre választja, akkor a közösség a lehető legrosszabbul teljesítő állapotba kerül, amit a játékelméletben a közösség tragédiájának hívnak. Az intelligensnek feltételezett játékosok ismerik saját és társaik lehe-

tőségét, és mindegyik lehetőségnél tudják, hogy az mekkora egyéni haszonnal jár a játékosok számára. Ennek ismeretében kell kiválasztaniuk az egyéni érdekeik szerinti legelőnyösebb döntést, miközben természetesen azt is figyelembe veszik, hogy társuk ismeri az „én tudom, hogy Ő tudja, ...” hatásokat is.

A játékelmélet kidolgozásában elévülhetetlen érdemeket szerzett *Neumann János*, akinek *Oscar Morgenstern*nel együtt írt könyve fogalmazta meg elsőként a játékelmélet kereteit, a játékok széles skáláját és alkalmazási lehetőségeit. Ennek eredményeképpen a játékelmélet ma már megalapozott matematikai háttérként szolgál a közgazdaságtan, illetve akár a katonai és a politikai döntéshozók számára is [4]. Ugyanakkor napjainkban a játékelmélet már messze túlnötte a hagyományos kereteit, összefonódott sok más tudományterülettel, és ezzel együtt a céljai is változtak. A terület fejlődésében fontos hajtóerőt jelent a társadalmi dilemmák gyökereinek a mélyebb megértése és az igény arra, hogy megoldásokat találjunk a nemkívánatos következmények elkerülésére. A biológia különösen inspiráló szerepet töltött be ebben a folyamatban, lényegesen kibővítve a vizsgálható rendszerek körét [5,6]. A darwini szelekció szellemében feltételezhetjük, hogy a sikeresebb döntés (vagy magatartás) fog elterjedni a sikertelenebb döntést követők kárára, ahogyan ez történik a biológiai természetes szelekciónál, amikor a játékelmélet nyereményfogalmát a fajok utódlétrehozó képességének leírására használjuk. Az 1970-es évek közepén így született meg az evolúciós játékelmélet, amit a hagyományos játékelmélet és a darwini alapeszmék társulásának tekinthetünk.

Az utóbbi gondolat szellemében az eredeti kérdésünket még élesebben fogalmazhatjuk meg: ha az élethelyzet és a sikeresebb, nagyobb hasznot hozó stratégia, akkor annak elterjedésével, illetve annak az utánzásával a darwini elv alapján az együttműködésnek ki kellene pusztulnia, ami éles ellentétben áll a valósággal! Elsőre talán meglepő lehet, de az emberiség a leginkább együttműködő faj. Gondoljon csak bele a kedves olvasó, aki esetleg az olvasás mellett teáját kortyolgatja, hogy hány ember

együttműködésére volt szükség ahhoz, hogy az olvasás és teázás együttes élvezetét megtehesse. Éppen az emberek együttműködésre való fokozott hajlama az, ami segítette a hatékony kommunikáció és munkamegosztás kialakulását, amelynek révén kiemelkedtünk a többi faj közül.

Visszatérve a kiindulási ponthoz, célunk célja annak tudatosítása, hogy a fizika, ami a tradicionálisan az élettelen természet tudománya, mennyire hatékonyan tud hozzájárulni az említett kérdéskör vizsgálatához. A fizika alapvető jellemzője a matematika nyelvén megfogalmazott összefüggések feltárása és a jósátság összehasonlítása a kísérleti eredményekkel, majd ezen ismeretek segítségével mindannyiunk számára hasznos technológiák kidolgozása. Ennek köszönhető, hogy ma szinte mindenki számára elérhető a mobiltelefon és a számítógép, utazhatunk repülőgéppel vagy konyhánkban jó hatásfokú mikrohullámú sütő és hűtőgép működik.

Evolúciós játékelméleti modellek

Az említett eszközök kifejlesztése éppen annak volt köszönhető, hogy megértettük az anyag belső szerkezetét, és feltártuk a mikroszkopikus alkotóelemek közötti kölcsönhatás jellegét. Ennek ismeretében már képesek vagyunk befolyásolni, illetve intelligensen tervezni a céljainkat szolgáló tulajdonságú anyagot.

A társadalmi és biológiai rendszerek „kölcönható részecskéi” az egyének, illetve az egyedek. A játékelmélet kifejezéseit használva a továbbiakban ők a játékosok, akik különböző stratégiát (döntést, magatartást) képviselve meghatározzák egymás nyereményét (haszonfüggvényét vagy utódlétrehozó képességét, ami a biológiai rendszerben a szomszédságban lévő fajoktól is függ). A rendszer fejlődését, vagyis a stratégiák térbeli és időbeli változását egy matematikai alakban megfogalmazott mozgástörvény írja le. A leggyakrabban használt mozgástörvény szerint a nagyobb nyereményt elérő, azaz a sikeresebb játékos nagy valószínűséggel terjesztheti környezetében a saját stratégiáját úgy, hogy a szomszédok átveszik azt. Ugyanakkor a való-

sághoz közelítő leírás érdekében kis valószínűséggel megengedjük az egyébként sikertelenebb stratégia átadását is. Ezen a módon figyelembe vehetjük a nyeremény becslésében fellépő esetleges bizonytalanságot, illetve a játékos téves döntési lehetőségét, de ez a hatás írja le egy hosszú távú haszon érdekében felvállalt kockázatvállalási hajlamot is.

A különböző konfliktusok leírására számos játék szolgál. A társadalmi dilemma játékok képviselik az olyan kölcsönhatást, amikor közösség számára legelőnyösebb együttműködés fenntartása nem esik egybe az egyéni érdekekkel. Ilyen élethelyzetet ír le például a fogolydilemma játék vagy az azzal ekvivalens adományozó játék. Mindkét esetben két ekvivalens játékosunk van, akiknek egyszerre kell választaniuk két lehetséges stratégia közül. Az együttes nyereményük akkor a legnagyobb, ha mindketten az együttműködést választják. Ugyanakkor ettől a stratégiaválasztástól megéri egyoldalúan eltérni, mert akkor magasabb egyéni nyereményt érhetünk el a játékos-társunk kárára. Ha viszont mindketten az élösködést választják, akkor mindketten a második legalacsonyabb nyereményt kapják. A játék érdekessége az, hogy a játékos-társ választásától függetlenül az élösködés biztosítja a magasabb egyéni nyereményt. Az adományozó játéknál a játékosok arról döntenek, hogy befizetnek-e egy c összeget, amelynek révén a társ egy $b > c$ összeget kap. Az adományozó játék többszereplős változata az ún. közlegelő játék. (A játék elnevezése arra az élethelyzetre utal, amikor egy falu gazdái közösen használnak egy véges méretű legelőt. Ha túl sok tehenet hajtának ki a legelőre, akkor az egy tehenre jutó tejhozam jelentősen csökken. Egy gazda viszont úgy is gondolkodhat, hogy ez nem akkora baj, ha cserébe neki több tehenre jut, így az ő összes tejtermelése nő. Ennek az önző magatartásnak az elharapózása esetén előállhat az a helyzet, amikor már mindenki rosszabbul jár. Hasonló problémával szembesülhetünk az adófizetés vagy a társadalombiztosítás esetén is.) Az említett szituációkat megragadó absztrakt játékban a játékosok felösszegzett egyéni befizetéseit megszorozzuk egy $r > 1$ faktoral, majd azt egyenlően szétosztjuk a csoport tagjai között, függetlenül az egyéni hozzájárulásuk mértékétől. Az r paraméter kifejezheti az együttműködésből származó előnyöket, vagy akár a specializációból (munkamegosztásból) fakadó hatékonyabb termelést is. Az egyéni befizetés különösen nagy csoportméretnél csökkenti jelentősen a tiszta egyéni nyereséget, ezért az önző egyéni érdek könnyen társadalmi tragédiába kényszeríti a játékosokat.

Mindennapi életünkben naponta többször játszunk adományozó vagy közlegelő játékot anélkül, hogy ennek tudatában lennénk. Ha például vállaljuk az egyéni költ-

ségét az egészséges életmódnak vagy a környezetkímélő magatartásnak, akkor annak többszörös hasznán osztozik a társadalom azáltal, hogy mindenki számára csökken az egészségügyi ellátás, illetve a környezet tisztításának a költsége. De ide sorolható az a döntésünk is, hogy a ránk háruló feladatot tisztességesen (vagyis a köz javára) vagy önző módon (az egyéni fáradtságot minimalizálva, nem kellő alapossgággal) végezzük-e el.

Ha játékosaink egy olyan globális környezetben vannak, amikor a játékos-társ kiválasztását véletlen folyamatnak tekinthetjük, akkor a stratégiaváltást leíró evolúciós folyamat a közösség tragédiáját fogja eredményezni. A valós helyzetekben azonban a játékosok ténylegesen csak a közösség egy kis hányadával vannak kapcsolatban, és ezek a kapcsolatok többé-kevésbé tartósak. Már régen tudjuk azt, hogy a tartós kapcsolatok (állandó szomszédok) révén az ismételt kölcsönhatás lehetővé teszi az együttműködők részére azt, hogy egymást segítve versenyképes nyereményt érjenek el az élösködőkkel szemben [7].

Az anyagtudományi kutatások alapvető kérdése az, hogy a szerkezet milyen módon befolyásolja az anyag mérhető makroszkopikus tulajdonságait. Ugyanez a kérdés a társadalmi dilemmáknál úgy szól, hogy a játékosok kölcsönhatását meghatározó szomszédosági gráf topológiája miként befolyásolja az együttműködés szintjét, vagyis az együttműködő stratégia gyakoriságát. Kiderült, hogy az együttműködési szint jelentősen javul, ha a játékosok kapcsolatait egy erősen inhomogén kapcsolati eloszlással modellezhetjük [8]. A matematikai modellek vizsgálata feltárta azt is, hogy a kapcsolatrendszer inhomogenitása mellett számos egyéb egyéni tulajdonság (pl. a stratégiaátadó képesség, társadalmi tekintély különbözősége stb.) heterogenitása is segítheti az együttműködés kialakulását. Az említett hatásoknak a hátterében az áll, hogy a valamilyen értelemben befolyásos, vagy vezető játékos a sikeressége révén mintául szolgál a környezetének, akik a stratégiáját másolni fogják. A befolyásos játékos környezete átveszi a vezető játékos magatartását, ami hátrány az élösködők számára, egyúttal óriási előnyt biztosít az együttműkö-

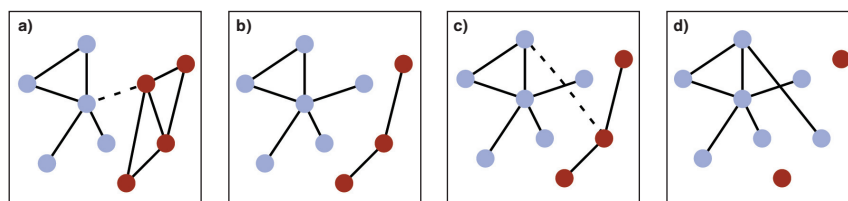
déshez. A folyamat eredményeképpen az egész populáció számára az együttműködés válik vonzó példává. Belátható az is, hogy az együttműködést segítő „egyenlőtlenség” spontán módon is ki tud alakulni, amennyiben nem csupán a stratégiaeloszlás változhat, hanem pl. a sikeres stratégiaátadás egyben a mintául szolgáló játékosnak egy olyan paraméterét (pl. tekintély) is változtatja, amelyik hatással lehet a későbbi utánpótlásra. Ilyenkor nemcsak a stratégiák, hanem más egyéni jellemzők is fejlődnek. Az egyenlőtlenség megléte, illetve bizonyos fokú hierarchia fennállása a szereplők között természetes dolog lehet, ami nem csupán az emberi társadalmakban figyelhető meg, hanem ma már jól dokumentált példákat ismerünk az állatvilágból is [9].

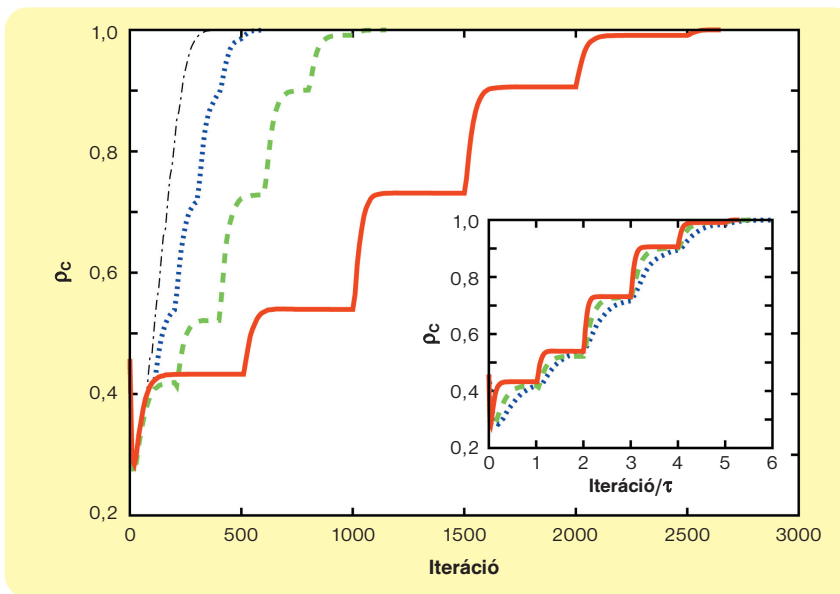
A kizárólagosság kockázata

Az említett „együttes fejlődés” egyben illusztráció arra a kutatási irányra is, amely a tartósan fennálló körülmények következményeinek a tisztázásán túl olyan mikroszkopikus mechanizmusok feltárását célozza meg, amelyek révén az együttműködő stratégia a látszólagos hátránya ellenére is legyőzheti az élösködést. (Talán érdemes megemlíteni, hogy egy jelenséget megérteni kívánó modell több paraméterének az együttes változásából fakadó komplexitás sokkal általánosabb jelenség, amelyre egy érdekes illusztrációt mutat Jánosi Imre cikke ugyanabban a számban [10].) Az általunk vizsgált esetben természetesnek látszik pl. azt feltételezni, hogy a kapcsolatok sem állandóak, hanem folyamatosan megszünek és keletkeznek. Ennek szellemében a játékosok, akik kölcsönhatását a fogolydilemma játék írja le, τ lépésként új kapcsolatokat létesíthetnek és a stratégiák összehasonlítása során a „megtérített” játékos az új stratégia átvételével együtt elveszti a korábbi kapcsolatait is. Az alapfolyamatot az **1. ábra** illusztrálja.

Mindez egy biológiai példában azt jelenti, hogy a vesztes (kevésbé sikeres) stratégiát képviselő játékos helyére kerülő „új” játékos kezdetben csak a mintául szolgáló „szülő” egyeddel lesz kapcsolatban. Ugyanakkor egy szociológiai helyzetben a fenti

1. ábra. A dinamikai szabály illusztrációja abban a modellben, ahol a stratégiaváltás a kapcsolati rendszer változását is maga után vonja. A szaggatott vonallal jelölt, újonnan létrehozott él mentén a sikeresebb stratégiát átvevő játékos a stratégiaváltással együtt elveszti más, korábbi kapcsolatait is





2. ábra. Véletlen kezdőállapotból indítva az együttműködő stratégia p_c sűrűségének változása az iterációk számának függvényében. A vonalak különböző gyakoriságú új kapcsolati frissítést feltételeznek: $\tau = 30$ (fekete), 100 (kék), 200 (zöld), 500 (piros vonal). Elegetően ritka kapcsolati építés felett ($\tau > 50$) a görbék egy lépcsőfüggvényhez konvergálnak (inzert ábra), ami az egyes csoportok időszakos homogenizálására utal

szabály megragadhatja azt a szituációt is, amikor egy főnök fokozatosan a hozzá lojális, az ő értékrendjét képviselő új beosztottakkal cseréli fel korábbi munkatársait. Bár ez első látásra nem nyilvánvaló, de a fenti szabályok mentén fejlődő rendszer hosszú távon az együttműködés térhódítását fogja eredményezni. Ezt mutatja a 2. ábra, ahol különböző τ értéket feltételezve ábrázoltuk az együttműködő stratégia gyakoriságának az időbeli fejlődését akkor, ha kezdetben véletlenszerű stratégia és kapcsolati eloszlás jellemzi a rendszert. Az említett szociológiai példánál maradván könnyen megérthető, hogy ez a dinamikai szabály az esetleges kezdeti helyi sikerek ellenére az önzést képviselő csoportok hosszú távú megszűnését eredményezi, hiszen a kisebb csoportok dinamikus uniformizálódása felszínre hozza az együttműködő csoportok jobb teljesítményét. (A lépcsőfüggvényre hasonló időbeli fejlődés az időszakosan elszigetelt csoportok homogenizálására utal.)

Stratégiák együttéléseként létrejövő újabb „stratégiák”

Természetesen más módon is elérhető az együttműködés fenntartása. Például úgy, hogy az önző stratégia nyeresiményét csökkentjük, így az kevésbé lesz vonzó magartatás a többiek számára. Ezt tekinthetjük büntetésnek, ami a játékelméleti modellben egy új stratégia bevezetését jelenti, és az számos további kérdést vet fel. A büntetés-

nek mindenképpen van valamilyen költsége, pl. fenntartani a szükséges intézményt, vagy egyéni büntetés esetén a társak folyamatos ellenőrzése stb. Ha csak passzívan együttműködünk, tehát a büntetést elkerüljük, de annak költségéhez nem járunk hozzá, akkor ez a stratégia a büntetés szempontjából ún. „másodrendű élősködésnek” számít, hiszen az ilyen játékos saját hozzájárulás nélkül kíván hasznot húzni a büntetés előnyeiből. Az alapkérdésünket átfogalmazva megvizsgálhatjuk azt, hogy evolúciósan fennmaradhat-e büntetés, illetve milyen feltételek szükségesek annak a „racionális” alapokon történő fenntartásához.

A közlegelő játék keretein belül végzett vizsgálatok arra világítottak rá, hogy a játékosok közötti térbeli struktúrának itt is meghatározó szerepe van. Ennek hiányában a büntetés nem életképes megoldás, ha csak nem adunk további esélyt azzal, hogy büntethető lesz a büntetést nem vállaló játékos is, vagy azzal, hogy játékosaink „kiszállhatnak” a játékból. Ugyanakkor, a korlátos kapcsolati rendszer „automatikusan” megoldhatja a másodrendű élősködés problémáját, mert a büntető játékosoknak lehetőségük van elkülönülni a büntetés költségét nem vállaló passzív együttműködőktől. Ezt követően az említett stratégiák már függetlenül küzdenek, és az élősködővel szemben hatékonyabb büntető stratégia kiszorítja a passzív stratégiát. Így a magartatásformák térbeli szétválása folytán fennmaradó élősködők aránya a társadalom számára elfogadható szinten marad.

A kapcsolati struktúra figyelembevétele

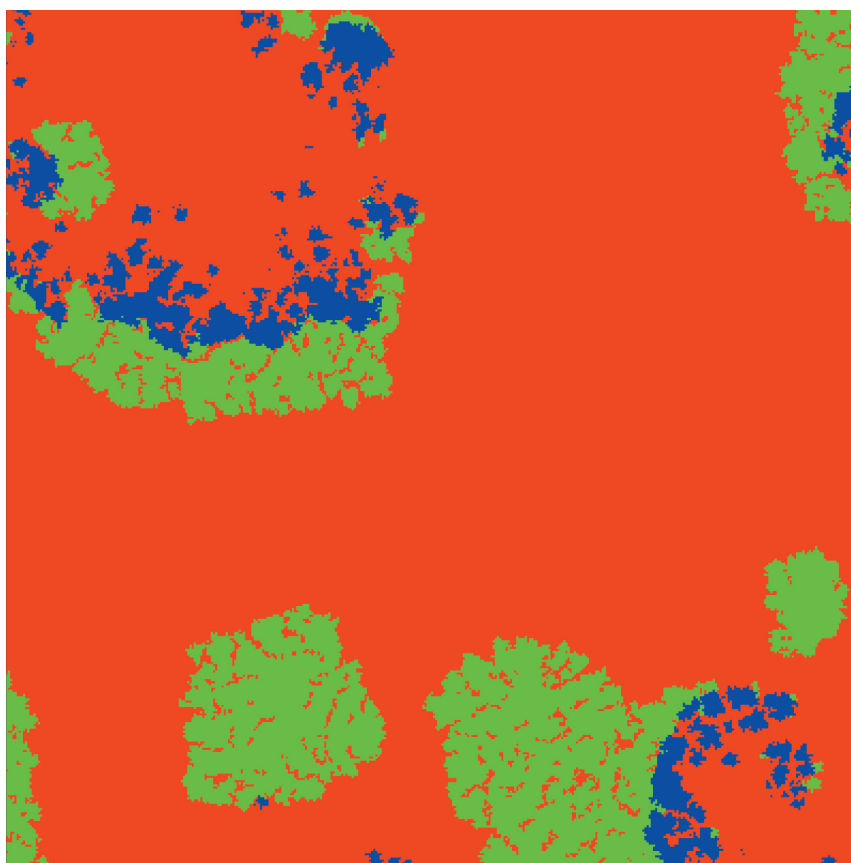
további olyan megoldások felbukkanását teszi lehetővé, amilyenre nem találunk példát egy olyan „globális” rendszerben, ahol a párcapcsolatok létrejöttét a véletlenszerűség uralja. Például az intézményes büntetésnél a büntető stratégiának van egy állandó költsége, ami független az élősködők számától (ez lehet pl. az igazságszolgáltatásnak az adókon keresztül történő fenntartása, vagy egyfajta biztosítási forma használata). A büntetés költségének és nagyságának a függvényében akár ciklikus dominancia (vagy közismertebb szóhasználattal körbeverés) is kialakulhat a büntető, a passzív együttműködő és az élősködő stratégiák között. Ezt a viszonyt a játékelméletben a kő-papír-olló játék illusztrálja. A ciklikus viszonyban lévő stratégiák általános jellemzője, hogy mindhárom stratégia jelen lehet úgy, hogy a stratégiák aránya térben és/vagy időben állandóan változik, pontosabban oszcillál egy átlagérték körül. Ilyen esetekben a rendszer a laikusok számára meglepő módon reagál a külső beavatkozásokra. Konkrétan, ha valamelyik stratégiát a paraméterek (törvények) változtatásával segítjük, akkor ennek a beavatkozásnak az igazi haszonélvezője a megsegített stratégia „ragadozója” lesz [11].

Az egyszerű ciklikus dominancia következményein túl további érdekes jelenségek léphetnek fel, ha több vetélkedő stratégia is jelen van a rendszerben. Ha ilyenkor a stratégiák egy szűkebb halmaza önmagában is képes az együttélésre, akkor az együttélésük mint önálló „elem” is része lehet egy ciklikusan domináns viszonyoknak, ami még bonyolultabb önszervező mintázatok kialakulását eredményezheti. *Boza Gergely* és *Scheuring István* ugyanebben a számban megjelenő cikkében számos biológiai rendszert láthatunk két faj együttélésére [12]. A bonyolultabb ciklikus dominanciára mutat példát a 3. ábra, ahol az élősködő, a passzív együttműködő és az intézményesen büntető stratégiák stabil együttélése valósul meg. Itt a sötétzöld színnel jelölt intézményesen büntető stratégia képes elterjedni a pirossal jelölt élősködők tartományában. Ugyanakkor az élősködő legyőzi a kék színnel jelölt passzív együttműködő stratégiát. Ez utóbbi a kisebb költsége miatt viszont képes elterjedni az elsőként említett megoldás rovására, amely a büntető és az élősködő stratégiák együttélését engedi meg [13].

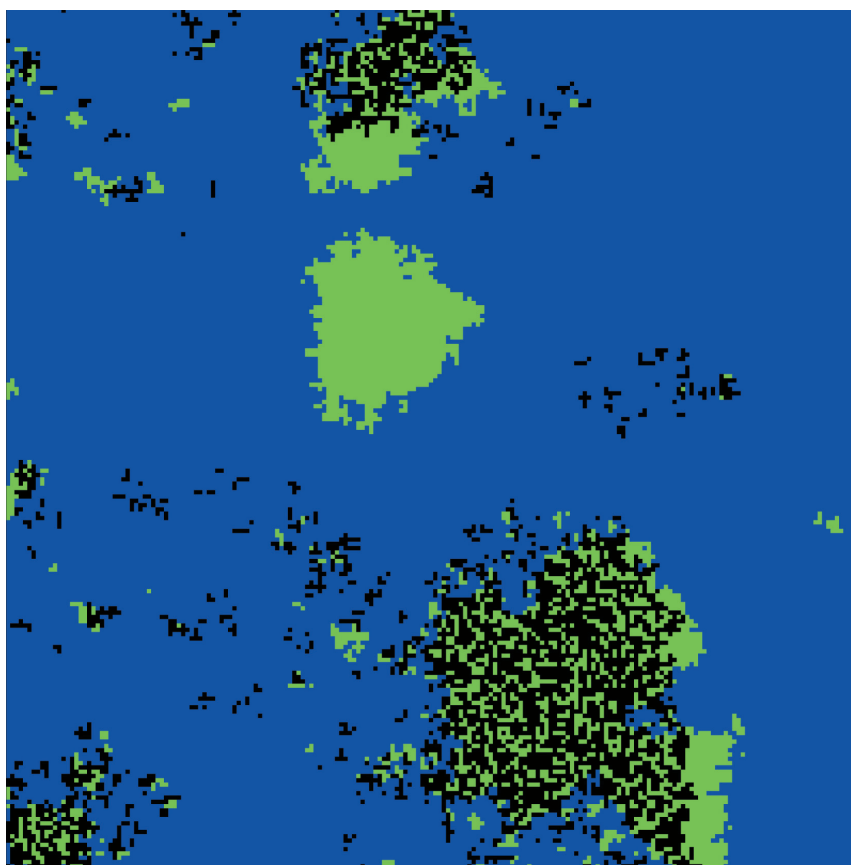
Fontos hangsúlyoznunk, hogy az ilyen társulásokat is tartalmazó ciklikus dominanciák lehetősége sokkal általánosabb, és a stratégiák számának növelésével szinte törvényszerűen bukkannak fel egy komplex rendszerben, mint a leginkább stabil megoldások.

Az utóbbi állítás illusztrálására tekintsük a kétszemélyes ultimátumjátékot, ami gyakran használt modell a testvériség tanulmányozására. A játék során az egyik játékos ($0 \leq p \leq 1$) ajánlatot tesz arra, hogy

Komplex viselkedés társadalmi dilemmákban

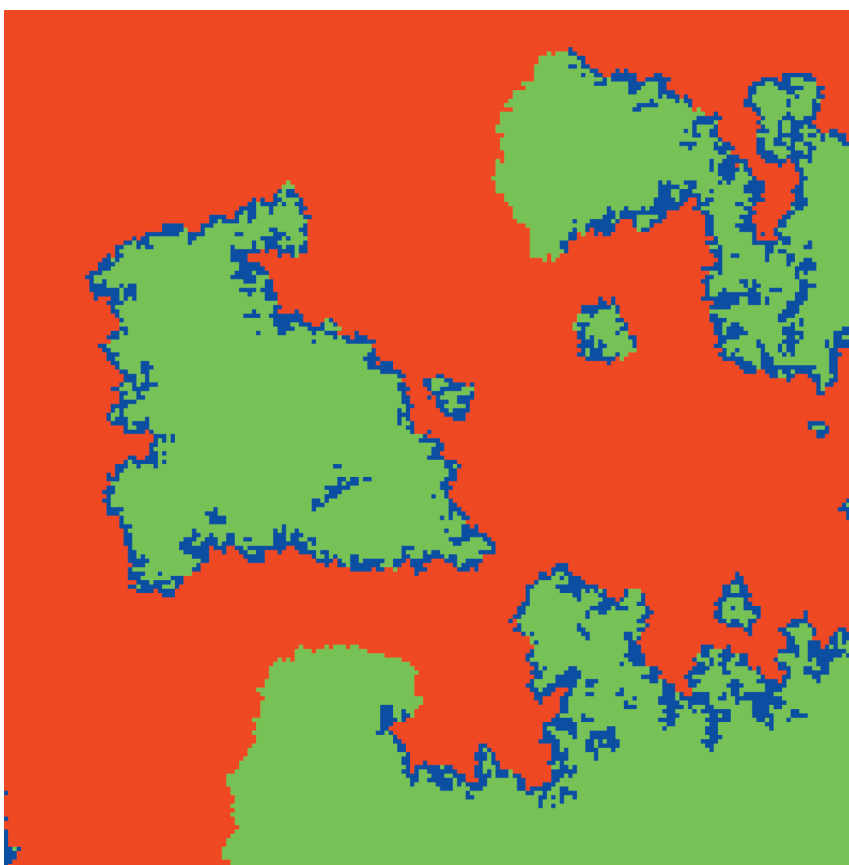


3. ábra. Három megoldás ciklikus dominanciájaként előálló, önszervező térbeli mintázatról készült pillanatfelvétel. Az idő múlásával az élőködő stratégia (piros) legyőzi a passzív együttműködőt (kék), aki ugyanakkor sikeres az élőködő és az intézményesen büntető (sötétzöld) stratégia együttélésével szemben. A kép egy 400x400-as négyzetet mutat [13]



milyen arányban osztozkodjanak egy egységnyinek tekintett összegben. A másik játékos csak egy bizonyos ($0 \leq q \leq 1$) szint feletti ajánlatot hajlandó elfogadni. Ha $p < q$, akkor az utóbbi visszautasítja az ajánlatot, és mindketten üres kézzel zárják a játékot, ellenkező esetben az ajánlat szerint osztozkodnak. A racionális magatartás ilyen esetben: felajánlani egy minimális p összeget, amit a másikkal célszerű elfogadnia (azaz alacsony q értéket választani), mivel az is több, mint a semmi. Ezzel szemben a kísérletek azt mutatják, hogy embertársaink zöme az ajándék igazságos megfelezésétől nem nagyon tér el. Ugyanakkor az aránytalanul alacsony p ajánlatot a kísérleti alanyok nagy része visszautasította, amit értelmezhetünk úgy is, hogy vállalta a másik megbüntetésének a „költségét”, ami az elmaradt haszon formájában jelentkezik. Ilyen helyzetben egy (p, q) ajánlati-elfogadási paraméter pár segítségével jellemezhető a játékosok stratégiája, amit a kísérletekben és a számítógépes modellezésekben lehet diszkrét értékekkel közelíteni. Például az „empatikus” játékosnál, aki ugyanazt várja el a társától, amit önmagától, az ajánlati és elfogadási értékek megegyeznek ($p = q$). Ez természetesen még nem határozza meg egyértelműen ezt a stratégiát, hiszen az ajánlott (és elvárt) szint lehet alacsony, de akár magas érték is. Az egyszerűség kedvéért a $[0, 1]$ ajánlati-elfogadási intervallumot n részre osztva E_1, \dots, E_n diszkrét empatikus stratégiát tekinthetünk. Nem túl sok ($n=5$) versengő stratégiát választva belátható, hogy a játék térbeli változatában ugyanolyan ciklikus dominanciára jellemző mintázat alakulhat ki, ha egy rögzített $R = (p, q)$ stratégia versengését vizsgáljuk kétféle (E_1 és E_2) empatikus stratégia jelenlétében. A 4. áb-

4. ábra. Három stratégia térbeli együttélése az ultimátum játékban. A 240x240-es rendszerről készült pillanatfelvételen a feketével jelölt $R = (0,645; 0)$ stratégia képes elterjedni a zölddel jelölt E_2 stratégia tiszta tartományában egy jellemző arányú keveréket hozva létre (a kép alján elhelyezkedő fekete-zöld folt ilyen megoldásra utal). Ugyanakkor a kizárólagosan az E_2 stratégiából álló zöld folt elhódítja a kék színnel jelölt E_1 stratégia tartományát. Az utóbbi viszont az elsőként említett társulást győzi le, megvalósítva a stratégiák és társulások ciklikus dominanciáján alapuló együttélést [14]



5. ábra. Az inváziós frontok harca egy büntető stratégiát tartalmazó térbeli modellben. A 200×200 -as rendszerről készült pillanatképről az élőködőket (piros) az egyéni büntetők (világoszöld) képesek legyőzni, de a kékkel jelölt passzív együttműködő stratégia életképes lehet az előbbieket határolva. A passzív együttműködő a büntető stratégia rovására terjed, de annak hiányában maga is védtelen lesz az élőködőkkel szemben, így végül a három stratégiát tartalmazó fronton az élőködők kerülnek nyertes helyzetbe [15]

rán feketével jelölt R stratégia képes elterjedni a zölddel jelölt E_2 stratégia tiszta tartományában, és azzal egy olyan keveréket alkotni, ahol az arányuk az R -t jellemző (p, q) értékektől függ. Ugyanakkor önmagában az E_2 képes elhódítani a kék színnel jelölt E_1 stratégia által uralt területet. Az utóbbi viszont az elsőként említett $(E_2 + R)$ társulást győzi le, ilyen módon zárva a kört [14].

A mór megtette kötelességét, a mór mehet

Visszatérve a büntető stratégiával kiegészített közlegelő játékra, végezetül egy olyan vizsgálatunkat szeretnénk bemutatni, ahol még a stratégiákat elválasztó frontvonal inváziójának az iránya sem egyértelmű, köszönhetően a további stratégiák esetleges jelenlétének. Az előző modellel ellentétben itt a büntető stratégia ún. „egyéni” büntető, ami azt jelenti, hogy a passzív együttműködőhöz képes nincs állandó többletköltsége: extra terhet csupán akkor visel, ha élőködő van a környezetében [15]. Megfelelően válasz-

tott büntetési költség, illetve büntetési tétel esetén elérhető, hogy a büntető stratégia képes legyőzni az élőködőket. Ennek megfelelően az 5. ábra bal alsó szélén a zöld színnel jelölt büntető stratégiát jelentő tartomány folyamatosan csökkenti a pirossal jelölt élőködők létszámát. Ugyanakkor, ha a találkozásuknál kék színnel jelölt passzív együttműködő stratégia is jelen van, akkor a büntetés viszonylagos nagy mértéke és nagy költsége miatt a passzív együttműködés választása előnyösebb, így a másik két stratégia frontvonala mentén egy vékony réteget alkotva képes továbbterjedni. Ez a sikere lesz egyben a végzete is, mert ha a passzív stratégia túlságosan feldúsul a frontvonal mentén, akkor a büntetők hiányában az eddigi „győztes” sebezhetővé válik az élőködőkkel szemben, akik így megfordítják a pirossal és a zölddel jelölt tartományok közötti invázió irányát, és végül már csak élőködők maradnak a rendszerben. Röviden, a passzív stratégia a „mór” szerepét játszva, pillanatnyi sikerükkel az élőködők érdekét szolgálják, de végül ők is áldozattá válnak. Ez az utóbbi eredmény ismételt rá-

világít arra, hogy a stratégiák viszonyának az általános elemzésénél nem szabad azt a módszert követni, hogy a stratégiák egyenkénti párba rendezésével határozzuk meg azok viszonyát: egy további stratégia jelenléte lényegesen befolyásolhatja a nyilvánvalónak tűnő viszonyokat.

A statisztikus fizikában gyakran emlegetett tapasztalatot idézve, „a több más”, ami különösen igaz akkor, amikor a stratégiák számának növelésével rohamosan növekszik a kialakuló megoldások (mintázatok) lehetséges száma és komplexitása. Terjedelmi okokból nem térünk ki rá, de további szabadsági fokot jelenthet az, ha a játékosok bizonyos, a stratégiaátadást befolyásoló képessége időfüggő (pl. fiatalon az ember is könnyebben tanul, könnyebben vesz át egy másik stratégiát, mint később). Ez implicit módon tovább növeli a játékosok állapotainak a számát, ami azt is eredményezheti, hogy akár pusztán két stratégia esetén is megfigyelhetünk a ciklikus dominanciára jellemző hullámfront-mintázatokat.

Összefoglalásul azt a tanulságot fogalmazhatjuk meg, hogy a sokszereplős rendszerekben a játékosok társulásából fakadó időfüggő mintázatok lényegesen komplexebbek lehetnek annál, mint ami az egyes stratégiák nyereségértékeinek az összehasonlításából következne. Emiatt számos olyan lehetőség van az együttműködés fenntartására, amelyeket az evolúciós folyamatok képesek felszínre hozni, vagy amelyeket mi emberek is érvényesíthetnénk a társadalom javára, ha kellő mélységben értenénk az egymással összefonódó hatások következményeit. Ehhez azonban arra van szükség, hogy a különböző tudományterületek kutatói összefogjanak, illetve szót értenek egymással a matematika nyelvén, és eredményeik hasznosítására a döntéshozók is törekedjenek. ✨

A szerző köszöni az OTKA K-101490 támogatását.

IRODALOM

- [1] Sigmund, K., *The Calculus of Selfishness*, Princeton Univ. Press, 2010.
- [2] Szabó Gy., *Magyar Tudomány*, **173**, 6. sz., 642, 2012
- [3] Pennisi, E., *Science*, **309**, 93, 2005.
- [4] Forgó F., *Magyar Tudomány*, **170**, 5. sz., 515, 2009.
- [5] Czárán T., *Magyar Tudomány*, **171**, 4. sz., 396, 2010.
- [6] Boza G., Kun A., Scheuring I., Dieckmann U., *PLoS Comput. Biol.*, **8**, e1002660, 2012.
- [7] Nowak, M.A., May, R.M., *Nature*, **359**, 826, 1992
- [8] Santos, F.C, Pacheco J.M., *Phys. Rev. Lett.*, **95**, 098104, 2005.
- [9] Nagy M., Ákos Zs., Biró D., Vicsek T., *Nature*, **464**, 890, 2010.
- [10] Jánosi Imre: *Energia és társadalom: drasztikus fázisátalakulás küszöbén állunk?* (Természet Világa mostani különszáma).
- [11] Hofbauer J., Sigmund K., *Evolutionary Games and Population Dynamics*, Cambridge Univ. Press, 1998.
- [12] Boza G., Scheuring I.: *Diverzitás és komplexitás a mutualista kapcsolatokban* (Természet Világa mostani különszáma).
- [13] Szolnoki, A., Szabó, G., Perc, M., *Phys. Rev. E*, **83**, 036101, 2011.
- [14] Szolnoki, A., Perc, M., Szabó, G., *Phys. Rev. Lett.*, **109**, 078701, 2012.
- [15] Szolnoki, A., Szabó, G., Czákó, L., *Phys. Rev. E*, **84**, 046106, 2011.

VUKOV JEROMOS

Csalni vagy nem csalni?

Matematikai komplexitás az emberi kapcsolatokban

Azt gondolnánk, hogy az emberi kölcsönhatásokat nem lehet leírni matematikai képletekkel. Meglepő módon a legtöbb konfliktushelyzet főbb jellemzői leírhatóak néhány paraméter segítségével. Azonban az egyszerű modell nem jelenti azt, hogy maga a jelenség összes tulajdonsága is egyszerűen kiszámolható, illetve megadható. Gondoljunk például a kettős ingára vagy akár a mágneses ingára [1]: a felépítésük végtelenül egyszerű (és a mozgásukat leíró egyenletek is egyszerűek), azonban a mozgásuk kaotikus, tehát nagyon nehéz megjósolni, hogy az ingák milyen pozícióban lesznek egy idő után, mivel a rendszer nagyon érzékeny a kezdeti feltételekre.

A Neumann János és Oskar Morgenstern által elsősorban közgazdasági problémák tanulmányozására létrehozott játékelmélet biológiai és szociológiai kiterjesztése kiválóan alkalmas az emberi kölcsönhatások leírására. Az egyik legismertebb és leginkább tanulmányozott „játék” az ún. *fogyódydilemma* helyzet. A furcsa név magyarázata a problémakör illusztrálására először felhasznált élethelyzetben keresendő. A történet szerint a seriff elkap két bűnözőt, akiket a közeli bolt kirablásával gyanúsítanak, azonban nincs ellenük megdönthetetlen bizonyíték, mert a lopott holmitól időben megszabadultak, és a szemtanúk is eltűntek. Emiatt a seriff külön-külön próbálja meg rávenni a két jómadarat, hogy tanúskodjanak egymás ellen a büntetésük csökkentésének reményében. Mivel a tolvajok nem tudják, hogy a társuk milyen döntést hoz, négy különféle kimenetel lehetséges. Ha mindketten hallgatnak a betyárbeszűlet kódexének megfelelően, akkor a seriffnek el kell engednie őket a vizsgálati fogság (2 hét) után. Ha egyikőjük beárulja a társát, míg a partner tartja a száját, akkor az árulkodót azonnal szabadon eresztik, a kijátszott társ viszont megkapja a büntetést járó teljes büntetést (1 év börtön). Ez az eset természetesen fordított sze-

reposztásban is megtörténhet. És végül, ha mindketten beárulják a másikat, akkor a seriff dörzsölheti a kezét, hiszen leültetheti mindkettőt, igaz, hogy a büntetési tétel ez esetben valamivel kisebb (8 hónap börtön), mert megoszlik a két rabló között (1. ábra). A helyzet azért állítja dilemma elé a bűnözőket, mert a másik beárulása a partner döntésétől függetlenül kedvezőbb kimenetelt eredményez: ha a másik tartja a száját, az „árulkodó” azonnal szabadul a bajtársiassággal járó vizsgálati fogság helyett; míg ha a partner árulkodik, akkor a kölcsönös árulkodással járó közös büntetés nyilvánvalóan jobb eredmény az egyedüli büntetésnél. Azonban, ha mindketten ezt a logikát követik, akkor mindketten börtönbe kerülnek hosszabb időre a kölcsönös hallgatással járó rövid vizsgálati fogság helyett.

A fenti helyzet különlegesnek tűnhet, de valójában az élet nagyon sok területén megjelenik nap mint nap. Amikor két félnek egyidejűleg kell döntést hoznia a másik döntésének ismerete nélkül, és a döntésektől függő kimenetek rangsora megegyezik a fogyódydilemma játék során felsoroltakéval, akkor a két résztvevő újra és újra a fenti dilemmával szembesül. A dilemma egyik leg-híresebb példája a nukleáris fegyverek leszerelési tárgyalása volt az Egyesült Államok és a Szovjetunió között, amelynek során a feleknek azt kellett eldönteniük, hogy a tárgyaláson elhangzottakkal egy időben *valóban* megsemmisítsék-e a nukleáris fegyvereket. A kölcsönös leszerelés jó dolog, azonban ha a másik fél megsemmisíti a bombákat, akkor a „csaló” fél abszolút világhuralmi pozícióba kerülhetett volna. Kölcsönös „csalás” esetén természetesen fennmarad a korábbi világpolitikai feszültség. A tárgyalások folyamán ugyan csökkentették a nukleáris fegyverarzenált, de a teljes leszerelés nem valósult meg, a dilemmahelyzet még mindig fennáll.

		2. játékos	
		Hallgatás (Együtműködés)	Árulás
1. játékos	Hallgatás (Együtműködés)	3 / 3 2 hét fogság / 2 hét fogság	0 / 5 1 év börtön / szabadon elmehet
	Árulás	5 / 0 szabadon elmehet / 1 év börtön	1 / 1 8 hónap börtön / 8 hónap börtön

1. ábra. A fogyódydilemma különböző kimenetelei a hozzájuk tartozó eredeti „nyereményekkel” és a kísérletekben használt értékekkel

Kisebb horderejű, hétköznapi példák a tolakodó vagy udvarias közlekedés, a tisztességes munkavégzés kontra élösködés, tisztességes vagy tisztességtelen magatartás, a személtelés vagy éppen a környezetszennyezés kérdése. Ezekben a példákban általában valamilyen befektetés (odafigyelés, energiabefektetés, munkavégzés stb.) elkerülésével juthatnak a csaló magatartást választók kisebb-nagyobb előnyökhöz, ha a partner(ek) a befektetés véghezvitele mellett dönt(enek). A természetben előforduló többszereplős fogolydilemma játékokra láthatunk példát a **2. ábrán**.

Az együttműködés matematikája

A téma könnyebb elemzésének érdekében a fogolydilemma játékban előforduló különböző fogalmak jelölésére általános kifejezéseket vezetünk be, és ezután a vizsgált szituációtól függetlenül ezeket a kifejezéseket fogjuk használni [2]. A kölcsönhatásban részt vevő feleket *játékosoknak* fogjuk hívni. Mint a fenti példákban is látszott, a játékosok nemcsak személyek lehetnek, hanem országok, cégek is. ökológiai példa esetén pedig állatok vagy akár baktériumok. A kölcsönhatás – *játék* – során a játékosok két döntés közül választhatnak: *együttműködés* vagy *árulás*, ahol az együttműködés természetesen a cinkostárs hallgatással való fedezése az eredeti fogolydilemma szituációjában. A döntések következtében elért személyes eredmény a *nyeremény*. Természetesen ez is nagyon sokfajta lehet a vizsgált esettől függően: az eredeti játék esetében a büntetés hossza (rövidsége), a leszerelési tárgyalásoknál a békésebb környezet vagy nagyobb világpolitikai befolyás, más szituációknál időbeli vagy anyagi nyereség. A lényeg, hogy fogolydilemma helyzet esetén a nyeremények sorrendje a következő: együttműködő partner ellen elkövetett árulás a „legjövendőzebb”. Ezt követi a kölcsönös együttműködés nyereménye, majd a kölcsönös árulásé. A legkisebb nyereményt pedig az elárult együttműködő „balek” kapja. A negatív nyeremények kiküszöbölésére a kísérletek folyamán általában az **1. ábrán** is látható 5, 3, 1, 0 értékeket használják.

Mint láttuk, egyszeri fogolydilemma találkozás során a racionális játékosok az árulást választják, azonban a természetben és az emberi kapcsolatokban ilyen szituációkban nagyon sok helyütt széles körű együttműködést figyelhetünk meg. Hogyan egészíthetjük ki a modellt, hogy az valóság-hűbb képet adjon? Az erre irányuló egyik próbálkozás a játék ismételt tétele volt. Azt gondolhatjuk, hogy a jövőbeli találkozásból elérhető kölcsönös együttműködési nyeremény reményében a játékosok nem fogják egymást korai árulással kölcsönös árulásokkal járó „adok-kapokba” hajszolni. Ez azonban nem teljesen igaz. Ha a játék ismétléseinek száma előre ismert, akkor racionális ellenfelek esetén, mindketten a teljes ismételt játék során következetesen árulást fognak elkövetni. Hiszen ebben az esetben az utolsó forduló egyenértékű egy egyfordulós fogolydilemmával, aminek az eredménye már ismert: kölcsönös árulás. Mivel ennek a fordulónak az eredményét már tudjuk, ez nem fogja befolyásolni az előző fordulók döntéseit. Hasonló logikát követve, az utolsó előtti fordulóban is árulást fognak elkövetni a játékosok. Így végül a teljes ismételt játék során folyamatos, kölcsönös árulások sorozatát figyelhetjük meg.

A helyzet azonban gyökeresen változik, ha az előre meghatározott számú találkozás helyett a következő forduló bekövetkezésének van valamekkora valószínűsége. Ebben az esetben ugyanis a játékosok nem tudják, hogy melyik lesz az utolsó forduló, és az előző bekezdésben ismertetett érvelés nem alkalmazható. Az összesen lejátszott játékok átlagos száma természetesen megegyezhet a fix számú ismétléses esettel. Megfelelően nagy valószínűségű ismétlés esetén itt már igaz lehet, hogy jövőbeli találkozásból elérhető nyeremény reményében a játékosok inkább a nagyobb összenereményű kölcsönös együttműködést választják az esetleges egyszeri, nagyobb haszonnal járó árulás helyett, amit valószínűen a partner szintén árulással torolna meg.



2. ábra: Többszereplős fogolydilemmák a természetben. Felső kép: A préda elejtéséhez közel kell merészkedni a bölényhez, vállalva az ezzel járó veszélyt. A távolabb maradás felel meg az áruló magatartásnak. Alsó: A pezsmatulok kört formálva védik kicsinyeiket. A kör megbontása és ezzel a kicsinyek kiszolgáltatása az áruló magatartás megfelelője. Ez a viselkedés természetszerűleg növeli az egyéni életben maradási esélyeket, azonban a közösségre nézve meglehetősen káros hatással van

Kölcsön kenyér visszajár?

A racionális, nyereménymaximalizáló (másképp fogalmazva: önző) játékosok számára értelemszerűen adódik a kérdés: milyen stratégiát célszerű használni az ismételt fogolydilemma játékban? A válasz nem triviális. Ugyanis a forduló számának növekedésével exponenciálisan nő a lehetséges döntések száma. A lehetséges stratégiák figyelembe vehetik az összes információt: a nyeremények nagyságát, a következő forduló valószínűségének értékét, a játékos és partnerének korábbi fordulóiban hozott döntéseit és ezek egymásutánját.

Az „ideális” stratégia megtalálására *Robert Axelrod* az 1980-as évek elején kiírt egy számítógépes versenyt. A pályázóknak számítógépes algoritmus formájában kellett beküldeniük a szerintük legjobb stratégiát. A verseny folyamán az összes (szám szerint 14) beküldött stratégia körmérkőzéses formában játszott egymás ellen, és a győztes az lett, amelyik a legnagyobb nyereményt gyűjtötte össze. A legjobb eredményt a kölcsön kenyér visszajár (KKV) stratégia érte el. Ez a stratégia az első fordulóban együttműködik, majd a partner előző fordulóbeli döntését viszonzozza. Egyszerűsége ellenére több nagyon fontos tulajdonsága van. Elsősorban *barátságos*, azaz soha nem követ el elsőként árulást. *Megtörő*, tehát azonnal árulással válaszol, illetve büntet, ha az ellenfél megpróbálja kihasználni. Viszont ezzel együtt *megbocsátó* is: ha a partner jó útra tér és újra együttműködni kezd, akkor ő is azonnal visszatér az együttműködéshez. És végül a KKV *egyszerű*. Ez azért fontos, mert az ellenfél könnyen kiismerheti, és mivel tudja, hogy nem

Csalni vagy nem csalni?

érhet el ellene előnyt árulással, logikus döntésként folyamatosan együttműködhet vele.

A verseny eredményének részletes ismertetése után Axelrod újabb versenyt hirdetett. Erre már 62 pályázó küldte el a stratégiáját. Több olyan stratégiát is beküldtek, amelyek megnyerte volna az első bajnokságot, többek között a KKV egy „javított” változatát. A stratégiák újra összemérték hatékonyságukat a körmérközésen, a végelszámolásnál azonban nagyon meglepő eredményt mutatott a kijelző: a versenyt újra a KKV nyerte! Az újonnan beküldött stratégiák megnyerték volna ugyan az első bajnokságot, de az új környezetben kioltották egymás hatását, és a sűrű pályán újra a KKV futott be az első helyre. Ez nem azt jelenti, hogy a KKV a legjobb stratégia minden esetben. Hiszen a második bajnokságon is lehetett volna alkotni nála hatékonyabb stratégiát a többi ellenfél *előzetes* ismeretében. Azonban ismeretlen környezetben a fent leírt tulajdonságainak köszönhetően nagyon kiemelttel teljesítményre képes.

A KKV jobb megértése érdekében hangsúlyozni kell, hogy ez a stratégia soha nem nyeri meg az összecsapásait egyetlen másik stratégia ellen sem, a legjobb esetben is döntetlenre végez, hiszen nem követ el először árulást, így nem tud „előnybe” kerülni a riválissal szemben. Amint azonban látható, nem is ez a fontos az eredményesség szempontjából. A balek ellen elkövetett árulás eredményezi ugyan a legnagyobb egyéni bevételt, de ezt hosszú távon nagyon kevés stratégia ellen lehet realizálni, így, úgy tűnik, hogy a második legnagyobb nyeresémet ésszerű célba venni, hiszen ennek megszerzése során nincs ellenérdekltség a játékosok között.

Az ismételt fogolydilemma játék későbbi tanulmányozása során is bebizonyosodott a KKV stratégia robusztussága, egyfajta speciális „joker” stratégiaként tekintettek rá, illetve a nagylelkű KKV-re, amely ugyanúgy működik, mint az egyszerű KKV, csak megbotcsátóbb: időnként a partner árulása után is együttműködik, bekegyjött nyújtva neki [3]. Karakterisztikus tulajdonságai miatt nagyon sokszor interpretálták az erkölcs matematikai kifejeződésékként, hiszen a „szemet szemért, fogat fogért” elvet követi, megbotcsátó és alapvetően a pozitív értékrendet követi. Emellett a nyeresémet is egyenlően osztja el, tehát „fair” stratégia.

Zéró-determináns stratégiák

A közelmúltban meglepő felfedezést tettek az ismételt fogolydilemma játék területén, azon a területen, amiről régóta azt hitték, hogy már mindent kiderítettek. *William Press* a végtelenszer ismételt fogolydilemmát tanulmányozta egy egyszerű problémát vizsgálva, amikor a számítógépes programja többször egymás után, különböző helyeken nullával való osztás miatt lefagyott. A játékosok nyeresémet ebben az esetben az egy fordulóban szerzett átlagos nyeresémet jellemzik végtelen sokszor ismételt játékot feltételezve. Az eredmények jó közelítéssel igazak nagyon hosszú idejű kölcsönhatásokra is. A probléma elemzése során *Press* a fizika sokféle területén – kvantumelektrodinamika, szilárdtestfizika, csillagászat, nukleáris technika – maradándó alkotó *Freeman Dyson* segítségét kérte, így derült végül fény a *zéró-determináns stratégiák* (ZD) létezésére, amelyek alapvetően változtatták meg a klasszikus, kétszereplős fogolydilemmáról kialakult képünket [4].

A vizsgálatok folyamán különféle hosszúságú memóriával bíró stratégiákat elemeztek. Egy n memóriájú stratégia a legutóbbi n fordulóban hozott döntések alapján hoz döntést arról, hogy az aktuális fordulóban mit cselekszik. Első pillantásra az ember úgy gondolná, hogy a hosszabb memóriájú játékos eredményesebb stratégiát képes kialakítani, hiszen több adat áll rendelkezésére, ki tudja elemezni a partner viselkedését stb. *Press* és *Dyson* közös cikkének első meglepő eredménye azonban ellentmond ennek a logikusnak tűnő feltételezésnek. Ugyanis a végtelenszer ismételt fogolydilemmában a rövidebb memóriájú játékos akarata érvénye-

sül. Ez azt jelenti, hogy az ellenfelénél hosszabb memóriájú stratégia alkalmazásával nem lehet előnyhöz jutni, mindig létezik egy olyan – a partner memóriaméretének megfelelő – stratégia, amelyik a hosszabb memóriájú változattal megegyező eredményeket produkál. Ez elsőre meglepően hangzik, de ha belegondolunk, hogy a partner csak az utolsó n lépésre alapozza a döntését, akkor semmi értelme régebbi eseményeket figyelembe venni a saját döntés meghozatalakor. Ennek egy speciális esete, amikor a partner nem használ memóriát. Ebben az esetben triviálisan mi sem érünk semmit hosszabb távú memória alkalmazásával.

A memória nélküli stratégiák meglehetősen egyszerűek, egy paraméterrel jellemezhetőek: valamikor valószínűséggel együttműködnek, egyébként pedig árulást követnek el minden körülménytől függetlenül. Ezért *Press* és *Dyson* olyan stratégiákat vizsgált, amelyek az utolsó forduló kimenetelére emlékeznek (azaz $n=1$). Ezek a stratégiák négy paraméterrel írhatóak le: a következő fordulóbeli együttműködés valószínűségével az előző forduló eredményének függvényében. Mivel a fogolydilemma négyféle kimenete lehetséges, ezért szükséges a négy valószínűség-érték. Ezek között a stratégiák között bukkantak rá a már korábban említett zéró-determináns stratégiákra. A név onnan származik, hogy a játékosok átlagnyeresémetének kiszámolása során egy mátrix determinánsát kell meghatározni, és ez a determináns válik nullává bizonyos stratégiák esetén.

A felfedezés lényege, hogy egy ilyen stratégia választásával az adott játékos befolyásolhatja saját és partnere nyeresémetének lineáris kombinációját *a partner stratégiaválasztásától függetlenül!* A stratégiák meghatározásánál egy egyenletrendszer megoldása adja a stratégiákhoz tartozó együttműködési valószínűségeket, azonban nem minden érték eredményez valós megoldást, hiszen a valószínűségeket 0 és 1 közé kell esniük. Ezért sorban megvizsgálták, hogy a különböző speciális esetekhez valóban léteznek-e megfelelő stratégiák.

Elsőként azt vizsgálták, hogy lerögzítheti-e az egyik játékos az ellenfél nyeresémetét. Meglepő módon ez bizonyos határok között lehetséges. Megfelelő stratégiaválasztással az ellenfél nyeresémeté beállítható a kölcsönös árulásért és a kölcsönös együttműködésért kapható nyeresémetek között bármilyen értékre. Ekkor a partner bármilyen stratégiát követhet, bármilyen döntést hozhat, ha „megszakad”, akkor is az első játékos által előre eldöntött nyeresémetet fogja elérni. Ezek a stratégiák ugyanakkor nem mondanak semmit a saját nyeresémetéről, ez függ az ellenfél stratégiájának választásától, lehet kisebb és nagyobb is, mint az ellenfelnek beállított érték.

Ezután megvizsgálták, hogy a saját nyeresémet előre meghatározása is lehetséges-e, de ebben az esetben a megoldások nem adtak valós valószínűség-értékeket, így – ahogy előre sejtethető volt – a saját nyeresémet nem fixálható.

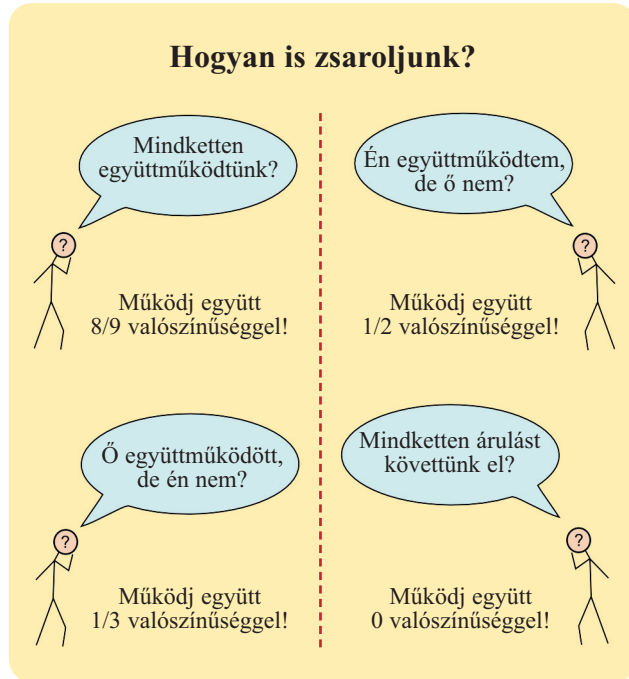
A zsarolás diadala?!

Az ellenfél nyeresémetének fixálása érdekes eredmény ugyan, de az érdekességen kívül nem használható fel sok mindenre, nem alkalmas a játék megnyerésére, hiszen nem garantálja, hogy a saját nyeresémet nagyobb lesz az ellenfelénél. Ezzel ellentétben az úgynevezett *zsaroló ZD* stratégiák már képesek erre. Ezek a stratégiák a kölcsönös árulásért járó nyeresémet fölötti extraprofitból követelnek (és kapnak) a partnernél jóval nagyobb részt. A kizsákmányolás mértékét a stratégia zsarolási faktora határozza meg: ez egytől végtelenig változhat. Például kettes zsarolási faktor esetén a zsaroló ZD stratégiát játszó játékos az ellenfél extra nyeresémetének kétszeresét kapja. A végtelen sok lehetséges ilyen stratégia közül egynek a „receptje” látható a **3. ábrán**. Elsőre ugyan gyanúsak tűnhet a dolog, hiszen a játék szimmetrikus. Hogyan lehet akkor az egyik játékos kitüntetett szerepben? Ha mindketten ilyen stratégiát használnak, akkor mindketten a másik extra

Vukov Jeromos: Csalni vagy nem csalni?

nyereményének kétszeresét kapják. Ez természetesen csak úgy lehetséges, ha ez az extra nyeremény nulla. Tehát ebben az esetben mindketten a kölcsönös árulásért járó, alacsony nyereményt kapják.

A zsaroló ZD stratégiák akkor nagyon hatékonyak, ha az ellenfél nincs tudatában, hogy zsarolni akarják, vagy szimplán nem is-



3. ábra. A zsarolás receptje: az ábrázolt stratégia kétszer annyi extra nyereményt zsebel be, mint a partnere. A konkrét valószínűség-értékek arra az esetre érvényesek, ha a fogolydilemmában a klasszikus 5, 3, 1, 0 nyereményeket használjuk

meri az ilyen stratégiákat (ami nem is túl meglepő, lévén, hogy ilyen sokáig nem derült fény a létezésükre). Ekkor a partner csak a saját nyereményének maximalizálásával van elfoglalva. Ezt a zsaroló ZD kiválóan ki tudja használni: minél együttműködőbb stratégiát választ a partner, annál nagyobb nyereményhez jut, ugyanakkor annál nagyobb extra nyereményhez juttatja a ZD stratégiát használó játékost is. Tehát a ZD stratégia mintegy alamizsnát dobva ösztönzi az ellenfelet a mindig együttműködő, alárendelt stratégia használatára, ami számára a legnagyobb hasznot hozza. A partner végül is saját kapzsiságának áldozatává válik és elősegíti a saját kiszákmányolását.

A helyzet gyökeresen megváltozik, ha a partner figyelemmel kíséri, hogyan változnak a nyeremények, miközben ő megpróbálja a saját nyereményét maximalizálni. Rájöhet, hogy a nyereménye lassanként növekszik ugyan, de ezzel extraprofit hoz juttatja a másik játékost. Ekkor az ismételt fogolydilemma átlényegül ultimátum játékká.

Az ultimátum játékban az egyik játékos egy nagyobb összeget kap kézhez, és el kell osztania játékos társa és önmaga között. Ha a játékos társa beleegyezik a felajánlott elosztásba, mindketten megtarthatják a nekik jutó részt, azonban ha visszautasítja az elosztási arányt, akkor egyikőjük sem kap semmit. A racionális, játékelméleti megoldás itt azt mondja, hogy bármilyen elosztást el kell fogadnunk, hiszen a minimális pénz is több az elosztás visszautasításával járó nulla összegnél. Emberekkel végzett kísérletek ugyanakkor azt mutatják, hogy a játékosok büntetésekké visszautasítják

az olyan elosztást, amely jelentősen eltér a fair ötven százalékától. A zsaroló stratégiát játszó játékos ellenfele olyan, mint az ultimátum játék során a felajánlott elosztásról döntő játékos. Észrevette, hogy az elosztás során partnere jóval nagyobb részt tart meg magának a közös bevételből. Ekkor eldöntheti, hogy a „kevés is jobb, mint a semmi” alapon elfogadja a tisztességtelen ajánlatot, vagy folyamatosan árulást választva megbünteti a zsarolni próbáló ellenfelet, ezzel mindkettőjüknek a minimális nyereményt juttatva.

Ha mindkét játékos ismeri a zsaroló ZD stratégiákat, akkor kezdetben valószínűleg megpróbálják egymást kihasználni, azonban ez a fentebb leírt módon nulla nyereményt eredményez mindkettőjüknek. Így, ha ésszerűen állnak a problémához, tárgyalni kezdenek, és lecsökkenthetik a zsarolási faktorokat az igazságos egyre, azonos, magasabb nyereményt juttatva mindkettőjüknek. Megjegyezzük, hogy a KKV stratégia is a zsaroló ZD stratégiák családjába tartozik, igaz, egyfajta „fehér báránként”, ugyanis a zsarolási faktora 1, tehát nem akar előnybe kerülni az ellenfélhez képest, csak fair elosztást követel. A ZD stratégiák tanulmányozása megmutatta azt is, hogy végtelen sok, a KKV-hoz hasonló fair stratégia létezik, azonban ezek nem írhatóak le olyan egyszerű „működési elvvel”, a stratégiaparamétereket látva nem nyilvánvaló, hogy igazából KKV-ként funkcionálnak.

Mint látható, a zsaroló ZD stratégiák akkor a leghatékonyabbak, ha az áldozat nincs tudatában annak, hogy ki akarják használni, és adaptív módon a saját bevételének maximalizálásával foglalkozik. Ha azonban olyan környezetben vizsgáljuk őket, ahol a partnerek egy fix stratégiát játszanak (mint pl. az Axelrod-féle versenyen), akkor nagyon könnyen kölcsönös árulási sorozatokba kerülhetnek, nagyon lecsökkentve ezzel a nyereményüket. Ugyanez az eredmény, ha önmaguk másolatával játszanak, ami például evolúciós környezetben fordulhat elő, ahol a sikeres egyedek (stratégiák) önmagukhoz hasonló utódokat hoznak létre a populációban.

Ezeket a problémákat küszöbölhetik ki a *nagylelkű ZD* stratégiák. Ezek a stratégiák a kölcsönös együttműködésért járó értékhez rögzítik a nyereményeket, méghozzá úgy, hogy a partnernek arányaiban nagyobb nyereményt biztosítanak saját maguknál. Valójában a veszteségből „követelnek” nagyobb részt maguknak. Elsőre ez öngyilkos stratégiának tűnik, hiszen az ilyen játékos matematikailag is feladja a lehetőségét annak, hogy megnyerjen egy összecsapást. Azonban végül is ugyanazon, önző, adaptív mechanizmussal, mint a zsaroló stratégiák esetében, a partner a saját nyereményének maximalizálása közben eljut a mindig együttműködő stratégiáig. Így, a „tartsd oda a másik orcádat is” mentalitású, ugyanakkor a másik játékos önzőségére is építő stratégia segítségével mindkét játékos a közösségileg elérhető legnagyobb nyeremenyvel gazdagodik. Adott esetben ugyan a zsaroló stratégiák nagyobb nyereményt érhetnek el a nagylelkű stratégiáknál, de általános, ismeretlen környezetben sokkal inkább kifizetődő a partnerek kölcsönös együttműködésre serkentése, mint a megtorlást is kockáztató kizsákmányolás.

Press és Dyson tavalyi felfedezése alapjaiban változtatta meg a fogolydilemmáról alkotott képünket, növelte az amúgy is komplex viselkedések körét. Szembesültünk a komplex rendszerek azon tulajdonságával, hogy bizonyos kérdések és jelenségek csak az elméleti, ill. matematikai keretek kiterjesztésével válnak tanulmányozhatóvá. A közeljövőben minden bizonnyal nagy figyelem fogja övezni az ilyen irányú további kutatásokat is. ❄

IRODALOM

- [1] Gruiz M., *Kaosz mint komplexitás*. Természet Világa jelen szám.
- [2] Axelrod R., *The evolution of cooperation*, Basic Books. New York, 1984.
- [3] Nowak M. A., *Sigmund K., Tit for tat in heterogeneous populations*. Nature 355, 250, 1992.
- [4] Press W. H., Dyson F. J., *Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent*. PNAS 109(26), 10409–13, 2012.
- [5] Szabó Gy., *Az együttműködés természete*. Magyar Tudomány 2012/6, 642–652, 2012.

MURAKÖZY BALÁZS

Teremtő rombolás közelnézetből

Az egyéni viselkedés és a gazdaság egészét jellemző mutatók közötti kapcsolat leírása sok nehézséget okoz a közgazdászoknak. Vannak olyan fogalmak, amik nem is léteznek egyéni szinten: ilyen például az infláció vagy a munkanélküliség. Vagy gondoljunk az elmúlt évek súlyos válságára: pontosan meg lehet érteni sok egyéni döntést, de a kialakult ingatlanbuborék vagy a kereslet visszaesése olyan jelenségek, amelyek csak a gazdaság szintjén léteznek. Ezek a jelenségek is rámutatnak arra, hogy a gazdaság egy komplex rendszer.

Az, hogy milyen nehéz a két szint összekapcsolása – vagyis milyen jelentősége van a komplexitáshoz kapcsolódó jelenségeknek –, nagyban függ attól, hogy kielégítő módon modellezhető-e az egész gazdaság egy hatalmas, de döntéseit tekintve „tipikus” egyén viselkedésével, vagy a gazdaság működésének megértéséhez egyenként kell modelleznünk a szereplők viselkedését és azok kölcsönhatását. Az elmúlt évtized kutatásai arra mutattak rá, hogy sok lényeges gazdasági jelenség és folyamat nem érthető meg egy egyszerűbb, „tipikus” egyént feltételező keretben.

Az említett probléma sokféle arcot ölt és sokféle megoldást javasoltak rá. Az egyik ilyen út az evolúciós közgazdaságtan [1]. Egy másik irány a sokszereplős modellezés [2]. Ebben a cikkben egy harmadik megközelítést emelek ki, amelyet a XX. század egyik legeredetibb közgazdásza, Joseph Schumpeter „teremtő rombolás”-nak nevezett. A teremtő rombolás az a folyamat, amelyben a hatékonyabb vállalatok gyorsabban növekednek, mint a kevésbé hatékony, például elmaradott technológiát alkalmazó vállalatok, így idővel nagyra nőnek, és kiszorítják kevésbé hatékony társaikat. Ez a vállalati szintű, a vállalatok sokféleségét hangsúlyozó folyamat az egész iparág hatékonyabbá válásához vezet, így nagy hatással van a gazdaság fejlődésére.

Schumpeter sok évtizede fogalmazta meg a teremtő rombolás alap gondolatát, de csak az elmúlt években jelentek meg olyan ada-

tok és modellek, amik segítenek mélyebben megérteni a teremtő rombolás működését és hatásait. Ezekben a kutatásokban az erőforrások vállalatok közötti „újraelosztását” reallokációnak nevezik (l. **1. blokk**).

Ebben a cikkben három példán keresztül mutatom be, hogy miért fontos a sokféle szereplő és a reallokáció vizsgálata a gazdaságban. Az első példa azt taglalja, hogyan kerültek át az erőforrások a hatékonyabb vállalatokhoz a rendszerváltást követően. A második példában azt vizsgáljuk, hogy mennyivel nőhetne meg India és Kína nemzeti jövedelme, ha nem akadályozná sokféle állami beavatkozás a hatékony vállalatok növekedését. Végül arról lesz szó, hogy milyen érvekkel szolgálnak a legújabb külkereskedelem-elméletek a vámk csökkentése mellett. Először azonban szólnom kell pár szót a versenyképesség legfontosabb mérőszámáról, a termelékenységéről.

A versenyképesség mérőszáma

A versenyképesség egyik alapvető mutatója a *termelékenység*. A termelékenység azt méri, hogy mennyi terméket képes előállítani adott mennyiségű erőforrás felhasználásával egy vállalat, egy iparág vagy egy ország. Mivel egyszerre többféle erőforrást használnak fel a termelésnél, többféleképpen lehet mérni a termelékenységet, de ezek közül a legfontosabb a munkatermelékenység, ami azt mutatja meg, hogy egy munkás átlagosan mennyi terméket¹ állít elő.²

A vállalati termelékenység szoros kapcsolatban áll azzal, hogy milyen költséggel tud termelni egy vállalat, ezért a termelékenyebb vállalatok előnyt élveznek versenytársaikkal szemben. A magasabb termelékenységű országok szintén előnyt élveznek a nemzetközi versenyben, és az egy főre jutó nemzeti jövedelmet is elsősorban a termelékenység határozza meg (**1. ábra**).

Az iparági termelékenység alakulása nagyon jól mutatja egy-egy iparág időbeli fejlődését. A **2. ábra** néhány iparág termelékenységének változását mutatja a rendszerváltást követő másfél évtizedben.³ A legfontosabb azt látunk, hogy mennyivel termelékenyebbé vált a gazdaság: a legtöbb iparágban 2–3-szorosára nőtt a dolgozók termelékenysége, ami hasonló mértékű növekedést hozott a nemzeti jövedelemben is. Fontos azonban a különböző iparágak eltérő teljesítménye is: miközben a textilipar termelékenysége nagyjából kétszeresére, a gépiparé ötszörösére nőtt. A következő fejezetben részletesebben is összehasonlítjuk ezt a két iparágat.

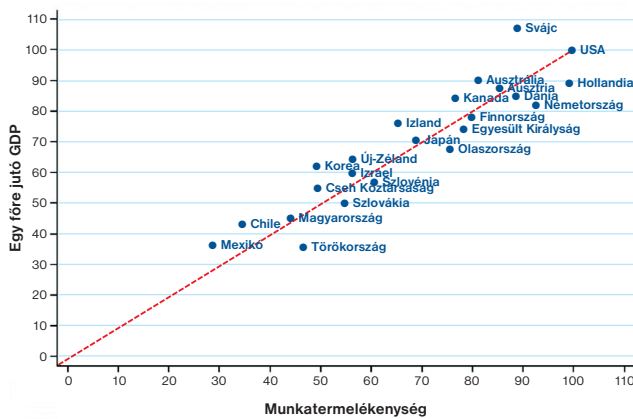
1. Reallokáció

A reallokáció mögötti fő gondolatot egy igen egyszerű példán keresztül is megérthetjük. Tegyük fel, hogy horgásztársaságként minket – valamiért – leginkább a tavunkban található halak átlagos testsúlya érdekel. Sajnálatos módon megjelenik a tavunkban egy aligátor, amelyik annál nagyobb eséllyel eszi meg a halakat, minél kövérebbek. Hogyan változik a halak átlagos testsúlya? Nos, ha egy „tipikus” hal viselkedésével modelleznénk a tó összes halát, akkor elkezdhetnénk vizsgálni, hogy diétába kezdenek-e az életüket féltő halak. Ezt persze nem lehet kizárni, de ennél sokkal lényegesebb az, hogy a kövérebb halakat megeszi a hulló, így megnő a kis halak aránya mind a tóban, mind az átlagos súly képletében. Közgazdasági értelemben az erőforrások „átcsoportosulására”, reallokációjára kerül sor a kisebb halak javára, ami alapvető hatással van az átlagos súly mutatójára.

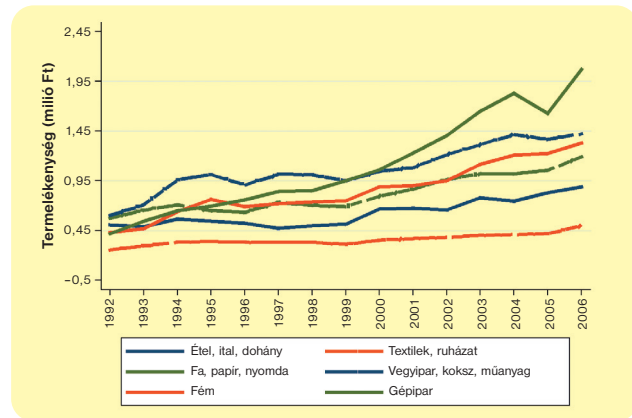
¹ Mivel a legtöbb vállalat vásárol alapanyagokat is, ezért a munkatermelékenység pontosan azt mutatja, hogy mekkora hozzáadott értéket állít elő (forintban mérve) egy-egy dolgozó egy év alatt.

² A termelékenység más mértékei figyelembe veszik azt is, hogy mennyi más erőforrást, például tőkét, használ fel a vállalat termelése során, vagy számol a dolgozók képzettségével is. A továbbiakban viszont az egyszerűség kedvéért eltekintünk a többi termelési tényezőtől, és a munkatermelékenységre koncentrálunk. A termelékenység másik lényeges problémája az, hogy a különböző vállalatok termékei eltérő minőségűek lehetnek, amelynek kezelésére különböző – bár nem tökéletes – módszerek léteznek.

³ Mértékegysége millió Ft hozzáadott érték/fő, az inflációval korrigálva.



1. ábra. Termelékenység (egy munkaóra jutó GDP) és az egy főre jutó GDP közötti kapcsolat 2011-ben (USA=100) (Forrás: OECD)



2. ábra. Különböző iparágak munkatermelékenységének alakulása Magyarországon (Forrás: [3])

Miért lettek termelékenyebbek a vállalatok a rendszerváltás után?

A reallokáció fogalmát szemléltető példa jól mutatta, hogy az elérő túlélési esélyek nagy hatással lehetnek az egyedeket jellemző különböző mutatók átlagára. Ez nemcsak az aligátorral teli tóban, hanem a gazdaságban is így van, itt azonban az átlagos termelékenység változását nem csupán az befolyásolja, hogy melyik vállalatok maradnak életben, hanem az is, hogy gyorsabban növekednek-e a termelékeny vállalatok.

Egy iparág termelékenysége – definíciónk szerint – az iparág termelésének és az iparágban dolgozók számának a hányadosa, ami egyben az egyes vállalatok termelékenységének a dolgozók számával súlyozott átlaga. A súlyozott átlagnak érdekes tulajdonsága (1. 2. blokk), hogy felírható két tag összegeként a következőképpen:

$$\text{iparági termelékenység} = (\text{súlyozatlan}) \text{ átlagos termelékenység} + \text{kovariancia a termelékenység és méret között}$$

A súlyozatlan átlag szerepe világos: minél nagyobb a vállalatok termelékenységének (súlyozatlan) átlaga, annál nagyobb lesz a súlyozott átlag is. Ezt viszont korrigálnunk kell a második – kovarianciának nevezett – taggal. A termelékenység és a méret közötti kovariancia azt mutatja meg, hogy nagyobbak-e a termelékenyebb vállalatok, vagyis együtt mozog-e ez a két változó. Ha a termelékenyebb vállalatok egyben nagyvállalatok is, akkor a kovariancia pozitív, és így a súlyozott átlag – vagyis az iparági termelékenység – nagyobb a vállalati termelékenység súlyozatlan átlagánál. Ha viszont általában kisebbek a termelékenyebb vállalatok, akkor negatív a kovariancia, és ezért lefelé kell korrigálnunk a súlyozatlan átlagot.

Ezt a gondolatot illusztrálja a 3. ábra. A grafikonokon látható, hogy a kovariancia értékének növekedésével egyre szorosabbá válik a kapcsolat a méret és a termelékenység között (egyre közelebb kerül a pontfelhő a 45 fokos egyeneshez), és ezzel párhuzamosan egyre nagyobb mértékben tér el a súlyozott és súlyozatlan átlag egymástól.⁴

2. A súlyozott átlag felbontása

A súlyozott átlagot a következőképpen írhatjuk fel:

$$\text{iparági termelékenység} = \sum_i s_i \text{ termelékenység}_i$$

ahol i a vállalatokat jelöli, s_i az i vállalat súlya – az általa foglalkoztatott dolgozók aránya az iparági dolgozók számához viszonyítva – és termelékenység_i az i vállalat termelékenysége.

Ezt könnyen beláthatjuk:

$$\begin{aligned} \text{iparági termelékenység} &= \frac{\sum_i \text{termelés}_i}{\sum_i \text{dolgozók}_i} = \\ &= \frac{\sum_i \text{dolgozók}_i \frac{\text{termelés}_i}{\text{dolgozók}_i}}{\sum_i \text{dolgozók}_i} = \sum_i \frac{\text{dolgozók}_i}{\sum_i \text{dolgozók}_i} \text{termelékenység}_i \end{aligned}$$

A súlyozott átlag érdekes tulajdonsága, hogy nagyon szemléletes módon bontható fel két részre. Jelöljük \bar{s} -al a vállalatok súlyának (súlyozatlan) átlagát, és termelékenység -el a vállalati termelékenység súlyozatlan átlagát. Ekkor az iparági termelékenységet felírhatjuk a következőképpen:

$$\begin{aligned} \text{iparági termelékenység} &= \overline{\text{termelékenység}} + \\ &+ \sum_i (s_i - \bar{s}) (\text{termelékenység}_i - \overline{\text{termelékenység}}) \end{aligned}$$

Az összeg első tagja egyszerűen a termelékenység súlyozatlan átlaga. A második tag egy összeg – ezt nevezzük *kovarianciának* –, amely a méret és a termelékenység együttmozgását írja le. Ha minden vállalat mérete egyenlő az átlagossal, akkor $s_i - \bar{s} = 0$ minden vállalat esetében, tehát a kovariancia egyenlő nullával. Ilyenkor a súlyozott és a súlyozatlan átlag egyenlő egymással.

Nézzük, mi történik, ha a termelékenyebb vállalatok nagyobbak, mint a kevésbé termelékenyek! Tegyük fel például, hogy iparágunk két vállalatból áll: az egyik 40 embert foglalkoztat, és 10 a termelékenysége, a másik 60 embert és 20 a termelékenysége. A termelékenység súlyozatlan átlaga 15. A kovariancia két pozitív szám összege: $(-0,1) \cdot (-5) + 0,1 \cdot 5 = 1$. A termelékenység súlyozott átlaga tehát $15 + 1 = 16$. (Ellenőrzésképpen kiszámolhatjuk a súlyozott átlag eredeti képletével is: $0,4 \times 10 + 0,6 \times 20 = 16$) A két változó közötti pozitív összefüggés pozitív kovarianciát jelent, ezért az iparági termelékenység nagyobb, mint a vállalati termelékenység súlyozatlan átlaga. Ha ennél is erősebb a méret és termelékenység közötti kapcsolat (például 20–80 a dolgozók aránya), akkor még nagyobb a kovariancia, és még nagyobb a súlyozott és súlyozatlan termelékenység közötti különbség.

Teremtő rombolás közlézetből

A termelékenység és a méret összefüggésének hatását felelősen, hogy a vállalatok méretének és termelékenységének eloszlása – a hálózatokhoz és más komplex rendszerekhez hasonlóan – nem felel meg a haranggörbének, hanem „váratlanul” nagy méretű vállalatokat is találhatunk a különféle iparágakban.⁵ Ez a néhány vállalat nagymértékben befolyásolja a vállalati termelékenység súlyozott átlagát, és ezen keresztül egy-egy iparág vagy ország versenyképességét is.

A termelékenység súlyozatlan átlagát és a méret és termelékenység közötti kovarianciát eltérő tényezők határozzák meg. Míg a súlyozatlan átlag talán a gazdaságban rendelkezésre álló tudástól függ leginkább, a kovariancia azt tükrözi, hogy több erőforráshoz jutnak-e a termelékenyebb vállalatok, vagy akadályozza-e növekedésüket valamilyen, például adózási vagy adminisztratív tényező.

Gondoljunk például a szocialista és a kapitalista rendszerek közötti különbségekre. Ha a szocialista vállalatok általában kevésbé hatékonyan működtek, az azt jelenti, hogy a termelékenységük súlyozatlan átlaga kisebb volt. Másik oldalról a szocialista rendszernek az is lényeges tulajdonsága volt, hogy a vállalatok számára – például új gépek vásárlására – rendelkezésre álló forrásokat nem a termelékenységgel arányosan, hanem politikai szempontok alapján osztotta el. Sőt, a szocialista rendszer gyakran elvonta a forrásokat a kisebb, nyereségesebb vállalatoktól, és odaadta őket a hatalmas, veszteséges vállalatoknak [4].

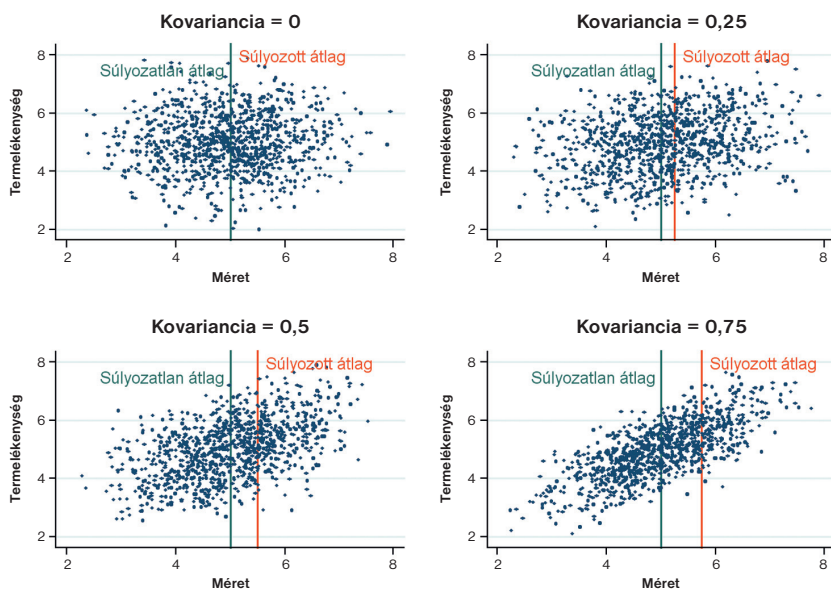
Egy kapitalista gazdaságban alapvetően más a helyzet. A teremtő rombolás modelljének megfelelően a piaci versenyben azok a vállalatok kerülnek előnybe, amelyeknek magasabb a termelékenysége.⁶ Mivel ezek a vállalatok több nyereséghez jutnak és a bankoktól is könnyebben kapnak hitelt, könnyebben tudnak terjeszkedni is. Ha ez így van, akkor a rendszerváltást követően erősödhetett a vállalatok mérete és a termelékenysége közötti kapcsolat, vagyis megnövekedett a két változó közötti kovariancia.

Ezt a hipotézist vizsgálja a 4. ábra, amely a magyar textilipar és a gépipar termelékenységét bontja fel minden évben 1992 és 2006 között [3]. A gépipar esetében hipotézisünk látványosan beigazolóódik. 1992-ben a – szocializmusból megmaradt – nagyobb vállalatok átlagosan kevésbé voltak termelékenyek, mint a kisebb cégek. A kapitalista körülmények között viszont gyorsan nőtt a két változó közötti kovariancia, olyannyira, hogy 2006-ban már több mint egyharmadát ez adta az iparági termelékenységnek. A textiliparban ezzel szemben a legtöbb vállalat kicsi maradt, és így a

⁴ Az ábrán egy kétváltozós normális eloszlásból szimuláltam a pontokat, ahol mindkét változó várható értéke 5 és szórása 1. Így a kovariancia megegyezik a korrelációval is.

⁵ Más szóval a vállalati méret és termelékenység eloszlás jobb szélén jobban írható le egy hatványfüggvénnyel, mint a normális eloszlással.

⁶ Lényeges aspektusa a teremtő rombolásnak az is, hogy a kapitalista vállalatok innovatívabban és sokkal több új vállalkozás indul. Eppen ezek a folyamatok biztosítják azt, hogy egyáltalán megjelenjenek a magas termelékenységű vállalatok.



3. ábra. A súlyozott és súlyozatlan átlag alakulása a kovariancia különböző értékeire

tényezők akadályozzák vagy segítik a hatékonyabb vállalatok gyorsabb növekedését. Az ilyen kérdésekre csak elméleti modellre építve tudunk válaszolni.

Ebből a szempontból nagyon fontos munka Hsieh és Klenow [5] tanulmánya, amely Kínát és Indiát vizsgálja. A kiindulópont az, hogy ideális – gazdaságpolitikai torzítások nélküli – esetben minden (amúgy hasonló) vállalat ugyanakkora kamatot fizetne tőkéje után, és az összes cég nyereségadója hasonló volna. Ilyenkor a termelékenyebb vállalatok alacsonyabb költség szintjüknek köszönhetően több terméket adhatnak el, ezért érdemes több dolgozót és tőkét alkalmazniuk. Az ilyen vállalatok tehát, ideális esetben, nagyobbra nőhetnek, ami a korábban elmondottak alapján magas iparági termelékenységhez vezet.

A valóságban azonban Kínában vagy Indiában nem teljesülnek ezek a feltételek, mert bizonyos – sokszor nem túl termelékeny – vállalatok különféle kedvezményben részesülnek. Az egyik ilyen kedvezmény az, hogy egyes vállalatok – mondjuk politikai kapcsolataiknak köszönhetően – alacsonyabb költséggel jutnak tőkéhez, mint mások. A második típusú kedvezményt az jelenti, hogy egyes vállalatok kevésbé tudnak terjeszkedni, mint amennyire

kovariancia nem járult hozzá jelentős mértékben az iparági termelékenységhez.

Mennyivel lehetne gazdagabb India, Kína vagy Japán?

A termelékenység egyszerű felbontása jól illusztrálja, hogy milyen fontos szerepet játszik a teremtő rombolás a gazdasági növekedésben. Ez az elemzési mód azonban meglehetősen elnagyolt, nem vizsgálja az egyes vállalatok viselkedését és, ami még fontosabb, nem tudja megmutatni, hogy milyen

4. ábra. A textilipar és a gépipar termelésének felbontása Magyarországon, 1992–2006 (Forrás: [3])



termelékenyséjük alapján optimális volna nekik. Ennek egyik oka lehet, hogy az állam jobban megadóztatja a nagyobb vállalatokat, mint a kisebbeket. Az ilyen típusú állami beavatkozásra nagyon jó példa Brazília, ahol a kisebb cégek alapvetően az informális gazdaságban működnek, így nem fizetnek adót. Ez azt jelenti, hogy a nagyobb cégeknek magas adókat kell fizetniük, hogy az állam fedezni tudja a költségvetési kiadásokat [6]. Ezeket – a vállalatok bizonyos csoportjainak adott kedvezményeket – egyfajta gazdaságpolitikai torzításnak nevezhetjük (l. **3. blokk**), hiszen egyenlőtlen versenyfeltételekhez vezetnek.

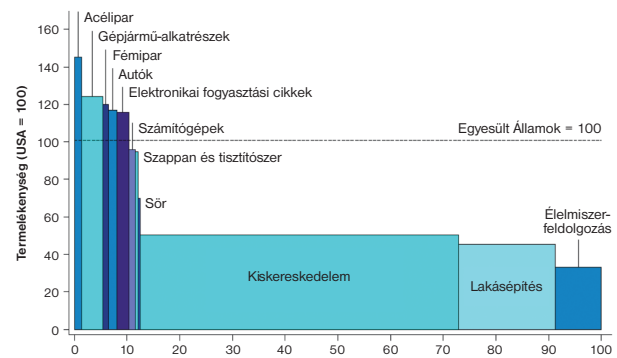
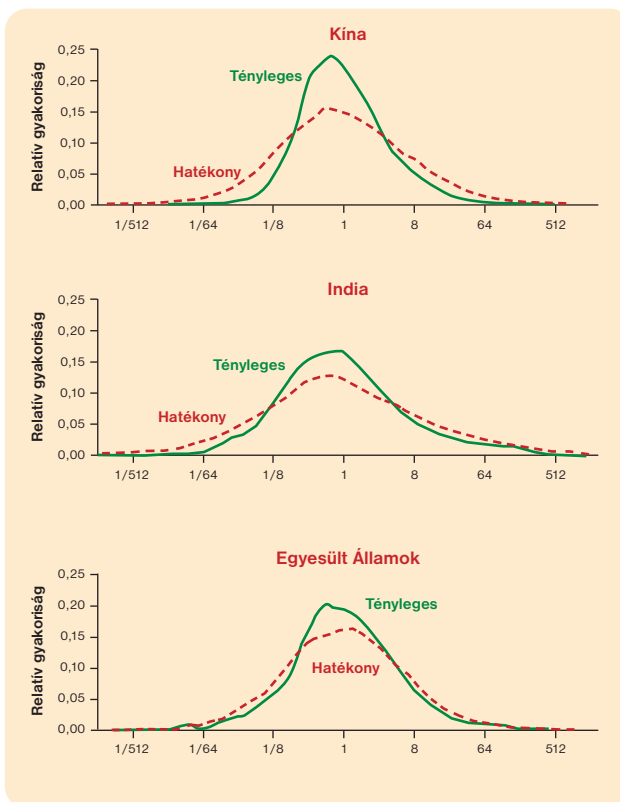
Hsieh és Klenow modelljük segítségével megmutatják, hogy mindkét típusú torzítás igen jelentős Indiában és Kínában is: mind-

3. A torzítások mérésének módszeréről

Az első típusú torzítást úgy értelmezhetjük, hogy a tőke költsége (a kamat) eltérő a vállalatok között, a másikat pedig úgy, hogy ugyanakkora mennyiségű terméket különböző áron tudnak eladni a különböző vállalatok.

Hsieh és Klenow modellje [5] – a közgazdaságtanban elfogadott feltevést használva – abból indul ki, hogy a vállalatok nyereségüket maximalizálják. Megvizsgálják, hogyan viselkednének a vállalatok torzítások nélküli esetben, és ezt összehasonlítják a tényleges viselkedésükkel. Ennek alapján megmutatják, hogy a kamatban megfigyelhető különbségek mértékét jól méri az, hogy mennyire eltérő az egy dolgozóra jutó tőke aránya a vállalatok között⁷, az adózás eltéréseit pedig azzal mérik, hogy ugyanannyi dolgozót foglalkoztatnak-e a hasonló bevételt termelő vállalatok.⁸

5. ábra. A termelékenység hatékony és tényleges eloszlása Kínában, Indiában és az Egyesült Államokban (Forrás: [5])



6. ábra. Különböző iparágak termelékenysége Japánban, 1999-ben (Forrás: [6])

két országban sokkal kevesebb kis- és nagyvállalat van, mint az Egyesült Államokban vagy a modell szerinti optimális gazdaságban (**5. ábra**). Ennek alapján úgy tűnik, hogy az állam által létrehozott torzítások támogatják a kisvállalatok növekedését egy bizonyos szintig, utána viszont akadályozzák azt, hogy igazán nagy és nemzetközi szinten is versenyképes vállalatok jöjjenek létre. Ennek következtében minden vállalat mérete hasonlóvá válik, és a méret és termelékenység közötti alacsony kovariancia alacsony iparági termelékenységhez vezet.

A modellépítés azért is fontos, mert lehetővé teszi a torzítások nemzetgazdasági hatásainak kiszámítását. A szerzők számításai szerint, ha a torzítások arra a szintre csökkennének, ami az Egyesült Államokban figyelhető meg, akkor (a 2005-ös adatok alapján) Kína termelékenysége 30, Indiáé pedig 59 százalékkal lenne magasabb. A szerzők bemutatják azt is, hogy a torzítások csökkentésével együtt járó reallokáció jelentősen, évi 2 százalékponttal járult hozzá Kína termelékenységének növekedéséhez 1998 és 2005 között, elsősorban az állami vállalatok privatizációjának és bezárásának köszönhetően. Eközben Indiában újabb korlátozásokat vezettek be, magasabb terhekkel sújtva a nagyobb vállalatokat, így egyenesen csökkent a méret és termelékenység közötti kapcsolat erőssége. Ez talán az egyik oka annak, hogy Kína ebben az időszakban jelentősen gyorsabban növekedett, mint India.

Nem szabad azonban azt gondolnunk, hogy nagy méretű torzítások csak szocialista vagy fejlődő gazdaságokban vannak jelen. Lewis [6] könyvében számos ország esetében megvizsgálja, hogy milyen mértékben marad el az Egyesült Államoktól a különböző iparágak termelékenysége. Rámutat arra is, hogy ezek a különbségek nagyon jól megmagyarázhatók olyan tényezőkkel, amelyek akadályozzák a teremtő rombolást.

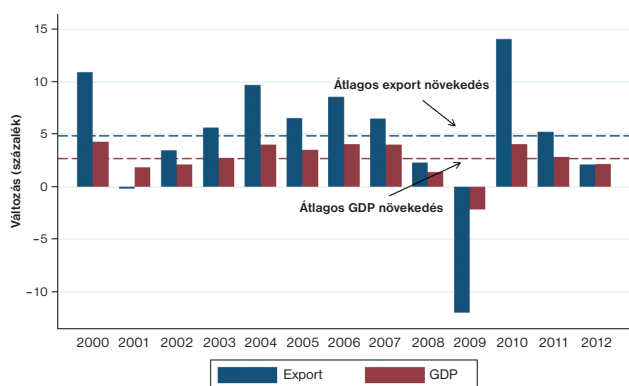
A **6. ábra** különböző iparágak termelékenységét mutatja Japánban. Az oszlopok magassága a termelékenységet (az Egyesült Államokhoz képest), szélessége pedig az iparágban dolgozók számát mutatja. Az ábra alapján elmondhatjuk, hogy Japánban duális gazdaság működik. Az exporttal versenyző, de egyébként viszonylag kevés embert foglalkoztató szektorok termelékenysége magasabb, mint az Egyesült Államokban. A kiskereskedelemben, az építőiparban és az élelmiszer-feldolgozásban viszont rengeteg szabály és támogatás akadályozza a hatékony vállalatok növekedését, és segíti az alacsony hatékonysággal működő kis cégek túl-

⁷ A modellben a vállalatok bizonyos mértékben helyettesíteni tudják egymással a tőkét és a munkát, ezért ha bizonyos vállalatoknak magasabb kamatot kell fizetniük, akkor ők kevesebb gépet vásárolnak egy-egy dolgozó számára.

⁸ Ha nem lenne torzítás, akkor azonos termelékenységű vállalatoknak azonos mennyiségű terméket kellene eladni. Ha viszont az egyik vállalatnak magasabb adót kell fizetnie, akkor kisebb marad. A szerzők modellje lehetővé teszi az eltérő méretű vállalatok összehasonlítását is.

⁹ Ez természetesen egyben azt is tükrözi, hogy ezeknek az iparágaknak a termékeivel nehéz nemzetközi kereskedelmet folytatni, így a torzító gazdaságpolitika hosszabb ideig maradhat fenn.

Teremtő rombolás közelnézetből



7. ábra. A reál GDP és külkereskedelmi növekedése 2000 és 2012 között Japánban (Forrás: külkereskedelmi volumenindex – WTO, reál GDP – Világbank)

élését.⁹ Ezekben a szektorokban a termelékenység átlaga és a vállalatmérettel mutatott kovarianciája is nagyon alacsony. A szektorok súlya egyben azt is jelenti, hogy ezek a korlátozások nagymértékben csökkentik a gazdaság termelékenységét és az egy főre jutó nemzeti jövedelmet is.

A globalizáció nyertesei és vesztesei

Japánban – más országokhoz hasonlóan – lényeges különbség rajzolódik ki tehát a nemzetközi kereskedelemben aktívan résztvevő iparágak és más szektorok között. A külkereskedelem kutatása során az egyik fontos feladat annak megértése, hogy milyen csatornán keresztül fejti ki ezt a hatást a nemzetközi verseny. A kérdést különösen lényegessé teszi az a tény, hogy az elmúlt évtizedekben – még a súlyos gazdasági válság ellenére is – dinamikus és a nemzeti jövedelemnél jelentősen gyorsabban nőtt a külkereskedelmi volumene, ahogy ezt a 7. ábra is mutatja.

A külkereskedelmi elmélet klasszikus eredményei az iparágak közötti reallokációt vizsgálják, így azt feltételezik, hogy az iparágakon belül ugyanolyanok a vállalatok. Ez a megközelítés sok fontos eredményt adott, és lényeges szerepet játszott annak bemutatásában, hogy milyen előnyökkel járhat a vámok csökkentése: az országok olyan termékek előállítására specializálódhatnak, amelyek termelésében előnyük van. Ezen modellek segítségével levonható fontos következtetés, hogy ez az országok közötti specializáció minden ország számára előnyös lehet. Ha az Egyesült Államok számítógépet tud hatékonyabban gyártani, Kína pedig cipőt, akkor mindkét ország jobban jár, ha mindegyik csak azt gyártja, amiben (relatív) hatékonyabb, mert így összesen több cipőt és számítógépet gyárthatnak.

Japán példája azonban azt mutatja, hogy a nemzetközi versenyben részt vevő iparágak általában termelékenyebbek. A részletesebb elemzés arra is rámutatott, hogy nem csupán arról van szó, hogy néhány iparág, amiben Japánnak eleve előnye van, egyben exportál is, hanem maga az exportálás is jelentősen növeli az iparágak termelékenységét. Erre a jelenségre a leginkább elegáns magyarázatot Melitz [7] modellje adja, melyben megmutatható, hogy a vámok csökkentése az iparági termelékenységet növelő reallokációhoz vezet. Ez a modell számos további kutatást indított be, amelyekről egy friss összefoglalót ad [8].

Melitz modelljének kulcseleme, hogy egy iparágon belül a vállalatok termelékenységét különbözőnek tételezi fel. A 8. ábra fel-

⁹ Az egyszerűség kedvéért feltesszük, hogy a vállalatok termelékenység szerinti eloszlása egyenletes. A valóságban az eloszlás inkább Pareto-típusú, így sokkal több alacsony termelékenységű vállalat van.

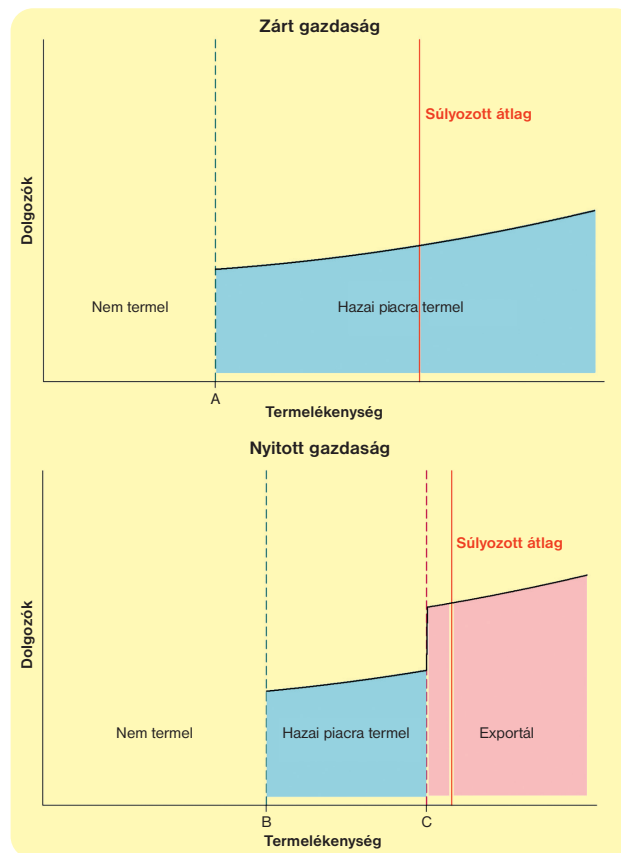
ső része egy zárt, vagyis más országokkal nem kereskedő gazdaságban működő vállalatok jellemzőit mutatja: a vízszintes tengelyen a termelékenység látszik, a függőlegesen pedig az, hogy hány embert foglalkoztatnak a különböző termelékenységű cégek. Az A pontnál alacsonyabb termelékenységű vállalatok nem tudnak nyereségesen működni, így nem is dolgozik náluk senki. Az A-nál termelékenyebb vállalatok már termelnek, és minél termelékenyebbek, annál versenyképesebbek, és annál nagyobbra nőnek. A függőleges vonal mutatja a vállalatok átlagos súlyozott termelékenységét.¹⁰

Melitz modelljének kulcskérdése az, hogy mi történik a vámok csökkentése után. A modell feltevése az, hogy viszonylag nagy költséggel jár az exportpiacra történő belépés, hiszen például az idegen nyelven beszélő dolgozók felvétele, a külföldi üzletfelek megkeresése és a reklámkampány megszervezése költséges. Ezt a feltevést az is alátámasztja, hogy minden országban viszonylag kevés vállalat exportál: ha nem járna nagy egyszeri költséggel a külföldi piacra lépés, minden vállalatnak megérné megtenni.

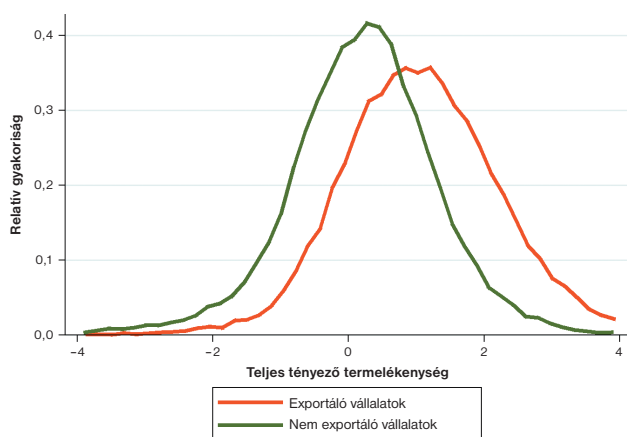
A 8. ábra alsó része mutatja azt, hogy mi történik a globalizáció során. Először is a leginkább termelékeny – a C ponttól jobbra lévő – vállalatok elkezdnek exportálni. Mivel ezek a vállalatok nem csak a hazai, hanem az exportpiacra is termelnek, újabb dolgozókat vesznek fel. Másodsor, a közepesen termelékeny vállalatok ezután is csak a hazai piacra termelnek. Harmadszor, a B ponttól jobbra lévő vállalatok tönkre mennek. Ennek két oka van. Egyrészt a külföldi vállalatok is megjelennek a hazai piacon, így erősödik a verseny. Másrészt, az hogy az exportáló vállalatok több dolgozót vesznek fel, ami béremelkedéssel jár, és megnöveli ezeknek a vállalatoknak a költségeit.

Az ábra jól mutatja azt is, hogy mi történik az iparág termelékenységével. A legkevésbé termelékeny vállalatok csődjé növeli a

8. ábra. Vállalatok és termelékenység a Melitz-modellben



Muraközy Balázs: Teremtő rombolás közelnézetből



9. ábra. Az exportáló és a nem exportáló vállalatok termelékenységének eloszlása Magyarországon, a teljes tényező termelékenység függvényében, 1992–2003 (Forrás: [11])

súlyozatlan átlagot, a leginkább termelékeny vállalatok növekedése pedig növeli a méret és a termelékenység közötti kovarianciát. Mivel a súlyozott termelékenységet mutató összeg mindkét tagja növekszik, az iparág termelékenysége is növekszik.

A Melitz-modell nagy újdonsága tehát az, hogy rámutat: a külkereskedelmi nyitás nemcsak iparágak közötti, hanem iparágakon belüli átrendeződéssel, reallokációval is jár. A szabadabb külkereskedelem a leginkább termelékeny vállalatok növekedésével és a legkevésbé termelékeny cégek kilépésével jár. Az átlagos termelékenység növekedése alacsonyabb árakhoz és magasabb bérekhez vezet, ami a külkereskedelemből származó haszon újabb formája.

A Melitz-modell empirikus szempontból is jól működik. Egyrészt számos tanulmány bizonyította, hogy az elmélet által előre jelzett „önszelekció” – vagyis az, hogy csak viszonylag kevés, nagyon termelékeny vállalat exportál – igaz a különböző országokban. (Az Egyesült Államokról [9], az európai országokról [10], a magyar vállalatokról pedig [11] ad áttekintést.) Ez a jelenség látható a 9. ábrán is, ami a nem exportáló és az exportáló vállalatok termelékenységének eloszlását mutatja Magyarországon.

Másrészt a vámok csökkentése is az előre jelzett hatással jár. Treffler [12] azt vizsgálta meg, hogy milyen módon befolyásolta az Egyesült Államok, Kanada és Mexikó közötti szabadkereske-

delmi megállapodás (a NAFTA) a kanadai iparágak termelékenységét. Az eredmények azt mutatják, hogy azokban a szektorokban, ahol leginkább csökkent a kanadai vállalatok által fizetett vám, átlagosan 14 százalékkal nőtt meg a termelékenység. Treffler ezután bemutatta, hogy az egyes vállalatok termelékenysége nem változott lényegesen. A vámcsökkentés hatása tehát teljes egészében a reallokáción keresztül érvényesült, vagyis azáltal, hogy ezekben az iparágokban a leginkább termelékeny vállalatok bővítették tevékenységüket, miközben a legkevésbé termelékeny vállalatok csökkentették foglalkoztatottaik számát vagy akár be is zártak.

A sokféle vállalatot feltételező külkereskedelmi modellek – a termelékenység-növekedés mellett – arra is rámutatnak, hogy mennyire eltérően élhetik meg a külkereskedelmi nyitást az egyes vállalatok és dolgozók. A reallokáció egyben azt is jelenti, hogy sokan elvesztik a munkájukat, és nehezen találnak másikat. A 10. ábra bemutatja, hogy milyen szoros kapcsolat volt a vámok csökkentése és a munkahelyek megszűnése között iparági szinten: minél jobban csökkent a kanadai vám, annál több állás szűnt meg.

Következtetések

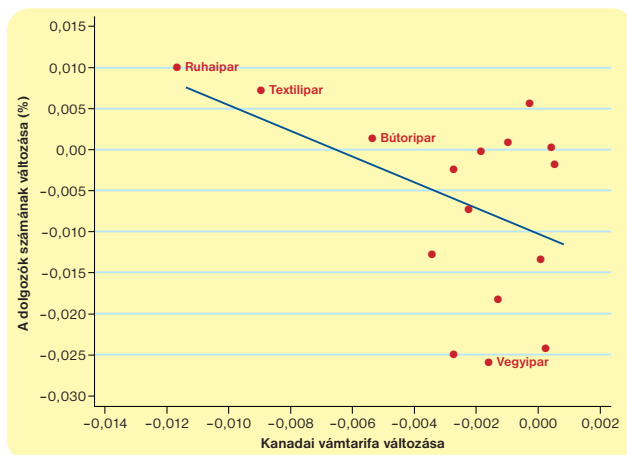
A cikkben taglalt példák mindegyike egy-egy olyan esetet mutat be, amelyben az egyéni döntéshozók – esetünkben vállalatok – heterogén viselkedésének fontos szerepe van az iparág vagy a teljes gazdaság teljesítménye szempontjából. Ezekben a példákban fontos szerepet játszik az iparágakon belüli reallokáció, az erőforrások vállalatok közötti átrendeződése.

Megmutattuk, hogy a szocialista rendszerben vagy az ázsiai gazdaságokban megfigyelhető torzítások ahhoz vezetnek, hogy a hatékonyabb vállalatok gyakran nem tudják elérni azt a méretet, amit egy kisebb súrlódásokkal működő gazdaságban elérnének. Azt is láttuk, hogy a torzítások hatása jelentős: Kína elmúlt évtizedbeli növekedéséből például évi két százalékpont magyarázható az erőforrások átrendeződésével.

Az elmúlt évtized külkereskedelem elméletei ugyanennek a gondolatnak egy másik alkalmazását jelentik. A külkereskedelmi nyitásra a különböző vállalatok eltérően reagálnak: a leginkább termelékenyek bővítik tevékenységüket, míg a magasabb béreknek köszönhetően a kevésbé termelékeny vállalatok iparágon belüli súlya csökken. Ez az átrendeződés nagymértékben növelheti a termelékenységet, de felhívja arra is a figyelmet, hogy a zsugorodó vállalatok vagy iparágak dolgozóit nagyon hátrányosan érintheti a globalizáció. ❄

A tanulmány az OTKA K-81947 sz. kutatás támogatásával készült. Köszönöm Holler Zsuzsa kitudó asszisztensi munkáját.

10. ábra. A kanadai vám csökkenése és a munkahelyek megszűnése közötti kapcsolat (Forrás: [12])



IRODALOM

- [1] R. Nelson, S. G. Winter: *Evolutionary Theorizing in Economics*. The Journal of Economic Perspectives, 16(2), 23–46, 2002.
- [2] J. D. Farmer, D. Foley: *The economy needs agent-based modelling*. Nature 460, 685–686, 2009.
- [3] Békés G., Halpern L., Muraközy B.: *A teremtő rombolás szerepe a vállalati termelékenység alakulásában Magyarországon*. Közgazdasági Szemle, LVIII. évf., 111–132, 2011.
- [4] Kornai, J.: *A szocialista rendszer: Kritikai politikai gazdaságtan*. HVG Kiadó, Budapest, 1993.
- [5] C. Hsieh, P. J. Klenow: *Misallocation and Manufacturing TFP in China and India*. The Quarterly Journal of Economics, 124(4), 1403–1448, 2009.
- [6] W. W. Lewis: *A termelékenység ereje*. GVH, 2008.
- [7] M. J. Melitz: *The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity*. Econometrica, 71, 1695–1725, 2003.
- [8] M. J. Melitz, R. J. Redding: *Heterogeneous firms and trade*. NBER Working Paper 18652, <http://www.princeton.edu/~reddings/papers/NBERw18652.pdf>, 2012.
- [9] A. B. Bernard, J. B. Jensen, S. J. Redding, P. K. Schott: *Firms in International Trade*. Journal of Economic Perspectives, 21(3), 105–130, 2007.
- [10] T. Mayer, J. Ottaviano: *The Happy Few: The Internationalisation of European Firms*. Intereconomics, 43(3), 135–148, 2008.
- [11] Békés G., Harasztosi P., Muraközy B.: *Firms and products in international trade: Evidence from Hungary*. Economic Systems, 35(1), 4–24, 2012.
- [12] D. Treffler: *The Long and Short of the Canada-U. S. Free Trade Agreement*. The American Economic Review, 94(4), 870–895, 2004.

GULYÁS ATTILA

Az agykérgi hálózatok szerkezeti és működési komplexitása

Az agyműködést az EEG-mintázatok komplex összjátéka jellemzi. Az elmúlt 10 évben a számítástechnika és a mérési rendszerek teljesítményének fejlődése lehetővé tette, hogy az agy aktivitását akár több száz csatornán keresztül vizsgáljuk egyszerre [1–3]. Ez megengedi, sőt nélkülözhetetlenné teszi, hogy az idegrendszer aktivitását egyre bonyolultabb statisztikai és matematikai módszerekkel vizsgáljuk. Az agykutatás most ért oda, ahova a fizikusok jó néhány évtizeddel ezelőtt a nagy gyorsítók, például a CERN felépítésekor eljutottak. A mérésekben keletkező hatalmas mennyiségű adatot számítógépek elemzik hetekig, illetve az adatokat és az elméleteket modellek segítségével illesztik.

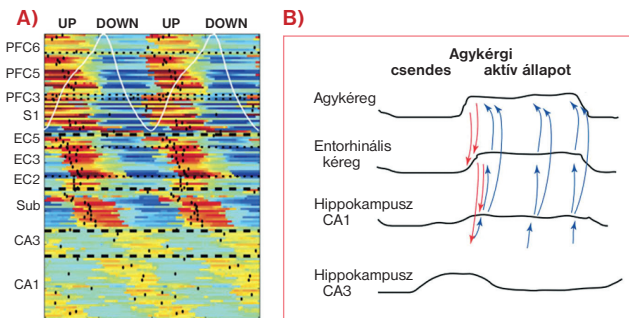
Ennek eredményeként az idegsejtek vagy agyterületek közötti kölcsönhatások magasabb szintjét érthetjük meg. A szerveződési szinteken felülről lefelé (top-down, a viselkedés felől az idegi folyamatokig) hatoló pszichológusok, neurológusok és megismerésemelvések, valamint a szerveződési szinteken alulról felfelé (bottom-up) haladó molekuláris biológusok, neuro-anatómusok és neuro-fiziológusok megismerése között egyre vékonyabb lesz a két alagutat elválasztó fal. A megismerés valahol az idegsejti makrohálózatok, a finom felbontású fMRI szintjén fog összeérni, de ehhez meg kell érteni a rendszer komplexitását.

Sokcsatornás EEG-vizsgálatok eredményeként egyre részletesebben kirajzolódik, hogy az agy különböző területeinek aktivitása különböző magatartási állapotokban eltérő módokon csatlózik össze, és az egyes területek a feladat függvényében különböző EEG-mintázatokot generálnak [1]. Ez annak a megnyilvánulása, hogy egy idegsejtekből álló hálózatnak számtalan feladatot kell megoldania: információt kódolni, tárolni, előhívni, feldolgozni. Mindezt egy időben egyetlen hálózat sem képes optimálisan végrehajtani, ezért (a híradástechnikából kölcsönzött szót használva) multiplexeli, azaz időben gyorsan váltogatva egymás után végzi el a különböző típusú feladatokat. Az egyes feladatokra jellemző aktivitásokat jelző elektromos jeleket, különböző dinamikájú (ritmusú) EEG-mintázatokként mérjük (1. ábra).

Kutatócsoportunkat elsősorban az érdekli, hogyan áll össze az egyes idegsejtek kölcsönhatásából az agykérgi aktivitásmintázat. A következő kérdésekre próbálunk választ kapni: Milyen idegsejtekből állnak az agykérgi hálózatok és hogyan viselkednek? Ezekből a sejtekből mennyi van és hogyan kapcsolódnak össze? Mitől marad az idegsejtek aktivitása optimális, se túl alacsony, se túl magas szinten? Hogyan alakul ki egyfajta dinamikával jellemezhető aktivitásmintázat? Mitől és hogyan kapcsol egy hálózat egy másik feladatot ellátó állapotba?

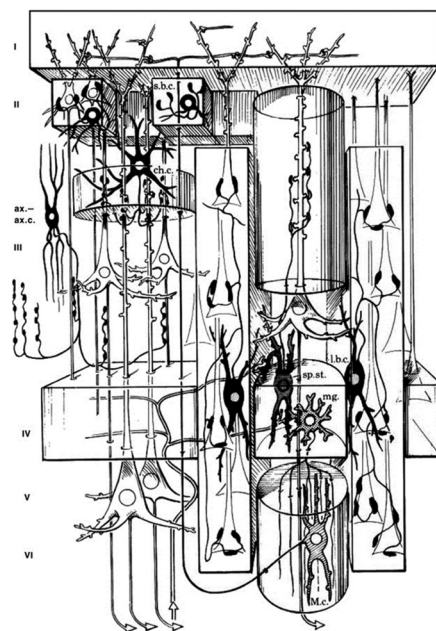
A serkentő és gátló idegsejtek, az agykéreg oszlopai és csavarjai

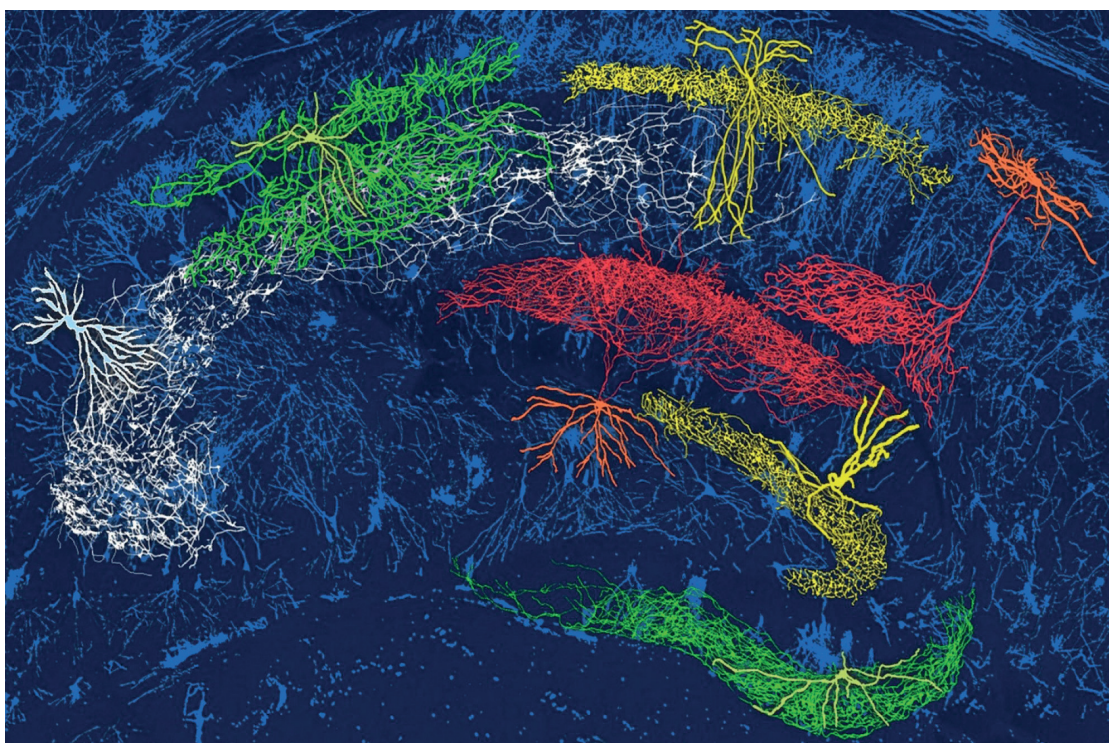
Szentágothai János agykérgi oszlopokba szerveződő idegsejteket ábrázoló gyönyörű rajza tökéletesen összefoglalja az agykérgi idegsejthálózatok szerveződési alapelveit (2. ábra). A sejtek 80%-át kitevő serkentő sejtek kapcsolatainak, a serkentő szinapszisoknak



1. ábra. Az agy különböző területeiről elvezetett EEG-jelek a magatartás függvényében szigorú hierarchia szerint összehangolva változnak. A) A hippokampuszból elvezetett theta-oszcilláció (fehér vonal az ábra felső részén) fázisaihoz csatolva változik az egyes agyterületeken található idegsejtek aktivitása (egy vízszintes vonal egy sejt, szaggatott vonal választja el az egyes agyterületeket, piros: magas, kék, alacsony aktivitás). Az egyes agyterületek aktivitásmaximuma időben eltolva követi egymást. B) Az egyes agyterületeken az idő függvényében változnak az aktivitási szintek (fekete görbék). A nyilak azt mutatják, hogy ennek alapján milyen lehet az információáramlás iránya

2. ábra. Szentágothai János rajza az agykérgi oszlopok szerveződéséről. Az ábrán fehér színnel jelölt, kúp alakú serkentő piramisok kapcsolatai alkotják az információ-feldolgozó és -raktározó hálózatot, a feketével jelölt változatos gátlósejtek pedig a rendszer működését stabilizálják





3. ábra. A hippocampusz gátlósejtjeinek változatossága. A sejtek zömét alkotó serkentő piramisisejtek mellett (nyúlványrendszerüket fehérrel mutatja a rajz) számos különböző gátlósejt típus stabilizálja a hippocampusz működését. A sárgával jelölt kosérsejtek fejtik ki a leghatékonyabb gátlást, időzítik a sejtek aktivitását. A zölddel és pirossal jelölt gátlósejtek a sejtek közeli és távoli dendritjeire érkező bemenetek hatását hangolják

az erősségében tárolódik az információ, amit korábban megtanultunk. Ezen sejtek aktivitásának mintázata kódolja a külvilágból érkező információt. Aktivitásuk változása, melyet a szinapszisokon keresztül terjedő jelek befolyásolnak, képezi az emlékezés és gondolkodás alapjait. Tanulás során pedig ezeknek a kapcsolatoknak az erőssége változik meg, így tárolódik az információ. Egy csupán serkentő elemekből álló hálózat azonban nem mutathat megbízható, kiegyensúlyozott működést. Pozitív visszacsatolás útján az aktivitás megállíthatatlanul növekedne, a zenei koncerteken előforduló hangmérnöki balesetekhez hasonló erős gerjedés következne be, amely nem más, mint egy epileptikus roham. Ennek elkerülésére az evolúció során a serkentő sejtek információ-feldolgozó állványzatába mint csavarok, gátlósejtek épültek be, hogy stabilá tegyék a vázat, a rendszer aktivitását.

Egy hálózat aktivitási szintjét egy, az információ raktározásának maximumát biztosító aktivitási szinten viszonylag szűk tartományban kell tartani (a legjobb, ha egy időben a sejtek 1–10%-a aktív). Az egyik véglet a kóma, a másik véglet az epilepszia (mindkettő tudatvesztéssel jár). Mivel az idegrendszerbe a környezetből, illetve az egyik agyterületről a másikra igen eltérő mennyiségű aktivitás érkezik, ez nem könnyű feladat. Feltételezhető, hogy ennek eredményeként alakult ki az evolúció során az a sok gátlósejt típus, amivel az anatómusok és élettanászok találkozhatnak. Egy idegsejt működésének számos aspektusát szabályozhatják a különböző típusú gátlósejtek. Az idegsejtek jelfelvező nyúlványrendszerén, a dendriteken végződő gátlósejtek csoportjai azt befolyásolják, hogy a sejt hogyan összegzi a rá érkező jeleket. De a dendritekre érkező gátlás befolyásolja azt is, hogy a különböző agyi területekről érkező jeleket hogyan súlyozza a sejt, vagy milyen fontos egy adott pillanatban érkező jel, és mikor kell tanulni, azaz a szinaptikus kapcsolatok erősségét változtatni [4]. Az idegsejt sejttestén keletkezik az akciós potenciál, mely a sejt kimenetét jelen-

ti. Az akciós potenciál kialakulását és időzítését is számos erre specializálódott a sejttest közelében ható gátlósejt képes befolyásolni.

A sejtek eltérő részein végződő gátlósejtek az aktivitás szabályozását előre és visszacsatolt gátló hálózatokon keresztül végezhetik. Mindkét elrendeződés fontos abban, hogy a bemenő aktivitási szinttől valamelyest független legyen az aktív sejtek mennyisége és így stabil legyen az aktivitás.

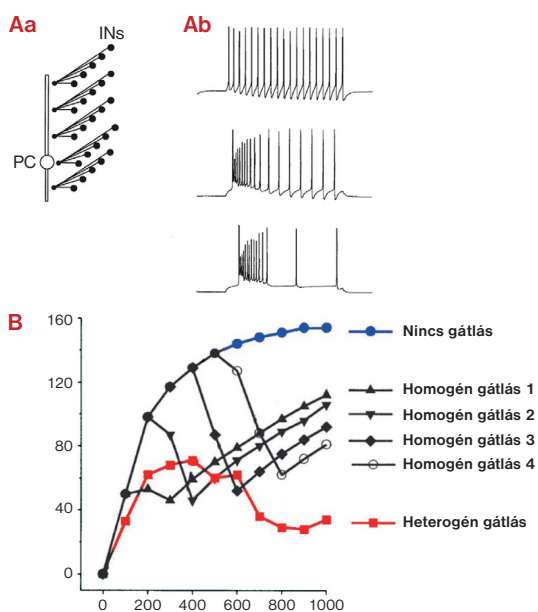
Eddig tehát azt lehetett látni, hogy az idegrendszert sokféle sejt-típus alkotja meghatározott rend szerint kapcsolódva. De a komplexitásnak még számtalan rétege van. Mint látni fogjuk, az elemek tulajdonságainak eloszlása is fontos kérdés, illetve az elemek dinamikus tulajdonságaiban és a tulajdonságok hangolásában is megjelenik a komplexitás.

Az alkotóelemek változatossága nemcsak gyönyörködtet, hanem stabilitást is ad

Ha a gátlósejteket részletesebben megvizsgáljuk, akkor azt találjuk, hogy mind anatómiai tulajdonságaikban (azaz, hogy milyen az alakjuk, mely meghatározza milyen és mennyi bemenetet kapnak, illetve milyen a kimenetük, más sejtekkel alkotott kapcsolatrendszerük), mind elektrofiziológiai tulajdonságaikban (hogyan összegzik a rájuk érkező jeleket, hogyan alakul rövid és hosszú távon aktivitásuk), illetve farmakológiai tulajdonságaikban (hogyan képesek különböző agyban található jelátvivő anyagok, hormonok áthangolni működésüket) jelentős változatosságot (3. ábra) mutatnak [5–8].

Hogy mire jó ez a nagy változatosság, arra *Aradi Ildikó* és *Soltész Iván* [9] egyik közleménye mutat rá. Neuronhálózat-modellezéssel összehasonlították az egyforma tulajdonságú gátlósejtek-

Az agykérgi hálózatok szerkezeti és működési komplexitása



4. ábra. Változatos tulajdonságú gátlósejtekből felépülő hálózat kifinomultabb szabályozást biztosít. Az Aa ábrarészen bemutatott modellhálózatban az Ab részen látható eltérő aktivitást mutató gátlósejtek közül helyeztek el hasonló vagy vegyes csoportokat, és vizsgálták a hálózat stabilitását. A B ábrán látszik, hogy növekvő bemenetre gátlás nélkül (kék) az aktivitás korlátlanul nőtt. Amennyiben különböző tulajdonságú, de homogén gátlósejteket (fekete vonalak) használtak, akkor az aktivitás egy ideig nőtt, majd eltérő szintekre állt be. A hálózat a bemeneti értéktől való legnagyobb függetlenséget (stabil aktivitást) akkor mutatta, amikor gátlósejtjei változatosak voltak (piros)

ből felépülő hálózat működését egy olyan, amelyikben a gátlósejtek tulajdonságai különbözőek voltak (egy lehetséges spektrumon belül véletlenszerűen kombinálódtak), de az átlagértékek megegyeztek az előzővel. A heterogén gátlósejteket tartalmazó hálózat jóval robusztusabb, stabilabb volt. A gátlósejtek a bemenő aktivitás sokkal nagyobb tartományát voltak képesek ellenőrzésük alatt tartani. A jelenséget könnyű megmagyarázni. Ha csak egyféle gátlósejt-csoport létezik, akkor egy küszöbérték felett bekapcsolódik a gátlás, alatta pedig nem. Ha az egyes sejtek más-más szinten kapcsolnak be, akkor sokkal finomabb lépcsőkben indul be a gátlás és kiegyensúlyozottabb lesz a működés szabályozása (4. ábra).

A helyzet hasonlít a városi legendaszerűen ismert tényhez, miszerint a fajgazdag esőerdő sokkal ellenállóbb a zavarásokkal (szárazság, kártevők) szemben, mint a monokultúras kukoricaültetvény. A háttérmechanizmus az volna, hogy egy bonyolult kapcsolati hálózattal leírható rendszer számtalan alternatív útvonalon ki tud térni a zavarásból, ellentétben egy igen redukált rendszerrel.

A gátlósejtek tulajdonságaiban megfigyelt változatosság a kutatók között is nagy változatosságot generál. Egy részük szerint a sejtek jól meghatározott, ha nem is kicsi, de megszámlálható számú csoportba sorolhatók, vagyis az anatómiai, élettani, farmakológiai és egyéb tulajdonságaik egymással átfednek. Saját tapasztalataim ennek részben ellentmondanak, mert a gátlósejtek bizonyos tulajdonságai meghatározott eloszlás szerint szabadon kombinálódnak. Ennek következményeképpen például az anatómiai és élettani tulajdonságaik alapján nagyon nehéz jól elkülöníthető csoportokba rendezni őket [8]. Az immunrendszer érése során egy aktív, véletlen generáló génszakasz-kombináció során alakulnak ki az egyes antigéneket felismerő sejtcsoportokban az immunoglobulinokat kódoló gének. Elképzelhető, hogy az idegrendszerben sem pontosan előre meghatározottak az egyes gátlósejtek tulajdonságai, hanem kombinatorikusan rendeződnek egymás mellé a tulajdonságok, per-

szek nem teljesen függetlenül, hanem egy evolúciós folyamat során legjobbnak bizonyuló eloszlás alapján. A fenti „heterogenitás stabilitást eredményez” példa támogatja, hogy az evolúció során ne egy szigorú terv, hanem egy sztochasztikus algoritmus alakítsa ki a sejtek típusait.

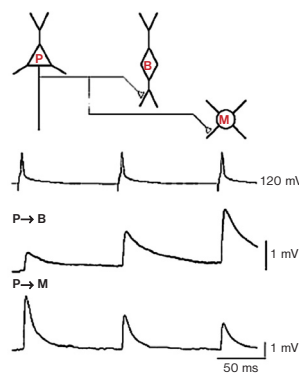
A kapcsolatok változatossága finomítja a dinamikát

Egy hálózat leírásánál nemcsak az alkotóelemeket, hanem a közöttük levő kapcsolatokat is jellemezni kell. Nem meglepő módon, az egyes sejtípusok közötti kapcsolatokban is igen nagy komplexitást és heterogenitást írtak le [10,11]. Serkentő (S) és gátló (G) sejtek között 4 irányú kapcsolat lehetséges: $S \rightarrow S$, $S \rightarrow G$, $G \rightarrow S$ és $G \rightarrow G$. Mivel azonban többféle serkentő sejt és jelentősen többféle gátlósejt létezik, egy agykérgi területen a lehetséges kapcsolattípusok száma sokkal nagyobb. Megvizsgálva, hogy egy jelsorozat hogyan jut át különböző sejt párok közötti kapcsolaton, azt találjuk, hogy ez sejt pártípusonként más és más. Vannak olyan kapcsolatok, amelyeken keresztül a sorozat egymás utáni tagjai egyre nagyobb (erősödő kapcsolat), és vannak olyanok, ahol egyre kisebb intenzitással jutnak át (gyengülő kapcsolat). Abban is különbség van, hogy milyen sebességgel nő vagy csökken az átvitel, illetve mennyi ideig tart a növekedés vagy a csökkenés. Ennek eredménye az, hogy egy agyi területre érkező jelsorozat a különböző típusú sejtekben más jelsorozatként jelenik meg [12]. Egy hosszabb jelsorozatból az erősödő kapcsolatot kapó sejtek csak a végét, míg a gyengülő jelet kapók csak az elejét „látják”. Különböző frekvenciájú jelsorozatok, mint egy jelszűrő esetében, más erősséggel jutnak át az eltérő tulajdonságokkal rendelkező kapcsolatokon (5. ábra).

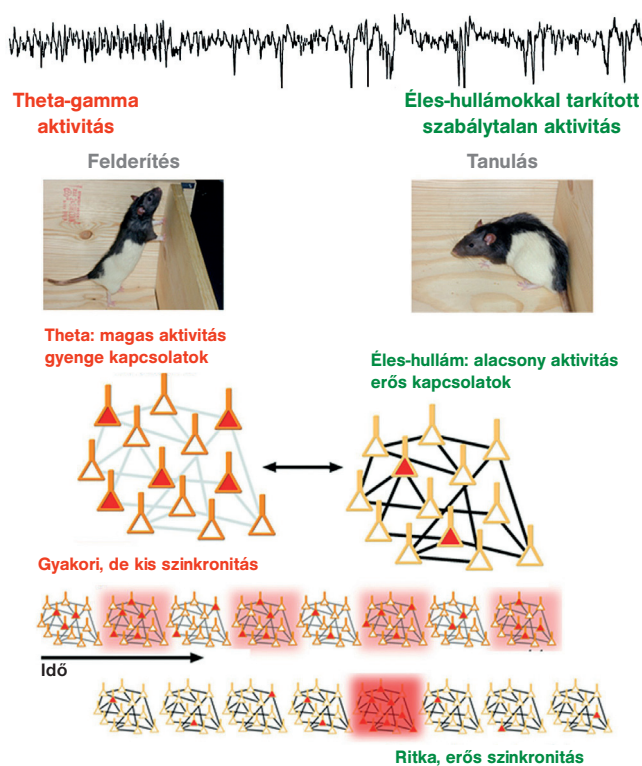
Az ilyen időérzékeny tulajdonságok magyarázzák meg, hogyan alakulnak ki az EEG-ben megfigyelhető többszörösen egymásba ágyazott periodikus mintázatok az idegsejtek visszacsatolt hálózataiban.

Egyetlen hálózat, többféle dinamika

Az írás elején felvetett kérdést: Hogyan tud egy hálózat többféle típusú aktivitásmintát, dinamikát előállítani? – csoportunk legújabb eredményei mutatták meg. Már régóta ismert, hogy az agy ősbibb területeiről felszálló pályarendszerek finomhangoló (moduláló) hatással átvivő anyagokat szabadítanak fel az egy magasabb területre. Az egyes ilyen anyagok: acetilkolin, szerotonin, dopamin, noradrenalin és hisztamin összefüggően különböző magatartási állapotokkal. Részt vesznek a napi ritmus, az éves ritmus, a jutalmazás és motiváció szabályozásában. A felszabadulásukkal kapcsolatos zavarok neurológiai problémákat, depressziót, alvászavarokat, Parkinson-kórt vagy játék- és kábítószer-függőséget okoznak. Ezek



5. ábra. A szinaptikus átvitel érzékenysége a beérkező jel mintázata. Egy serkentő sejt különböző típusú gátlósejtekre máshogy adja át az információt. A P-vel jelölt sejt jelsorozata (a sematikus ábra alatti első sor) a B jelű sejtben egyre erősödő jelként jut át, míg az M jelű sejtben ennek fordítottja történik, a jelek egyre gyengyebbek lesznek. Ennek eredményeként a B jelű sejt a jelsorozat végén lesz aktív, míg az M jelű az elején



6. ábra. Ugyanaz az agyterület (itt a hippokampusz) különböző dinamikus és eltérő feldolgozást végző állapotokba kapcsol magatartási állapot függvényében, amint azt a változó dinamikájú EEG is jelzi (felső sor). Amíg az állat a környezetét deríti fel, theta-gamma aktivitás mérhető (bal oldal), mely alatt az információ előhívása és rövid távú raktározása történik. Amikor az állat számára fontos dolgot kell tanulni, a hálózat éles-hullám generáló állapotba kapcsol (jobb oldal). Az átkapcsolást az okozza, hogy az agykéreg alatti területekről származó acetil-kolin koncentrációja leesik. Ezáltal csökken a sejtek serkenthetősége (középső sor: narancs vs. sárga sejt körvonalak), viszont megnő a közöttük levő kapcsolatok erőssége (vékonyabb vs. vastagabb vonalak). Ennek hatására a gamma aktivitás alatt megfigyelhető gyakori aktivitásnövekedést (4. sor), ritkán bekövetkező, de magasabb szinkronitását események, éles hullámok váltják fel (5. sor)

az anyagok egyrészt diffúz módon felszabadulva lassan modulálják a serkentő és gátlósejtek bizonyos paramétereit, másrészt, mint Freund Tamás korábban kimutatta [13], közvetlenül és gyorsan ki- vagy bekapcsolnak bizonyos gátlósejt típusokat. Ennek eredményeként ugyanaz az idegsejthálózat eltérő magatartási állapotokban (amikor más típusú jelfeldolgozásra van szükség) különböző aktivitás dinamikát mutat (6. ábra), és ezáltal feltételezhetően más funkciót lát el.

A hippokampusz az agykéreg olyan ősi területe, amely igen fontos szerepet tölt be az emlényomok kialakulásában, elraktározásában és előhívásában. Buzsáki György elmélete szerint [14] az emlényom kialakulása két lépésben történik meg. Amikor az állat felderíti környezetét, a hippokampuszból egy periodikus aktivitást, úgynevezett theta-modulált gamma aktivitást lehet elvezetni. Amikor valami fontos történik az állattal, jó (pl. táplálék) vagy rossz (áramütés) tapasztalat éri, a hippokampusz éles hullámokat generál, melyek során a viszonylag alacsony aktivitást másodpercenként egyszer előforduló, 0,1 másodpercig tartó magas, szinkronizált aktivitás tagolja. Az elmélet szerint a theta-gamma aktivitás alatt rövid időre elraktározott emlényom-csírák az éles-hullám-aktivitás alatt visszajátszódnak és hosszú távra elraktározódnak az agykéreg más területein.

Kimutattuk, hogy amikor az acetilkolin- (ACh) szint leesik a hippokampuszban, a sejtek serkenthetősége csökken, ezzel egy időben a sejtek közötti jelátvitel erőssége nő. Ennek hatására lesz a felderítő magatartást kísérő periodikus gamma aktivitásból a tanulást jelentő éles-hullám aktivitás, mivel a hálózati paraméterek áthangolása miatt csökken annak az esélye, hogy a hálózati aktivitás újrainduljon, viszont amikor ez bekövetkezik, akkor az erős kapcsolatokon keresztül az aktivitás magas szintre épül fel, így az emlényom visszajátszása bekövetkezik. A két állapotban más a magas aktivitási szakasz hossza, azaz a hálózaton belül a visszacsatolt feldolgozás ideje. A gammaaktivitás egyetlen, viszonylag rövid, magas aktivitási fázisú ciklusa alatt a sejtek állapotát elsősorban a külvilágból érkező bemenetek határozzák meg, és csak kevésbé szól bele a feldolgozásba a sejtek közötti kapcsolatokról származó visszacsatolt jel. A hosszabb ideig tartó éles-hullám alattiban a külső bemenetek viszonylag kis súllyal játszanak szerepet, mivel az aktivitás alakulását elsősorban a sejtek közötti kapcsolatokon keresztüli visszacsatolt feldolgozás dominálja. Látható tehát, hogy a két állapotban a hálózat másféle módon dolgozza fel az információt. Felderítés (gamma) alatt a hippokampuszban sejt állapotát elsősorban a külvilág vezérli, míg az emlényom raktározási fázisban (éles-hullám) a sejtek visszacsatolt kapcsolatrendszerének szinapszisaiiban hagyott időleges változások (rövidtávú emlényomok) irányítják a sejtaktivitás fejlődését.

Összefoglalva: ahhoz, hogy az agy komplex feladatokat tudjon megoldani, szükség van egyrészt arra, hogy részhálózatai többféle dinamikára és ezáltal feldolgozási módra legyenek alkalmasak, másrészt ezeknek a részeknek nagyon pontosan összehangolt és együtt változó működést kell mutatniuk. Mindezt az teszi lehetővé, hogy az agyi idegsejt-hálózatokban egyrészt sokféle különböző tulajdonságú idegsejt található, melyek között a kapcsolatok tulajdonságai is pontosan meghatározott változatosságot mutatnak, másrészt, hogy mind a sejtek tulajdonságait, mind a közöttük levő kapcsolatok tulajdonságait moduláló hatású átvivő anyagok hangolják, melyek meghatározott mintázatban szabadulnak fel az agy különböző területein a magatartási állapot függvényében. ✨

Kutatásainkat az OTKA 83251 számú pályázata támogatta.

IRODALOM

- [1] Battaglia, F. P., Benchenane, K., Sirota, A., Pennartz, C. M. A., and Wiener, S. I. *The hippocampus: hub of brain network communication for memory*. Trends in Cognitive Sciences 15, 7 (2011), 310–318.
- [2] Luczak, A., Bartho, P., Marguet, S. L., Buzsáki, G., and Harris, K. D. *Sequential structure of neocortical spontaneous activity in vivo*. Proc Natl Acad Sci USA 104, 1 (2007), 347–52.
- [3] Montgomery, S. M., Sirota, A., and Buzsáki, G. *Theta and gamma coordination of hippocampal networks during waking and rapid eye movement sleep*. J Neurosci 28, 26 (2008), 6731–41.
- [4] Miles, R., Toth, K., Gulyás, A. I., Hajos, N., and Freund, T. F. *Differences between somatic and dendritic inhibition in the hippocampus*. Neuron 16, 496196761 (1996), 815–23.
- [5] Freund, T. F. and Buzsáki, G. *Interneurons of the hippocampus*. Hippocampus 6, 497072927 (1996), 347–470.
- [6] Gulyás, A. I., Miles, R., Hajos, N., and Freund, T. F. *Precision and variability in postsynaptic target selection of inhibitory cells in the hippocampal CA3 region*. Eur J Neurosci 5, 12 (1993), 1729–51.
- [7] Klausberger, T. and Somogyi, P. *Neuronal diversity and temporal dynamics: the unity of hippocampal circuit operations*. Science 321, 5885 (2008), 53–7.
- [8] Parra, P., Gulyás, A. I., and Miles, R. *How many subtypes of inhibitory cells in the hippocampus?* Neuron 20, 5 (1998), 983–93.
- [9] Aradi, I. and Soltesz, I. *Modulation of network behaviour by changes in variance in interneuronal properties*. J Physiol 538, Pt 10 (2002), 227–251.
- [10] Gupta, A., Wang, Y., and Markram, H. *Organizing principles for a diversity of GABAergic interneurons and synapses in the neocortex* [see comments]. Science 287, 545120102820 (2000), 273–8.
- [11] Losonczy, A., Zhang, L., Shigemoto, R., Somogyi, P., and Nusser, Z. *Cell type dependence and variability in the short-term plasticity of EPSCs in identified mouse hippocampal interneurons*. J Physiol 542, Pt 122090830 (2002), 193–210.
- [12] Glickfeld, L. L. and Scanziani, M. *Distinct timing in the activity of cannabinoid-insensitive and cannabinoid-insensitive basket cells*. Nat Neurosci 9, 6 (2006), 807–15.
- [13] Freund, T. F. and Antal, M. *GABA-containing neurons in the septum control inhibitory interneurons in the hippocampus*. Nature 336PBS Record: 249, (1988), 170–173.
- [14] Buzsáki, G. *Two-stage model of memory trace formation: A role for "noisy" brain states*. Neuroscience 31, (1989), 551–570.

CSÉPE VALÉRIA

Kognitív fejlődés, agyi komplexitás, matematika

Idegtudományi tanulságok a természettudomány oktatásához

Az oktatás és nevelés, s annak eredménye sok tekintetben boldogulásunk, egészségünk, jólétünk és jóllétünk eredője. A formális oktatás során szerzett tudás és készségek, s ezeknek a megszerzés körülményeitől független alkalmazhatósága, a kompetencia tette, teszi lehetővé, hogy az ember a biológiai evolúció korlátain túllépjen. Nem újdonság, hogy a jó oktatás, s különösen a természettudományok oktatása a gyakorlatra, a tapasztalatokra építve működik a legjobban, s ezek nem függetlenek az igencsak komplex agyi folyamatoktól, jóllehet ezekről még ma is jóval kevesebbet tudunk, mint szeretnénk. Az elmúlt évtizedben az idegtudomány minden korábbinál több adattal szolgált ahhoz, hogy megértsük a tanulás mibenlétét, az oktató, fejlesztő módszerek hatását. A korlátlan lelkesedés és a kételkedésre épülő elutasítás jelenti azt a két végletet, amellyel az oktatás az idegtudomány eredményeit, s az idegtudomány a pedagógiai tudományokat és gyakorlatot figyeli, értékeli. Mégis, sokak számára elfogadható, hogy az idegtudomány gyors fejlődése olyan új felismerésekhez vezetett, amelyek a tanulás és tanítás megértésének a korábbiaktól radikálisan eltérő útjait jelenti.

Az idegtudomány és a pedagógiai tudományok közötti interakciók kezdetén fogalmazott úgy Mayer [1], hogy az igazán sikeres megközelítés kétirányú kapcsolatra épül; az idegtudományi kutatások befolyásolják az oktatástudományiakat, s az oktatáskutatók az idegtudományi kutatásokat. Ahhoz azonban, hogy ebben a kétirányú hatásrendszerben megbízható és használható eredmények szülessenek, elengedhetetlen az interdiszciplináris megközelítés és felkészültség [2, 3]. Az alig egy évtizedes oktatás- vagy pedagógiai idegtudományi eddigi eredményei szerint úgy tűnik, hogy ez a kétirányú út járhatóbb, mint az egyirányú, azaz a kognitív idegtudomány eredményeinek a tanításban történő feltételek nélküli alkalmazása. Tény, hogy a kognitív idegtudomány kutatási arzenálja ma már igen jelentős, ám eredményeinek akkor van valódi hatása az oktatásra, ha a kísérletek kérdésfeltevése a kognitív fejlődés-pszichológiai és pedagógiai elméletekből indul ki. Kulcsfontosságú tehát e kutatásokban az empirikus kutatásokban alkalmazott interdiszciplináris perspektíva. Mielőtt azonban az agyi változások és tanítási, tanulási módszerek egymásra hatását elemeznénk, ki kell térnünk arra, hogy mi jellemzi a biológiai és környezeti interakcióknak azt az igen komplex rendszerét, amelyben a kognitív fejlődés megvalósul.

Kognitív fejlődés és agyi komplexitás

A megismerő funkciók fejlődése a biológiai előfeltételek s a környezeti hatások olyan kölcsönhatásában alakul ki, amelyet jellegzetes dinamika jellemez, bármely szinten vizsgáljuk is a biológiai és a környezeti. Kognitív szempontból ez a fejlődés követhető az alkalmazkodó (adaptív) neuro-kognitív folyamatok során kialakuló

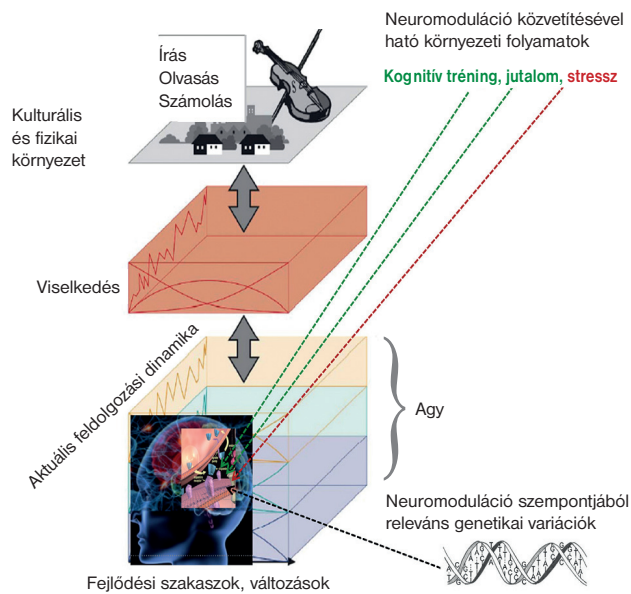
ló hosszú távú leképeződés (reprezentáció) vizsgálatával. Az alkalmazkodást lehetővé tevő alkalmazkodó rugalmasságnak, a plaszticitásnak sok szintje van a mikro szinttől a makróig; az egyént jellemző genetikai, neuronális, agyi és viselkedési plaszticitástól a társas és kulturális alkalmazkodásig.

Az agy egyik lenyűgöző működési tulajdonsága, hogy az idegsejtek olyan vegyületeket (neurotranszmitterek) bocsátanak ki, amelyek közöttük a jelátvitelt szabályozzák (bővebben lásd [4])¹. Számos neurotranszmitter-rendszer, így például a dopamin, szerotonin, noradrenalin, kiterjedt idegsejt-hálózatokra hatva, nagy szerepet játszik a megismerő funkciók agyi folyamataiban. Az elmúlt évtized egyre több adattal szolgált például arról, hogy az egyes transzmitterrendszereknek milyen szerepe van az ember figyelmi, emlékezeti rendszereiben, a kognitív kontrollban, hogy csak néhányat említsünk. Ezek az adatok azonban a felnőttekre vonatkoznak, hiszen a receptor képződő eljárások alkalmazása erősen korlátozott a gyerek és a serdülő korú kísérleti személyek esetében. Van azonban néhány olyan vizsgálat, amely azt mutatja, hogy az agykérgi dopaminrendszer elhúzódó fejlődése felelhet a figyelmi rendszer és végrehajtó funkciók lassú kialakulásáért [5]. Mindkét kognitív rendszer olyan általános funkciókért felelős, amelyekre az iskolai tanítás, tanulás erősen támaszkodik. Mi több, az agykérgi és kéreg alatti dopaminrendszer eltérő fejlődése, a serdülőkorban megváltozott érzékenység (stressz, szociális jutalom) tovább módosítja a kognitív fejlődést.

A molekuláris genetika legújabb eredményei még tovább árnyalják azt a képet, amelyet a tudomány az agy, a viselkedés és a kognitív fejlődés viszonyáról tudott. A tudomány mai ismeretei szerint ugyanis a fejlődő agy anatómiai és mindenekelőtt neurokémiai változásai az egyes életkorokban eltérően módosítják a genetikai adottságok (genotípus) kifejeződését (fenotípus). Az agyi mechanizmusok ezért úgynevezett „köztes fenotípusnak” tekinthetők [6] a viselkedéses fenotípus és az idegrendszer genetikai kifejeződésében. A különböző szinteken megnyilvánuló plaszticitás az agyi hálózat változásához vezet, s minden környezeti hatás a neuromoduláción keresztül hat a fejlődésre. A fizikai és kulturális környezet, a neuromoduláción keresztül érvényesülő környezeti hatások (kognitív tréning, jutalom, stressz) a társas és megfigyeléses tanulás, az explicit tanulás, a fejlesztés, a genetikai kifejeződés, viselkedés és agyi fejlődés modellje [7], a kognitív fejlődés komplexitásának egyik jó példája. A feltárandó folyamatokat jól szemlélteti az 1. ábra.

Minden olyan fejlesztés, amely a kognitív architektúrának olyan összetevőjére irányul, amelynek funkcionális agyi hálózata valamelyik nagy transzmitterrendszerhez tartozik, biztató eredményeket ígér. Itt most nem térünk ki az elmúlt 2–3 év kutatási eredmé-

¹ Szerkesztők megjegyzése: a különszámban Gulyás Attila cikke foglalkozik ezzel a témakörrel.



1. ábra. A kognitív fejlődés, a környezeti hatások és neuromoduláció modellje ([7] alapján)

nyekre, csupán azt emeljük ki, hogy a gyermekek és serdülők tanulási folyamatainak mélyebb megismeréséhez jelentősen járulhat majd hozzá a transzmitterrendszerek és a kognitív fejlődés összefüggéseinek a feltárása. Azt érdemes azonban figyelembe venni, hogy az eddigi eredmények a kognitív pszichológia fogalmai szerint terület-általános, azaz valamennyi kognitív funkcióban szerepet játszó rendszerekre (figyelem, emlékezet, motiváció) vonatkozik. Kevesebbet tudunk a terület-specifikus rendszerekről, jóllehet ezek komplexitása így is igen nagy. Ezek közül itt a matematikát, mindenekelőtt annak elemi folyamatait emeljük ki, s mutatjuk be, hogy milyen összetett az idegtudományi eredményeinek alkalmazása a tanítási módszerek kialakításában.

A számoló agy

Több közvetett idegtudományi adatot ismerünk arra vonatkozóan, hogy milyen kihívást jelentenek a fejlődő agy számára a természettudományok, itt azonban, egyszerűsítési és terjedelmi okokból egyaránt, a matematika területére szűkítjük az elemzést. Tény, hogy a mentális aritmetika kognitív idegtudományi kutatásának jelentős részét a felnőtteken végzett vizsgálatok teszik ki, egyre több fejlődési adatunk is van, ezért elsősorban ezeket az eredményeket emeljük ki.

A matematikai gondolkodás kialakulásának megértéséhez olyan komplex és dinamikus rendszert kell feltárnunk, amelyben az agyi komplexitás, a megismerő (kognitív) rendszer, s ezen belül a tanulás fejlődési változásai egyaránt számításba veendőek. A matematika és az agyi működések összefüggéseinek kutatásában a kognitív fejlődés-idegtudományi furcsa, bár nem szokatlan útja (lásd az olvasás fejlődésének idegtudományi vizsgálatairól Tóth és Csépe [8] tanulmányát), hogy a számérzék, a mentális aritmetika, s a matematikai gondolkodás megismerésében az atipikus fejlődés, a diszkalkulia áll a középpontban, s a tipikus fejlődést egyelőre kevesen vizsgálják szisztematikusan.

Tudjuk azonban, hogy a számok felfogására és az azokkal végzett műveletek elvégzésére úgymond elhuzalozott az emberi agy. A kognitív modellek [9, 10] egy úgynevezett számfeldolgozó modult feltételeznek, s idegtudományi adatokkal is alátámasztják azt, hogy a számérzék számolási képességünk olyan veleszületett alap-

ja, amely erősen független a többi kognitív képességtől. E modul része a kis számosságok számlálás nélküli megragadása (a későbbi szubitizáció lehetséges alapja), összeadása, kivonása, s nagyságviszonyuk megértése [11]. E képességek agyi – s különösen a fejlődésben lévő – folyamatainak feltárása több forrásból is fakad. A számfeldolgozás komplexitása neuronális szinten adódik, hogy a rendszert funkcionális és neuroanatómiai szinten is elkülönülő, s más feldolgozó hálózatok részét is képező területek alkotják. Kihívást jelent az is, hogy a számolás fejlődésmenete nem egyenes irányú folyamat, s az egyéni eltéréseknek igen jelentős a szerepe.

A matematika, hasonlóan más kognitív rendszerekhez, biológiailag elsődleges, azaz veleszületett, univerzális képességekre, valamint másodlagos, az elsődlegesekre támaszkodó, a kultúra, a formális oktatás által közvetített készségekre támaszkodik. *Grabner* és munkatársai [12,13] az agyi aktivitást mérve, a matematikai kompetencia és a tanítás, fejlesztés kölcsönhatását vizsgálták. Egész konkrétan, egy rövid aritmetikai tréninget követően, az agyi aktivitást mérték alacsony és magas matematikai kompetenciát mutató tanulóknál. Azt találták, hogy a matematikai kompetencia szintjével összefüggő agyi aktivitásváltozás (az agyi számfeldolgozó rendszerhez tartozó bal oldali anguláris tekervény) a tréninget követően eltűnt. Ez tanítási szempontból azért is fontos, mert egy adott aritmetikai problémára létrehozott feladatsor gyakorlásával olyan szint érhető el, amely alkalmas a kompetencia-különbségek enyhítésére.

Érdeemes megjegyeznünk, hogy míg a veleszületett elsődleges képességek elhívása gyors és egyszerű, ezek kiterjesztése, azaz a másodlagos készségek elsajátítása lassú, komplex, s erőfeszítést igényel. Ez utóbbiakhoz tartozik többek között a számrendszerek megértése, az aritmetikai tények és műveletek elsajátítása, a matematikai problémamegoldás.

A két komplex rendszer, vagyis az agy és a matematika specifikus feldolgozó rendszerének összekapcsolása régi törekvése a kognitív pszichológiának, így több modellel is operálnak a kutatók. A legismertebb, s legelfogadottabb talán azért, mert humán és állatkísérletes agykutatási adatokra is támaszkodik *Stanislas Dehaene* hármaskód modellje [10]. Ennek lényege, s neve is innen származik, hogy a számolási feladatok három elkülönülő rendszerre támaszkodnak. A mennyiségek szimbolikus kódja a szám, s ennek hang-sorként tárolt neve (auditoros-verbális szókeret), vizuális szimbóluma (vizuális arabszám-formátum), s betűkkel leírt formája (számszó-formátum) ugyanazon jelentést hordozza. Ez a jelentés az analóg mennyiségprezentációra, vagyis a számok nagyságértékére támaszkodik, s egy úgynevezett mentális számeqyenesen működik.

A veleszületett képesség, a számérzék és a kultúra elsőként a számlálásban, s ehhez a számszavak kötött sorrendben történő elsajátításban jelenik meg. Az iskoláskorig bekövetkező változásokra itt most részleteiben nem térünk ki, csupán röviden összefoglaljuk, hogy melyek a legfontosabb lépések [14]. A kezdet az egyesével történő számlálás, ezeknek a halmazok elemszámához, a tárgyak és események sorához rendelése a kardinális számok megjelenéséhez vezet. Ez csak a fejlődés egy adott szakaszában bővíülhet új stratégiával, s általában hétéves kortól tudnak a gyermekek kettesével, hármasával számlálni; kevesebben a '+n', többen az összeadó stratégiát alkalmazzák, vagyis az elemeket alcsoportonként számolják össze.

A stratégiák hatékonysága eltérő ütemben javul; az egyesével számlálás 5 és 7 éves kor, a '+n' 11 és 15 éves kor között. Lassú, sok gyakorlást igénylő folyamatok ezek, mégpedig olyanok, amelyek igencsak távol vannak a matematikától. A matematika alapozása szempontjából azonban nem mindegy, hogy mit tudunk arról, hogy a számok milyen kihívást jelentenek a fejlődő agy számára.

Szinte valamennyi kutató egyetért abban, hogy a számfogalom kialakulását, a matematikai készségek fejlődését a mennyiség feldolgozására specializálódott, veleszületett rendszer segíti, nevezük akár biológiailag elsődleges képességnek [15], akár számér-

Kognitív fejlődés, agyi komplexitás, matematika

zéknek [10]. A fentebb említett, a számok nevének hallott és írott, valamint szimbólummal jelzett jelentését meghatározó analóg reprezentációhoz a térre vonatkoztatott asszociációk köthetők, nem véletlen tehát a metaforaként használt szakkifejezés, a mentális számegyenes. Ezen a mentális számegyenesen a számok jobbra haladva nőnek, balra pedig csökkennek (a mi kultúránkban). Számos adat – s köztük nem egy agykutatósi – arra utal, hogy a számok nagyságviszonyainak megítélésénél is arra a mentális számegyenesre támaszkodunk, amelyen a numerikus távolság viszonylagos. Talán nem meglepő, hogy a számok diszkriminációja is a Weber-törvényt követi; két szám összehasonlításának ideje a numerikus távolság függvénye, azaz minél nagyobb két szám különbsége (távolsága), annál rövidebb a válasz. Ezt nevezzük távolsághatásnak. Ugyanakkor, minél nagyobbak a számok (azonos különbséggel), annál hosszabb a válasz (nagysághatás, első közlésként lásd [16]). Ennek oka *Dehaene* [10] szerint, hogy a számok leképezése logaritmikus skálán történik, bár mások szerint párhuzamosan működik egy lineáris számegyenes is.

A nagyság- és a távolsághatás két olyan működési jellemző, amely a reprezentáció tulajdonságaiból, s a feladatban résztvevő agyi területek működéséből következik. A kutatót itt leginkább az agyi mechanizmusok feltárása érdekli, a pedagógust viszont az, hogy milyen tanítási stratégiát kövessen. Bár mindenki egyetért azzal, hogy a mentális számegyeneshez való hozzáférés a számok elsajátításának kezdetétől automatikus, nincs konszenzus abban, hogy a számok reprezentációja miként változik a tapasztalatok hatására.

A számokkal végzett műveletek tipikus fejlődésének kognitív idegtudományi vizsgálatában olyan összefüggések tárhatók fel, amelyek a viselkedéses módszerekkel nem ragadhatók meg. Ez a pedagógiai kutatások számára olyan új magyarázati szintet jelent, amely a tanításban eredményesen használható. Az alábbiakban néhány példát emelek ki ezek közül.

Példaként érdemes megemlíteni *Stavy* és munkatársai kutatásait [17], amely a geometriai feladatokban az intuitív interferencia jelenségét vizsgálta. Kísérleti paradigmájuk tudományelméleti szempontból sem lényegtelen, hiszen a klasszikus geometria alapját alkotó intuíción, pontosabban az intuíción és a tudományosan megismerhető összefüggés egyezését, illetve ütközését tanulmányozta. A pedagógia szempontjából ez akkor válik érdekessé, amikor az euklidészi geometria tanítását követően a nemeuklidészi geometriát kell megértetni. Ez utóbbival kezdődik ugyanis a geometria elszakadása a tapasztalattól, az intuíción, s ez tovább erősödik a modern algebrai felfedezésekkel. A geometria új területein a megértéshez és a tudás megszerzéséhez az absztrakt gondolkodás megjelenése szükséges, különösen, hogy a modern fizika a geometria egyes részterületeinek ismereteire épít. *Stavy* és munkatársai vizsgálatukban az agyi aktivitást mérték úgy, hogy a kísérletben résztvevőknek két geometriai forma kerületét kellett összehasonlítani. Az egyik feltételben az intuitív következtetés a helyes eredményt adta, a másikban viszont ütközött a helyes megoldással, amit interferenciának nevezünk. A feladat intuíciónak ellentmondó, azaz a kézenfekvővel éppen ellentétes megoldását jelentősen megnőtt aktivitás kísérte az agy pre-frontális területein. A szerzők kiemelik, hogy mindez a gátlási kontroll megnövekedett szerepére utal az ilyen feladatokban. A tanulság pedig az, hogy az intuíciónak ellentmondó geometriai, s feltehetően valamennyi természettudományi törvény, ismeret a gátlási kontroll adott fejlettségi szintjén tanítható jól, s az interferencia leküzdése a példák gyakorlásával lehetséges.

Egy további példa talán még érthetőbben rávilágít arra, hogy az idegtudomány más összefüggéseket képes látni, mint a viselkedéses vizsgálat, s a gyermekek problémamegoldásának a felnőttektől való lényeges eltérését képes kimutatni. *Kaufman* és munkatársai [18] egy egyszerű szám-összehasonlítás feladatot végeztek gyerekekkel és felnőttekkel. A meglepetést az jelentette, hogy miközben a feladatmegoldás viselkedéses mutatókkal követett eredményei nem különböztek, az agyi aktivitás mintázata jelentős

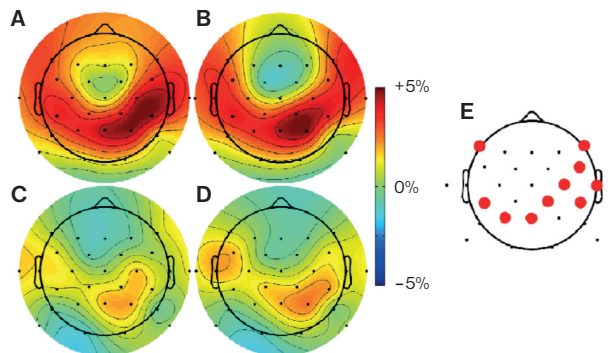
eltérést mutatott. Csak a gyerekeknel jelent meg az agy megnövekedett aktivációja azokon a területeken, amelyekhez a tárgyak megragadása, s a kéz ujjainak kérgi reprezentációja köthető (jobb szupramarginális tekervény, posztcentrális tekervény). A számok nagyságának összehasonlításánál tehát a gyerekek az ujjak reprezentációjára támaszkodtak, azaz olyan stratégiára, amelyet a felnőttek nem alkalmaznak.

A tanulság az, hogy bármily egyszerű is legyen egy számokkal végzett feladat, a viselkedéses eredmények nem informálnak arról, hogy milyen szinten tart a feldolgozó rendszer.

Amint arról már szó esett, az idegtudomány és a pedagógiai kutatás kapcsolata ideális esetben kétirányú, azaz a pedagógiai kutatások is inspirálhatják az idegtudományi vizsgálatokat. Ennek egyik példáját ugyancsak a matematika tanításának területéről vesszük. A pedagógiai kutatások az elmúlt évtizedben különleges figyelmet fordítottak azokra a megoldási stratégiákra, amelyekre a gyerekek támaszkodnak olyan egyszerű aritmetikai feladatokban, mint a számítási eredmények előhívása, számokra bontás, csoportosítás, elemenkénti kivonás. A vizsgálatok a meglehetősen nagy egyéni variabilitásra, s az egyén próbáról próbára mutatott teljesítményének változékonyságára hívták fel a figyelmet [19]. Idegtudományi szempontból ez igen fontos adat, hiszen a jelek kiértékelésének egyik alapja, hogy az egyéni agyi aktivitást (agy elektromos jelek), illetve azok mutatóit (aktív terület, térfogat) összevonjuk, átlagoljuk, s ezzel figyelmen kívül hagyjuk a megoldási stratégiák variabilitását. Az elmúlt években, részben a feladatmegoldási stratégiák pedagógiai kutatásainak köszönhetően, előtérbe került e stratégiák mibenlétének és agyi korrelátumainak vizsgálata. Kimutatták, hogy az agyi aktivitás változik az aritmetikai művelet elsajátításával [20], s ez összefüggésbe hozható a stratégiaváltással, vagyis az erőfeszítést igénylő műveletről az aritmetika tények automatikus előhívására való áttéréssel. Fialat felnőttekkel végzett vizsgálatokban azt is sikerült bemutatni, hogy ez a túltanulás miként jellemezhető, az új tények elsajátításakor aktív területekről miként tevődik át a feladat az automatikus előhívás agyi területeire. A másik igen tanulságos példa a két tudományterület interakciójára az aritmetikai tanulás egyéni különbségeinek feltárása.

A kognitív idegtudomány ezt a szempontot sokáig figyelmen kívül hagyta, ám az okok részletezése helyett inkább két vizsgálat eredményeire utalunk. Munkacsoportunkban a hatékonyságot és az alkalmazott stratégiák egyéni eltéréseit vizsgáltuk elektrofiziológiai módszerekkel. Az egyik vizsgálatban [21] a résztvevő egyetemi (nem matematika szakos) hallgatók feladata az volt, hogy a képernyőn megjelenő, összeadásokat tartalmazó egyenletekről igaz-

2. ábra. Az algebrafeladatok megoldásában kevésbé hatékony (A és B) és hatékony (C és D) csoport agyi elektromos aktivitásának szinkronizáció térképe. Az első oszlop (A és C) 3, a második oszlop (B és D) 4 összeadás esetén mutatja a szinkronizáció mértékét. Az E térképen a piros pontok a szignifikáns csoporteltérést mutató válaszok elvezetési helyeit jelölik ([21] alapján)



hamis ítéletet hozzanak. Korábbi viselkedéses vizsgálatok alapján két csoportot alakítottunk ki, olyanokat, amelyek az összeadási feladat megoldásának pontosságában nem, sebességében viszont eltértek. Egyszerűbben fogalmazva, az egyik csoport hatékonyabban, de nem eredményesebben dolgozott, mint a másik. A hatékonyabb csoportot a másik csoportban nem látható agyi aktivitás megjelenése (folyamat-specifikus gátláshoz kapcsolható parietális alfa) jellemezte (2. ábra). Mindezt olyan folyamat-specifikus gátlás jeleként értelmeztük, ami a gyakorlás során erősödött. Mára számos további adat igazolja, hogy az erősödő gátlás az ingerek hatékony átalakítását segíti, s ez az összeadási feladatok megoldását támogató stratégia.

Az alkalmazott stratégiák szerepét egy másik vizsgálatunkban vak kísérleti személyeknél a numerikus távolsághatás elektrofiziológiai korrelátumait mérve követtük [22]. Azt találtuk, hogy ez a hatás a vizuális tapasztalattal nem rendelkezőknél is megfigyelhető, s az evolúciósan behuzalozott számfeldolgozó rendszer, s annak agyi hálózata megfelelően működik. Kimutatható volt az is, hogy a vakok a feldolgozás kezdeti szakaszában egy más, az emlékezeteti és végrehajtó funkciókért felelős, a számfeldolgozásnál általánosabb feladatokat ellátó agyi területre támaszkodnak (3. ábra).

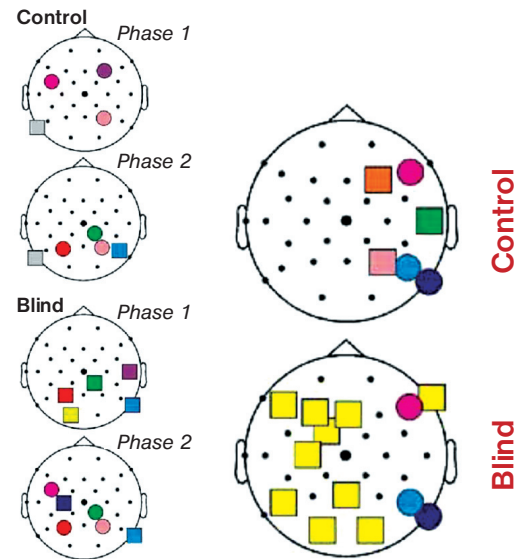
Végül, e rövid válogatást követően, nézzük meg, hozott-e újat az idegtudomány a tanítási módszerekre vonatkozóan. Sajnos meg kell állapítanunk, hogy a tanítási módszerek az agyi aktivitásra gyakorolt hatását meglehetősen kevesen vizsgálták. Delazer és munkatársai [23] két tanítási módszert hasonlítottak össze. Az egyik a gyakorlást, és a számítási tények memorizálását alkalmazta, a másik az alapelvek és az aritmetikai relációk megértetésére fektette a hangsúlyt. A viselkedéses mutatók azt jelezték, hogy a két csoport teszteredményeiben nincs különbség, azaz ebből a szempontból mindegy, hogy milyen tanítási módszert alkalmaznak. Az agyi aktivitásmutatók azonban lényeges különbséget jeleztek a két csoport között; Lee és munkatársai [24] az algebrafeladatok megoldására a szingapúri iskolákban alkalmazott két módszert hasonlította össze. Az egyik az úgynevezett diagram módszer (a tanulóknak a szöveges feladatok megértéséhez diagramokat kell rajzolniuk, innen a módszer neve), a másik pedig a szimbolikus (a szöveges feladatot szimbólumok használatával egyenletté kell alakítani). Lee és munkatársai is azt találták, mint már oly sokan: a viselkedéses mutatókban nem volt különbség a két csoport között. A szimbolikus módszert alkalmazó feladatnál az agyi figyelmi rendszerhez tartozó területek (szuperior parietális lebenyke, és a precuneus nevű terület) aktivitása jelentősen megnőtt, jelezve, hogy ez a módszer sokkal nagyobb figyelmi bevonódást igényel. A viselkedéses és idegtudományi módszerek egymást kiegészítő alkalmazás azt jelzi, hogy azonos teljesítmény nagyon eltérő agyi aktivitással járhat együtt. Ez azért is lényeges, mert az életkornak és a lehetséges stratégiáknak megfelelően alkalmazott módszerek a matematika, s feltételezhetően a feladatrendszerét tekintve nagyobb agyi hálózatra támaszkodó természettudományi kompetencia tanításában is jobb és tartósabb eredmények hozhatnak.

Olvasás és matematika

Kulturális találmányaink, vagyis az olvasás és az elemi matematika formális keretek között történő elsajátítása azonos időben, az iskolába lépéssel kezdődik. Az olvasástechnika fejlődésének vizsgálata, majd a szövegértési készségek mérése során az elmúlt években egyre gyakrabban vetődött fel az a kérdés, hogy van-e szoros összefüggés a szövegértési teljesítmény és a szöveges feladatok megoldása között. Egy egészen friss, közlés alatt álló közlemény [25] megállapítása szerint az olvasási teljesítmény egyes mutatói eltérően képesek előre jelezni a szöveges matematikai feladatok sikeres megoldását. Finn kutatók 99 nyolcadik osztályos tanulót vizsgálva azt találták, hogy a fiúknál a szövegértésben mutatott telje-

sítmény és a matematikai feladatok sikeres megoldása között szoros volt az összefüggés, magának az olvasástechnikának szintje (folyékony, jó tempójú olvasás) viszont mindkét nemnél jól jelezte mind a számolási, mind a szövegesfeladat-megoldási készségeket.

Talán meglepő lehet, de az adatok egyértelműen bizonyítják, hogy a matematikatanulás sikerét befolyásolja az olvasástechnika jó szintje. Ahhoz, hogy ezt megértsük, röviden ki kell tennünk, ha csak röviden is, az olvasás agyi feldolgozó rendszerére.



3. ábra. Távolság- (A) és nagysághatás (B) vakoknál (blind) és az illesztett kontrollnál (control). A körök azokat a helyeket jelölik, ahol a két csoport agyi válasza hasonlóak, a négyzetek azt, ahol eltérőek. Jelentős különbség látható a távolsághatás esetében a feldolgozás első szakaszában (Phase 1), valamint a nagysághatás esetében ([22] alapján)

Az olvasás agyi hálózata tanulás, az agyi plaszticitás miatt lehetővé tett funkcionális szakosodás eredménye. A feladatösszetevőkre specializálódó feldolgozókörok fejlődését követve könnyen belátható, hogy az egyes alapfunkciókat egyre jobban specializáló területek látják el, más agyi területek viszont egyre több, korábban erősen specializáltak tartott funkcióért is felelősek.

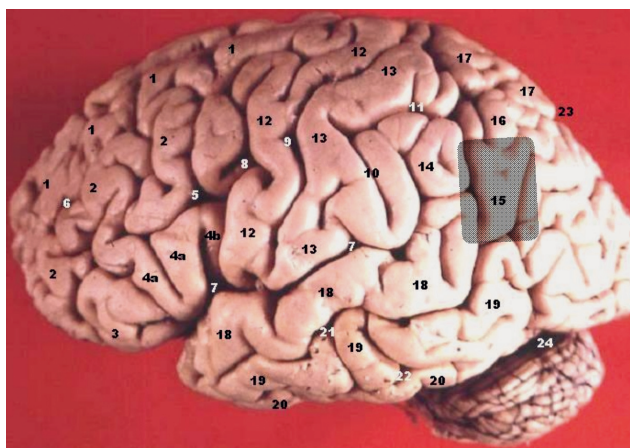
Az olvasásban több agykérgi feldolgozó kör vesz részt, s ezek közül a legnagyobb változás a látási (pl. betűk) és a hallási (pl. beszédhangok) ingerek összekapcsolódásában meghatározó területen (dorzális feldolgozó kör) figyelhető meg. Olvasási feladatokban például az e körhöz tartozó kérgi területeken a modern képalakító eljárásokkal mérhető aktivitás változása a gyerekeknél igen magas, a felnőtteknél ennél jóval alacsonyabb. A felnőttek és a gyerekek agyi aktivitásmintázatainak összehasonlításával egy specializáló agyi hálózatot tudunk változásában, átszerveződési folyamatában megragadni. Az említett finn adatok arra utalnak, hogy lehetnek olyan kérgi területek, amelynek változása mind az olvasás, mind a szöveges matematikai feladatok megoldásában érintett, így lehetséges magyarázatát adja a talált összefüggésnek. Valójában egy olyan terület van, amely mind az olvasásban, mind a számfeldolgozásban kitüntetett szerepet tölt be, s ez a dorzális feldolgozó körhöz tartozó anguláris tekervény (4. ábra). Egy nemrégiben publikált eredmény megerősíti ezt a feltevésünket. Rosenber-Lea és munkatársai [26] 2. és 3. osztályos gyerekek mentális aritmetikai teljesítményét agyi képalkotó eljárásos vizsgálva azt találták, hogy a két rendszerhez egyaránt tartozó anguláris tekervény mutatta az egyik legnagyobb változást. Fontos megjegyez-

Kognív fejlődés, agyi komplexitás, matematika

nünk, hogy az olvasás fejlődésének agyi vizsgálatai, ezekre itt most nem térünk ki, ugyanitt találják a legnagyobb változást. Bár nagyon új területen jár itt az idegtudomány, s adataiból sem lehet a gyakorlatra vonatkozóan azonnal következtetést levonni, egy megállapítható; az olvasástechnika s a szövegértés fejlődése befolyásolhatja az elemi matematika fejlődését.

Kihívások és aggodalmak

A kognív idegtudomány és az oktatáskutatók egy évtizedes párbeszéde új korszakot nyitott. Ennek egyik pozitív hozama, hogy a tanítási módszerek hatásának eddig nem látható összetevőit a módszerfejlesztők egy része egyre jobban ismeri, s az eredményeket saját szakmai szűrőjén áteresztve értékeli. Az agykutatási eredmények végletes leegyszerűsítése azonban tévutakra vezet, s az agyra való hivatkozás (jobbféltekés tanítás, agytorna, agykutatásra alapozott oktatás stb.) nem egyszer megalapozatlan elvárást gerjeszt. Ez különösen igaz a matematika és a természettudományok területén. Az, hogy az idegtudomány az egyes szaktudományok tudásrendszerének agyi mechanizmusait még alig ismeri, óvatosságra int. Ugyanakkor az idegtudomány által feltárt összefüggések az oktatási módszerek kialakításában jól használhatók. Az idegtudomány és a matematikaoktatás területén végzett interdiszciplináris



4. ábra. Az olvasás és számfeldolgozás, aritmetika agyi hálózatához egyaránt tartozó terület, az anguláris tekervény (Forrás: Wikimedia Commons)

kutatások eredményei ma még elsősorban a számok reprezentációjáról, az aritmetikai műveletek agyi korrelátumairól nyújtanak megbízható ismereteket, s egyelőre jóval kevesebb adatunk van a komplexebb, magasabb szintű matematikai készségek agyi jellemzőiről. A magasabb szinten valódi kihívást jelentenek a kutatók számára, hiszen az illesztendő szaktudományi ismeretek komplexitása egyenként is igen nagy, együtt pedig olyan méretű, amelyhez új kutatási stratégia szükséges. Sajátos kihívást jelent már önmagában az is, hogy a kognív pszichológiai és az oktatástudományi elméleteknek olyan mélységben kellene azonosítani az adott tudásterület, esetünkben a matematika szempontjából meghatározó kognív folyamatokat, hogy azok az idegtudomány empirikus feltételeinek megfeleljenek. Ez természetesen nem csupán az idegtudományi kutatások céljait szolgálná, hanem azt is, hogy az oktatástudomány önmaga számára is feltárja a komplex matematikai készségek rendszerét.

További kihívást jelent az idegtudományi eredmények külső megbízhatósága, ami gyakori kérdése ma a kognív idegtudománynak, hiszen a feladatok a mérhetőség szempontjait követik, így igen

egyszerűek, a vizsgálatokban résztvevők száma pedig korlátozott. Az iskolai tanulás azonban nagyon más, hiszen befolyásolja számos környezeti tényező. Ha azt szeretnénk, hogy a hangzatos „idegtudományra alapozott oktatás” megvalósuljon, olyan vizsgálatokra lenne szükség, amelyek azt tárják fel, hogy az idegtudományi kutatások megállapításai mennyire érvényesek az osztályban folyó tanulásra, esetleg képesek-e azt előjelezni. ❄

Kutatásainkat az OTKA K101087-es számú pályázata támogatja.

IRODALOM

- [1] Mayer, R. E. (1998). *Does the brain have a place in educational psychology?* Educational Psychology Review, 10, 389–396.
- [2] Ansari, D., Coch, D. (2006). *Bridges over troubled waters: Education and cognitive neuroscience.* Trends in Cognitive Sciences, 10, 146–151.
- [3] Szűcs, D., Goswami, U. (2007). *Educational neuroscience: Defining a new discipline for the study of mental representations.* Mind, Brain and Education, 1, 114–127.
- [4] Vizi, S. E., Lajtha, A., 2008. *Handbook of Neurochemistry and Molecular Neurobiology: Neurotransmitter Systems.* 3rd ed. Springer, New York.
- [5] Diamond, A., Briand, L., Fossella, J., Gehlbach, L., 2004. *Genetic and neurochemical modulation of prefrontal cognitive functions in children.* American Journal of Psychiatry 161,125–132
- [6] Meyer-Lindenberg, A., Weinberger, D.R., 2006. *Intermediate phenotypes and genetic mechanisms of psychiatric disorders.* Nat. Rev. Neurosci. 7, 818–827.
- [7] Li, Shi-Chen (2013). *Neuromodulation and developmental contextual influences on neural and cognitive plasticity across the lifespan.* Neuroscience and Biobehavioral Reviews, in press
- [8] Tóth, D., Csépe, V. (2008). *Az olvasás fejlődése kognív pszichológiai szempontból.* Pszichológia, 28(1), 35–52.
- [9] Butterworth, B. (1999). *The mathematical brain.* London: Macmillan.
- [10] Dehaene, S. (2003). *A számérzék: miként alkotja meg az elme a matematikát?* Budapest: Osiris Kiadó.
- [11] Xu, F., Spelke, E., Goddard, S. (2005). *Number sense in human infants.* Developmental Science 8, 88–101.
- [12] Grabner, R. H., Ansari, D., Koschutnig, K., Reishofer, G., Ebner, F., Neuper, C. (2009a). *To retrieve or to calculate? Left angular gyrus mediates the retrieval of arithmetic facts during problem solving.* Neuropsychologia, 47, 604–608.
- [13] Grabner, R. H., Ischebeck, A., Reishofer, G., Koschutnig, K., Delazer, M., Ebner, F., et al. (2009b). *Fact learning in complex arithmetic and figural-spatial tasks: The role of the angular gyrus and its relation to mathematical competence.* Human Brain Mapping, 30, 2936–2952.
- [14] Camos, V. (2003). *Counting strategies from 5 years to adulthood: Adaptation to structural features.* European Journal of Psychology of Education, 18, 251–265.
- [15] Geary, D. (1995). *Reflections of evolutions and culture in childrens cognition – implications for mathematical development and instruction.* American Psychologist, 50, 24–37.
- [16] Moyer, R., & Landauer, T. (1967). *Time required for judgments of numerical inequalities.* Nature, 215, 1519–1520.
- [17] Stavy, R., Goel, V., Critchley, H., Dolan, R. (2006). *Intuitive interference in quantitative reasoning.* Brain Research, 1073, 383–388.
- [18] Kaufmann, L. (2008). *Dyscalculia: Neuroscience and education.* Educational Research, 50, 162–175.
- [19] Siegler, R. S. (2007). *Cognitive variability.* Developmental Science, 10, 104–9.
- [20] Zamarian, L., Ischebeck, A., Delazer, M. (2009). *Neuroscience of learning arithmetic – Evidence from brain imaging studies.* Neuroscience and Biobehavioral Reviews, 33, 909–925.
- [21] Csuha J., Csépe V. (2009) *Elemi algebrafeladatok hatékonyabb megoldást kísérő elektrofiziológiai változások.* Pszichológia, 29, 101–118.
- [22] Szűcs, D., Csépe, V. (2005) *The parietal distance effect appears in both the congenitally blind and matched sighted controls in an acoustic number comparison task.* Neuroscience Letters 384, 11–16.
- [23] Delazer, M., Ischebeck, A., Domahs, F., Zamarian, L., Koppelstaetter, F., Siedentopf, C. M., et al. (2005). *Learning by strategies and learning by drill. Evidence from an fMRI study.* Neuroimage, 25, 838–849.
- [24] Lee, K., Lim, Z. Y., Yeong, S. H. M., Ng, S. F., Venkatraman, V., Chee, M. W. L. (2007). *Strategic differences in algebraic problem solving: Neuroanatomical correlates.* Brain Research, 1155, 163–171.
- [25] Kyttilä, M., Björn, P. M. (2013). *The role of literacy skills in adolescents' mathematics word problem performance: Controlling for visuo-spatial ability and mathematics anxiety.* Learning and Individual Differences, in press.
- [26] Rosenberg-Lee, M., Barth, M., Menon, V. (2011). *What difference does a year of schooling make?: Maturation of brain response and connectivity between 2nd and 3rd grades during arithmetic problem solving.* Neuroimage, 57, 3 796–808.

www.termeszetvilaga.hu
KLIKKELJEN RÁ!

ERCSEY-RAVASZ MÁRIA-TOROCZKAI ZOLTÁN

A döntéshozatal és a Sudoku káosza

Amikor döntéseket hozunk napi teendőinkkel kapcsolatban, bizonyos korlátozásokat kell kielégítenünk. Míg ezek a döntések elég egyszerűek, a korlátozások kielégítése egészen nehéz feladat is lehet. A tudomány minden területén előfordulnak ilyen feladatok, és az optimális megoldás megtalálására komoly matematikai és számítógépes módszerekre van szükség. Külön tudományág foglalkozik az ilyen számítási feladatok komplexitásának a felmérésével, de ezek a komplexitás-elméletek a digitális számítógépeken kidolgozott algoritmusok hatékonyságát tanulmányozzák [1,2]. Mi ebben a tanulmányban az algoritmusokat olyan folytonos idejű mozgást leíró matematikai, ún. analóg rendszerekkel helyettesítjük, amelyek önmaguktól kielégítik a korlátozásokat, és a végső egyensúlyi állapotuk megadja a feladat megoldását. A feladatok nehézsége véges élettartamú (tranziens) káoszként jelentkezik, és az ún. szokási ráta alapján egy, a Richter-skálához hasonló nehézségi mérték definiálható, amely egyedi feladatok nehézségét képes jellemezni.

A korlátozás-kielégítési feladatok fontosságát mi sem mutatja jobban, mint széles körű alkalmazásaik, szinte bármilyen tudományterületen és ipari ágazatban előfordulnak. Talán a legismertebb ilyen nehéz feladat az utazó ügynök problémája [3]. Képzeljük el, hogy adott néhány város, és az ügynöknek el kell jutnia mindegyik városba egyszer és csakis egyszer (**1A. ábra**). Melyik a legrövidebb út, amely mentén ezt megteheti? A problémának van egy döntési feladatként való megfogalmazása is: létezik-e egy adott távolságnál rövidebb útvonal, amelyen az utazó ügynök eljut mindegyik városba egyszer és csakis egyszer? Ha kevés város adott, mi magunk is könnyen találunk egy jó megoldást, de sok város esetén számítógépes programokkal is költséges a legrövidebb utat megtalálni (**1B. ábra**).

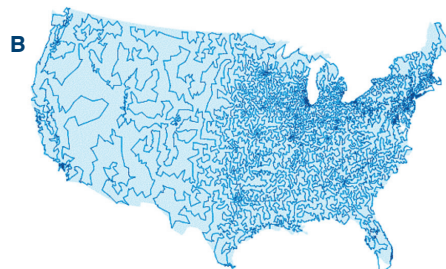
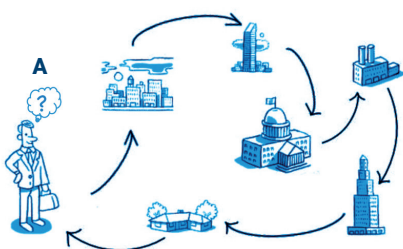
Korlátozás-kielégítési feladatokkal találkozhatunk például a gyógyszerkutatásban is: úgy kell megtalálni a kémiai anyagokat, hogy a gyógyszernek meglegyenek az elvárt hatásai, de minél kevesebb káros mel-

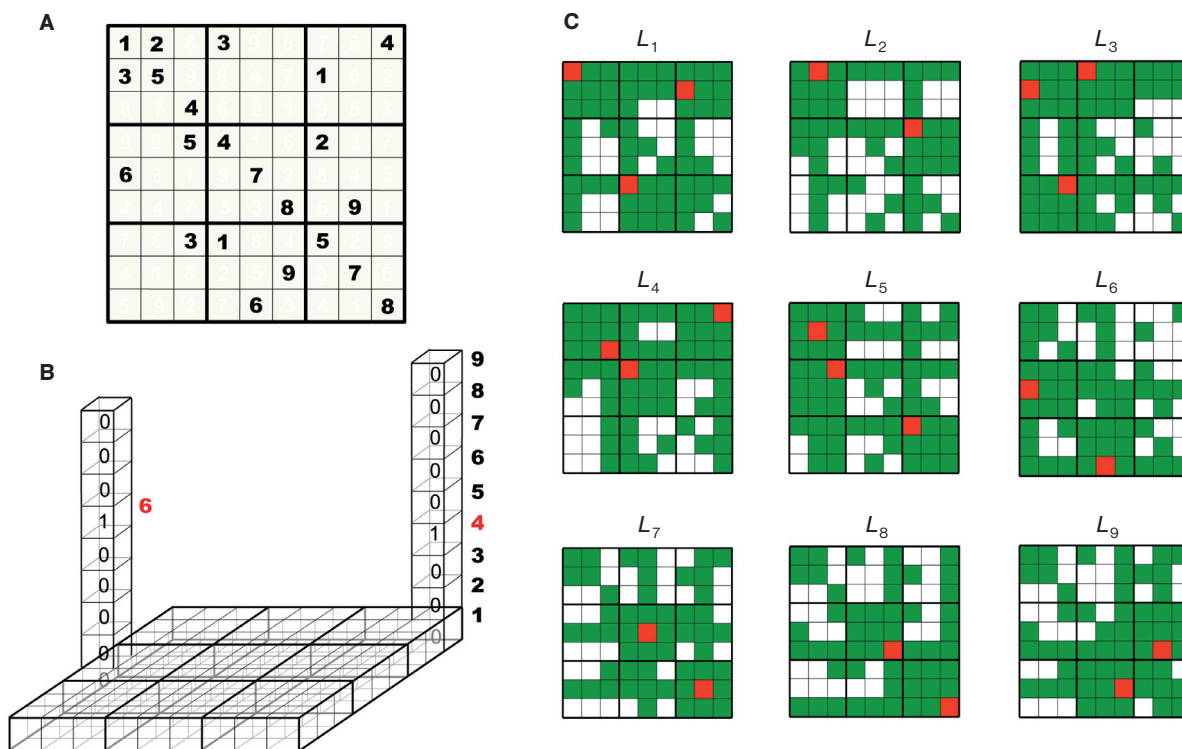
lékhatása legyen. A mérnökök áramkörök tervezésekor találkoznak hasonló feladatokkal, a fizikában például spinűveg-modelleknél van szükség optimalizációs módszerekre: a szomszédos spinek közti kölcsönhatások jelentik a korlátozásokat, és meg kell találni a legkisebb energiájú állapotot [4]. A matematikában logikai korlátozás kielégítésére irányuló feladatokban ÉS, VAGY és NEM műveletekből felépített logikai feltételeket kell kielégíteni [5]. Ezt az irodalomban a SAT-feladatként rövidítik (az angol satisfiability szóból eredően). Sokak által ismert és ugyanakkor matematikailag is jól megfogalmazható korlátozás-kielégítési feladatok a rejtvények. Napjaink egyik legdivatosabb rejtvénye a Sudoku [6] (**2A. ábra**). Ebben egy 9×9 -es négyzetrácsot kell kitölteni számokkal úgy, hogy minden sorban, minden oszlopban és mind a kilenc 3×3 -as elkülönített négyzetben is 1-től 9-ig minden számjegy egyszer és csakis egyszer szerepeljen. Az előre megadott számok korlátozzák, hogy hova milyen számjegyet írhatunk. Ezeket a rejtvényeket úgy terve-

zik, hogy mindig csak egy megoldásuk legyen.

A számítógép-elméletnek külön ága foglalkozik a számítási problémák és döntési feladatok nehézségének/komplexitásának tanulmányozásával, ennek célja a problémák osztályozása a megoldásukhoz szükséges számítógépes források alapján. A mai digitális számítógépeink minden információt $\{0,1\}$ értékekkel tárolnak, és diszkrét lépésekben végeznek el minden műveletet [1]. A költséget a megoldáshoz szükséges idő, pontosabban az algoritmus által végzett lépések száma jelenti. A költség sok mindentől függ, a legnyilvánvalóbb természetesen a feladatok mérete (pl. az utazó ügynöknek hány várost kell bejárnia, vagy hány logikai változó van a SAT-feladatban), de ugyanannyi változós feladatok között is vannak könnyűek, illetve sokkal nehezebbek. Ezért egy feladattípus nehézségét az határozza meg, hogy a *legrosszabb esetben* (legnehezebb feladatoknál) a megoldáshoz szükséges idő hogyan nő a probléma méretével. Ha létezik olyan algoritmus, amelyben a lépések száma polinomiálisan függ a feladat méretétől, azt mondjuk, hogy az hatékonyan megoldható, és a feladat az ún. P osztályba tartozik. Léteznek ellenben olyan döntéshozatali problémák, amelyekre nem ismerünk hatékony algoritmusokat, annak ellenére, hogy a megoldás helyessége polinomiális idő alatt ellenőrizhető le. A hatékonyan leellenőrizhető megoldású problémák osztálya az ún. NP osztály, amibe a P osztály is beletartozik. Viszont az NP-nek van egy alosztálya, az NP-teljes problémák, amelyekre nem ismert hatékony megoldó algoritmus: minden ismert algoritmus exponenciálisan növekedő lépésszámot igényel. Az NP-teljes osztályba tartozó feladatok között szerepel a SAT és az utazó ügynök problémája is (döntéshozatali formában). Az

1. ábra. A. Az utazó ügynöknek meg kell találnia a legrövidebb utat, amely minden célpontot pontosan egyszer érint. **B.** Sok város esetén a feladat számítógépekkel is nagyon költséges. Az ábrán 13509 amerikai várost érintő legrövidebb út látható, ennek megoldása 1998-ban rekordot jelentett [3]. A számítógépek teljesítménye sokat nőtt azóta, a módszerek is fejlődtek, de kvalitatív változás nem történt, a nagy feladatok megoldása most is komoly gond





2. ábra. A. Ebben a Sudoku rejtvényben $d = 22$ számjegy adott. B. Minden számjegyet egy 9 változóból álló logikai tömbként tárolunk, ezek a függőleges oszlopok. C. A 9 vízszintes réteg mindegyike megfelel egy számjegynek L_i , $i = 1, \dots, 9$. Az előre megadott számjegyek által rögzített változók négyzeteit megszíneztük (piros IGAZ (1), zöld HAMIS (0)). A fehér négyzetekben maradnak a feladat ismeretlen változói. Ebben az esetben $N = 232$ változó és $M = 236$ feltétel marad

NP-teljes problémák fontos tulajdonsága, hogy egyetlen ilyen feladatra kidolgozott hatékony algoritmus biztosítaná az összes NP feladat hatékony megoldását (a „teljesség” innen ered). Például a Cook–Levin-tétel alapján az összes NP feladat hatékonyan átalakítható SAT formába – megfogalmazható logikai korlátozásokkal – ezért a SAT hatékony megoldása az összes NP feladat hatékony megoldását biztosítaná. Továbbra is nyitott kérdés, hogy léteznek-e ilyen algoritmusok. A Clay Matematika Intézet 2000-ben kitűzött Millennium-díj hét feladata között – amelyek megoldásáért egyenként egymillió dolláros jutalmat ígér – ez is szerepel. Az általános nézet az, hogy a P osztály nem egyenlő az NP-vel (vagyis nem léteznek hatékony algoritmusok az NP-teljes feladatokra), ellenben valószínű, hogy ennek bizonyítása még sokat vár magára.

Talán furcsának tűnhet, hogy a fizika segíthet ilyen feladatok megoldásában, de megfelelő tulajdonságú fizikai rendszerek pillanatok alatt végeznek el olyan optimalizációs feladatokat, amelyeket mi számítógépekkel évek alatt sem tudnánk kiszámolni. Jó példa erre a fehérhajtogatás. Ahhoz, hogy egy több száz aminosavból álló lánc

a sejtben funkcióképes legyen, fel kell vennie a megfelelő háromdimenziós alakot (alapállapot). Csak speciális struktúrájú fehérjék képesek gyorsan és következetesen megtalálni ezt az optimális alakot, de még ezek esetében is akár 10^{16} -szor is lassúbbak a számítógépes szimulációk, mint maga a fizikai folyamat. Mivel a digitális (CMOS) technológia fejlődése lassan eléri határait, egyre több kutatás irányul analóg, folytonos idejű számítógépek tervezésére, amelyek folytonos időbeli mozgásokat (folyamszerű dinamikát) próbálnak számításokra használni [7,8]. Annak ellenére, hogy az ilyen rendszerek sokkal gyorsabbak, mint digitális számítógépeink, vigyázni kell arra, hogyan definiáljuk a költséget, amikor a feladatok komplexitását tanulmányozzuk. Itt már nemcsak az időt, hanem a befektetett energiát is vizsgálni kell a költség mérésékor (a digitális gépeken ez nem számít, mivel minden lépés azonos energiát fogyaszt függetlenül a feladattól).

Mi egy olyan tulajdonságokkal jellemezhető, folytonos idejű mozgást leíró matematikai rendszert dolgoztunk ki, amellyel SAT-feladatokat lehet megoldani: a rendszer egyensúlyi állapotai megfelelnek a SAT-feladat megoldásainak, és a mozgást bármilyen kezdeti állapotból indítva, ez addig „keres”, amíg be nem ragad egy egyensúlyi állapotba, amely már egy megoldásnak felel meg [9]. Érdemes megjegyezni, hogy bi-

zonyos fizikai rendszerekkel elérhető az, hogy növekvő energiabefektetéssel NP-teljes (vagyis nehéz) feladatokat folytonos időben hatékonyan (polinomiálisan) megoldjunk [9]. Nehéz feladatoknál a megoldás elérését megelőző dinamika (mozgási folyamat) egészen bonyolult lehet, és a folytonos idejű rendszerben fellépő tranzienst káosz tükrözi a feladat nehézségét. A tranzienst káosz* élettartama, vagy ennek közel reciproka, a szökési ráta alkalmazható mint a feladat nehézségének mértéke [10]. Mivel a SAT egy NP-teljes feladat, a korlátozáskielégítési feladatoknak egy nagy osztálya hatékonyan átalakítható SAT-feladattá, ezért a módszer sok helyen alkalmazható. Itt a Sudoku rejtvényen szemléltetjük a módszert és a tranzienst káosz tulajdonságai alapján definiált nehézségi mérték érvényességét.

A Sudoku mint logikai SAT-feladat

Annak ellenére, hogy a Sudokuban 1-től 9-ig vannak a számok, könnyen átalakítható logikai SAT-feladattá, amelyben csak 1 vagy 0 (IGAZ vagy HAMIS) logikai változók szerepelnek. Azt, hogy egy kis négyzetben melyik számjegy van, tárolni tudjuk egy 9 logikai változóból álló tömbbel,

* Erről a jelenségről szól *Gruiz Márton* írása is.

¹ Ha N a probléma mérete, akkor annak valamely hatványával (N^n) arányosan nő az idő.
² Ebben az esetben N a hatványkitevőben szerepel: pl. 2^n vagy 10^n szerint nő.

amelyben mindig egy és csakis egy 1-es áll. Pl. ha a számjegy 6, akkor a hatodik helyen levő változó lesz igaz és az összes többi hamis. Úgyis elképzelhetjük, mintha a 9×9 -es négyzetrács helyett most $9 \times 9 \times 9$ kockánk lenne (**2B. ábra**), a Sudoku minden kis négyzetéhez tartozik egy függőleges oszlop, és minden kis kockában $\{0,1\}$ logikai változók vannak. Ezeket jelöljük az $x_{i,j}^a$ tömbbel, ahol (i,j) a kis négyzet koordinátája a Sudokuban (a bal felső sarokból számozva), és a jelöli, hogy hányadik rétegben (L_a) levő logikai változóról beszélünk (lentől felfelé számozva). A **2A. ábrán** a Sudoku (1,9) négyzetében 4-es számjegy szerepel, ez azt fogja jelenteni, hogy $x_{1,9}^4 = 1$ és $x_{1,9}^a = 0$ minden $a \neq 4$ értékre.

Hogyan fogalmazhatjuk meg a Sudoku rejtvény szabályait mint logikai korlátozásokat? Például az a feltétel, hogy a 4-es számjegy egy adott sorban csak egyszer fordulhat elő, azt jelenti, hogy az L_4 rétegben ebben az adott sorban pontosan egy igaz logikai változó lehet. Így az L_4 rétegben 9×3 ilyen feltétel lesz, mert mind a kilenc sorra, oszlopra és 3×3 -as négyzetre teljesülnie kell, hogy a 4-es számjegy pontosan egyszer szerepelhet. Ezeket a feltételeket az összes számjegyre (mind a kilenc rétegben) fel kell írni, így összesen $9 \times 9 \times 3$ ilyen feltételünk lesz. Ezek mellett mind a 9×9 függőleges oszlopban (toronyban) is pontosan egy változó lehet igaz, hiszen a Sudoku egy kis négyzetébe mindig egyetlen jól meghatározott számjegy van (nem lehet egy oszlopban egyszerre több számjegynek megfelelő logikai változó igaz). Így összesen a Sudoku megfogalmazható mint $9 \times 9 \times 4 = 324$ logikai feltétel.

Amikor adott egy rejtvény, az előre beírt számjegyek elég sok logikai változót le-rögzítenek a $9 \times 9 \times 9$ -es kocka 729 eleméből. Ezeknek a számától és elhelyezésétől függ, hogy hány ismeretlen változó marad, és pontosan emiatt ez a Sudoku nehézségével is összefügg. A **2C. ábrán** kirajzoltuk a Sudoku változótömbjének 9 rétegét. A rejtvényben megadott számok miatt le-rögzített logikai változók helyeit megszíneztük: piros az IGAZ és zöld a HAMIS. Az ismeretlen változókat fehéren hagytuk. Például a **2A. ábrán** mutatott rejtvényben az (1,9) helyen megadott 4-es számjegy igaz értéként rögzíti az L_4 -es rétegben az (1,9) helyen szereplő logikai változót, ugyanakkor hamis értéket kötelez az L_4 -es réteg adott sorában, oszlopában és 3×3 -as kockájában az összes többi változónak. A többi rétegben az (1,9) helyen szereplő változóknak is hamisaknak kell lenniük. Ebben az adott rejtvényben összesen $d = 22$ szám volt megadva, így a **2C. ábrán** 22 piros négyzet van. Ezenkívül 475 zöld négyzetünk van, és a fehér négyzetek, vagyis az ismeretlen változók száma $N = 232$. A logikai feltételek száma is lecsökken, mert lesznek már olyan sorok, oszlopok és 3×3 -as

négyzetek, amelyekben nem marad fehér négyzet (ismeretlen változó). A **2. ábrán** mutatott példában végül $M = 236$ feltétel maradt. Ugyanakkor a feltételekben szereplő ismeretlenek száma is jóval kevesebb, mint 9 (pl. az L_1 réteg harmadik sorában csak két fehér négyzet van, ami azt jelenti, hogy ebből a két változóból csak az egyik lesz igaz). Annak ellenére, hogy a d összefügg a rejtvény nehézségével, látható, hogy a megadott számok elhelyezésétől is függ, hogy éppen mennyi ismeretlen változó és hány feltétel marad.

Ahhoz, hogy az általunk tervezett időbeli mozgás meg tudja oldani ezt a logikai feladatot még egy átalakításra lesz szükség. A Sudokuban szereplő korlátozásokat a SAT feladatokban használt standard formába kell alakítani, vagyis csak ÉS, VAGY, illetve NEM műveletekből álló logikai feltételeket használhatunk. Mivel minden logikai feltételt át lehet alakítani ilyen formába, ez nem jelent problémát (a technikai részletek az **1. blokkban**).

A SAT-megoldó rendszer

Ha eldobunk egy labdát, annak mozgását (dinamikáját) a labda három koordinátájának időbeli változásával írhatjuk le. Ezt a mozgást fizikai törvények határozzák meg, úgymint gravitáció, közegellenállás stb. Egy idő után a labda leáll, eljut egy egyensúlyi állapotba, ahol a koordinátái már nem változnak tovább. Mi a SAT-feladatot egy olyan virtuális térbeli mozgással oldjuk meg, amelyben a koordinátáknak megfelelnek a SAT ismeretlen változói, és az egyensúlyi állapotban ezeknek a koordinátáknak (változóknak) az értékei megadják a feladat megoldását. Ellenben a mozgást leíró törvényeket mi határozzuk meg, és a kihívás úgy megtervezni ezt a mozgást, hogy bármilyen kezdeti állapotból is indulva olyan egyensúlyi állapotba jusson a rendszer, amely egy megoldásnak felel meg. Egy általános SAT-

feladatban adott N számú logikai változó $x_i = \{0,1\}$, $i = 1, \dots, N$, de mivel egy folytonos időben és folytonos térben történő mozgást tervezünk, ezért ezeket megfeleltetjük folytonos $s_i \in [-1,1]$ intervallumban mozgó változóknak. Az $s_i = -1$ azt jelképezi, hogy a logikai változó hamis, ha $s_i = +1$ akkor igaz, de nekünk fontos, hogy halagunk az értékek folytonosan változhatnak a $[-1,1]$ intervallumban, mert így definiálni tudunk egy analóg költségfüggvényt minden korlátozásra. Ha ez a költség nulla, akkor a változók értékei olyan logikai értékeknek felelnek meg, amelyek kielégítik az adott korlátozást. A feladat megoldását megadó értékeknel az összes korlátozás költségfüggvénye nulla kell, hogy legyen (részletek a **2. blokkban**).

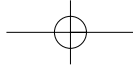
A fizikai mozgások analógiájával élve ez a költségfüggvény egy energiának felel meg, ezért természetes lenne olyan mozgást definiálni, amely ezt az energiát a lehető leghatékonyabban csökkenti. Az s változók terének minden pontjában kiszámolható egy energiaérték ($E(s)$), ez egyfajta felületet határoz meg, és az ezen mozgó pont mindig a legmeredekebb lejtőn haladna a legközelebbi gödör felé (**3A. ábra**). Amikor már nem tudja tovább csökkenteni az energiát, beragad a gödör mélyébe, vagyis a rendszer egy egyensúlyi állapotba jutott. Ellenben a sokdimenziós, bonyolult felületeken sok olyan gödör van, amelynek a mélyén az energia nem nulla. Ezeket az egyensúlyi állapotokat nem tekinthetjük olyan megoldásoknak, amelyek kielégítenék az eredeti célunkat. A cél elérése érdekében *segédváltozókat* vezetünk be: minden korlátozás kap egy súlyt a költségfüggvényben (a_m , $m = 1, \dots, M$, ahol M a korlátozások száma), és ez a súly időben változik. Minél hosszabb ideig nem volt kielégítve egy korlátozás, annál jobban nő a súlya, annál fontosabbá válik, és ezt a súlyozott költséget (jelöljük ezt a t pillanatban $V(s)$ -sel) leghatékonyabban csökkentő mozgás olyan irányba fog haladni, hogy ezeket a legfontosabb korlátozásokat minél ha-

1. A Sudoku feltételeinek átalakítása standard SAT-formába

A Sudoku korlátozásait úgy kell átfogalmazni, hogy minden feltételben belül csak VAGY műveletek szerepelhetnek (\vee jel) a változók, illetve azok tagadott formái között. Pl. egy ilyen feltétel lehetne: $C_1 = (x \vee y \vee \bar{z})$, ahol a fölülvonás a változó tagadását jelöli. A teljes feladat ilyen feltételek ÉS-sel (\wedge jel) összekötött sorozata, $F = C_1 \wedge C_2 \wedge \dots \wedge C_M$ teljesítése, vagyis az összes M feltételt egyszerre ki kell elégíteni. Bármilyen logikai feltétel átalakítható ilyen formába. A mi esetünkben egy konkrét korlátozás azt mondta ki, hogy a benne szereplő változók közül egy és csakis egy lehet igaz. Pl. a **2C. ábrán** látható, hogy az L_3 réteg harmadik sorában 3 ismeretlen változó van: $x_{3,7}^3, x_{3,8}^3, x_{3,9}^3$. Azt, hogy ezek közül pontosan egy lehet igaz, négy ilyen feltétellel lehet felírni a következőképpen:

$$(x_{3,7}^3 \vee x_{3,8}^3 \vee x_{3,9}^3) \wedge (\bar{x}_{3,7}^3 \vee \bar{x}_{3,8}^3) \wedge (\bar{x}_{3,7}^3 \vee \bar{x}_{3,9}^3) \wedge (\bar{x}_{3,8}^3 \vee \bar{x}_{3,9}^3). \quad (1)$$

Az első feltétel azt fejezi ki, hogy legalább egy változó igaz, a másik három pedig, hogy bármelyik kettő közül nem lehet egyszerre mindkettő igaz.



2. A SAT-megoldó mozgás

Először az $x_i = \{0,1\}$ logikai változók helyett folytonos értékeket vezetünk be: $s_i \in [-1,1]$. $s_i = -1$ felel meg a 0 (HAMIS) értéknek, az $s_i = 1$ pedig az 1 (IGAZ) értéknek, ellenben a kettő között bármilyen értéket megengedünk a dinamika során. A SAT feladatot egy mátrixszal tudjuk kódolni: $c_{mi} = 1$, ha az i változó a tagadás nélküli alakjában szerepel a C_m feltételben (l. **1. blokk**), $c_{mi} = -1$, ha a tagadása szerepel, és $c_{mi} = 0$, ha egyáltalán nem része a C_m feltételnek. Legyen k_m a feltételben szereplő változók száma. Így a $K_m(\mathbf{s}) = 2^{-k_m} \prod (1 - c_{mi}s_i)$ költségfüggvény $[0,1]$ közötti értékeket vehet fel. Nulla lesz, ha van olyan változó, ami kielégíti a feltételt, egy, mikor egyik sem elégíti ki. Így a teljes SAT-feladat költségfüggvénye $E(\mathbf{s}) = \sum_m K_m(\mathbf{s})^2$.

A legegyszerűbb mozgás, amelyet az \mathbf{s} változóknak definiálhatunk az a gradiens irányú mozgás (piros nyilak a **3A. ábrán**), ellenben nehéz feladatoknál a költségfüggvénynek nagyon bonyolult alakja lehet az \mathbf{s} térben, könnyen megesik, hogy a változó egy lokális optimumba ragad be. Ahhoz, hogy ezt elkerüljük, segédváltozókat vezetünk be, amelyek időben változtatják ezt az energiafelületet, és nem engedik beragadni az \mathbf{s} változókat egy lokális minimumba. Minden C_m feltételhez rendelünk egy időfüggő $a_m(t)$ segédváltozót, és az új energiafüggvény: $V(\mathbf{s}, \mathbf{a}) =$

$\sum_m a_m K_m(\mathbf{s})^2$. A rendszer időbeli változását a következő differenciálegyenlet-rendszer adja meg [9]:

$$\frac{ds_i}{dt} = -\frac{\partial V(\mathbf{s}, \mathbf{a})}{\partial s_i}, \quad i = 1, \dots, N \quad (2)$$

$$\frac{da_m}{dt} = a_m K_m(\mathbf{s}), \quad m = 1, \dots, M, \quad (3)$$

ahol a kezdeti $t = 0$ állapotban $s_i(0) \in [-1,1]$ minden i -re és $a_m(0) > 0$ minden m -re. Láthatjuk, hogy az \mathbf{s} változók a $V(\mathbf{s}, \mathbf{a})$ energiát csökkenthetik, az \mathbf{a} változók viszont csak növekedni tudnak. Ha egy C_m feltétel ki van elégítve, akkor a költségfüggvénye $K_m(\mathbf{s}) = 0$, és az a_m változó nem nő. Ellenben ha a költségfüggvény $K_m(\mathbf{s}) > 0$, akkor a megfelelő a_m segédváltozó exponenciálisan növekszik. Míg az $E(\mathbf{s})$ energiának az \mathbf{s} tér (a $[-1,1]^N$ hiperkocka) minden pontjában időben rögzített értéke van, a $V(\mathbf{s}) \equiv V(\mathbf{s}, \mathbf{a}(t))$ energiafelület időben változik a segédváltozók dinamikája miatt. Valahányszor az értékük eljut egy olyan térrészbe, a következő pillanattól lokális minimum (egy völgy a **3A. ábrán**), azok az a_m segédváltozók kezdenek növekedni, amelyeknek megfelelő korlátozások abban a térrészben nincsenek kielégítve. Ez megnöveli ott az energiát, amíg végül a változó kimozdulhat a völgyből (a **3A. ábra** t_3 pillanatában).

marabb kielégítse. Mivel a megoldásnak megfelelő eredeti energia (költség) nulla volt, a súlyozott költség is nulla marad. Úgy lehet elképzelni, hogy az új $V(\mathbf{s})$ energiafelület időben változik, mint egy hullámzó szőnyeg, a gödrök erre-arra tolódnak, egyedül a megoldásnak megfelelő nulla energiájú gödör marad rögzítve (**3A. ábra**). Minden olyan esetben, amikor a mozgás egy olyan völgy mélyéhez közeledik, amelyik nem igazi megoldás, akkor lesznek olyan korlátozások, amelyek súlya növekedik, ezért az adott pontban egyre nő a $V(\mathbf{s})$ energia, és a völgy lassan eltűnik (**3A. ábra** t_3 pillanata), így a mozgás nem áll le. Csak azok az egyensúlyi állapotok léteznek, amelyekben az energia nulla, vagyis megoldásnak felelnek meg (részletek a **2. blokkban**).

Ezt a mozgást a **3B. ábrán** szemléltetjük egy $N = 50$ változós SAT-feladaton (ez a példa nem egy Sudoku rejtvény, hanem egy véletlenszerűen generált SAT-feladat). A mozgást az s_1, s_2 síkon követjük. A kezdeti időpillanatban $s_i(0) = 0$ minden i -re (fekete kereszt a térképeken). Egy t időpillanatban szintérréppel ábrázoljuk a $V(\mathbf{s})$ energiát, az $(s_1, s_2) \in [-1,1] \times [-1,1]$ négyzet minden pontjában. A szín a V értéket jelöli (lásd a színskálát). Úgy kell elképzelni, mintha a sokdimenziós energiafelületet az (s_1, s_2) sík felett vizsgálnánk. A fekete vonal mutatja a mozgáspályát a $[0, t]$ időintervallumban, vagyis, hogy az (s_1, s_2) változók milyen értékeket vettek fel az eltelt időben. A világoskék pont a pálya $(s_1(t), s_2(t))$ végpontját mutatja. A **3B. ábrán** ilyen térképeket mutatunk néhány kiragadott időpillanatban. Látható, hogy pl. $t \in [81,8, 83,1]$ időintervallumban a pálya beragad egy „energia-

völgybe” az $s_j = -1$ érték környékén, mert ott a V energiaérték kisebb, mint körülötte, de ahogy telik az idő, egyre nő abban a pontban a V (pirosodik a térkép), amíg végül $t = 84,2$ körül kiszabadul abból a térrészből, és átmegegy egy másik völgybe, az $s_j = -1, s_2 = 1$ pont környékén. Ott ismét meg kell várnia, amíg az energia eléggé megnő, de $t = 88,1$ körül onnan is megszökik. Ez az egész folyamat során így működik, és egy nagyon bonyolult pálya leírása után végül $t = 94,6$ -ra eljut a megoldásba, ahol $V(\mathbf{s}) = 0$. Ez már egy stabil fixpont, egyensúlyi állapot, ahol a pálya véglegesen leáll, és a koordináták (\mathbf{s}, \mathbf{a} változók) nem változnak többet.

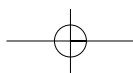
Láthatjuk a **3. ábrán**, hogy a pálya nagyon bonyolult lehet. Valóban, nehéz feladatoknál a mozgás kaotikus viselkedést mutat [11,12]. Mivel a bonyolult pálya ellenére egy idő után a rendszer megtalálja a megoldást, és elér egy stabil egyensúlyi állapotba, a káosz mindenképpen véges élettartamú, ún. tranzien káosz [13]. A pályák bonyolultsága függ a feladat nehézségétől. SAT-feladatoknál, ha kevés feltételünk van, akkor nagyon sok megoldás létezik, és ezeket könnyű megtalálni, így a pálya is gyorsan eljut egy stabil egyensúlyi állapotba. Ha sok korlátozás van, kevesebb a megoldás, nő a pályák bonyolultsága, ugyanakkor a kaotikus viselkedés miatt a kezdeti feltételekre is nagyon érzékeny a mozgás. Ezt szemléltetjük a **4. ábrán**. Legyen a $t = 0$ pillanatban $\mathbf{s}(0) = \{s_1(0), s_2(0), s_3(0), \dots, s_N(0)\}$ egy véletlenszerű kezdeti állapot. Minden $s_i(0)$ értéket lerögzítünk, kivéve az $s_1(0)$ és $s_2(0)$ értékeit, amelyeket egy négyzetrács mentén változtatunk a $[-1,1] \times [-1,1]$ síkban. A térképen egy adott pont színe jelöli,

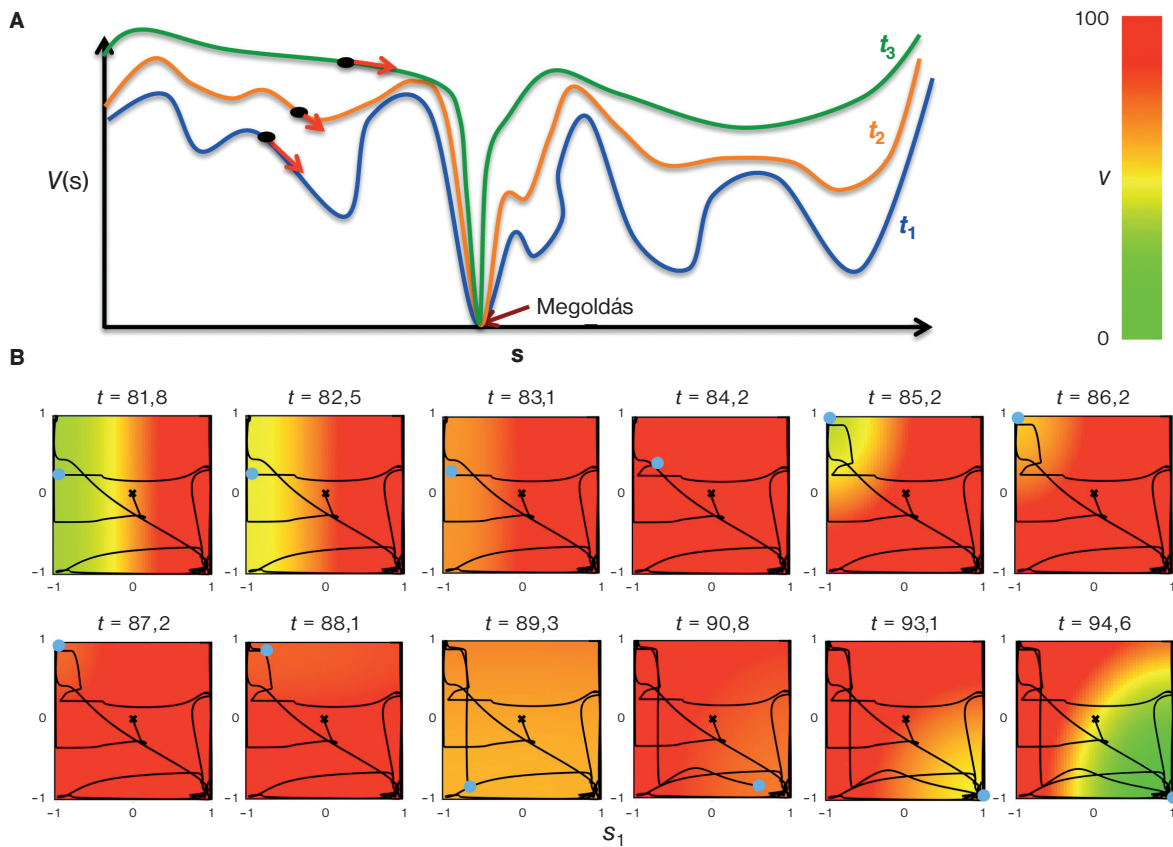
hogy abból a pontból indítva a mozgást melyik megoldást találta meg (melyik egyensúlyi állapotba jut el). Az első térképen egy $N = 50$ változós és $M = 150$ korlátozásból álló feladatot választottunk, a következők térképeken pedig ezt a feladatot bővítjük újabb és újabb feltételekkel (növeljük az M -et). A kezdeti $\mathbf{s}(0)$ tömb ugyanaz marad. Amikor nagyon kevés korlátozás van, a pályák nem érzékenyek a kezdeti állapotra: annak ellenére, hogy sok megoldása van a feladatnak, csak az $s_1(0)$ és $s_2(0)$ értékét változtatva nem változik sokat a pálya, és mindig ugyanabba az egyensúlyi állapotba (megoldásba) jut el, az egész térkép egyszínű lesz. A korlátozások számának növelésével bonyolódik a térkép is, egyre több megoldás színe jelenik meg, és a vonzási tartományok határai is bonyolódnak. Amikor túl sok feltétel van, akkor egyre kevesebb megoldás marad, és mind a mozgás, mind a vonzási tartományok fraktálstruktúrája* egyre bonyolultabb lesz (**4. ábra**).

Tranziens káosz a Sudokuban

Visszatérve a Sudoku rejtvényre, itt is azt tapasztaljuk, hogy a pályák bonyolultsága függ a feladat nehézségétől. A **2. ábrán** látható Sudoku rejtvény elég nehéz feladat. Az **5. ábrán** kirajzoltuk a rejtvény 2. sorának és 9. oszlopának megfelelő változók időfejlődését. A **2C. ábrán** láthattuk, hogy az L_1, L_3, L_4, L_5, L_8 rétegekben az ennek a négyzetnek megfelelő változók rögzítve

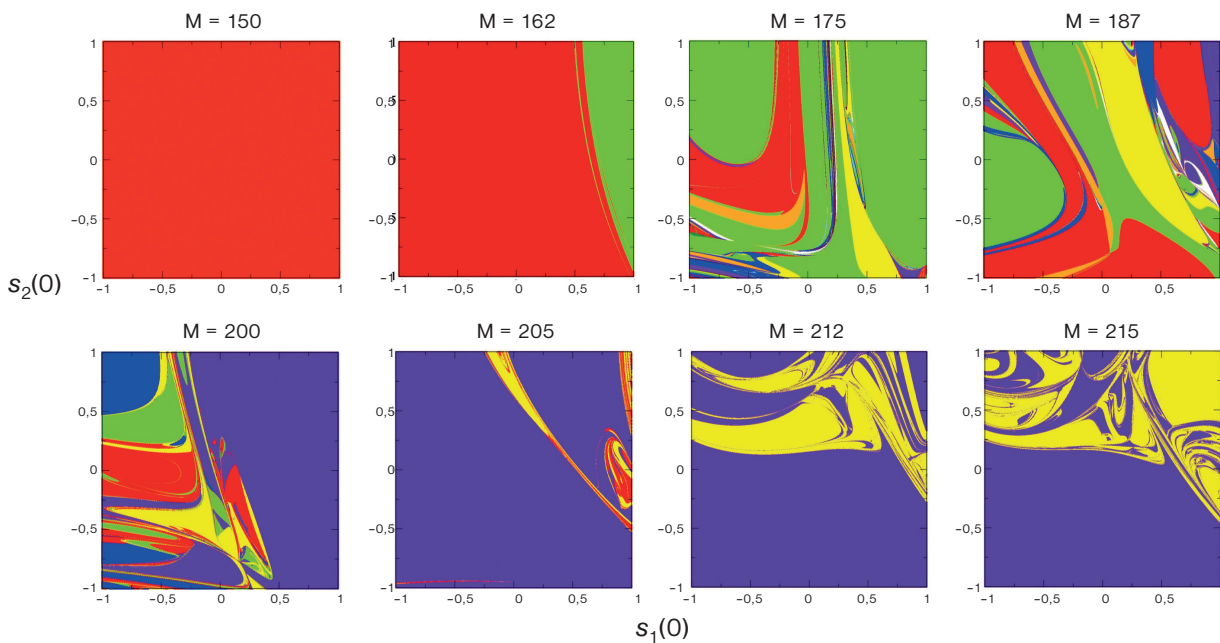
* Hasonlóan a mágneses inga vonzási tartományaihoz, l. *Gruiz Márton* 5. ábrája.





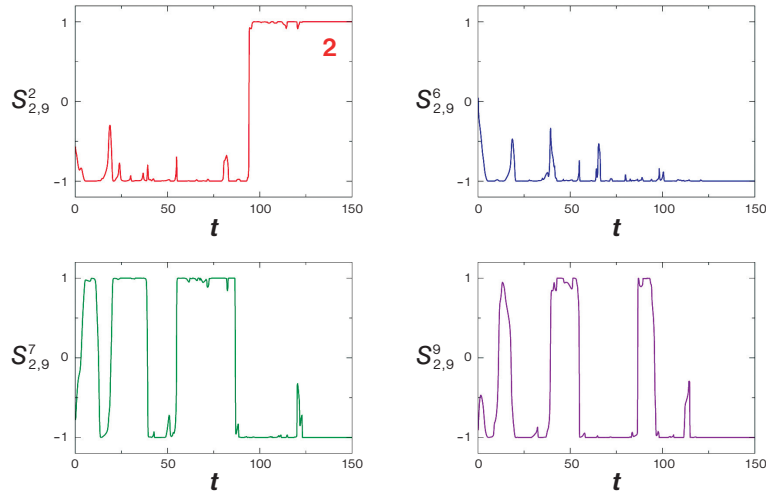
3. ábra. A. Sematikus ábra: A $V(s)$ energiafelület az s térben időben (a t_1, t_2, t_3 pillanatokat ábráztuk) változik a segédváltozók dinamikája miatt. A megoldás pontjában az energia mindig nulla, de a lokális minimumok (völgyek) változnak. Ezen a változó energiafelületen az s változók mozgása mindig a legmeredekebb irányba haladva csökkenti az energiát (piros nyilak). B. Egy $N = 50$ változós SAT-feladatnak az (s_1, s_2) síkban leírt pályája az $s_i(0) = 0$ (minden i -re) kezdő állapotból indítva (fekete kereszt). A szintérvonal mutatja az adott pillanatban a V energia értékét a sík különböző pontjaiban (színskála). A fekete vonal az eddig leírt pályát, a világoskék pont pedig az (s_1, s_2) változók pillanatnyi értékeit jelöli

4. ábra. Egy $N = 50$ változós SAT-feladatot bővítünk újabb és újabb korlátozásokkal (M). Rögzítünk egy kezdeti feltételt $s(0)$, és ennek csak az első két elemét változtatjuk egy négyzettrács mentén: $s_1(0) \in [-1, 1]$, $s_2(0) \in [-1, 1]$. A térképek egy adott pontjából indítva a mozgást, a szín jelöli, hogy melyik megoldásba (egyensúlyi állapotba) jutott el a rendszer



A döntéshozatal és a Sudoku káosza

1	2	6	3	9	5	7	8	4
3	5	9	8	4	7	1	6	2
8	7	4	6	2	1	9	5	3
9	8	5	4	1	6	2	3	7
6	3	1	9	7	2	8	4	5
2	4	7	5	3	8	6	9	1
7	6	3	1	8	4	5	2	9
4	1	8	2	5	9	3	7	6
5	9	2	7	6	3	4	1	8



5. ábra. A 2. ábrán mutatott Sudoku megoldása. A rejtvény 2. sorában és 9. oszlopában levő kis négyzetnek 4 ismeretlen változója van. Ezeknek a pályája látható egy véletlenszerűen megválasztott kezdeti állapotból indítva. Mikor a rendszer megtalálja a megoldást az $s_{2,9}^2 = 1$ és $s_{2,9}^6, s_{2,9}^7, s_{2,9}^9 = -1$, mutatva, hogy a (2,9) négyzetben a megoldás a 2-es számjegy

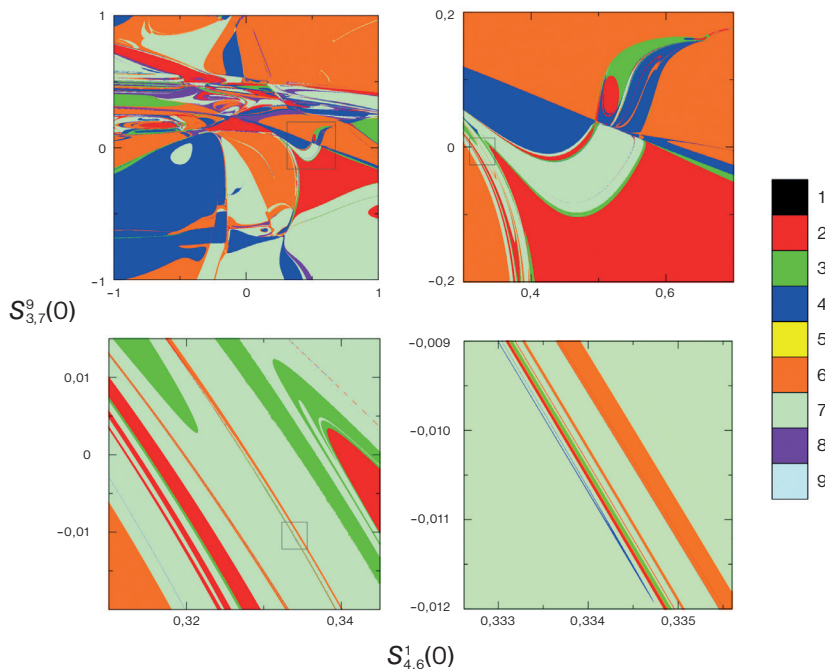
vannak a már megadott számok miatt. Az ismeretlen változók, amelyek részt vesznek a mozgásban: $s_{2,9}^2, s_{2,9}^6, s_{2,9}^7, s_{2,9}^9$ (5. ábra). Amikor a (2), (3) rendszer kb. 150 időegység után megtalálja a megoldást, az $s_{2,9}^2 = 1$ IGAZ-nak megfelelő értéket mutat, míg $s_{2,9}^6, s_{2,9}^7, s_{2,9}^9 = -1$ HAMIS-nak megfelelő értékű. Az 5. ábrán mutatjuk a rejtvény megoldását is, valóban a (2,9)-es négyzetben a

2-es számjegy lesz a megoldás. Egy hasonlóan nehéz rejtvény megoldásának folyamatát láthatják filmen is a [14] linken.

A káosz megjelenését itt is mutatja a kezdeti állapotra való érzékenység. A 6. ábrán a 4. ábrához hasonló térképeket mutatunk a világ egyik legnehezebb Sudoku rejtvényére (lásd [10]). Ismét rögzítve van egy kezdeti feltétel $s(0)$, és abban csak két

értéket változtatunk egy négyzetrács mentén. Itt a két választott változó az $s_{4,6}^1$ és $s_{3,7}^9$, de szinte bármilyen választás hasonlóan bonyolult térképeket eredményezne. Mivel itt egyetlen megoldás van, a pályák érzékenységét most másképp szemléljük. Itt a pontok színe azt jelöli, hogy a Sudoku (6,8)-as elemében a $t=20$ időpillanatban éppen melyik számjegy szerepel, mint igaz változó (lásd a színskálát). Ebben a rejtvényben a (6,8)-as elemnek megfelelő változók közül az 1,5,8,9-es változók vannak rögzítve, ezért ezek a színek nem is jelennek meg a térképeken. A megoldásban pedig a 4-es számjegy szerepel, látható, hogy $t=20$ -nál a térképnek már viszonylag nagy része sötétkék. Az egymás utáni térképeken a halványan bejelölt fekete négyzetek újabb és újabb kinagyítását láthatjuk. A nagyításokkal mindig új, addig nem látható medencehatárok jelennek meg, és ez arra utal, hogy a színek határa fraktál jellegű.

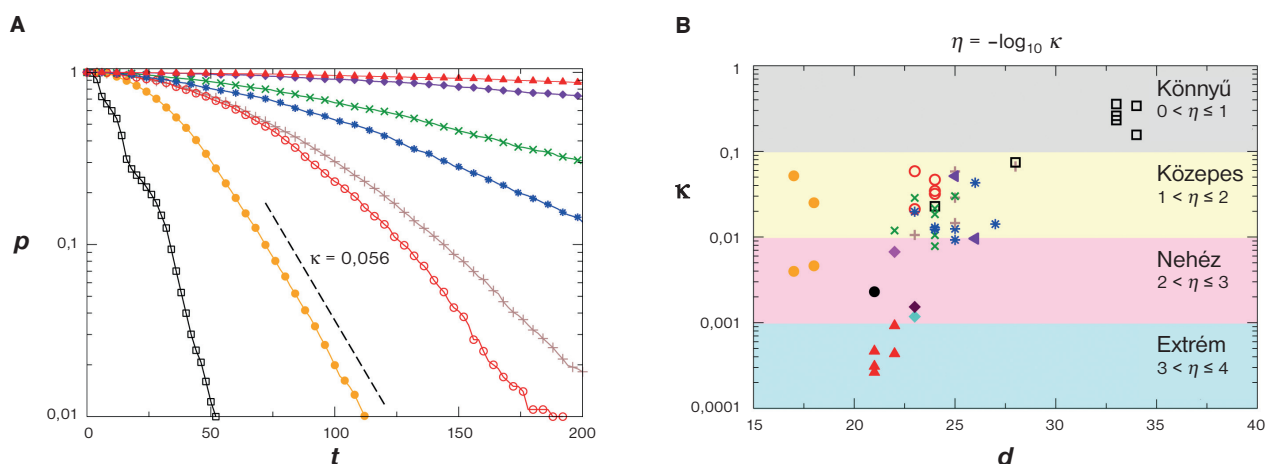
6. ábra. A Sudoku rejtvényt megoldó folytonos idejű mozgásnak a kezdeti állapotában minden változó értékét rögzítjük, kivéve az $s_{4,6}^1$ és $s_{3,7}^9$ változókat, amelyeket egy négyzetrács mentén változtatunk. Minden pont színét az határozza meg, hogy a $t=20$ időpillanatban a rejtvény (6,8)-as elemének megfelelő változók (számjegyek) közül éppen melyik igaz (lásd a színskálát). Míg az első térkép a teljes $[-1,1] \times [-1,1]$ tartományt mutatja, a következők mindig az előző térképen bejelölt fekete négyzet kinagyításai



Sudoku a Richter-skálán

A tranzien káosz jellegzetes tulajdonsága, hogy ha egy adott térrészből véletlenszerű kezdeti állapotból elindítunk egy pályát, annak a valószínűsége, hogy t idő után a pálya még nem hagyta el az adott térrészt, időben exponenciálisan csökken [13,15]. A mi esetünkben ez úgy jelentkezik, hogy ha sok véletlenszerű kezdeti feltételből indítunk el mozgásokat, azoknak a pályáknak a száma, amelyek még nem találták meg a megoldást, exponenciálisan csökken: $p(t) = ce^{-\kappa t}$ (7A. ábra)³. A κ -t szökési rátának nevezzük, és ennek reciproka tekinthető a tranzien káosz átlagos élettartamának. Minél lassab-

³ e a természetes logaritmus alapszáma, l. Gruiz Márton írásának 5. blokkját.



7. ábra. *A.* A megoldást t időig még nem talált pályák számának $p(t)$ eloszlása log-lin skálán több különböző Sudoku rejtvényre. Az egyenes szakaszok exponenciális csökkenésre utalnak. A meredekség abszolút értéke a szökési ráta, pl. az okkersárga görbe esetében ez 0,056. A színek és szimbólumok a különböző helyről gyűjtött rejtvényeket jelölik, a referenciák a szövegben találhatóak. *B.* A szökési ráta (κ) log-skálán a megadott számok, d , függvényében a különböző rejtvényekre

ban oldja meg a rendszer a feladatot (hosszabbak a tranziensek), annál kisebb a szökési ráta, ezért ez nehézségi mértékként alkalmazható bármilyen konkrét/egyedi feladatra. A **7B. ábrán** mutatjuk a κ értékeit különböző Sudoku rejtvényekre. Az x tengelyen a d paraméter szerepel, amely megmondja, hogy a rejtvényben hány számjegy volt adott. Látható, hogy van korreláció a szökési ráta (ezáltal a nehézség) és d között, ellenben nem föltétlenül a legkisebb d értékeknél a legnehezebb a feladat.

A szökési ráta több nagyságrenden keresztül is változhat, ezért a rejtvények nehézségének mérésére egy, a Richter-skálához⁴ hasonló mértéket definiáltunk:

$$\eta = -\log_{10} \kappa. \quad (4)$$

Négy kategóriába soroltuk a feladatokat: könnyű ($0 < \eta \leq 1$), közepes ($1 < \eta \leq 2$), nehéz ($2 < \eta \leq 3$) és extrém ($3 < \eta \leq 4$).

Sokan foglalkoznak azzal, hogy a Sudoku rejtvények nehézségét méri. Általában ezek a módszerek azt próbálják felbecsülni, hogy az emberek számára mennyire nehéz egy feladat, tehát valamennyire szubjektíven próbálják nehézségi kategóriákba osztani a rejtvényeket. Az általunk definiált mérték egy fizikai, objektív paraméter, csakis a mozgást leíró rendszer (ezáltal a feladat) struktúrájától, bonyolultságától függ, és fontos előnye, hogy egyetlen számmal jellemzi a feladat nehézségét. A szakirodalomból és a különböző weboldalokról gyűjtött rejtvényeken teszteltük a módszert (**7B. ábra**) és az általunk adott mérték jól korrelál az emberi megítélésével. A [16] weboldal például négy kategóriába osztja a feladatokat, mindegyikből kiválasztottunk néhány rejtvényt, és teszteltük a módszerünket: könnyű (fekete négyzet a **6. ábrán**), közepes (pi-

ros kör), nehéz (zöld kereszt) és abszurd (kék csillag). A beosztásuk jól korrelál az η mértékkel, ellenben a legnehezebb feladataik is csak $\eta \approx 2$ közelében (az általunk definiált közepes és nehéz határán) vannak. Ehhez hasonló a helyzet a [17] weboldalon talált Sudoku rejtvényekkel is (barna + jel). Egyes napilapok évente kinevezik az év legnehezebb rejtvényét, valóban ezek szökési rátája általában jóval kisebb, mint az említett weboldalon talált rejtvényeké: *Cave-man Circus 2009* ($\eta = 2,93$; türkiz \blacklozenge), *Guardian 2010* ($\eta = 2,82$; sötétbarna \blacklozenge). A legnehezebb rejtvényeket a Wikipedia toplistáján találtuk [18] (piros háromszögek). Egyedül ezek kerültek az extrém kategóriába ($\eta \in [3,0; 3,6]$). Néhány $d = 17,18$ megadott számmal jellemzett rejtvényt is teszteltünk (nemrég bizonyították, hogy ahhoz, hogy csak egy megoldás legyen $d \geq 17$ [19]), de meglepően a kicsi d érték még nem biztosítja azt, hogy a rejtvényt nagyon nehéz (okkersárga pontok).

Következtetések

Láthatuk, hogy a számítási problémák komplexitásának tanulmányozása fontos ágazata a matematikának és számítógép-elméletnek. Mivel a diszkrét korlátozás-kielégítési (döntési) feladatoknak egy elég nagy és alkalmazások szempontjából jelentős osztályára, az NP-teljes feladatokra [1,2] nincsenek eléggé hatékony algoritmusok, ezek nehézségének jellemzése különösen fontos probléma. Általában feladatcsoportokat és típusokat tudunk nehézség szempontjából jellemezni, úgy, hogy megnézzük, a változó számának növelésével hogyan nő a megoldáshoz szükséges idő. Ilyenkor mindig a legrosszabb esetet kell figyelembe venni. Ellenben egyedi feladatok nehézsége nagyon

változó lehet még a legnehezebb feladat-osztályokban is. Itt megmutattuk, hogy a diszkrét döntéshozatali feladatokat mindig meg lehet oldani speciális, folytonos idejű mozgásokat definiálva [9]. Az optimalizációs nehézség tranzienk káoszként jelentkezik, és a szökési ráta alapján egy jól meghatározott skaláris nehézségi mérték definiálható, amely jellemezni tudja *egyedi* feladatok nehézségét is [10]. A módszert a Sudoku rejtvényen szemléltettük, de a SAT-feladat NP-teljesége miatt az optimalizációs feladatok egy tag osztályán alkalmazható. \star

IRODALOM

- [1] Garey M. R., Johnson D. S.: *Computers and Intractability: A Guide to the Theory of NP-completeness*, W.H. Freeman & Co., New York, USA, 1990.
- [2] Fortnow L.: *ACM Communications* **52**, 78, 2009.
- [3] Applegate D. L., Bixby R. E., Chvatal V., Cook W. J., *The Traveling Salesman problem: a Computational Study*, Princeton University Press, 2007.
- [4] Barahona F.: *J. Phys. A: Math. Gen.*, **15**, 3241, 1982
- [5] Cook S.: *Proc. Third Ann. Symp. Theory of Computing ACM*, 151, 1971.
- [6] Rosenhouse J., Taalman L.: *Taking Sudoku Seriously: The Math Behind the World's most popular Puzzle*, Oxford University Press, New York, 2011.
- [7] Liu S.-C., Kramer J., Indiveri G., Delbruck T., Douglas R.: *Analog VLSI: Circuits and Principles*, MIT Press, 2002.
- [8] Chua L. O., Roska T.: *Cellular Neural Networks and Visual Computing: Foundations and Applications*, Cambridge University Press, 2005.
- [9] Ercsey-Ravaszi M., Toroczkai Z.: *Nature Physics*, **7**, 966, 2011.
- [10] Ercsey-Ravaszi M., Toroczkai Z.: *Scientific Reports*, **2**, 275, 2012.
- [11] Tél T., Gruiz M.: *Chaotic Dynamics: An Introduction*, Cambridge University Press, 2006.
- [12] Ott E.: *Chaos in Dynamical Systems 2nd ed.*, Cambridge University Press, 2002.
- [13] Lai Y.-C., Tél T.: *Transient Chaos: Complex Dynamics on Finite-Time Scales*, Springer, 2011.
- [14] <http://obelix.phys.nd.edu/~toro/archive/sudoku.mp4>
- [15] Kadanoff L.P., Tang C., *PNAS*, **81**, 1276, 1984.
- [16] <http://www.sudokuoftheday.co.uk>.
- [17] <http://www.extremesudoku.info/sudoku.html>.
- [18] http://en.wikipedia.org/wiki/Algorithmically_of_sudoku#Exceptionally_difficult_Sudokus_.28hard.
- [19] McGuire G., Tugeman B., Civarío G.: <http://arxiv.org/abs/1201.0749>.

⁴ A Richter-skála a földrengések energiájának logaritmusát adja meg.