

# Mérünk vagy értékelünk?

FATALIN LÁSZLÓ

*Napjainkban különösen aktuálissá vált az alkalmazott értékelési rendszerek és mérési eljárások újragondolása. A tanári szabadság növekedésével együttjárnak olyan döntési feladatok is, amelyeknél alapvetően csak szubjektív megítélésünkre támaszkodhatunk. A tanárképzésből ezidáig kimaradt az értékelési rendszerek és mérési eljárások tudományos alapjainak átfogó áttekintése, pedig ezek ismerete hasznos segítséget jelenthetne gyakorlati munkánk során. A mérés – értékelés általános folyamata, tudományos alapjainak körvonalai hagyományos értelemben vett elméleti keretbe nem zárhatóak be, az interdiszciplináris fogalmakhoz hasonlóan ellenállnak az ilyen irányú törekvéseinknek. E témakörök egyes fejezetei meglehetősen jól kidolgozottak, más területein jelenleg is folynak az alap kutatások.*

A hétköznapi gyakorlatban lépten-nyomon alkalmaznunk kell valamilyen pedagógiai értékelési rendszert függetlenül attól, hogy milyen mélységű tudományos ismerettel rendelkezünk a mérés-értékelés általános folyamatáról. A többnyire csak a pedagógusvényainkra alapozott ítéleteink meghozatala során számos esetben találjuk szembe magunkat megválaszolhatatlannak tűnő kérdésekkel akár úgy, hogy ösztönösen vetődik fel bennünk, akár úgy, hogy diákok szegeznek nekünk. A szakmai rutin megszerzésével egyidejűleg persze azt a készséget is elsajátítjuk, hogy ezeken az eldönthetetlen, következtetésképpen „értelmetlen” kérdéseken hogyan kell elegáns módon átsiklani. A probléma persze ettől függetlenül élő marad akkor is, ha tabuként kezeljük és akkor is, ha lelkiismereti konfliktust okoz nekünk.

Milyen tényezők hatnak egy-egy osztályzat megállapításakor? Mennyire reális, objektív az értékelésünk? Esetenként egy tanuló feleletét úgy ítéljük meg, hogy az öt skatulya közül egyikbe sem esik bele, valamelyik kettő között van, miért van csak öt skatulya? A pozitívabb illetve negatívabb értékelésnek milyen várható hatása lesz a tanuló további teljesítményére? A matematika 4-es jegy az ország egész területén ugyanazzal a mércével kerül megállapításra, vagy netalántán a különböző helyeken szerzett ugyanolyan érdemjegy különböző tudást takar? Ismereteink mai szintjén ugyan átfogó, megnyugtató választ adni e kérdésekre még nem tudunk, mégis foglalkoznunk kell e problémákkal. A modellmódszer alkalmazásával közelítve a mérés-értékelés folyamatához meglehetősen nyilvánvaló, ámde sokszor figyelmen kívül hagyott igazságra derül fény.

## Objektivitás vagy komplexitás?

E cikk címében feltett kérdés megválaszolásához először célszerű megvizsgálni, hogy mit takar a mérés és az értékelés fogalma, mi a közös és mi a különbség e két fogalom jelentésében. Különböző lexikonok megfogalmazásai alapján megállapítható, hogy mindkét esetben összehasonlításról van szó, mérés esetén egy rögzített etalonra hasonlítjuk össze a mérendő mennyiséget és eredményül többnyire egy számot kapunk, míg értékelés esetén egy többé-kevésbé pontosan meghatározott szempontrendszerhez viszonyítunk. Már e két fogalom felszínes körülírása is tükrözi azt a hozzájuk tapadó hétköznapi értelmezést, hogy a mérés egy objektív tudományos eljárás, az értékelés viszont szubjektív tényezőket is tartalmazhat.

Az értékelési rendszerekkel szemben mindig felmerülhet a szubjektivitás gyanúja, objektivitásának megkérdőjelezhetősége. A gyakorlatban egyre inkább terjedő tesztrend-

szerek látszólag megoldják az objektivitás kérdését, hiszen a tudományos igényességgel összeállított tesztlapok objektív mérőeszközként funkcionálnak és eredményül is egy számot adnak, azaz a tesztlapos számonkérés mérésnek minősül. A tesztlapok ezen óriási előnye okozza egyre szélesebbkörű elterjedésüket, helyenként jól megfigyelhető fetiszizálásukat, pedig alkalmazásuk óta számos negatív tapasztalat és észrevétel is felgyűlt már.

A tesztrendszerek kialakulása egybefonódott az intelligencia vizsgálatok történetével. A híres-hírhedt IQ-fogalom kialakítása szinte teljesen kielégíti még a fizikai mennyiségek mérésére vonatkozó Carnap-féle kritériumokat is, beleértve a skálatörvények meghatározását és a skálatörvénynek a vonatkoztatási rendszertől való függését is. A kisebb-nagyobb-egyenlő reláció is értelmezést nyer. Az összehasonlítás, a mérés tehát feltétlenül objektív, ugyanakkor korlátozott érvényességű is, hiszen a Carnap-féle kritériumok is csak az igazán egyszerű fizikai mennyiségek mérésénél érvényesek.

A számszerűsítésre való törekvés a vizsgálatok nagy részére jellemző. Egy felmérési dolgozat eredménye egy osztályzat, amit esetenként az elért összpontszám alapján állapítunk meg, az intelligenciatesztek kiértékelésének eredményeként pedig az IQ értéket adjuk meg, a népszerűségi vizsgálatokban pedig az ún. tetszési indexet határozzuk meg. Modellelméleti szempontból az ilyen mutatószámokkal történő jellemzők tulajdonképpen a számok halmazát próbáljuk meg modellként felhasználni. Az eredmények értelmezésében a számok közötti rendezési reláció gyakran kiemelkedő szerepet kap. Tapasztalataink alapján a számok halmaza modellként széleskörben jól alkalmazható és emiatt esetenként megfelelünk e modell érvényességi körének vizsgálatáról.

Sokszor találkozhatunk az iskolai osztályzat, vagy az IQ érték olyan értelmezésével, amely azt a tudás illetve az értelmi képesség abszolút mutatószámának állítja be. E fogalmak tartalma azonban meglehetősen bonyolult ahhoz, hogy a láncszerűen elrendezett számok valamely részhalmaza megfelelő pontosságú modelljük lehessen. A szín jellemzésére is legalább három számot (trikromatikus mérőszámot) használunk, a zenei hangok jellemzéséhez pedig a hangmagasságon kívül még a hangerősséget és a hangszínt is meg kell adnunk. A tudás és az értelem legalább összetett fogalmak, ezért az olyan állítások igazságtartalmát jogosan kérdőjelezhetjük meg, amelyek szerint egy-egy mutatószám elegendő megfelelő pontosságú jellemzésükhöz.

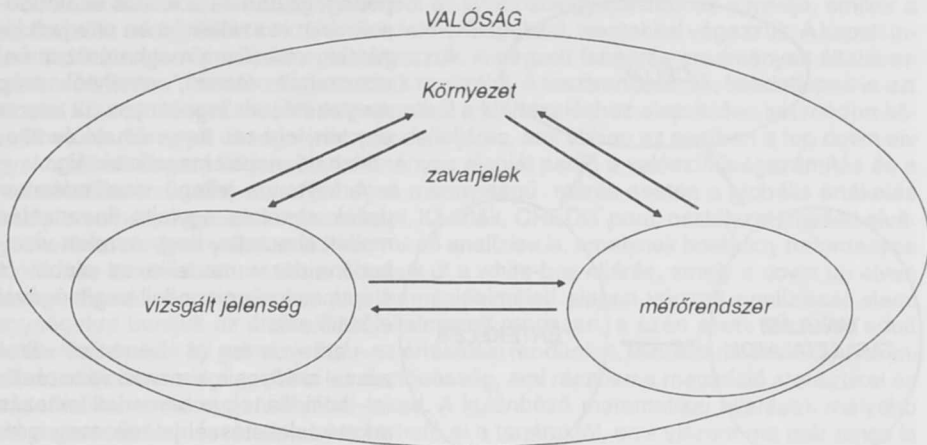
Az előzők alapján kitűnik, hogy a mérési eljárásnak van egy óriási hátránya is, hiszen az összehasonlítás számszerű eredménye csak bizonyos dolgok jellemzésére megfelelő. Ez talán tömörebben úgy is megfogalmazható, hogy a hagyományos értelemben vett mérési eljárások önmagukban nem alkalmasak komplex rendszerek jellemzésére. Ez esetenként legalább akkora hátrány, mint amekkora előny rejlik az objektivitásában. Az értékelési rendszereknél a komplexitásra való törekvés mindig magától értetődik. A mérés-értékelés folyamata alatt mindig olyan eljárást értünk, amelyben az objektivitás és a komplexitás együttesen érvényesül. A mérés-értékelés folyamatának megtervezésekor mindig tekintettel kell lenni a legáltalánosabban alkalmazott ponderábilis (számszerű, mérhető) mennyiségek mellett a tanítás-tanulás folyamatában fellelhető imponderábilis (nem számszerűsíthető) tényezők strukturális vizsgálatára is.

## A mérés rendszerteknikai modellje

A mérési eljárások általános rendszerteknikai modelljét az 1. ábra tünteti fel. A mérési eljárás szemléltetéséhez vázolt modellben a valóságot három részre vágtuk:

- a vizsgált jelenség
- mérőrendszer
- környezet.

Ez a három rész páronként kölcsönhatásban áll egymással. A környezet hatásait általában zavarjeleknek nevezzük. Tudományos vizsgálatokban a mérések során ezeket a zavarjeleket, ezek hatásait igyekszünk minél kisebb mértékűre szorítani, hiszen az eredő mérési hibát illetve a mérés reliabilitását ezek jelentősen befolyásolhatják. A mérőrendszer és a vizsgált jelenség, rendszer közötti kapcsolat is jelentős hibaforrás lehet. A problémát itt az jelenti, hogy a mérőrendszer rákapcsolása a vizsgált jelenségre mindig



1. ábra

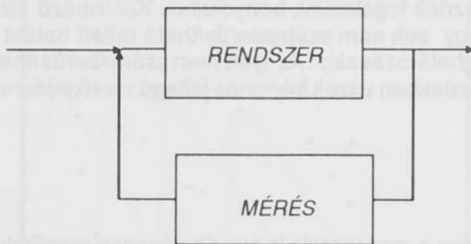
A mérési eljárás rendszerteknikai modellje

megváltoztatja az eredeti állapotot. A műszaki-természettudományos vizsgálatokban ezen kölcsönhatást is minimálisra igyekszünk csökkenteni, nagy pontossági igények esetén igen elterjedt a kompenzációs elv különböző megvalósítása. A pedagógiában viszont mind elméleti, mind gyakorlati megközelítésben az értékelés szervesen kapcsolódik a tanítási tevékenységhez.

A tanítási-tanulási folyamat modelljei között nagy számban fordulnak elő ún. viselkedésmoდეllek, hiszen alapvetően célirányos tevékenységről van szó. Az adott oktatási rendszerbe bekapcsolódó tanulóra olyan hatásokat igyekszünk gyakorolni, hogy a végző tanuló a kitűzött céloknak megfeleljenek. Ennek eléréséhez különböző értékelési rendszereket alkalmazunk, amelyeknek vitathatatlanul jelentős visszahatása van a tanulóra, azaz a különböző értékelési rendszerek alkalmazása illetve elhagyása jelentősen megváltoztatja a tanítás-tanulás folyamatát. (Az emberekre is érvényes az energiaoptimum elve, azaz teljesítményük igazodik a támasztott követelményekhez.) Az értékelés visszahatását is figyelembevevő elemi viselkedésmoდეlel tömbvázlata a 2. ábrán látható.

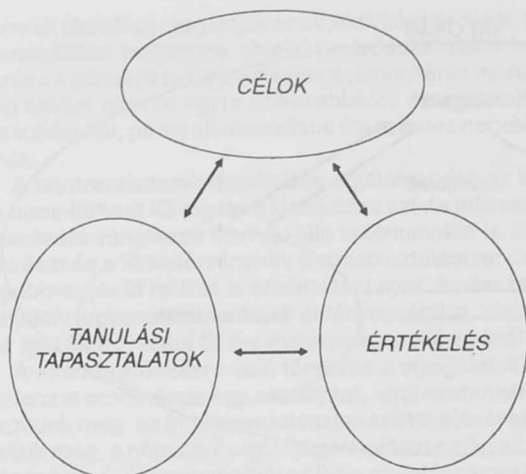
Bíráható az itt vázolt koncepció, mely szerint viselkedésmoდეlellekkel közelíthető a mérés-értékelés folyamata, hiszen első megközelítésben csak a tanulóra irányuló hatásokat veszi figyelembe. Úgy tűnik, hogy ezt a megközelítési módot mind a külföldi, mind a hazai törekvések már jelentősen túlhaladták. E változást jól szemlélteti a Tyler-féle értékelési moდეlel (3. ábra), amely talán a legnagyobb hatást gyakorolta a kutatásokra. (A Tyler-féle moდეlelben a három alapelem – célok, tanulók tanulási tapasztalatai, értékelés – között fellépő kölcsönhatásokat dinamikus kölcsönkapcsolatként kell értelmeznünk.)

E jogos kritikai észrevételek sem indokolják azonban a viselkedésmoდეlel való megközelítés elvetését, hiszen más szaktudományokban sikeresen alkalmazták ezeket és ennek eredményeként kialakultak egzakt leírási módjuk is. Ez lehetővé teszi, hogy megpróbálkozzunk a mérés-értékelés folyamatának fizikai és matematikai modellezésével is. Egy fizikai vagy matematikai moდეlel alkotása nyilván messze magasabb megismerési szintet tükröz, mint a



2. ábra

Visszacsatolt elemi viselkedés moდეlel



3. ábra  
A Tyler-féle értékelés modell

pedagógiában és sok más szaktudományban is széleskörben elterjedt kizárólag verbálisan meghatározott ún. absztrahált modell, amelytől még meglehetősen ingoványos út vezet egy ténylegesen használható és ellenőrizhető modell megalkotásáig.

A folytonos jellegű viselkedésmo-  
dellek elmélete egyfajta linearizálás  
mellett kimutatta, hogy minden visel-  
kedőrendszer modellje az alábbi öt  
alaptípusú elemi modell segítségével  
előállítható:

- arányos tag (a bemeneti változással arányos a kimeneti változás)
- holtidős tag (a bemeneti változás adott késleltetéssel jelenik meg a kimeneten)
- integráló tag (a bemeneti változások időbeli mennyiségi felhalmozódása, összegeződése jelenik meg a kimeneten)
- tárolás tag (a bemenet által meg-

határozott exponenciális változás jelenik meg a kimeneten)

- differenciáló tag (a bemenet változásának sebességével arányos a kimenet).

Ezen elemi modellek soros, párhuzamos kapcsolásával kialakított általában visszacsatolást is tartalmazó hatásláncú viselkedésmo-  
dellek megalkotásához nemcsak elméletileg közelíthetünk, hanem aktív kísérletezéssel a bemeneti változásokra nyert válaszokból e modellek egzakt módon előállíthatók. E fegyvertár alkalmazása jelentős segítséget nyújt a mérés-értékelés folyamatának jobb megismeréséhez, ráirányítja a figyelmet az értékelési rendszer szabályozóként való működésére, valamint az oktatási rendszer stabilizását biztosító illetve felborító funkciójára.

A mérés-értékelés folyamatához jól illeszkedő viselkedésmo-  
delld kidolgozása előtt még jelentős akadályok tornyosulnak. Az alapvető nehézséget a validitásnak (érvényesség) nevezett mérésmetodológiai követelmény támasztja. Ez rendkívül összetett, nehezen megfogható problémája a mérésnek. A felvetődő kérdés az, hogy az értékelés során azt mérjük-e, amit akarunk. Vizsgálhatjuk például azt, hogy egy intelligenciateszt valóban az intelligenciát méri-e. A válasz megkereséséhez persze szükségünk lenne az intelligencia definíciójára. (Komolytalan próbálkozás, bár többen megkíséreltek ezen úgy átsiklani, hogy az intelligencia az, amit a teszt mér.) Egy-egy fogalom tudományos értelmezése és az összehasonlítást biztosító mérési eljárások meghatározása elválaszthatatlan egymástól, együtt alakulnak ki meglehetősen hosszas, fáradtságos kutatómunka nyomán. A pedagógiában, pszichológiában ... használt fogalmaink bonyolultak. Különböző vizsgálatok már több esetben rámutattak, hogy sok nem számszerűsíthető rejtett hatást is figyelembe kell venni egyes fogalmak meghatározásakor. Az ilyen nem számszerűsíthető hatások feltárására és vizsgálatára az előzőekben vázolt folytonos jellegű viselkedésmo-  
dellek nem alkalmasak.

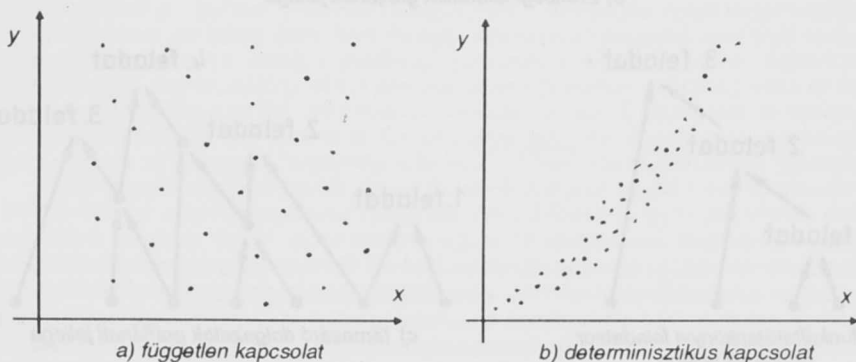
## Ponderábilis mennyiségek

A számszerűsítés és ezzel együtt gyakran a rangsorolás is egy általánosan megfigyelhető emberi törekvés, amit számos kedvező tapasztalat indukál. Az összehasonlítások során a számhalmaz sok esetben kitűnő modellteret bizonyul, ami nem meglepő, hiszen a számfogalom éppen a mennyiségi összehasonlítások tapasztalata alapján alakult ki minden emberben, sőt a számfogalom meghatározására kidolgozott matematikai elméletek is ezen alapulnak.

Századunkban széles körben elterjedt az információgyűjtésnek az a módja, amikor a tapasztalati adatgyűjtést kérdőívekkel, felmérőlapokkal, tesztekkel végezzük. A kapott információkat valamilyen módon feldolgozzuk. A végzett felmérés eredményeit általában „számszerűsített” formában igyekszünk realizálni. A tesztrendszerek, feladatlapok is ezt az utat követik. Modell elméleti szempontból a kiértékeléshez alapvetően két módon közelíthetünk. Az egyik út a black box modelleljárás, ami ebben az esetben a top down elv gyakorlati alkalmazását jelenti, tudományos alapját pedig a valószínűség számítás és a matematikai statisztika adja. Tömeges mennyiségű mérés esetén a globális értékelés (mely eredménye gyakran egy szám, pl. IQ érték, CREDIT pont, osztályzat) mellett elvégezhető az eredmények részletekbe menő analízise is, amelynek hatékony tudományos módszere lehet a faktoranalízis. A másik út a white-box eljárás, amely a down up elven alapul. Ennek alkalmazásakor az információelméleti alapon végzett analízissel elemi egységekre bontjuk az értékelésre alkalmazott rendszert, s ezen elemi részekre adott értékekből építjük fel szintézis útján az értékelési rendünket. Mindkét módszer gyakorlati alkalmazásában megfigyelhető az ösztönösség, ami részben a megfelelő statisztikai és matematikai ismeretek hiányából fakad. A különböző matematikai eljárások mélyebb megismerése természetesen nem várható el a tanároktól, erre viszont ma már nincs is szükség. A számítógépek segítségével megfelelő szoftverek birtokában a feladatok elvégezhetőek a mögöttük meghúzódó elméletek ismerete nélkül is. A tanárok ugyanúgy dolgozhatnak ilyen szoftverekkel, mint a titkárnők a számítógépes szövegszerkesztőkkel. Az iskolákból egyenlőre hiányoznak ezek a szoftverek és a szoftverpiacon is nehezen akadhatunk rá ilyen programcsomagra.

Ezek szélesebb körű kifejlesztésére azért is szükség lenne, mert gyakran még a tudományos vizsgálatokban is találkozhatunk egyes matematikai eljárások helytelen értelmezésével és alkalmazásával. A korrelációs együtthatót például gyakran úgy interpretálják, mintha az a két vizsgált komponens közötti sztochasztikus kapcsolat erősségét mérné. A korrelációs együttható nulla volta e felfogásban azt jelenti, hogy a két komponens egymástól teljesen független, ha pedig az értéke 1, akkor a két komponens között determinisztikus, azaz függvénykapcsolat van. Ez az értelmezés sajnos tévedésen alapszik. A matematikusok valóban törekednek, mind a mai napig sikertelenül, ilyen tulajdonságú mérőszám megalkotására, a korrelációs együtthatóról ugyanis már régóta kiderült, hogy valójában csak a két komponens között fennálló lineáris kapcsolat erősségét méri. Konkrét példákat lehet mutatni olyan esetre, amikor a két komponens között determinisztikus függvénykapcsolat van és a korrelációs együttható értéke mégis 0. E bűvös mérőszámok helyett egy ún. pontfelhős ábra gyakran lényegesen többet mond. A 4. ábra még a laikusok számára is többet mond az  $x$  és  $y$  mennyiségek közötti sztochasztikus kapcsolat erősségéről, mint egy szám.

A matematikában kevésbé jártas emberek számára is használható értékelési szoftverek elterjesztésével a kiértékelések ösztönös jellege jobban háttérbe szorulna.



4. ábra  
A sztochasztikus kapcsolat erősségét kifejező pontfelhős ábra

A black box és a white box eljárást alkalmazva ugyanazon dolog kiértékelésére általában különböző eredmény adódik, ami érthető ezen eljárások közelítő jellegére gondolva. E két eljárás iteratív alkalmazása esetenként ugyan megoldja a problémát, de bonyolultsága miatt ez a módszer is csak erősen korlátozottan használható. Az előző eljárások közelítő jellege a számszerűen nem mérhető, rejtett hatások, tényezők, az ún. imponderábilák következménye.

## Imponderábilák

A tanítás-tanulás, a mérés-értékelés folyamatában fellépő imponderábilák feltárására, jellemzésére a klasszikus matematikai módszerek általában nem alkalmasak. E tényezők vizsgálata napjainkban még az alapkutatások körébe tartoznak. E kutatások eredményei arra utalnak, hogy az imponderábilák megismeréséhez elsősorban a véges matematikai modellek nyújtanak hatékony segédeszközöket. A számítógépek megjelenésével és elterjedésével egyidejűleg a matematikának ez az ága is rohamos fejlődésnek indult. A közoktatásból és a tanárképzésből azonban ezidáig kimaradt az új szemléleti módok bevezetése, megismertetése.

A célnak megfelelő véges matematikai modellek megalkotása általában strukturális elemzések alapján történik. A strukturális vizsgálatok alapját pedig az adott jelenségkörben felismert relációk elemzésével végezhetjük el. Ennek egyik igen hatékony módszere a Galois-algoritmus, amely a vizsgált relációhoz egy gráfot állít elő. Ezen objektív értékelési módszer egyre szélesebb körű felhasználására az utóbbi két évtizedben jelentős erőfeszítések történtek. A műszaki kockázatelemzési feladatokhoz és a tananyagstrukturák elemzéséhez egyaránt jól használhatónak bizonyult ez a megközelítés. Pedagógiai alkalmazásai az atomerőművi biztonságtechnikai vizsgálatokban való alkalmazásával egyidőben jelentkeztek. Taneszköz tervezése, vagy egy tanulócsoporthoz tudásstrukturájának, vagy egy tankönyv fogalmi strukturájának elemzése egyaránt jól elvégezhető e módszerrel.

A Galois-algoritmus jól felhasználható feladatsorok strukturális elemzésére is. Megállapítva, hogy egy-egy feladathoz milyen alapinformációk tartoznak és ezt ún. megalapo-



5. ábra  
Feladatsorokhoz tartozó Galois-gráfok



zási relációnak tekintve elkészíthető a Galois-gráf. Ilyen elemzést elvégezve egy érettségi-felvételi dolgozat, egy témazáró dolgozat és egy feladatgyűjtemény kiválasztott fejezetének feladatai alapján az 5. ábrán feltüntetett típusú gráfokat kapjuk eredményül. Az érettségi feladatsorhoz tartozó gráf alapvetően széteső jellegű, míg a feladatgyűjtemények esetében erősen hierarchikus felépítettséget tapasztalhatunk. A témazáró dolgozatok feladatai mindkét jelleget tükrözik némileg. A white box eljárásban használt szintézis pontosabb végrehajtásához tulajdonképpen e szerkezetet szükséges feltárni, hiszen csak ennek ismeretében dönthetjük el objektív módon a pontozást.

Hasonló elemzés végezhető el egy-egy témazáró dolgozat kapcsán arra a relációra is, hogy melyik tanuló melyik feladatot oldotta meg. Az osztály tudását a kapott struktúra lényegesen jobban jellemzi, mint a numerikus adatok, hiszen a gráfon a ki-mit tud és ki-mit tudnak információ is megjelenik és így lehetővé válik a tanítandó ismeretek sorrendjére olyan tanítási stratégiák kidolgozása is, amelyek optimálisak.

E Galois-algoritmus előnye, hogy a különböző problémákhoz tartozó különböző típusú relációkat azonos módon, ugyanazon algoritmussal közelíti meg. Eredménye ugyan nem „számszerű”, de egységesen és objektív módon a gráfokat használja modellként. E gráfok összehasonlítása egy egyszerű rendezési relációval senkinek sem jut eszébe, míg numerikus jellemzés esetében két számot ösztönösen is összehasonlítunk „nagyságuk” szerint. A használt modellér éppen ezért alkalmasabbnak tűnik sok fogalom, jelenség pontosabb jellemzésére, mint az eddigiekben általánosan alkalmazott numerikus jellemzés. A strukturális elemzések kutatása során számos nehézséggel, gyakran nyitott matematikai problémával találkozunk szembe magunkat. Ezek megoldása, valamint számítógépes szoftver formában történő realizálásuktól általános elterjedésüket várhatjuk.

## IRODALOM

- Báthory Zoltán*: A pedagógiai értékelés és a közoktatás szabályozási mechanizmusai. Műv. Min. Vezetőképző és Továbbképző Int. – Veszprém Megy. Ped. Int., Veszprém, 1987.
- Csáki Frigyes – Bars Ruth*: Automatika. Tankönyvkiadó, Budapest, 1983
- Fatalin László – Varsics Zita*: A tudományos modellalkotás alapjai I-II. Calibra Kiadó, Budapest, 1993.
- Fáy Gyula*: An Algorithm for Finite Galois Connections (Technical Report), Institute for Industrial Economy, Organization Technique, Budapest, 1973.
- Golnhofer Erzsébet – M. Nádasi Mária – Szabó Éva*: Készülünk a vizsgáztatásra. Korona Kiadó, Budapest, 1993.
- Horváth György*: Az értelem mérése. Tankönyvkiadó, Budapest, 1991.
- Schnell László*: Jelek és rendszerek mérés technikája. Műszaki Könyvkiadó, Budapest, 1985.
- Takács Viola*: Two pedagogical application of Galois graphs Lecture. Darmstadt Technische Hochschule, 1984.