Naiv és optimalizált versenyértékelés az oktatásban

(Naïve and optimized evaluation in competitive education)

Pitlik László, Rikk János, Pitlik Marcell (MY-X team)

Kivonat: Aki mást értékelni mer, az vagy váljon az értékelés gurujává, vagy ne értékeljen senkit, mert a mindenki másként egyforma elv nem maradhat egyszerű szóvirág az évszázadok óta az értékelések egysíkságát tömegesen megszenvedők számára. Igaz, a politikai közbeszéd, ill. a joggyakorlatok sokasága gyorsan kijózanít mindenkit a soknézőpontúság kapcsán – de ez a józanodás motivációvesztő és nem személyiségmegtartó, mint a robot-alapú értékelés maga.

Kulcsszavak: optimalizálás, hasonlóságelemzés, objektivizálás

Abstract: The willingness for evaluating other objects (persons) have to involve the willingness to become a master/guru of the evaluation. Evaluators without the holistic knowledge about the anti-discriminative principle (where each object can have the same evaluation value) may not evaluate anybody/anything. The naïve evaluations cause a lot of pain/misfortune for crowds being affected through the simplicity of these evaluation logics. The political communication and/or the daily logic of the legal processes are trivial examples of the multi-layered interpretation – unfortunately with the consequence becoming demotivated - instead of motivated as in case of the complexity of the robot-based evaluations.

Keywords: optimization, similarity analysis, objectivism

# Bevezetés

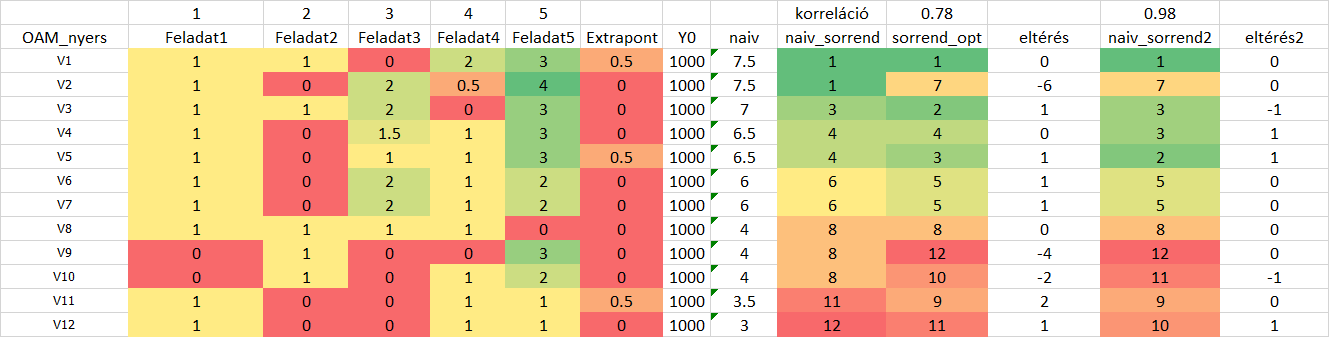
Előzmények:

* <https://miau.my-x.hu/miau2009/index.php3?x=e0&string=robottan.r>
* <https://miau.my-x.hu/miau2009/index.php3?x=e0&string=szumm>
* <https://miau.my-x.hu/miau/231/mtu2017/ertekkutatas_mtu_1_full_hu.docx>
* <https://miau.my-x.hu/miau/quilt/alternative_evaluations.docx>
* <https://miau.my-x.hu/miau/270/y0_alternative_approaches.docx>
* <http://miau.my-x.hu/miau/147/anonim_tdk.xls>
* <http://miau.my-x.hu/miau/197/5kerdes_6pont_rendszerkonzisztencia.xls>
* ...
* <https://hu.wikipedia.org/wiki/%C3%89l%C5%91-pontrendszer>

Aki mást értékelni mer (vö. oktatók, vezetők = célcsoport), az vagy váljon az értékelés gurujává, vagy ne értékeljen senkit naiv módszertani alapokon, mert a mindenki másként egyforma elv (=cél) nem maradhat egyszerű szóvirág az évszázadok óta az értékelések egysíkságát tömegesen megszenvedők számára. Igaz, a politikai közbeszéd, ill. a joggyakorlatok sokasága gyorsan kijózanít mindenkit a soknézőpontúság kapcsán – de ez a józanodás motivációvesztő és nem személyiségmegtartó, mint a robot-alapú értékelés maga (= hasznosság).

# Alapadatok és alternatívák

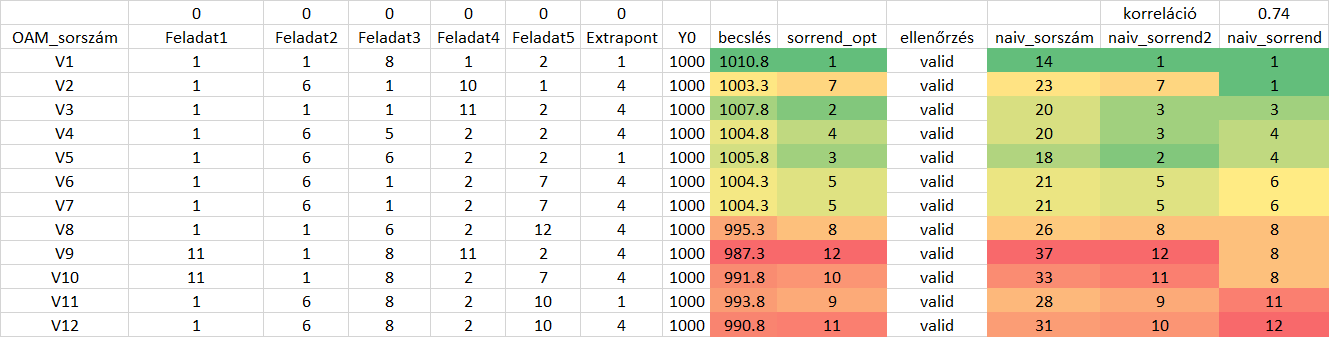
Képzeljünk el egy (itt és most anonim) versenyt, melyen 12 versenyző vesz részt, 5 reguláris feladat kapcsán és extrapontok szerzésének lehetősége mellett. A kanonizált folyamatban évszázadok óta a részpontok összege alapján illik győztest hirdetni. Ez a megközelítés (1. ábra: naiv, ill. naiv\_sorrend oszlopok) semmi veszi az alapvető pedagógia elvet, miszerint lehet minden versenyző másként egyforma.



1. Ábra: A nyers kiindulási adatok (forrás: saját ábrázolás)

Az éles adatok alapján látható, hogy a naiv (nem optimalizált) értékelés holtversenyre vezet (7.5-7.5 pont). A holtverseny alapja, hogy minden egyes tizedpont ugyanannyit ér, bármely versenyszámban (feladat esetén) sikerült megszerezni. A feladatokat nem kontextusok, hanem a versenyzők által ezek mentén szerzett pontok teszik nem feltétlenül azonos értékűvé: pl. ha 12 fő esetén (vö. Feladat1) 10 fő képes pontot szerezni – 2 fő nem, ill. sokkal értékesebb illene, hogy legyen Feladat2 esetén a pontszerzés, mert a versenyzők zöme nem szerzett itt pontot. Az sem mindegy, ki szerezte pl. a legnagyobb elérhető pontszámot…

Ha az online anti-diszkriminatív modellezési keretrendszerben (vö. <https://miau.my-x.hu/my-x-factor-y>, ill. <https://miau.my-x.hu/myx-free/coco/index.html> - Y0 modellek) valaki feltölti a 2. ábra szerinti sorszámnézetet, akkor kapható meg az 1. ábra optimalizált nézete, ill. naiv2-ként a sorszámok összege alapján számított alternatív sorrend, ahol már itt meg kell jegyezni, hogy ebben az esetben a sorszámok és nem ma tizedpontok azok, melyek ekvivalensek minden versenyszám minden sorszám-szintjén.



1. Ábra: A nyers pontszámok sorszám-nézete (forrás: saját ábrázolás)

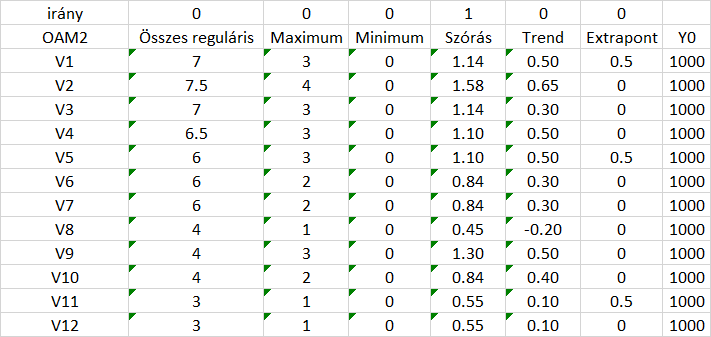
Az 1. és a 2. ábra alapján ki kell emelni, hogy:

* A holtverseny már nem holtverseny, sőt a V2 az 1. helyről visszaesik a 7. helyre, ami 12 verseny kapcsán 6 helyezés, azaz 50% eltérés.
* A pontszám-alapú (naiv1) és az optimalizált sorrendek korrelációja (többek között ezért) csak 0.78.
* Az optimalizált sorrend és a naiv2-höz tartozó sorrend közötti korreláció: 0.98, vagyis itt is érezhető, hogy a tizedpontok ekvivalenciája és a sorszámok ekvivalenciája közel sem azonos hatású.
* A két naiv sorrend között a korreláció: 0.74, mely erősíti a fentieket.

További részletek: <https://miau.my-x.hu/miau/273/naiv_optimalizalt_verziok2.xlsx>

# Kiegészítő elemzések és eredményeik

Az 1. ábra kapcsán már jelzett értelmezési rétegek (ki hol és hol mennyi pontot szerzett a többiekhez képest) vezetnek a 3. ábrához:



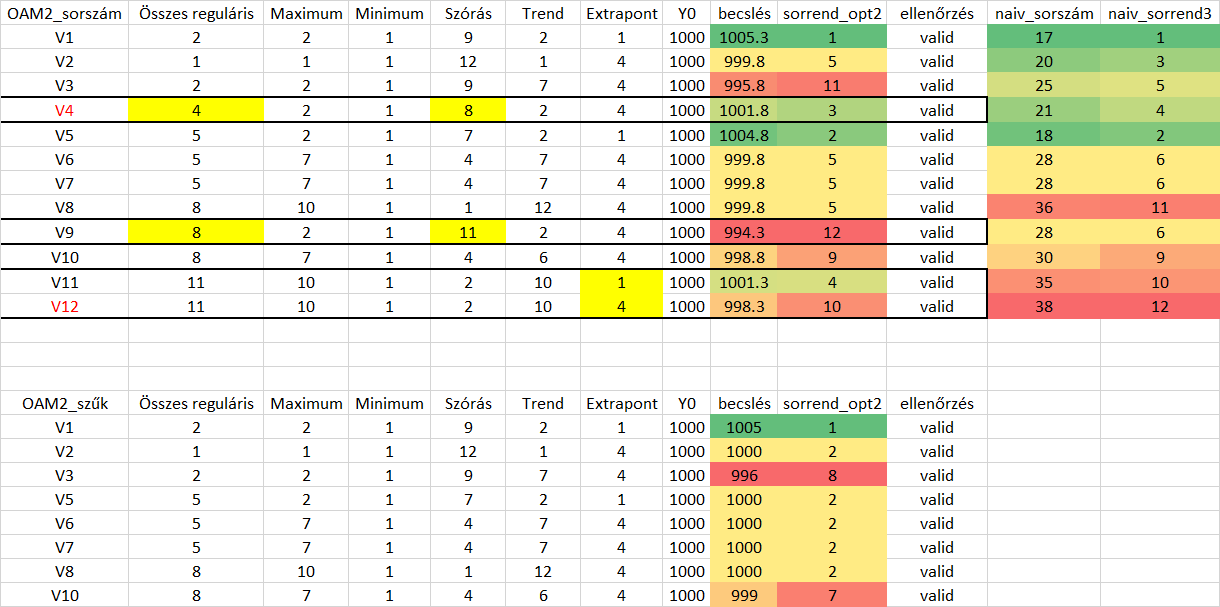
1. Ábra: Speciális értékelési szempontok (forrás: saját ábrázolás)

A 3. ábra azt sugallja, hogy

* Annál jobb végső helyezést érdemel egy versenyző, minél nagyobb a reguláris feladatok keretében szerzett pontjainak összege.
* Annál jobb végső helyezést érdemel egy versenyző, minél nagyobb az elért maximális pontszáma (hiszen ez azt jelzi, mennyire tud komplexen, hibátlanul dolgozni nagy kihívások esetén – feltételezve, hogy a pontszámok racionális arányokat tükröznek a feladatok nehézségéhez mérten).
* Annál jobb végső helyezést érdemel egy versenyző, minél nagyobb az elért minimális pontszáma (hiszen ez azt jelzi, mennyire stabil a versenyző tudása a vizsgált spektrum bármely részét illetően).
* Annál jobb végső helyezést érdemel egy versenyző, minél kisebb a szórása a részpontjainak (hiszen így kisebb az esélye annak, hogy pontot realizál valaki véletlen találat formájában).
* Annál jobb végső helyezést érdelem egy versenyző, minél nagyobb az 1-2-3-4-5 feladat mentén a pontszámai trendje (hiszen a maximális pontszám az 5. feladatnál, vélhetően a legnehezebb, legkomplexebb feladatnál látható – ill. feltételezhető, hogy a feladatok egymás után hajtandók végre, s így mérni lehet a fáradási potenciált is a tudáspotenciál mellett).
* Annál jobb végső helyezést érdemel egy versenyző, minél több extrapontot realizál (hiszen ezen keresztül kapnak lehetőséget a pontozók, olyan részleteket is értékelni, melyek alapvetően kívül esnek a tervezett fókuszon – pl. innovativitás, alternativitás – de ezek tényszerűen létező többletteljesítmények.

Ez az értelmezési tér azért szofisztikáltabb, mint a hagyományos (vö. 1. ábra), mert

* Lehetőséget kínál a véletlenszerű tudás, azaz a tipp-alapú pontszerzés feltárására,
* Hiszen a pl. szórás és a trend, de a max-min-távolság is arra mutat rá, hogy akárki akárhol nem veszíthet pontot, ha máshol már pontot volt képes szerezni.
* Így a tudás egyenszilárdsága, konzisztenciája és becsülhető, ami V3 esetén komoly következményekkel jár, mert V3 a naiv 3. és az ennek párját jelentő optimalizált 2. helyről a 11. helyre esik vissza, ami 9 helyezés eltérés egy 12 elemű rangsor-skálán.



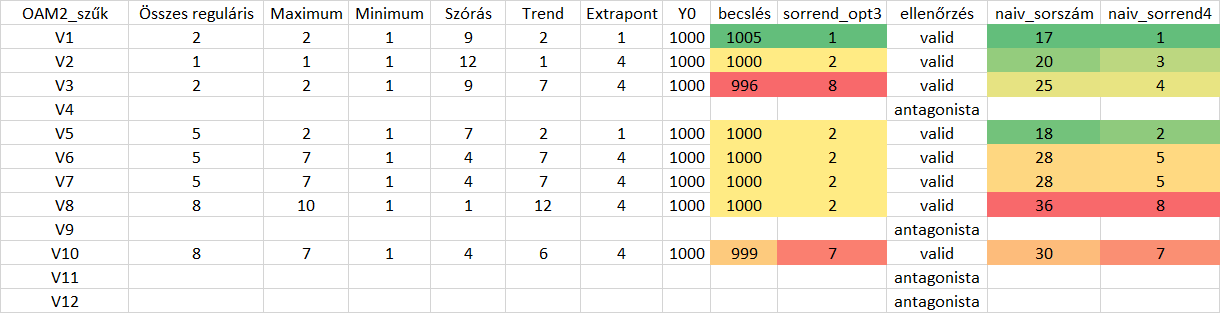
1. Ábra: A speciális OAM(2) – (forrás: saját ábrázolás)

A 4. ábra felső része alapján látható, hogy

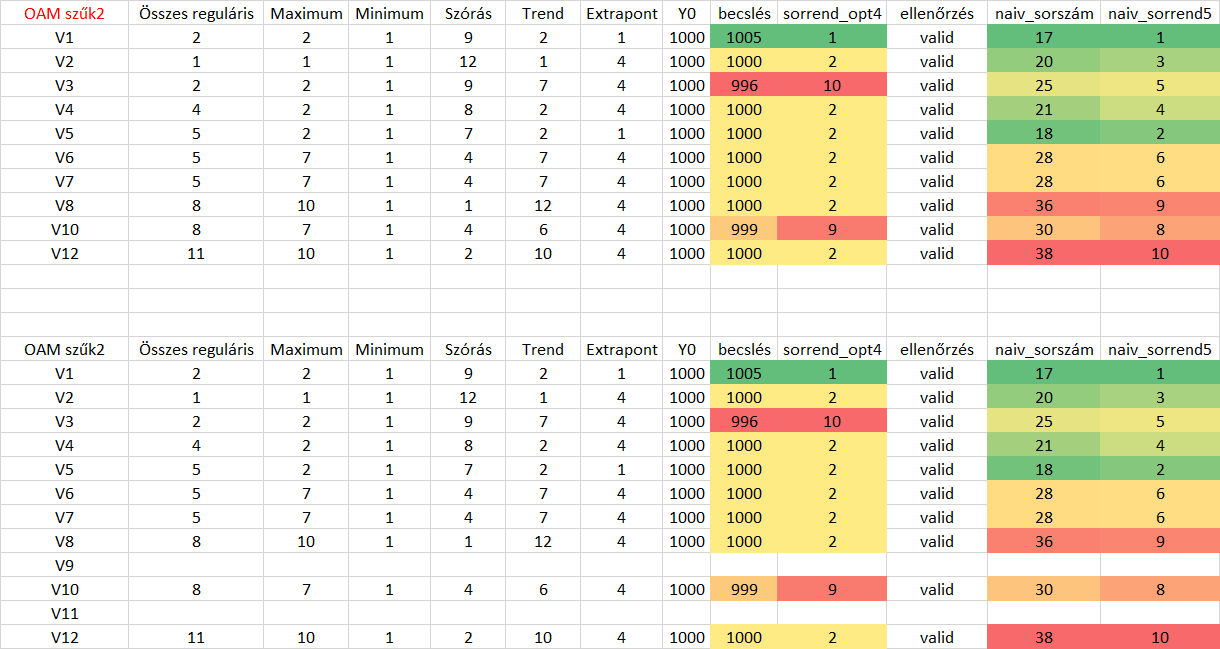
* A V11 és V12 egymás antagonistája, azaz egyikük mindenben jobb / semmiben nem rosszabb, mint a másik, s mégis azonos értékűnek kellene tudni értelmezni őket, ami lehetetlen.
* A V4 és a V9 hasonlóképpen antagonista.
* A V4>V9, ill. a V12>V11 – elemzés nélkül is beláthatóan.
* Az eredeti holtverseny (V1=V2) utáni V1>V2, s a „V2 csak a 7. helyezett illene, hogy legyen” durva kontraszt finomodik: a V2=5 (sorrend\_opt2)
* A V9 12. helyének újabb megerősítése a pontösszeg szerinti 8. hellyel szemben az antagonizmus felismerésén keresztül új szempontot hoz be a mindent összemosó összegzett pontszám-alapú értékelésbe.

A 4. ábra alsó része rámutat arra, hogy

* Az antagonisták kizárása mellett csökken a versenyzői létszám.
* Ahol a kizárás pl. közbeszerzés esetén azért lehet jogszerű, mert stróman-elv mentén ható objektumok rosszhiszeműsége alappal vélelmezhető.



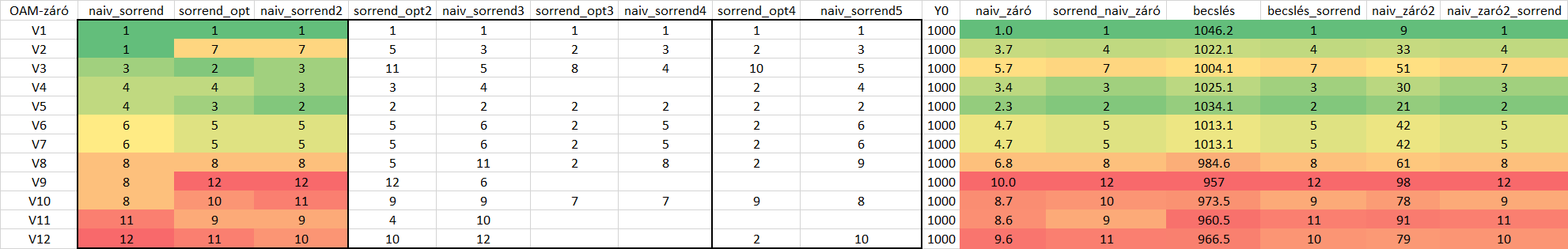
1. Ábra: Az erősen szűkített OAM és hatásai (forrás: saját ábrázolás)



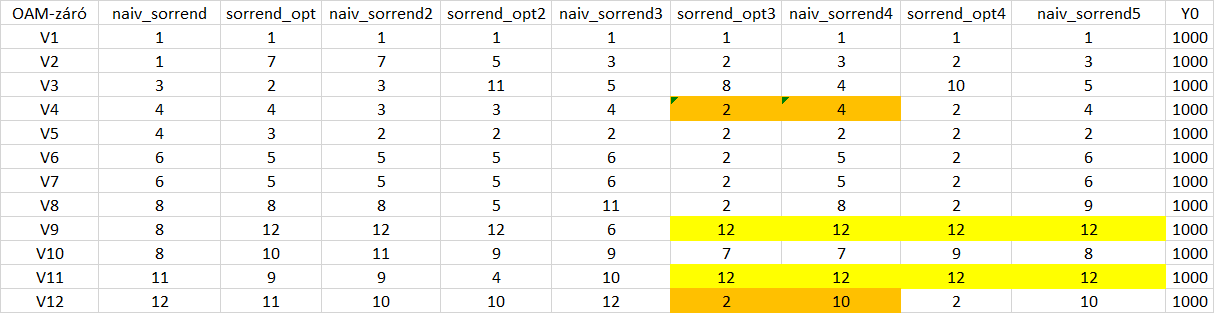
1. Ábra: A kevésbé szűkített OAM és hatásai (forrás: saját ábrázolás)

Az 5. és a 6. ábra tehát a 2+2 antagonista törlését, ill. csak a 2 gyenge antagonista törlését mutatja be, ahol a V2 immár akár 2. is lehetne a naiv holtverseny után, de közben ez a 2. helyezés normaszerűen bevonz sok további versenyzőt.

Itt kell megjegyezni, hogy az anti-diszkriminatív értékelés önkorlátozó, ismeri a valid/invalid önértékelési állapotokat, amikor is egy input OAM inverzének hatására nem inverz eredmény áll elő, akkor egyes objektumok (versenyzők) kapcsán a robot semmilyen értékelést nem fog mondani, vagyis úm. bölcsen hallgat ahelyett, hogy fecsegne, mint pl. egy regressziós egyenes minden pillanatban növekedést vagy csökkenést mutatva ott, ahol az alaplázgörbe per definitionem hullámzik…



1. Ábra: Az összes eddig naiv és optimalizált sorrend együttes hatása (forrás: saját ábrázolás)

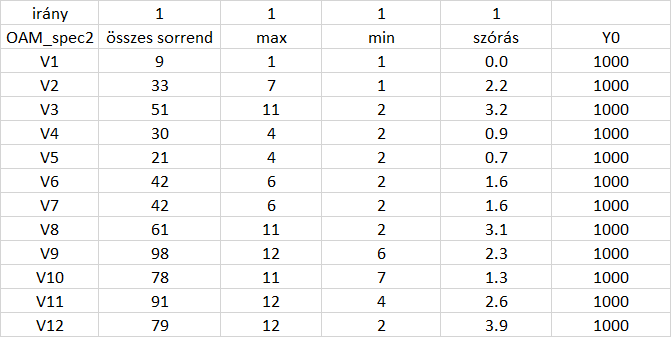


1. Ábra: A 7. ábra adathiányainak kezelése (forrás: saját ábrázolás)

A 7. és 8. ábra mintegy parlamenti patkó összefoglalja az összes létező (legitim) véleményt és keresi (ismét csak optimalizációval és naiv módon is) ennek eredőjét.

Az antagonisták kapcsán a kizárt gyengék a 12. helyet kapják a hiányzó sorszámaik helyére, míg a jobbak a szűkített kizárás eredményeit – mert a jobbakat nem lehet túlbüntetni a 12. hellyel, hiszen ők jobbak, mint a gyengébbek – éppen ezért létezik az antagonizmus.

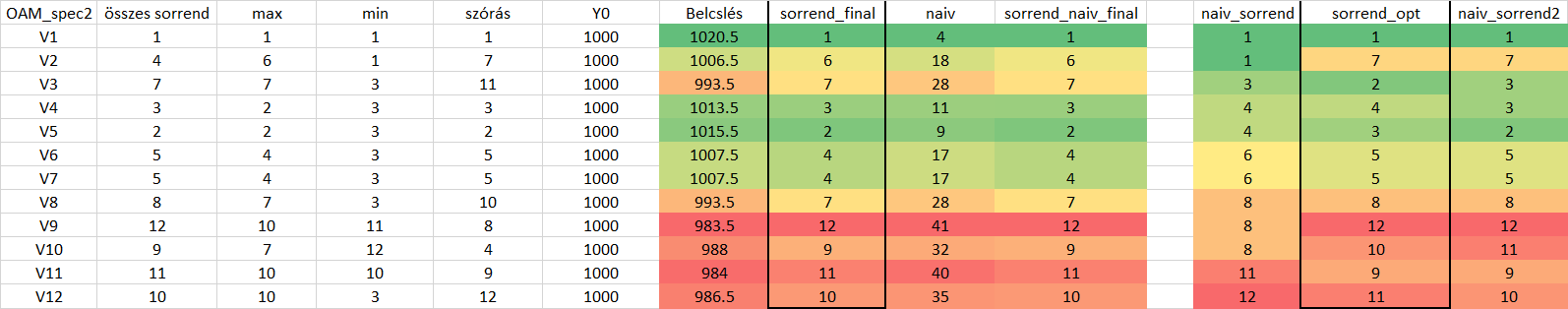
A konklúzió a V2 szempontjából követve már egy egyre racionális 4. helyezést mutat a kiindulási holtverseny és a „gyorsteszt” 7. helyezéséhez képest.



1. Ábra: A speciális nézet alkalmazása az összes sorrendre (forrás: saját számítások)

A 9. ábra kapcsán folytatódik a csipp-csipp-csóka / masztaba-épő játék: vagyis ismét lehet egy alternatív szemléletű jóság-leképezést készíteni, ahol

* A sorszámok összege legyen értelemszerűen minél kisebb, annál jobb…
* A sorszámok maximuma esetén is igaz, hogy a minél kisebb állapot esetén beszélhetünk az annál jobb helyezésről.
* A sorszámok minimuma esetén is igaz, hogy minél kisebb, annál jobb az alapelv.
* A sorszámok szórása kapcsán igaz, hogy a szórás (kiszámíthatóság, egyenszilárdság vs. a véletlen pontszerzés értékként való elismerése verseng egymással), így a minél kisebb, annál jobb elv itt is igaznak fogadható el.



1. Ábra: Az új speciális nézet eredményei (forrás: saját ábrázolás)

A 10. ábra nem tekinthető a végeredménynek, noha a végén áll a fenti lépéssornak. Az összes, a 7. ábra inputoldalához képest új sorrend új 7B állapotot eredményezhet, vagyis az értékelés egy fajta végtelen ciklusba keveredhet.

A V2 a 10. ábra alapján visszaesett újra a 6. helyre – vagyis a 10. ábra és az 1. ábra egy fajta kör megtételét sejtetik az esetlegesen végtelen spirálban…

A V1-2-3-4-5 az 1-7 sorszámintervallumban ingadozik megközelítéstől függően. A V1 első helye nem vitatható el. De a további helyezések kapcsán előzetesen kell deklarálni, mit tekint majd az értékelő robot valóban értékesnek (vö. speciális nézetek attribútumainak befogadni akarása).

# Konklúziók

Mint látható, egy/minden iskolai feladat objektív, optimalizált értékelése lényegében a Jóról alkotott filozófiai viták melegágya, noha ez a gondolati komplexitás egyszerűen említésre sem kerül a pedagógiai gyakorlatban az iskolai mindennapokban.

Még szerencse, hogy létezik a mindenki-másként-egyforma-elvre való utalás, egy fajta vigaszként azoknak, akik a buta, leegyszerűsítő, egysíkú értékelési rendszer károsultjaiként tekinthetők.

Az optimalizálás kapcsán azonban még fontosabbá válik, mit is akarunk JÓ-nak tekinteni: pl. a véletlenszerű találatok, a tudás inkonzisztenciájának jelei, ill. a tudás egyenszilárdságának gyanúja büntetendő-e?

# Kiegészítő feladat

Mi a hiba? <https://miau.my-x.hu/miau/273/naiv_optimalizalt_verziok.xlsx>

Elvileg hibátlan megoldás: <https://miau.my-x.hu/miau/273/naiv_optimalizalt_verziok2.xlsx>

A folyamat megértésének, ellenőrzésének önellenőrző gyakorlata az, ha valaki meg tudja találni a hibás xlsx-ben azt a logikai zavart, mely a V1-től elvette volna a győzelmet…

# Sokal-effektusok gyanúi

Jelen dokumentum tartalmaz apró zavarokat a gondolatmenetben/bemutatásban/elemzési folyamatban, melyek észre nem vétele, nem jelzése az Olvasók által kockázatok forrása: vö.

* Az antagonista (királycsináló objektumok kapcsán a gyengék jelölése)
* Az antagonista objektumok csak 2 objektum törlése

# Referenciák

…szövegközben…