

# Beosztások igazságosságának vizsgálata

Rónai-Kovács Martin

2022

## Kivonat

Ebben az összefoglalóban a beosztásokra vonatkozó igazságosság fogalmát próbáljuk megfogalmazni, mint általános mérték. Az ezzel kapcsolatos legnagyobb probléma, hogy a legtöbb beosztás esetén egyéni algoritmusok készítenek beosztásokat, egyéni be- és kimenetekre. Megállapítunk általánosítási eljárásokat és fogalmakat, és ezeket használva definiáljuk a hasonlóságot bemenetek között és az igazságosságot, mint az ettől való eltérés mértékét.<sup>1</sup>

## 1. Bevezetés

A beosztási problémák megoldása, lévén egy NP-teljes probléma, nagy körütekintést, hatékonyságot, és igen gyakran problémára specifikált megoldási módszert igényel. Emiatt kényszerűen el kell fogadnunk egy-egy elkészült beosztásról, hogy annál jobbat nem fogunk találni, megfelel a pszeudo-tökéletes eredmény is.

Ezzel az esetek nagy részében ki is lehet egyezni, azonban fontos kérdés, hogy milyen szempontból nem felel meg az adott beosztási eredmény. Egyes beosztásokban tapasztalható, hogy a beosztó algoritmus olyan szabályokat és egyszerűsítéseket tanul meg a már létező mintákból, amit mi nem szeretnénk figyelembe venni, például nemet, rasszot, kort, párkapcsolati státuszt, márkát, vagy bármi hasonlót. Ezt szeretnénk észrevenni.

A beosztási probléma sokrétűsége miatt azonban ennek minden problémára születhet egyéni meghatározása és kimutatási algoritmus, pont ugyanúgy, ahogy a beosztási algoritmus is általában egy-egy problémára születik meg.

Ezt szeretnénk mi áthidalni azzal, hogy megállapítunk egy általánosított eljárást, amivel az igazságosság (amiről fentebb esett szó) mértéke megállapítható egy tetszőleges beosztási problémára készült beosztásról.

A módszer megalkotása során következő célokat határozzuk meg:

1. A módszer legyen problémafüggetlen, ideértve a problémához tartozó be- és kimenetek szemantikáját és formátumát.
2. A módszer legyen algoritmusfüggetlen, tehát ne vegye figyelembe a kapott be- és kimenet közti logikai kapcsolatot.
3. A módszer a lehető legkevesebb emberi beavatkozást igényelje. Ez lényegében azért lesz fontos szempont, mert feltételezzük az igazságosság objektív létét, illetve el szeretnénk kerülni, hogy a megválaszolatlan kérdéseket problémáfüggő megoldásokkal helyettesítsük.
4. A módszer eredménye (kimenete) legyen számszerű; hiszen így állapítható meg, hogy egyes beosztások igazságosabbak, mint mások.

Amit mi keresünk, az egy  $f(\{I_S, O_S\}) = \varphi \in \mathbb{R}$  függvény, ahol  $\{I_S, O_S\}$  egy tetszőleges beosztási probléma be- és kimenete. Ezen  $f$  megalkotásakor kell a fenti szempontokat megvalósítanunk.

## 2. Beosztások közti reláció

**2.1. Definíció** (Beosztás). A beosztás egy tetszőleges beosztási probléma egy bemenete és az arra (egy tetszőleges beosztáskészítő által) adott kimenet.

$$S = \{I_S, O_S\}, \quad (1)$$

ahol  $I_S$  az  $S$  beosztás bemenete, míg  $O_S$  az  $S$  beosztás kimenete.

Tegyük fel, hogy a beosztásainkat elkészítő folyamat egy  $\sigma$  függvény, tehát  $\sigma(I_S) = O_S$ . Logikus lehet feltételezni azt, hogy eltérő bemenetekhez nem minden  $\sigma$  esetén lehet ugyanazt a kimenetet rendelni.

**2.1. Tétel.**

$$\exists \sigma : \sigma(I_S) \neq \sigma(I_{\Sigma}) \quad (2)$$

<sup>1</sup>Ezt előre fogalmaztam meg, nem ez lesz ideírva

Ennek jelentősége, hogy a beosztásokban az igazságosságot a kimenetek is szabják; ha egy kimenet eleve nem lehet elég megfelelő (mert olyan "szerencsétlen" bemenethez tartozik), akkor a hozzá tartozó igazságossági mérték sem az adott beosztásról ad nekünk információt: előfordulhat, hogy egy megadott bemenet a lehető legoptimálisabb beosztás igazságossága is rosszabb, mint egy másik bemenethez tartozó legrosszabb beosztásé.

**2.2. Definíció** (Izomorfia). Két beosztás pontosan akkor izomorf, ha azonos bemenetük van.

$$S \cong \Sigma \iff I_S = I_\Sigma \quad (3)$$

**Megjegyzés.** Két beosztás pontosan akkor izomorf, ha a beosztási algoritmus átalakításával elérhető, hogy a kimenetek azonosak legyenek.

**2.2. Tétel.** Két beosztás igazságossága között csakis akkor létezik reláció, ha a beosztások izomorfak.

$$\exists \theta \in \{<, >, =\} : f(S) \theta f(\Sigma) \iff S \cong \Sigma \quad (4)$$

Azaz a trihotómia nem teljesül bármely két beosztás igazságosságára.

### 3. Általánosítás

Határozzuk meg pontosabban a bemenetet. Egy beosztás általában entitások egymáshoz rendeléseként fogható fel, tulajdonképpen az előre ismert tulajdonságokhoz állapítunk meg relációkat és az ezekhez tartozó egyéb tulajdonságokat tudjuk még kiszámolni.

**Példa:** CPU-magok és feladatok esetén az entitásaink a magok és a feladatok. A kimenetünk az, hogy melyik feladatot melyik mag végzi el, illetve hogy melyik maghoz hány feladat tartozik (ez utóbbi egy számított tulajdonság). Kiszámítható még, hogy melyik magnak mekkora az összerhelése, de itt már számíthatunk idővel vagy feladatszámval is (Ez még problémákhoz vezet, lásd később).

Azonban a bemenetek leírása egyszerű: minden entitásunkat egy vektorként kell értelmezzük, melynek elemei a paraméterértékek. Mivel többféle vektorra lehet szükség (mert egy CPU-magot nem ugyanúgy írunk le, mint egy feladatot), ezért a vektorok mérete eltérő.

Először határozzuk meg a beosztási problémát. Minden problémában entitások (beosztandók) és követelmények szerepelnek. Így adódik a következő egyszerű leírás:

**3.1. Definíció** (Beosztási probléma). Egy beosztási probléma definíciója a példányok leírásainak halmaza és a követelmények halmaza:

$$D_S = \{E_S, C_S\}, \quad (5)$$

ahol  $E_S$  a probléma entitásainak leírásait (entitásosztályok),  $C_S$  pedig a probléma követelményeit tartalmazza.

**Megjegyzés.**  $E_S$  és  $C_S$  pontos formájával nem kell foglalkoznunk, mert nem használjuk őket (ezek a  $\sigma$  függvény definiálásához kellenek); elég tudnunk, hogy  $E_S$  az entitásosztályok egy *halmaza*.

**Megjegyzés.** Az entitások itt nem nyers adatként jelennek meg. Úgy kell őket elképzelnünk, mint egy paraméterszámítási *függvény* eredményét ( $p(r) = \varepsilon$ ), ami a nyers adatokból releváns információkat nyer ki számunkra. Erre azért van szükség, mert például zárvízsgabeosztásnál mindegy, mely hallgatókat vizsgáztatja egy konzulens, a kérdés az, hány hallgatót vizsgáztat egy konzulens.

**3.2. Definíció** (Csoport). Azon entitások, melyek egy osztályba tartoznak, egy csoportot alkotnak. Egy  $e \in E_S$  osztályhoz tartozó csoportot a következőképpen definiálunk:

$$G(e) = \bigcup_{\varepsilon : e} \varepsilon \quad (6)$$

Ahol az  $\varepsilon : e$  jelentése, hogy  $\varepsilon$  az  $e$  osztályleírásnak megfelel.

**3.3. Definíció** (Homogenitás). Két entitás homogén egymással, ha azonos osztályhoz tartoznak.

$$\varepsilon_1 \sim \varepsilon_2 \iff \varepsilon_1 : e \wedge \varepsilon_2 : e \quad (7)$$

Az egyes entitások leírása bővül a beosztás ( $\sigma$ ) hatteretése után, ezért érdemes megkülönböztetni a bemeneti és kimeneti állapotokat.

**3.4. Definíció** (Be- és kimeneti állapotleírás). Egy  $\varepsilon$  entitás bemeneti leírása  $\varepsilon_{in}$ , kimeneti leírása  $\varepsilon_{out}$ ; hasonlóan a be- és kimeneti osztályleírás  $e_{in}$  és  $e_{out}$ .

Ezen definíciók ismeretében már meg tudjuk határozni a bemenetet.

**3.5. Definíció** (Bemenet). A bemenet egy  $S$  beosztásra ( $I_S$ ) megegyezik a csoportok halmazával.

$$I_S = \bigcup_{e \in E_S} G(e_{in}), \quad (8)$$

tehát  $I_S$  azonos méretű vektorok halmazainak halmaza.

Az egyes entitásosztályok külön bemenetként való kezeléséhez határozzuk meg a csoportonkénti bemenetet:

**3.6. Definíció** (Csoportonkénti bemenet). Az  $e$  entitásosztályhoz tartozó csoportonkénti bemenet  $I_{S,e}$  megegyezik  $e_{\text{in}}$  csoportjával.

$$I_{S,e} = G(e_{\text{in}}) \quad (9)$$

Tehát meghatároztuk a beosztási problémánk bemenetének leírását. Tételezzük fel, hogy a kimenetet hasonlóan írhatjuk le, tehát a fenti definíciókban  $I_S$  felcserélhető  $O_S$ -sel, amennyiben az  $e_{\text{in}}$ -t is felcseréljük  $e_{\text{out}}$ -tal, stb.

**3.1. Tétel** (Normalizáció I.). Minden bemenetnek létezik normáltja, vagyis

$$\forall I_{S,e} : \exists \| I_{S,e} \| \quad (10)$$

*Bizonyítás.* Használjuk az alábbi lemmát:

**3.2. Lemma** (Entitás-vektor azonosság). Minden entitás leírható valós vektorként, azaz  $\varepsilon = \underline{\varepsilon} \in \mathbb{R}^n$ , ahol  $n$  az  $\varepsilon$  entitás paramétereinek száma (ideértve a számított paramétereket is).

$\| I_{S,e} \|$  azonos a benne lévő entitások normáltjaival, hiszen azok egymástól függetlenek. Mivel ezek valós vektorok  $\mathbb{R}^n$ -ben (3.2), létezik normáltjuk, és így  $I_{S,e}$ -nek is létezik normáltja.  $\square$

**Megjegyzés.** A tétel fontossága, hogy tetszőleges beosztási problémánál a bemenetet általánosítani tudjuk.

**Következmény.** Csakis akkor tudunk általánosítani, ha a bemeneti vektorok normálhatók.

## 4. Paraméterek függése

Az eddigiek alapján szeretnénk elérni, hogy az igazságossági függvény ( $f$ ) bemenetei általánosak legyenek, és ehhez meghatároztuk a befogadható formákat. A teljes általánosításhoz azonban el kell végeznünk a normalizációt is, hogy tényleg tetszőleges entitásokkal tudjunk dolgozni.

A probléma itt a normalizációval van, hiszen meg kell határoznunk, hogy hogyan lehetne tulajdonságvesztés nélkül, a paraméterek egymáshoz való relációjának megtartásával együtt módosítani az entitásokat.

Mivel

$$\| I_S \| = \bigcup_{e \in E_S} \| I_{S,e} \|, \quad (11)$$

a teljes bemenetet tudjuk normálni, amennyiben az csoportokat egyesével tudjuk normálni. Ezt meg tudjuk tenni, mivel a csoport entitásai egymástól függetlenek (ahogy azt a (3.1) tétel bizonyításakor feltételeztük).

Logikusnak tűnhet, hogy (mivel valós vektorokról van szó) ezeket egész egyszerűen euklideszi hosszúsággal leosszuk:

$$\| \underline{v} \| = \frac{1}{\|\underline{v}\|} \underline{v}$$

így kapva egy egység hosszú vektort tetszőleges  $\mathbb{R}^n$  térben. Azonban ezzel elkövethetjük azt a hibát, hogy két relatíve távoli pont távolságát 0-ra csökkentjük.

**Példa:**  $\underline{u} = (0, 1, 1)$ ;  $\underline{v} = (1, 1, 1)$ ;  $\underline{w} = (10, 10, 10)$  vektorok esetén normálás során  $\underline{v}$  és  $\underline{w}$  távolsága 15.588-ról 0-ra csökken, míg  $\underline{v}$  és  $\underline{u}$  távolsága 1-ről 0.605-re csökken; tehát a  $\underline{v} - \underline{w}$  és  $\underline{v} - \underline{u}$  közti reláció megfordul, ha ezeket egyesével normáljuk (jelen példában feltételeztük, hogy két vektor távolsága megegyezik azok különbségével).

A példában látható jelenséget szeretnénk elkerülni, így máshogyan kell a normalizációt definiálnunk.

**4.1. Definíció** (Normalizáció). Egy  $\underline{\varepsilon}$  entitás normáltja  $\nu(\underline{\varepsilon})$  (vagy  $\| \underline{\varepsilon} \|$ ). A normálás az azonos paraméterek közti relációt megtartja, de a vektort véges térbe transzformálja.

$$\nu : \mathbb{R}^n \rightarrow [r_1; r_2]^n, \quad (12)$$

ahol  $r_1, r_2 \in \mathbb{R}$  és  $r_1 < r_2$ .

**Megjegyzés.** Az azonos paraméterek közti relációt meg kell tartanunk, hogy ne forduljon elő a fenti példában látott eset (reláció megfordulása). További előny, hogy a paraméter normáltja ez esetben függeni fog a többi entitás értékskálájától, de nem függ a többi paramétertől.

**Megjegyzés.** A továbbiakban  $r_1 = 0$  és  $r_2 = 1$  értékekkel használjuk a normalizáció fogalmát.

Hogy a (4.1) definíciónak megfeleljünk, feltehetjük, hogy a paramétereket egyenként normálhatjuk. Ennek azonban van egy feltétele, ami az alábbi tételben olvasható.

**4.1. Tétel** (Normalizáció II.). Amennyiben az entitás paraméterei függetlenek, az entitás normáltja megegyezik az entitás paramétereinek normáltjával:

$$\forall i \neq j : \underline{\varepsilon}_i \not\propto \underline{\varepsilon}_j \implies \| \underline{\varepsilon} \|_i = \| \underline{\varepsilon}_i \| \quad (13)$$

**4.2. Tétel** (Visszaállíthatóság).  $\nu$  reverzibilis,  $\exists \nu^{-1}$ .

*Bizonyítás.* Mivel  $f$  csak adott bemenetre ad szám-szerű értéket (2.2), ismerjük a teljes lehetséges értékészletet minden paraméterre. Ebben az esetben (tehát mindig) belátható, hogy

$$\nu_i(\varepsilon) : [r_1; r_2] \rightarrow [0; 1],$$

ahol  $r_1 = \arg_{\varepsilon} \min(\varepsilon_i)$  és  $r_2 = \arg_{\varepsilon} \max(\varepsilon_i)$ . Tudjuk, hogy minden entításra igaz, hogy azonos paraméterek között a relációs viszony megmarad (4.1), így

$\nu_i^{-1}$  maga is megfelel a normalizáció fogalmának. Ekkor könnyen választhatunk egy olyan értékészletet, amely az eredeti paraméter értelmezési tartománya. Ezzel megalkottuk  $\nu$  inverzét. (És megállapítottuk, hogy  $\nu$  bijektív).  $\square$

**Megjegyzés.** Ezzel beláttuk, hogy normalizáció során nem történik információvesztés.

## 5. Normalizációs problémák

A normalizációval beláttuk, hogy tetszőleges beosztási probléma (mely leírható a (3.1) formában) általánosítható.

Ennek azonban feltétele volt, hogy az osztályok, az entitások, és azok paraméterei egymástól függetlenek legyenek. Ez főként a  $p$  függvény (3.1) megalkotásában jelentős szempont.

A másik aggály, ami felmerülhet, hogy (véges pontosságú eszközök miatt) egyes adatok irrelevánssá válhatnak a normalizáció során. Például lehetnek olyan paraméterek, melyeket nem lineárisan, hanem logaritmikusan kell normálni, mivel maguk a paraméterek exponenciálisak. Ez a probléma elkerülhető, ha a normalizációs ( $\nu_i$ ) függvényeket minden paraméterre körültekintően választjuk meg. Ez talán még automatizálható is regressziókkal.

Ha a normalizáció során elérjük, hogy a skálán elhelyezkedő értékek ugyanolyan eloszlásúak legyenek (lehetőleg lineárisak), akkor tetszőleges paraméter *tekinthető* lineárisnak. Ennek jelentősége a továbbiakban jelentős, hiszen ezzel elérjük, hogy két entitás távolságának számításakor a paraméterek homogének legyenek eloszlás szempontjából (tehát ne kelljen minden paraméterre meghatározni egy "távolságbefolyásoló tényezőt", vagy hasonlót).

## 6. Távolságok

A távolságok nagy szerepet játszanak az igazságosság megállapításában, hiszen a beosztás előtti és utáni távolságok közti eltérésből deríthetjük ki a legszembevetőbb anomáliákat (ha csak 1GHz-es CPU-ink vannak, rendben van, hogy mind ugyanannyi feladatot kap; de ha van mellette egy 10GHz-es is, akkor az egyenlő szétosztás nem igazságos).

Mivel a bemeneteken azok a paraméterek, melyeket csak a beosztás befolyásol nem lesznek egymástól távoliak (mindegyik alapértéken lesz), feltehető, hogy

<sup>2</sup>Szükség van linearitásra? Nem elég, ha mind ugyanolyan? Egyáltalán ez az engedmény valóban engedmény lenne (azon kívül, hogy általánosabb esetet fedne a definíció)?

a bemeneteken a távolságok kisebbek, mint a kimeneten:

$$d_{in}(x, y) \leq d_{out}(x, y)$$

Ugyanezt megfogalmazhatjuk egyszerűbben, ha a be és kimeneten ugyanaz a függvény méri a távolságot ( $d_{in} \Leftrightarrow d_{out}$ ):

$$d(x_{in}, y_{in}) \leq d(x_{out}, y_{out}) \quad (14)$$

Itt egy feladatunk maradt: határozzuk meg  $d$ -t.

**6.1. Definíció** (Távolságfüggvény). Az igazságossági távolságfüggvény két azonosan paraméterezett entitásról ad olyan arányszámot, melyből azoknak különbsége megállapítható. A beosztásban akkor állapíthatunk meg igazságtalanságot vagy igazságossági anomáliát, ha a be- és kimeneten mért távolságok jelentősen eltérnek.

$$d : \mathbb{R}^{2|e|} \rightarrow \mathbb{R} \quad (15)$$

ahol  $|e|$  az  $e$  entitásozstály paramétereinek száma (azaz egy  $\varepsilon : e$  vektor mérete).

Amennyiben valóban feltehetjük, hogy a paraméterek eloszlása azonos, minden két példány paraméterein ugyanazt a műveletet kell elvégeznünk, majd egyetlen összegző művelettel megállapíthatjuk, hogy az ezen paramétereket tartalmazó entitások távolsága mekkora.

**6.2. Definíció.** (Távolság) Két entitás távolsága az őket leíró vektorok euklideszi távolsága (lineáris eloszlású normáltak esetén <sup>2</sup>).

$$d(x, y) = \sqrt{\sum_{i=1}^{|e|} (x_i - y_i)^2} \quad (16)$$

ahol  $x \sim y$ .

Ezután állapítsuk meg, miért fontos ez számunkra. Igazságosnak akkor tekintünk egy beosztást, ha a beosztás elkészítése során nem torzítjuk azokat a távolságokat, melyek már a bemeneten is léteznek. Ezt felfoghatjuk úgy is, mint a Lipschitz-tulajdonság betartását (valamely  $M$  konstansra):

$$\forall x \sim y : d(x_{in}, y_{in}) \leq M d(x_{out}, y_{out}) \quad (17)$$

Egy beosztástól azt szeretnénk, hogy a kimenetek távolságai arányosságban álljanak a bemeneti távolságokkal, tehát legyen igaz, hogy

$$d(x_{out}, y_{out}) \propto d(x_{in}, y_{in}). \quad (18)$$

Egyáltalán ez az engedmény valóban engedmény lenne (azon

**6.3. Definíció** (különbségváltozás,  $\delta$ ). Két egymással homogén entitás különbségváltozása megegyezik a bemeneti és kimeneti különbségük közti eltéréssel. Jele:  $\delta$ .

$$\delta(x, y) = d(x_{\text{out}}, y_{\text{out}}) - d(x_{\text{in}}, y_{\text{in}}) \quad (19)$$

Az eddigiekből következik, hogy a fairness-függvény értéke az egyes párok  $\delta$ -jából lesz számítva; ezt az átszámítást végezze esetünkben egy  $u$  (unfairness) függvény,  $f_e$  jelezze az  $e$  entitásosztályon belüli igazságosságot:

$$f_e(S) \propto u_e \left( \bigcup_{x \sim y} \delta(x, y) \right) \quad (20)$$

ahol  $x : e$ . Ezután a teljes fairségről megállapíthatjuk az egyszerű összefüggést:

$$f(S) \propto \sum_{e \in E_S} f_e(S) \quad (21)$$

Az eddig megállapítottakat kell a fairness-függvény megalkotásakor figyelembe venni; ez alól kivétel a  $d$  pontos definíciója, hiszen az euklideszi távolság valódi távolságot mér két entitás között <sup>3</sup>.

## 7. Súlyozás

Felmerülhetnek az igazságossággal kapcsolatban további kérdések, melyek megválaszolásával optimálisabb vagy akár gyorsabban számítható eredményeket kaphatunk. Ilyen lehet például, hogy minden entitás közti igazságosság kell-e (pl. időszeletek egyformasága számít?), vagy vannak-e olyan paraméterek, amiket érzékenyebben kell kezelni, mint a többi (pl. a "rassz" vagy "párkapcsolati státusz" tulajdonság és a hitelbírálat sikeressége nem jó, ha korrelál, így ezt érzékenyebben vizsgálhatjuk  $f$ -hez).

Paraméterek fontosságának kijelentéséhez definiáljuk a súlyokat.

Innentől jelölje jobb felső indexben zárójelezett változó az entitás osztályát:  $x^{(e)} \Leftrightarrow x : e$ .

**7.1. Definíció** (Súly). Egy osztály egy példányának  $x_i^{(e)}$  paraméteréhez tartozó súly azt fejezi ki, mennyire befolyásolja a fairséget egy  $y \sim x$  példány  $y_i^{(e)}$  paraméterétől való erős eltérés a be- és kimenet között.

Ezen súly minden esetben 0 és 1 közötti értéket vesz fel.

$$\underline{w}^{(e)} \in [0; 1]^{|e|} \quad (22)$$

**7.2. Definíció** (Súlyozott entitás). Egy entitás súlyozása a paraméterenkénti szorzat számításából adódik:

$$X^{(e)} = w^{(e)} x^{(e)} \quad (23)$$

**Megjegyzés.** Ha  $w^{(e)}$  minden eleme 1, akkor a súlyozás eredeti állapotot hagy, míg ha minden eleme 0, akkor az összes létező beosztást "igazságosra" módosítja.

**7.1. Tétel** (Awareness - unawareness átjárhatóság). Bináris súlyok használatával elérhető, hogy a "fairness through unawareness" definíciót felhasználjuk, annak ellenére, hogy a "fairness through awareness" szerint haladt az eddigi gondolatmenetünk. Ekkor az érdektelen változók súlyát 0-ra, a többi súlyt 1-re kell beállítanunk.

## 8. Súlyozási problémák

...

## 9. Igazságosság megállapításának folyamata

...

## 10. Algoritmusfüggetlenség

...

<sup>3</sup>Ha egy paraméter pl.  $e^x$  jellegű értékeket vesz fel, azaz exponenciális eloszlású, akkor a paraméteren mért távolság torz; ezért figyelniük kell a normáltakra, hogy lineárisak legyenek.