Budapesti Corvinus Egyetem
Közgazdaságtudományi kar
Matematikai közgazdaságtan tanszék

Az egyensúly vélekedés-alapú finomításai szignáljátékokban

Stancsics Martin
Gazdaságelemzés és gazdaságmodellezés szekció
Gazdaságelemzés szak
2014.

Konzulens: Dr. Gömöri András

Tartalomjegyzék

BEVEZETÉS, MOTIVÁCIÓ

A szignáljátékokban, mint a nem teljes információs szekvenciális játékokban általában, a tökéletes bayesi egyensúly a leggyakrabban használt megoldáskoncepció. Azonban sajnos egy szignáljátékhoz gyakran sok ilyen egyensúly tartozik, melyek közül egy vagy több nem tűnik ésszerűnek. Kiterjedt szakirodalma van az egyensúly lehetséges finomításainak, illetve különböző kritériumoknak, melyek arra szolgálnak, hogy kizárják ezen egyensúlyok közül azokat, amik nem tűnnek intuitívnak. Használhatók ezen egyensúlyok eliminálására néhány esetben hagyományos Nash-egyensúly finomítások, mint pl. a remegő kéz egyensúly (trembling hand equilibrium (Selten, 1975)), előre indukcióra épülő finomítások és kritériumok, pl. stabil egyensúly (Kohlberg & Mertens, 1986), Mertens-stabil egyensúly (Mertens, 1989), intuitív kritérium (Cho & Kreps, 1987), divinity kritérium (Banks & Sobel, 1987) , veretlen egyensúly (Mailath, Okuno-Fujiwara, & Postlewaite, 1993), illetve történtek próbálkozások más területek koncepcióinak alkalmazására e probléma megoldására, mint. pl. az evolúciósan stabil egyensúly használata (Jäger, 2008).

Ezen dolgozatban bemutatok egy új kritériumot – amelyet komparatív kritériumnak fogok nevezni – a nem intuitív egyensúlyok szűrésére szignáljátékok esetén, ami koncepciójában leginkább a veretlen egyensúlyra hasonlít, azonban diszkrét játékokra (olyan játékokra, amelyben véges sok szereplő és véges sok lehetséges stratégia van) van definiálva. Bemutatom továbbá a legelterjedtebb előre indukción alapuló kritériumot, Cho és Kreps intuitív kritériumát. Ezután példákon keresztül összehasonlítom a két kritériumot, és ezeken keresztül megmutatom, hogy bár látszatra hasonló a két kritérium, a koncepció mögöttük egészen különböző.

SZIGNÁLJÁTÉKOK

Szignáljátékoknak nevezzük azokat a szekvenciális játékokat, ahol az egyik fél olyan, a játék kimenetele szempontjából releváns információval rendelkezik, amelyet a másik nem ismer (hívjuk ezt az információt A típusának), továbbá a jól informált fél (akit ezután A-nak nevezünk) lép először. A a rendelkezésére álló rejtett információ alapján választ a rendelkezésére álló akciók közül, ezért a lépéséből B következtetni tud A privát információjára, A lépése egy üzenet. B megfigyeli ezt az üzenetet, frissíti a vélekedéseit A rejtett információt illetően, és ez alapján hozza meg a döntését.

ÁLTALÁNOS MODELL, FORMALIZMUS

Tekintsünk most egy általános szignálmodellt két szereplővel, Cho és Kreps jelöléseit használva (Cho & Kreps, 1987). A két szereplőt hívjuk A-nak és B-nek. Először a természet lép, kiválasztja A típusát (t) egy véges T halmazból a π valószínűségi eloszlás által meghatározott módon. A megfigyeli ezt az információt, B azonban csak π-t ismeri, az köztudott tudás. Ezután A küld egy $m \in M$ (M véges halmaz) üzenetet B-nek. Jelölje $M(t)$ a t típusú játékos által küldhető üzenetek halmazát, $T(m)$ pedig azon típusok halmazát, akik küldhették az m üzenetet. Miután B megfigyeli az üzenetet, küld rá egy $r \in R$ (R véges halmaz) választ. Jelölje $R(m)$ azon válaszok halmazát, amelyeket B küldhet az m üzenet esetén. Ezzel a lépéssel a játék véget ér, A és B kifizetése rendre $u(t, m, r)$ és $v(t, m, r)$.

Jelölje továbbá A stratégiáját $\rho(m, t)$, ahol minden t-re $\rho(\cdot, t)$ egy valószínűségi eloszlás $M(t)$-n, B stratégiáját pedig $\phi(r, m)$, ahol minden m-re $\phi(\cdot, m)$ egy valószínűségi eloszlás $R(m)$-en. Ezt úgy kell értelmezni, hogy egy t típusú A játékos $\rho(m, t)$ valószínűséggel fogja az m üzenetet küldeni, B válasza pedig $\phi(r, m)$ valószínűséggel lesz r az m üzenetre.

B az alapján választ r-t, hogy A üzenete alapján következtet annak típusára. Legyen ez az úgynevezett korrigált vélekedés (posterior belief) μ, ami egy valószínűségi eloszlás $T(m)$-en. Jelölje továbbá $BR(\mu, m)$ B legjobb válaszát, ha a vélekedése az m üzenet után μ. Minden $I \subset T(m)$-re legyen $BR(I, m)$ válaszok halmaza, amik lehetnek legjobb válaszok valamely olyan vélekedés mellett, ahol kizárólag I-beli típusok szerepelnek pozitív valószínűséggel, tehát $BR(I, m) = \bigcup_{\{\mu|\mu(I)=1\}} BR(\mu, m)$. Legyen továbbá $MBR(\mu, m)$ B legjobb kevert válasza az m üzenetre és a μ vélekedésre, illetve $MBR(I, m) = \bigcup_{\{\mu|\mu(I)=1\}} MBR(\mu, m)$.

AZ EGYENSÚLYOK SZÁMOSSÁGA

Mertens és Kohlberg megmutatta, hogy egy véges sok játékossal és véges sok stratégiával rendelkező játékban a Nash-egyensúlyok halmaza (a tiszta stratégiákon értelmezett valószínűségi eloszlásoknak értelmezve a Nash-egyensúlyokat) véges sok összefüggő halmazból áll . Megmutatták továbbá, hogy generikus játékok esetén[1] a Nash-egyensúlyi kimenetelek halmaza véges sok pontból áll (Kohlberg & Mertens, 1986). Ez azt jelenti, az egyensúlyok végtelensége abból ered, hogy az egyensúlyi pályán kívüli vélekedések, ezáltal egyensúlyi pályán kívüli stratégiák különböznek. A kifizetésekben is különböző egyensúlyokat fogom a későbbiekben szignifikánsan különböző vagy jelentősen különböző egyensúlyként emlegetni. Nem

[1] A generikus játékokról bővebben: (Kreps & Wilson, 1982)

generikus játékoknál is elképzelhető, hogy csak véges sok szignifikánsan különböző egyensúly van csak, amint az a példák alapján látható lesz, azonban itt ezt nem biztosítja semmi, így a továbbiakban ismertetett kritériumok alkalmazása esetén meg kell bizonyosodni arról, fennáll-e ez a feltétel. Mivel a komparatív kritérium egyensúlyok összehasonlításán alapul, nagy szerepe lesz ennek, hiszen így csak véges sok összehasonlítást kell végezni.

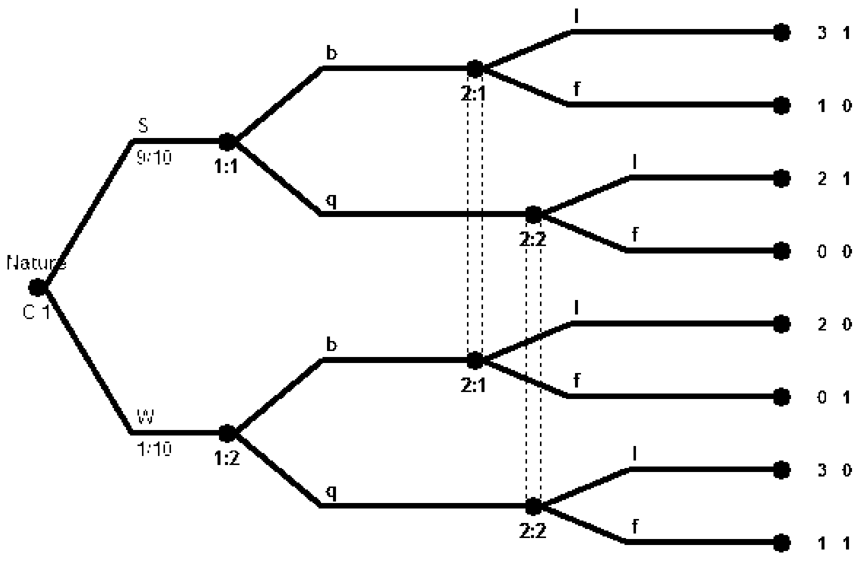
NEM INTUITÍV EGYENSÚLYOK

A nem teljes információs játékokban leggyakrabban alkalmazott megoldáskoncepció, a bayesi Nash-egyensúly (Harsányi, 1967/68), illetve ennek szekvenciális játékokra vonatkozó változata, a tökéletes bayesi egyensúly a játékosok stratégiának megadásán túl meghatározza a játékosok vélekedéseit is. A vélekedéseknek összhangban kell lenniük azzal az alapvető ténynyel, hogy a játékosok várható kifizetésüket maximalizálják. Ebből következően azokhoz az üzenetekhez, amiket a jól informált fél pozitív valószínűséggel küld az adott egyensúlyban, a Bayes-tétel alapján meg van határozva a másik játékos vélekedése. Azonban azokhoz az üzenetekhez, amelyeket a vizsgált egyensúlyban az egyensúlyi pályán soha nem választja az A játékos, B sokféle vélekedést rendelhet, csupán annak kell teljesülnie, hogy ezek legyenek konzisztensek az egyensúlyi stratégiákkal.

Ennek következtében a nem teljes információs, szekvenciális játékokban azt, hogy egy stratégiaegyüttes (a hozzá tartozó vélekedésekkel) egyensúlynak számít-e, nagyban befolyásolja, hogy a rosszul informált fél milyen vélekedéseket rendel az egyensúlyon kívüli üzenetekhez. Gyakran előfordul, hogy ezek a vélekedések valahogy nem tűnnek racionálisnak, annak ellenére, hogy az egyensúly definíciója értelmében B bármit gondolhat. Az ilyen egyensúlyok kiszűrésére az intuitív típusú kritériumok jelenthetnek megoldást. Ezeknek az a fő gondolata, hogy megpróbáljuk kiszűrni az olyan egyensúlyokat, ahol B nem gondolhatja racionálisan, hogy az m egyensúlyon kívüli üzenetet egy t típusú A küldte, mégis szerepel t B vélekedéseiben az m üzenet esetén. A dolgozat további részében arra mutatok lehetőségeket, hogy mit is jelenthet az, hogy „B nem gondolhatja racionálisan”, két intuitív típusú kritérium részletes bemutatásán keresztül.

A SÖR-KAKAÓ JÁTÉK

A következőkben Cho és Kreps – szabad fordításban – sör-kakaó (beer-quiche) játékát fogom bemutatni a nem intuitív egyensúly szemléltetésére (Cho & Kreps, 1987). Képzeljük el, hogy az A játékos a reggelijét fogyasztja egy étteremben, a parkolóban pedig egy kötekedős B játékos vár rá. A kétféle típusú lehet: kemény (s) vagy puhány (w). Azt hogy milyen típusú, a természet dönti el – A 0.9 valószínűséggel kemény, 0.1 valószínűséggel pedig puhány lesz. Ezen kívül kétféle italt választhat a reggelijéhez: sört vagy kakaót. A kemény típusú *A* a sört preferálja, a puhány pedig a kakaót. *B* nem tudja megfigyelni, hogy *A* milyen típusú, csak azt, hogy mit iszik a reggelije mellé. Miután befejezte a reggelijét, és elindul az autójához, *B* eldöntheti, hogy beleköt vagy békén hagyja. Amennyiben *A* kemény típusú, *B* békén szeretné hagyni, amennyiben puhány, szívesen belekötne. Amennyiben a megfelelő cselekvést választja, *B* kifizetése 1 lesz, ha pedig nem, akkor 0. *A* kifizetése aszerint alakul, hogy mit iszik, és hogy belekötnek-e. Ha nem kötnek bele, az +2 addicionális hasznosságot jelent neki, az pedig, ha a kedvenc italát fogyaszthatja a reggelijéhez, még egyet. A játék extenzív alakja a következőképpen néz ki:



1. ábra - A sör-kakaó játék

Először is azt kell észrevenni, hogy ebben a játékban – mint sok más szignáljátékban is – végtelen sok egyensúly van. Ha mindkét típusú *A* játékos ugyanazt a stratégiát játssza, tehát az egyensúly elvegyítő, akkor a másik akcióhoz *B* tetszőleges vélekedéseket rendelhet *A* típusára vonatkozóan addig, amik az ezekre adott legjobb válasza *B*-nek nem motiválja semelyik típusú *A*-t se arra, hogy eltérjen az egyensúlytól.

Nézzük először azt az egyensúlyt, amikor *A* mindig a sört választja. Ekkor ez az üzenet nem ad semmilyen plusz információt *B* számára, így a vélekedése *A* típusáról továbbra is az eredeti, azaz hogy 9/10 eséllyel A erős, 1/10 eséllyel pedig puhány. Erre B legjobb válasza az, hogy békén hagyja, ugyanis ez maximalizálja a várható hasznosságát. Mivel a kakaót soha sem választja A, B arra vonatkozó vélekedése tetszőleges lehet.

Vizsgáljuk meg, ezek közül melyeknél marad egyensúlyi stratégia A-nak a sör és B-nek a békén hagyás. Amíg B vélekedése olyan, hogy kakaó megfigyelése esetén harcolni fog, addig biztosan nem akar eltérni semelyik típusú A sem ettől az egyensúlytól – ez pedig akkor áll fenn, amikor B vélekedésére a legjobb válasz a harc. Jelölje p azt, hogy B szerint A mekkora eséllyel erős típusú, ha kakaót iszik – ekkor persze $1 - p$ eséllyel lesz gyenge típusú. Könnyű belátni, hogy amíg $p < 0{,}5$, addig B-nek a harc a legjobb válasz, A-nak nem éri meg eltérni, és valóban egyensúlyban vagyunk.

Összefoglalva az első egyensúlyt, A típustól függetlenül sört fog inni, B pedig ha sört figyel meg, úgy gondolja A-ról, hogy 9/10 valószínűséggel kemény típusú, ezért nem köt bele, ha pedig kakaót figyel meg, akkor úgy véli, hogy A csak $p < 0{,}5$ valószínűséggel kemény típusú, és beleköt. Senkinek sem áll érdekében egyedül eltérni ezektől a stratégiáktól, B vélekedései konzisztensek a játékosok profitmaximalizálásával, ez valóban egy tökéletes bayesi egyensúly.

Érdekességképpen gondoljuk most meg, mi történik, ha kakaó megfigyelése esetén B vélekedése A-ról az, hogy $p = 0{,}5$ valószínűséggel kemény, és ugyanekkora eséllyel puhány típusú. Ekkor, ha A kakaót iszik, B mindkét lehetséges akciója, és ezeknek minden lehetséges keverése legjobb válasz erre az üzenetre. Ezek közül csak azok alkotnak egyensúlyt az előbbi egyensúlyi pályával, amelyeknél B úgy kever, hogy semelyik típusú A-nak se érje meg eltérni. Tehát ha B q eséllyel nem harcol és $1 - q$ eséllyel harcol, akkor addig lesz ez egyensúly, amíg mindkét típusú A várható hasznossága nagyobb, ha sört iszik és nem harcol, mintha kakaót inna és $1 - q$ eséllyel harcolna. A kemény A-ra ez természetesen mindig igaz, ha $3q + q \leq 2$, azaz $q \leq 1/2$.

Eddig a vizsgált egyensúlyaink racionálisnak tűnnek. Nézzük meg most azonban azt az esetet, amikor mindkét típusú A kakaót iszik, így B ebben az esetben nem harcol, ha pedig A sört iszik, akkor B úgy gondolja, hogy A $p < 1/2$ eséllyel kemény típusú csak, így harcolni fog. Itt is teljesül a tökéletes bayesi egyensúly minden feltétele, mégis érezhető, hogy valami nincsen egészen rendben vele. Méghozzá az, hogy mégis miért gondolná B azt, ha sört figyel meg, hogy A puhány. Az egyensúly definíciója azt mondja, hogy miért ne? Egyensúlyon kívüli üzenetekre nem használható a Bayes-tétel, B vélekedése tetszőleges lehet. Azonban intuíciónk szerint mégse racionális az a feltételezés, hogy a söröző A puhány típusú, ugyanis egy puhány A csak alacsonyabb kifizetést érhet el, ha a sört választja. Az ilyen egyensúlyok kizárására szolgál az intuitív kritérium.

INTUITÍV TÍPUSÚ KRITÉRIUMOK

Most már minden adott ahhoz, hogy definiáljam azt a két kritériumot, amit vizsgálni fogok. Emlékezzünk vissza, hogy a sör-kakaó játékban az alapján zárhattam ki a kakaós egyensúlyt, hogy B-nek nem racionális azt gondolnia, hogy egy puhány A küldte az egyensúlyon kívüli sör üzenetet, hiszen ő semmiképpen nem járhatna jobban abban az esetben, mint a jelen helyzetben. Mindkét bemutatott kritériumnak ez lesz a fő gondolata: kizárjuk azokat az egyensúlyokat, amelyek olyan vélekedésre épülnek, ami nem racionális.

A módszer a következő lesz: az adott véges egyensúllyal rendelkező szignáljátékban rögzítsünk egy egyensúlyt, és legyen $u^*(t)$ a t típusú A játékos kifizetése ebben az egyensúlyban.

Első lépésként állítsunk egy kritériumot, ami alapján azt mondhatjuk, hogy egy m egyensúlyon kívüli üzenetet nem küldhetett racionálisan a t típusú játékos. Mondjuk továbbá azt, hogy egy m üzenetre B sohasem fog racionálisan r-rel válaszolni, ha $r \notin BR(T(m), m)$.

Ezután határozzunk meg egy sorrendet, amelyben ezt a két kritériumot (akár iterálva) alkalmazni fogjuk. Az elsővel $M(t)$-ből törlünk minden olyan m-et, ahol az (m, t) pár nem racionális. A másodikkal $R(m)$-ből törlünk olyan r-eket, ahol az (r, m) pár nem racionális.

Miután végeztünk, minden m egyensúlyon kívüli üzenethez lesz egy $T^s(m)$ halmazunk, amelyben azok a t-k vannak, akik racionálisan küldhették az m üzenetet (vagy ha úgy szeretnénk definiálni, van egy $S(m)$ halmazunk, amelyben azok a t-k vannak, akik racionálisan nem küldhették az m üzenetet, és a kettő között az összefüggés: $T^s(m) = T(m) - S(m)$). Most nézzük meg, hogy továbbra is egyensúly-e az eredeti stratégiapár. Ha igen, akkor az egyensúly átment a teszten, ha pedig nem, mert van olyan típusú A, akinek most már egyedül is megéri eltérni az eredeti stratégiájától, akkor megbukott.

Az eredeti intuitív kritériumban az szerepel, hogy amennyiben $T^s(m)$ üres, az szükséges a teszt sikeres teljesítéséhez, hogy minden $T(m)$ halmazon értelmezett μ valószínűségi eloszlásra létezzen $\phi \in MBR(\mu, m)$, ami nem ellenkezik az eredeti egyensúlyi kimenetellel. Ez működik Cho és Kreps intuitív kritériumával, mert ott egy kis egyszerűsítéssel élve eleve nem járhat senki jobban m küldésével a választól függetlenül, így nem is fognak eltértérni, mindegy, mi B korrigált vélekedése. A komparatív kritériumban ugyanakkor ez a megfogalmazás okozhatna problémákat. Mondjuk ezért inkább azt, hogy ha $T^s(m) = \emptyset$, akkor a vizsgált egyensúly nem bukhat meg az m üzeneten. Az intuitív kritérium ezzel nem változott meg, viszont így közös lehet a két kritérium második lépése (megjegyezném, hogy az összefoglaló, képlettel megadott formában nem kellett módosítani ezen a lépésen).

AZ INTUITÍV KRITÉRIUM

Ebben a fejezetben In-Koo Cho és David Kreps intuitív kritériumát mutatom be részletesen (Cho & Kreps, 1987). Rögzítsünk egy adott egyensúlyt, és ismét legyen $u^*(t)$ a t típusú A játékos által elért kifizetés ebben az egyensúlyban. Mondjuk azt, hogy a t típus nem küldheti racionálisan az m egyensúlyon kívüli üzenetet (azaz az (t, m) pár kizárható), ha

$$u^*(t) > \max_r u(t, m, r).$$

A teszt legyen a következő: zárjunk ki minden olyan (m, r), párt, ami az előző fejezet értelmében nem racionális. Ezután zárjunk ki minden olyan (t, m) párt, ami a fenti kritérium alapján nem racionális. Végül pedig alkalmazzuk az előző fejezetben ismertetett 2. lépést.

Összefoglalva az intuitív kritérium a következő:

- Minden m egyensúlyon kívüli üzenethez legyen $S(m)$ azon t típusok halmaza, akik nem küldhették azt racionálisan, azaz
$$u^*(t) > \max_{r \in BR(T(m), m)} u(t, m, r)$$
- Ha bármely m üzenethez létezik olyan $t' \in T$, hogy
$$u^*(t') < \min_{r \in BR(T(m) \setminus S(m), m)} u(t', m', r),$$
akkor az adott egyensúly megbukott az intuitív kritériumon.

A KOMPARATÍV KRITÉRIUM

Ebben a fejezetben egy olyan kritériumot szeretnék felvázolni, amely néhány helyzetben erősebb mint az intuitív kritérium, viszont nehezebben alkalmazható, és van vele néhány probléma, amiket a későbbiekben fogok részletezni. Az alapgondolat az, hogy a vizsgált egyensúlyt nem önmagában nézzük, hanem a többi kritériumhoz hasonlítjuk – ezért használom a komparatív kritérium elnevezést. Ennél a kritériumnál első lépésként nem csak azokat a t típusokat zárom ki, amelyek az m üzenet küldése esetén B semmilyen racionális válasza mellett nem tudnak nagyobb hasznosságot elérni, hanem azokat, amelyek egyik olyan egyensúlyban sem érnek el nagyobb hasznosságot m küldésével, ahol pozitív valószínűséggel küldik az m üzenetet. Mivel az első állítás implikálja a másodikat, úgy tűnhet, hogy a komparatív kritérium erősebb. Ez azonban megtévesztő, ennek megállapításához fontos még egy eset megvizsgálása.

Mivel ez a kritérium egyensúlyok összehasonlítására épül, egyből felmerül a kérdés, hogy mi történjen azokkal az m üzenetekkel, amelyeket az A játékos egyik egyensúlyban sem küld. Itt két lehetőséget tudok elképzelni. Az egyik, talán kézenfekvőbb, az, hogy azt mondjuk, hogy ezekben az esetekben $T^s(m) = \emptyset$. A másik opció az, hogy az ilyen üzenetekre az intuitív kritérium szerint vizsgáljuk az egyensúlyt. Mindkét megközelítésnek van előnye és hátránya is, amint az a későbbiekben ki fog derülni.

A komparatív kritérium tehát a következő lépésekből áll:

- Rögzítsünk egy egyensúlyt, legyen itt a t típusú A kifizetése $u^*(t)$
- Számoljuk ki minden típusú A játékos kifizetését minden más egyensúlyban is (a korábbi megjegyzés miatt véges sok esetet elég vizsgálni)
- Minden m egyensúlyi üzenetre vizsgáljuk meg, hogy melyik típusú játékosok járhatnak vele jobban egy másik egyensúlyban. Legyen ezek halmaza $T^s(m)$
    - 1. opció: amennyiben m semelyik egyensúlyban nem szerepel, legyen
    $$T^s(m) = \emptyset$$
    - 2. opció: amennyiben m semelyik egyensúlyban nem szerepel, legyen
    $$T^s(m) = \left\{t \in T \mid u^*(t) \leq \max_{r \in BR(T(m),m)} u(t,m,r)\right\}$$
- Alkalmazzuk az intuitív típusú kritériumok második lépését

Összesítve:

- Minden m egyensúlyon kívüli üzenethez legyen $S(m)$ azon t típusok halmaza, akik küldhették azt racionálisan, azaz létezik olyan, a vizsgálttól különböző egyensúly, ahol nem járnak rosszabbul, azaz
$$u^*(t) \leq u'(t,m,\mu)$$
ahol $u'(t,\rho(\cdot,t),\mu)$ egy olyan egyensúly, hogy $\rho(m,t) > 0$

- 1. opció: ha m semmilyen egyensúlyban nem szerepel pozitív valószínűséggel, akkor $T^s(m) = \emptyset$.
- 2. opció: ha m semmilyen egyensúlyban nem szerepel, akkor

$$T^s(m) = \left\{ t \in T \mid u^*(t) \leq \max_{r \in BR(T(m),m)} u(t,m,r) \right\}$$

- Ha bármely m üzenethez létezik olyan $t' \in T$, hogy

$$u^*(t') < \min_{r \in BR(T^s(m),m)} u(t',m',r)\,,$$

akkor az adott egyensúly megbukott az intuitív kritériumon.

A KÉT KRITÉRIUM ÖSSZEHASONLÍTÁSA

Bár a két kritérium hasonlónak tűnik, valójában az alapkoncepció között óriási különbség van. Míg az intuitív kritérium esetén az egyensúlyokat önmagukban vizsgáljuk, a komparatív kritériumnál a többi egyensúlyhoz hasonlítjuk őket. Talán a különbséget azzal lehet a legjobban megfogni, hogy mi az az üzenet, amit az A játékos át szeretne adni B −nek az m egyensúlyon kívüli üzenet küldésével. Mindkét kritérium alapgondolata az, hogy A az m egyensúlyon kívüli üzenet küldésekor úgy vélekedik, hogy végig fogja gondolni, hogy milyen típusú A-nak érné meg azt az üzenete küldeni, és a vélekedéseit leszűkíti ezekre az típusokra. Ha egy kicsit dinamikus értelmezést adunk a játéknak, a különbség ott rejlik, hogy az intuitív kritérium esetén B úgy gondolja, hogy A az egyensúlyon kívüli üzenettel egy nagyobb, nem feltétlen egyensúlyi kifizetés felé szeretne elmozdulni, a komparatív kritérium esetén azonban B úgy értelmezi az m üzenetet, hogy A egy számára jobb egyensúlyba szeretne átlendülni. mindkét érvelés gondolat megalapozottsága mellett lehet érveket találni. Statikus értelmezésben véleményem szerint a komparatív kritérium gondolatmenete tűnik ésszerűbbnek, hiszen ha a játékosok köztudott tudásnak tekintik azt, hogy ismerik az előre indukciót, logikus feltételezésnek tűnik tőlük az is, hogy úgy gondolják, hogy csak egyensúlyi kimenetel valósulhat meg. Dinamikus (úgy értve, hogy a játék ismételt) értelmezésben azonban az intuitív kritériumot tartom megalapozottabbnak, hiszen amíg a rendszer egy egyensúlyból átkerül egy másikra, jelentős időt tölthet egyensúlyon kívül is, folyamatosan változva.

ERŐSSÉG

Hasonlítsuk össze most a két kritériumot abból a szempontból, hogy egy adott játékban melyikkel lehet több egyensúlyt kizárni. Az alapgondolat az, hogy a komparatív kritérium jobban leszűkíti B játékos lehetséges korrigált vélekedéseinek halmazát, és így az ezekre adható legjobb válaszok halmazát is, ezért nagyobb eséllyel találunk olyan típust, aki minden ilyen válasszal jobban jár. Nézzük meg ezt formálisan felírva is. Először hasonlítsuk össze az intuitív kritériumot a komparatív kritérium második változatával (2. opció).

Rögzítsünk egy egyensúlyt, jelölje ebben t játékos kifizetését $u^*(t)$, m pedig legyen egyensúlyon kívüli üzenet. A $T^s(m), T(m)$ és $S(m)$ halmazoknál az alsó index jelölje azt, hogy a komparatív kritérium szerinti (c index), vagy az intuitív kritérium szerinti (i index) halmazokról van szó. Elsőként mutassuk meg, hogy $T_i^s(m) \supset T_c^s(m)$. Ehhez csak azt kell észrevenni, hogy annak, hogy ha valamely típusú A számára létezik olyan egyensúly, amelyben az m üzenetet küldi, és jobban jár, szükséges feltétele az, hogy B-nek van olyan r válasza, ami legjobb válasz valamire, és A a vizsgált egyensúlynál jobban jár, ha ő m-et játszik, B pedig r-t válaszol rá.

Innen két esetre kell bontanunk a vizsgálódást. Először nézzük meg azt, hogy mi történik az olyan üzenetekre, ahol $T_c^s(m)$ (és így persze $T_i^s(m)$ sem) nem üres halmaz. Azt kihasználva, hogy $X \subset Y \Rightarrow \min X \geq \min Y$, egyből láthatjuk, hogy

$$\min_{r \in BR(T_c^s(m),m)} u(t', m', r) \geq \min_{r \in BR(T_i^s(m),m)} u(t', m', r)$$

Ezért abból, hogy egy egyensúly az m üzenet és a t játékos miatt megbukik az intuitív kritériumon (tehát a jobboldali kifejezés nagyobb, mint $u^*(t)$), implikálja azt, hogy megbukik az komparatív kritériumon is (a baloldali kifejezés is nagyobb lesz $u^*(t)$-nél).

A második esetben azokat az üzeneteket kell vizsgálni, amelyekkel semelyik más egyensúlyban nem jár jobban az A játékos, típustól függetlenül. Ettől azonban még lehetséges, hogy van olyan típus, amely nagyobb kifizetést tud elérni m küldésével B valamely válasza esetén, még akkor is, ha ez legjobb válasz valamire. Ebben az esetben $T_c^s = \emptyset$, azonban $T_i^s \neq \emptyset$. Elképzelhető ezért az is, hogy (t, m) pár miatt ki lehet zárni a vizsgált egyensúlyt, ahol $t \in T_i^s$. A komparatív kritérium alapján ugyanakkor nem zárhatjuk ki az egyensúlyt egy m üzenet esetén, ha $T_c^s = \emptyset$.

Itt szeretném megemlíteni a különbséget a komparatív kritérium két különböző fajtája között. Egyértelmű, hogy a második opció erősebb, mint az első, mivel ott elméletben olyan m üzenetek esetén is ki lehet zárni egy egyensúlyt, amit semmilyen egyensúlyban nem küldenek, azáltal, hogy az intuitív kritériumra változik ilyen esetekben. Azonban még itt is elképzelhető olyan eset, hogy $T_c^s = \emptyset$, viszont $T_i^s \neq \emptyset$, olyankor, amikor az m üzenetet küldi valamely típus valamely egyensúlyban, azonban ezzel rosszabbul jár, mint a vizsgált egyensúly esetén.

Összefoglalva tehát, egyensúlytól függően lehet erősebb a komparatív kritérium és az intuitív kritérium is. Azonban fontos azt megjegyezni, hogy amint a példák alapján is látszani fog, a legtöbb helyzetben a komparatív kritérium minden olyan egyensúlyt ki tud szűrni, amelyet az intuitív, és sokszor többet is. Ahhoz, hogy legyen egy olyan egyensúly, amit csak az intuitív kritérium tud eliminálni, nagyon speciális feltételeknek kell teljesülnie. A komparatív kritérium első opcióját használva ennek az a szükséges és elégséges feltétele, hogy az intuitív kritériumot használva kizárólag olyan üzenetek alapján lehessen kiszűrni az adott egyensúlyt, amelyekkel A nem járhat jobban semelyik másik egyensúlyban, vagy ha a második opciót használjuk, az lesz a feltétel, hogy kizárólag a semmilyen egyensúlyban nem szereplő üzenetek alapján lehessen eliminálni a vizsgált egyensúlyt az intuitív kritériummal.

AZ EREDMÉNYEK EGYÉRTELMŰSÉGE

A komparatív kritérium alapötlete, és legnagyobb előnye az intuitív kritériummal szemben, egyben a legjelentősebb problémája is. Az, hogy csak az egyéb egyensúlyoktól függ, egyfajta lényegtelen alternatíváktól való függetlenséget is jelent, viszont azzal jár, hogy ha változnak az egyensúlyok, változhat a kritérium által hozható döntés egy-egy egyensúlyról. Ez pedig jelentős probléma, mert a kritérium alkalmazása során egyensúlyok záródnak ki.

Tovább súlyosbítja a helyzetet, ha egy egyensúlyt eliminálunk, és a kritérium újbóli alkalmazásánál nem vesszük figyelembe, azt sem tudjuk biztosan megmondani, hogy nő vagy csökken az elfogadható egyensúlyok száma. Emlékezzünk vissza, az előző fejezetre, ahol beláttuk, hogy a $T^s(m)$ halmaz egyre szűkebbé válásával nő az esélye, hogy az m üzenet alapján eliminálhatunk egy kritériumot, azonban ha üres lesz, hirtelen elveszíti az erejét. Tehát ha kizárunk egy egyensúlyt, az is lehet, hogy bizonyos egyensúlyok, amik addig átmentek az intuitív kritériumon, most már nem fognak, de az is elképzelhető, hogy más, kizárandó, vagy akár már kizárt egyensúlyok újra meg fognak felelni. Ez azt is maga után vonja, hogy elképzelhető, hogy a végül megmaradó egyensúlyok halmaza függ a vizsgálat sorrendjétől.

Ez a probléma megoldható azzal, ha azt mondjuk, hogy a technikai egyensúlyokat kell figyelembe venni az intuitív kritérium alkalmazásakor. Ez megszünteti az előbb felvázolt problémás dinamikát, ugyanakkor korlátozza is a komparatív kritérium hasznosságát. A későbbiekben a Spence-modellnél be fogok mutatni egy példát, ahol a komparatív kritérium iterált alkalmazása során szűkítjük le egyeleműre az egyensúlyok halmazát, ez lehetetlenné válna, ha rögzítve lennének az egyensúlyok.

Ha olyan játékot vizsgálunk, ahol a komparatív kritérium nem vezet iterálás nélkül eredményre, akkor érdemes fontolóra venni, milyen egyéb kritériumot vagy megoldáskoncepciót lehet alkalmazni, ami a fent leírt bizonytalanság nélkül is eredményre vezet. Ha pedig mindenképpen a komparatív kritériumot szeretnénk használni, a többszöri alkalmazás során fontos, hogy körültekintően járjunk el, és csak akkor fogadjuk el a végeredményt, ha sikerül róla megbizonyosodni, hogy egyértelmű, vagy minimum stabil (abban az értelemben, hogy megadott sorrendű végtelen iteráció esetén létezik a limese) a kapott egyensúlyok halmaza.

SZÁMÍTÁSI NEHÉZSÉG

További különbség a kritériumok között, hogy ahhoz, hogy az intuitív kritériumot alkalmazzuk egy adott egyensúlyra, nem szükséges ismernünk a többi egyensúlyi helyzetet. Ahhoz azonban, hogy a komparatív kritérium alapján akár elfogadjunk, akár elvessünk egy egyensúlyt, ismerni kell a játék összes egyensúlyát. Ez vitathatatlan előnye az intuitív kritériumnak, főleg számítógépek számára is nagynak számító játékok esetén, mivel jelenleg nincs olyan ismert algoritmus, amely polinomiális időben tudna Nash-egyensúlyokat keresni (ez az úgynevezett PPAD-teljes problémák közé tartozik (Daskalakis & Goldberg, 2009)), az intuitív kritérium pedig egy egyensúlyra maximum $|T| * |M| * |R|$ összehasonlítást igényel, ezért polinomiális időben számolható.

Ugyanakkor megjegyezném, hogy ha egy játék összes egyensúlyáról el szeretnénk dönteni, hogy elfogadjuk-e, akkor úgyis mindenképpen meg kell határozni mindet, azaz az intuitív kritérium elveszti ezt az előnyét. Sőt, miután meghatároztunk minden egyensúlyt, a komparatív kritérium alkalmazása lesz hatékonyabb, mivel kevesebb összehasonlítást kell végezni az első, és gyakran a második lépésben is. Olyan játékokban tehát, ahol viszonylag kevés az egyensúly és viszonylag sok a lehetséges lépés, a komparatív kritérium a hatékonyabb, a sok egyensúllyal rendelkező játékokban pedig az intuitív kritériumot egyszerűbb alkalmazni.

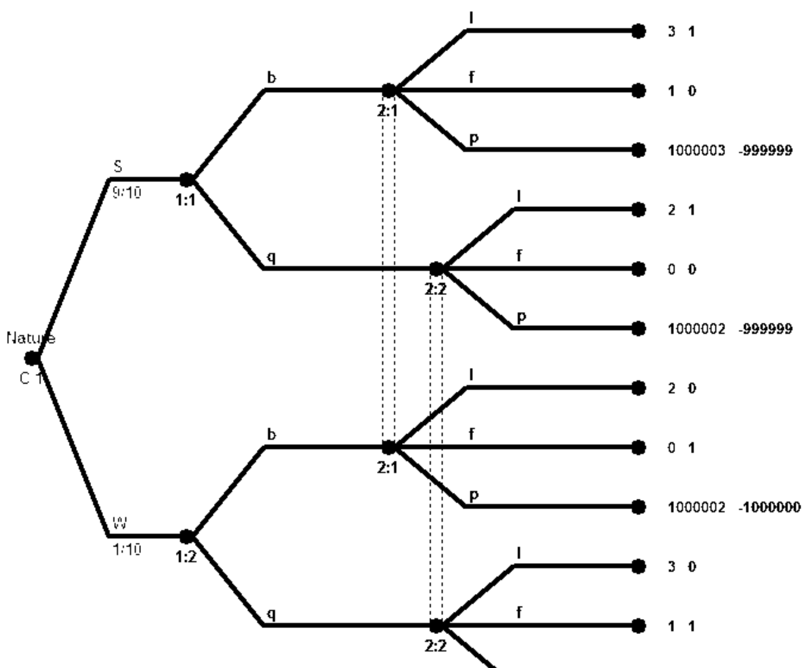
PÉLDÁK AZ INTUITÍV TÍPUSÚ KRITÉRIUMOK ALKALMAZÁSÁRA

Ebben a fejezetben példákon keresztül mutatom meg az előzőekben bemutatott kritériumok alkalmazását. Először visszatérek Cho és Kreps sör-kakaó játékának egy kissé módosított változatára, azután ennek egy bővített verzióját használom egy olyan eset bemutatására, ahol a komparatív kritérium segítségével ki lehet zárni egy nem intuitív egyensúlyt, az intuitív kritérium viszont nem alkalmas erre. Utolsó példaként Spence munkaerőpiaci modelljének egy változatát fogom használni.

ELSŐ VARIÁCIÓ A SÖR-KAKAÓ JÁTÉKRA

Emlékezzünk vissza a sör-kakaó játékra. Két egymástól szignifikánsan különböző egyensúlyunk volt, mind a kettő elvegyítő. Az egyik az volt, amikor az A játékos típustól függetlenül a sört választotta, B játékos pedig sör megfigyelése esetén békén hagyta A-t. A második egyensúly ennek az ellentéte volt, A játékos mindenképpen a kakaót választotta, B játékos pedig kakaó esetén nem kötött bele A-ba. Továbbra is a korábban használt jelöléseket használom annyi kiegészítéssel, hogy legyen most $u^b(t)$ a t típusú A játékos sörivós egyensúlyban elért kifizetése, $u^q(t)$ pedig a t típusú játékos kakaós egyensúlyban elért kifizetése.

Bővítsük ki annyival a játékot, hogy B-nek A minden lépésére van egy olyan válasza is, hogy :egymillió dollárt ad A-nak – jelölje ezt p (Cho & Kreps, 1987). Ekkor a játék gráfja a következőképpen fog kinézni:



2. ábra - Első variáció a sör-kakaó játékra

Természetesen az egyensúlyok nem változnak, mivel B soha nem fogja p-t választani, azonban ez a kis módosítás lehetőséget ad arra, hogy szemléltessük, miért van szükség a nem legjobb válaszok kizárására $R(m)$-ből az intuitív kritérium esetén.

INTUITÍV KRITÉRIUM

Először szemléltetésképpen gondoljuk meg mi történne, ha nem tennénk meg a legelső lépést, azaz nem zárnánk ki B olyan válaszait, amik nem legjobb válaszok semmire. Emlékezzünk, kakaós egyensúlyban vagyunk, és azt vizsgáljuk, kinek érheti meg a sör (b) egyensúlyon kívüli üzenetet küldeni. Azokat a t típusokat zárhatjuk ki az intuitív kritérium értelmében, akikre teljesül, hogy $u^q(t) > \max_r u(t,b,r)$, ahol most $r \in T(b)$. Puhány típusú A esetén az egyenlőtlenség bal oldalán szereplő érték a 3, jobb oldalon pedig a 1000002 áll, így $W \notin S(b)$. Hasonlóan egyszerűen látható az is, hogy a kemény típust sem tudjuk kizárni. Az előzőekből következik, hogy $S(b) = \emptyset$, azaz $T^s(b) = T(b)$. B korrigált vélekedését A típusáról nem sikerült szűkíteni, ezért például az eredeti vélekedéseit használva az egyensúly átmegy az intuitív kritériumon.

Alkalmazzuk most helyesen az intuitív kritériumot.

- Első lépés

  Mivel első lépésben megtiltjuk, hogy B nem legjobb választ válasszon, most már a következő egyenlőtlenségnek kell teljesülnie ahhoz, hogy azt mondhassuk, hogy a t típusú A nem fogja racionálisan a b üzenetet küldeni: $u^q(t) > \max_r u(t,b,r)$. A puhány típus esetén a $3 > 2$ egyenlőtlenség valóban teljesül, kemény típusú esetén viszont $2 > 3$ szerepel a két oldalon, tehát az egyenlőtlenség nem áll fenn.

  Formálisabban leírva:

$$S(b) = \left\{ t \in T \mid u^q(t) > \max_r u(t,b,r) \right\} = \{W\}, \qquad T^s(b) = T(b) \setminus S(b) = \{S\}$$

- Második lépés

  Most nézzük meg, hogy Létezhet-e olyan μ vélekedése B-nek A típusáról $T^s(b)$-n, hogy az eredeti egyensúly továbbra is egyensúly marad. Mivel $T^s(b)$ egyelemű, az egyetlen lehetséges vélekedés az, hogy B 1 valószínűséggel úgy gondolja, hogy A kemény típusú. Ekkor azonban ha sört figyel meg, nem fog kötözködni, A pedig azzal, hogy sört iszik kakaó helyett, 2-ről 3-ra növelheti a hasznosságát az egyoldalú eltéréssel, az eredeti egyensúlyi stratégiák már nem alkotnak egyensúlyt.

Képlettel leírva ez a következőképpen néz ki:

$$u^*(S) < \min_{r \in BR(\{S\})} u(S, b, r),$$

azaz az egyensúly megbukik az intuitív kritériumon.

KOMPARATÍV KRITÉRIUM

Első lépésként keressük meg az összes egyensúlyt, és írjuk fel a különböző típusú A-k kifizetéseit ezekkel az egyensúlyokban. Emlékezzünk, hogy korábban beláttuk, hogy két tartalmilag különböző egyensúly van csak, a kakaós elvegyítő és a sörivős elvegyítő egyensúly. Ezekben az egyensúlyokban a következő kifizetések járnak A-nak:

$$u^q(W) = 3, u^q(S) = 2, u^b(W) = 2, u^b(S) = 3$$

A kritérium értelmében azok t típusok kerülnek a $T^s(b)$ halmazba, amelyekhez létezik olyan, a vizsgálttól eltérő egyensúly, ahol nem járnak rosszabbul, és pozitív valószínűséggel választják b-t abban az egyensúlyban. Mivel csak két egyensúlyunk van, igen könnyű ezt a vizsgálatot elvégezni: egyértelmű, hogy az u^b-vel jelölt egyensúlyban a puhány típus rosszabbul jár, a kemény típus pedig jobban, tehát

$$T^s(b) = \{S\}$$

Innen teljesen ugyanúgy folyik a dolog, mint az intuitív kritérium esetében. Megnézzük, hogy ha B vélekedéseit az $\{S\}$ halmazra korlátozzuk, továbbra is lehet-e egyensúly az eredeti stratégiaegyüttes. Természetesen nem lehet, mivel $u^*(S) < \min_{r \in BR(\{S\})} u(S, b, r)$, így a kakaós elvegyítő egyensúly megbukik a komparatív kritériumon is.

Alkalmazzuk most a kritériumot a másik egyensúlyra is. Mivel a kifizetések A számára fordítva vannak, ebben az esetben az fog kijönni, hogy a q egyensúlyon kívüli üzenetre $T^s(q) = \{W\}$. Nézzük most meg, létezhet-e olyan vélekedése B-nek $T^s(q)$-n, amely mellett az eredeti egyensúlyi stratégiák továbbra is egyensúlyt alkotnak. Az egyetlen lehetséges vélekedés persze az, amikor B q megfigyelése esetén 1 valószínűséggel gondolja úgy, hogy A puhány, így bele fog kötni. Ebben az esetben semelyik típusú A-nak nem éri meg egyedül eltérnie az egyensúlytól, ugyanis csökkenne a kifizetése. Képlettel megfogalmazva

$$\nexists\, t \in \{S, W\} : u^*(t) < \min_{r \in BR(\{W\})} u(t, b, r).$$

EGY MÁSIK VARIÁCIÓ A SÖR-KAKAÓ JÁTÉKRA

Ebben a fejezetben egy olyan játékot fogok bemutatni, amelyben szintén két lényegileg eltérő egyensúly van, és szintén lehet amellett érvelni, hogy az egyik nem racionális. A játék érdekességét az adja, hogy az intuitív kritérium nem tudja kiszűrni a rossznak tűnő egyensúlyt, a komparatív kritérium pedig igen.

A játékhoz tartozó történet hasonlóan kezdődik, mint a sör-kakaó játéknál. Az A játékos éppen a reggelijét fogyasztja, B pedig a parkolóban várja, és eldöntheti, hogy belekössön-e. B nem tudja megfigyelni, A milyen típusú, csak azt, hogy sört vagy kakaót választ, és ez alapján kell eldöntenie, mit cselekszik. A különbségek ott kezdődnek, hogy ebben a játékban a jól informált fél háromféle típusú lehet: kemény (S), puhány (W) és vitatkozós (D), a természet ezeket rendre 0,7, 0,1 és 0,2 valószínűséggel választja. Miután A megfigyelte a típusát két lehetőség áll előtte: ihat sört (b) vagy kakaót (q). Ezután B lehetséges lépései attól függenek, hogy mit választott A. Amennyiben kakaót iszik, B most is aközött választhat, hogy békén hagyja (l), vagy rátámad (b). Ha azonban A a sört választja, B-nek van még egy lehetősége: meg is szólhatja A-t azért, amiért már reggel sört iszik. Ezt a választ jelölje v. Tömören összefoglalva az eddigieket, a játék fő jellemzői a következők:

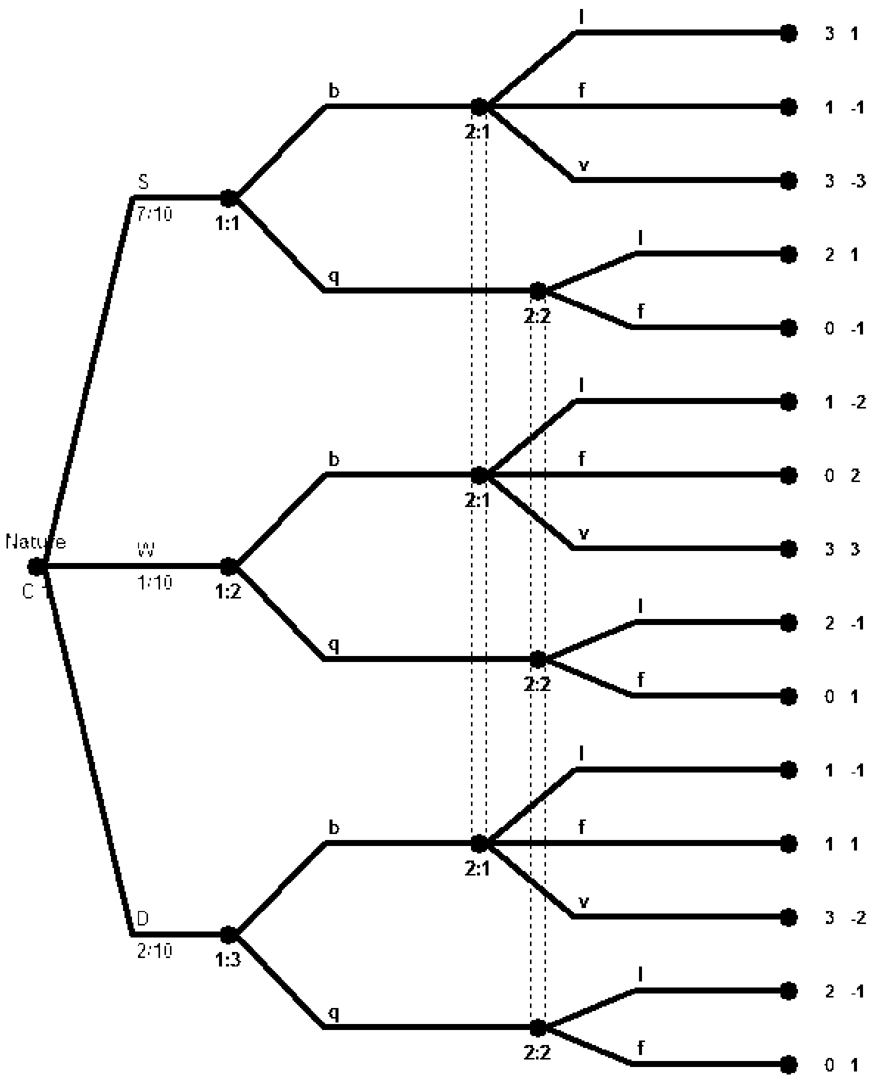
- $t \in \{S, W, D\}$, t privát információja A-nak
  az egyes típusok a priori valószínűségei pedig: $\pi(S) = 0{,}7, \pi(W) = 0{,}1, \pi(D) = 0.2$
- $M(S) = M(W) = M(D) = \{b, q\}$
- $R(b) = \{l, f, v\}, R(q) = \{l, f\}$

A kifizetések a következőképpen alakulnak: amennyiben A kakaót iszik, úgy típusától függetlenül csak annyit szeretne, hogy ne támadja meg B. Ha ez történik, hasznossága 2, ha nem, akkor 0. B pedig akkor szeretne belekötni A-ba, ha annak típusa puhány vagy vitatkozós, és akkor nem, ha kemény. Ha ezt eltalálja a kifizetése 1, ha nem, -1.

A sörös oldalon kissé bonyolultabb a helyzet. Vitatkozni mindenki szeret, aki sört iszik, ezért ebben az esetben minden típusú A hasznossága 3. Harcolni senki nem szeret, főleg a puhány típus nem: az ő kifizetése ebben az esetben 0, a másik két típusé 1. Azt, hogy békén hagyják, igazán csak a kemény típusú A szeretné: az ő kifizetése ilyenkor 3, a többieké 1. B kifizetése a következőképpen alakul: egy kemény A-val nem szeretne harcolni, ha mégis ezt teszi, -1 a kifizetése. A vitatkozós típussal már szívesen harcolna, ekkor 1 hasznosságot ér el, az pedig minden parkolóban kötözködős ember álma, hogy kifogjon magának egy gyenge, ittas embert

– ha a gyenge típust támadja meg 2 hasznosságot ér el. Amennyiben a békén hagyást választja, minden esetben ennek ellentettje lesz a kifizetése. Ha *B* úgy dönt, hogy beszól *A*-nak, akkor kemény típus esetén -3, gyenge típus esetén 3, kötözködős típus esetén pedig -2 hasznosságot ér el.

A játék extenzív alakja a következő:



3. ábra - Második variáció a sör-kakaó játékra

Egy fontos tényt egyből lehet látni a gráf alapján: B-nek egyik lehetséges válasza sem olyan, hogy soha nem legjobb válasz. A kakaós oldalon a harc D és W típusú A ellen lesz a legjobb válasz, a békén hagyás pedig S típusú A ellen. Ha viszont A sört választ, akkor a harc D típus esetén, a békén hagyás S típus esetén, a megszólás pedig W típus esetén lesz a legjobb válasz.

Ebben a játékban két, egymástól jelentősen eltérő egyensúly van: az egyik az, amikor mindhárom típusú A sört iszik, B pedig a békén hagyást választja, a másik pedig az, hogy mindhárom típus kakaót iszik, és ilyenkor B szintén békén hagyja A-t. Az egyensúlyon kívüli üzenetekhez egyszerűen lehet olyan vélekedéseket találni B-nek, hogy annak megfigyelése esetén olyan választ adjon, hogy semelyik A-nak ne érje meg eltérni. Azt sem nehéz belátni, hogy nem létezik részben szeparáló egyensúly, és olyan sem, amelyben A több üzenetet is pozitív valószínűséggel küld. Ezeknek a vizsgálata nem kötődik szorosan a tárgyalt témához, nem adna hozzá semmit a dolgozathoz ennek megmutatása, úgyhogy ettől most eltekintek. Ami számunkra fontos, az a különböző típusú A játékosok által elért kifizetések a két egyensúlyban. Az előző modellben használt jelölésekkel élve ezek a következőképpen fognak alakulni: a kakaós egyensúlyban $u^q(S) = u^q(W) = u^q(D) = 2$, a másik egyensúlyban pedig $u^b(S) = 3, u^b(W) = u^b(D) = 1$. Ezzel kapcsolatban máris megtehetjük a következő, fontos megfigyelést: a kakaós egyensúlyban a puhány és a vitatkozós A jár jobban, a sörivósban pedig a kemény típusú.

Nézzük most meg jobban a kakaós egyensúlyt. B milyen vélekedései mellett lesz ez valóban egyensúly? Ahhoz, hogy A ne térjen el egyoldalúan, szükséges, hogy B a harcot legalább 0,5 valószínűséggel válassza. Vegyük észre, hogy ez elégséges feltétel a többi játékoshoz is, ugyanis ilyenkor senki sem ér el 2-nél nagyobb várható hasznosságot. Azt nem fontos vizsgálni, hogy pontosan milyen vélekedések mellett fog B így cselekedni, elég azt megjegyeznünk, hogy B vélekedésében a kemény típusú A-n kívül valamely másik típusnak is pozitív valószínűséggel kell szerepelnie ehhez.

A kérdés most az, hogy milyen típusú A-ról képzelhető el racionálisan, hogy az egyensúlyon kívüli sörivás üzenetet választják. Mondhatjuk azt, hogy minden típusról elképzelhető, hogy hogy az egyensúlyon kívüli üzenetet küldi, ugyanis b küldésével mindegyikük érhet el nagyobb kifizetést. Vagy ennél tovább is mehetünk, és azt is mondhatjuk, hogy a játékosok úgy gondolkoznak, hogy valamely egyensúlyi kimenetel valósul meg mindenképp, ezért azt kell nézni, milyen típusok járhatnak jobban valamely másik (esetünkben az egyetlen másik) egyensúlyban. Ez esetben az egyensúlyon kívüli üzenet valami ilyesmi jelentéssel bír: „Mindketten tudjuk, hogy a játékban egyensúlyi kimenetel valósul meg. Azt is mindketten tudjuk, hogy csak én járok jobban a másik egyensúlyban, ezért a többieknek nem érné meg ezt az üzenetet küldeni. Kérlek frissítsd a vélekedéseidet, és cselekedj aszerint, hogy tudod, hogy ha a sört választom biztosan kemény vagyok.”.

Az, hogy ez mennyire hihető vagy intuitív gondolatmenet, vitatható, és valószínűleg megoszlanak róla a vélemények. Amennyiben elfogadjuk, hogy a játékosok ezt végiggondolják, akkor elfogadjuk a komparatív kritériumot is, ugyanis ez a döntő gondolat. Én személy szerint úgy gondolom, hogy ez attól függ, milyen játékra szeretnénk alkalmazni kritériumot. A legtöbb szignáljátékban túlzásnak tartom, hogy a játékosok az előre indukción túl még azt is végiggondolják, hogy csak egyensúlyi kimenetek között gondolkozzanak. Azonban nem tartom teljesen elképzelhetetlennek, hogy ha ehhez a játékhoz egy reggeli helyett egy két nagy vállalat stratégiai interakciójáról szóló történet tartozna, a játékosok eljutnának eddig a problémás egyensúly kizárásáig.

INTUITÍV KRITÉRIUM

Amint az előzőekben is utaltam rá, az intuitív kritérium segítségével ebben a játékban nem tudjuk B vélekedéseit szűkíteni a kakaós egyensúlyban. Ebben a részben ezt fogom megmutatni.

Rögzítsük azt az elvegyítő egyensúlyt, amelyben A típustól függetlenül kakaót iszik, B ebben az esetben békén hagyja, sör megfigyelése esetén pedig olyanok B vélekedései, hogy a harcot választja. Emlékezzünk, nincs olyan lépése B-nek, ami semmire sem legjobb válasz, ezért első lépésben semmilyen üzenetet nem tudunk kizárni. Nézzük most meg az

$$S(b) = \left\{ t \in \{S, W, D\} \mid u^q(t) > \max_{r \in BR(T(b), b)} u(t, m, r) \right\}$$

halmazt. Könnyű látni, hogy ez üres lesz, hiszen a kakaós egyensúlyban mindenki 2 hasznosságot ér el, sör választása esetén pedig az elérhető legnagyobb kifizetés mindenki számára 3 egység. A kritérium második lépésében vizsgáljuk meg, van-e olyan t' típusú A, aki az eltéréssel biztosan jobb kifizetést érhet el, azaz

$$u^q(t') < \min_{r \in BR(\{S, W, D\}, b)} u(t', b, r).$$

Természetesen nem lehet ilyen, ugyanis B vélekedései pont ugyanazok lehetnek, mint eredetileg, és ha lenne ilyen t', akkor már az sem lett volna egyensúly.

A fentiek értelmében azt kell mondanunk, hogy az az egyensúly, amikor A minden esetben kakaót iszik, átment az intuitív kritériumon, nem zárhatjuk ki.

Komparatív kritérium

Vizsgáljuk meg most ugyanezt az egyensúlyt a komparatív kritérium segítségével. Továbbra is legyen rögzítve az az egyensúly, ahol minden típusú A kakaót iszik. Csupán annyiban fog különbözni a gondolatmenet, hogy $T^s(b)$-be azok a típusok kerülnek, amelyekhez létezik olyan egyensúly, hogy pozitív valószínűséggel választják b-t és nagyobb a várható kifizetésük. Megint könnyű dolgunk van, hiszen csak egyetlen másik egyensúly létezik: az, amelyben A típustól függetlenül a sört választja, B pedig békén hagyja. Ebben az egyensúlyban kizárólag a kemény típus ér el nagyobb kifizetést, mindenki más alacsonyabbat. Ebből következően $T^s(b) = \{S\}$.

Nézzük most meg a második lépést. Van-e olyan t' típus, amelyre

$$u^q(t') < \min_{r \in BR(\{S\},b)} u(t', b, r)$$

teljesül. Ebben az esetben a $BR(\{S\}, b)$ halmaznak csak egy eleme van, méghozzá l, azaz az, hogy a B játékos békén hagyja A-t. A kérdés így a következő lesz: van-e olyan t', melyre

$$u^q(t') < u(t', b, l)$$

teljesül. A kemény típus ilyen, ezért neki megérné eltérni az egyensúlyi stratégiájától, és a sört választania. Ennek következtében a kakaós elvegyítő egyensúly megbukik az intuitív kritériumon.

Spence-féle munkaerőpiaci modell – kétszereplős változat

Ebben a fejezetben Spence munkaerőpiaci modelljének egy változatát mutatom be (Spence, 1973), és és a potenciális egyensúlyokat megvizsgálom az intuitív és a komparatív kritérium segítségével. Először is fel szeretném hívni a figyelmet arra, hogy az intuitív kritériumot és a komparatív kritériumot is olyan játékokra definiáltam, melyekben mindkét fél csak véges számú akció közül választhat, az itt bemutatott modellben viszont folytonos intervallumból választható az üzenet és a válasz is. Ennek ellenére alkalmazhatók a kritériumok, azonban fontos megjegyezni, hogy az ilyen játékokra sokkal hasznosabb például Mailath, Okuno-Fujiwara és Postlewaite veretlen egyensúlya (undefeated equilibria), ami az alapgondolatot illetően nagyon hasonlít a komparatív kritériumra, azonban pontosan az ehhez hasnonló, folytonos játékokra van kitalálva (Mailath, Okuno-Fujiwara, & Postlewaite, 1993).

Tegyük fel, hogy egy munkapiacon kétféle munkavállaló van: magas és alacsony produktivitású. Jelölje őket rendre $h \text{ és } l$. Van továbbá a piacon két, teljesen megegyező vállalat, akik Bertrand-versenyt folytatnak a munkaerőpiacon. Az egyszerűség végett tegyük még fel, hogy a vállalatok egyetlen inputja a munka, és konstans mérethozadékú technológiával dolgoznak. A munkavállaló típusát nem tudják megfigyelni a vállalatok, csupán a valószínűségét tudják annak, hogy egy munkavállaló h vagy l típusú $(\pi(h) = p, \pi(l) = 1 - p)$, mindkét típusnak van lehetősége iskolai végzettséget választani, amit $e \in \mathbb{R}^+$ jelöl. A munkáltató ezt megfigyelve valamilyen $w \in \mathbb{R}^+$ bért ajánl nekik.

A munkavállaló hasznosságfüggvénye a következő:

$$U(w, t, e) = w - v(e, t),$$

ahol $v(e, t)$ az iskolai végzettség költségfüggvénye, és feltesszük, hogy $v(0, t) = 0, v_e(e, t) > 0, v_{ee}(e, t) > 0, v(h) < v(l), v_e(h) < v_e(l)$ – az alsó indexek a parciális deriváltakat jelölik. Legyen továbbá a munkavállaló rezervációs hasznossága, így rezervációs bére is 0.

A munkáltatónak egy t típusú munkavállaló $f(t)$ bevételt termel (úgy, hogy $f(h) > f(l)$), így a kockázatsemleges vállalatok profitfüggvénye a következőképpen alakul:

$$B(w, t) = f(t) - w$$

Mivel a vállalatok Bertrand versenyt folytatnak, egyensúlyban a munkavállaló fizetése mindig az általa várhatóan termelt bevétel. Ha ennél kevesebb lenne, akkor mindkét cégnek megérné ennél ε-nal többet ajánlani, és hozzá menne a munkavállaló. Ha több lenne, a vállalatnak nem érné meg foglalkoztatni a munkavállalót, ugyanis kevesebb bevételt termelne a vállalat számára, mint amennyit fizetni kellene neki.

Nézzük meg azt a szeparáló egyensúlyt, amit Riley definiált 1979-es cikkében (Riley, 1979). Itt először az alacsony produktivitású munkavállaló választ egy olyan oktatási szintet, ami maximalizálja a hasznosságát, feltéve, hogy megkapja a produktivitásának megfelelő bért. Azaz a következő a feladata:

$$\max_e f(l) - v(e, l)$$

Az ezt maximalizáló oktatási szint természetesen a 0 lesz. Ezután a magas produktivitású munkavállaló az ő hasznosságát maximalizáló oktatási szintet választja azzal a feltétellel, hogy az alacsony produktivitású típus rosszabbul jár, ha megpróbálja magas produktivitásúnak kiadni magát, azaz:

$$\max_e f(h) - v(e, h)$$

$$fth.\ f(h) - v(e,l) \leq f(l) - 0$$

Jelölje az ezt maximalizáló oktatási szintet e^s[2]. Rövid megjegyzésként érdemes megemlíteni, hogy természetesen valójában szigorú egyenlőtlenségnek kellene állnia a feltételben, ekkor viszont nem lenne megoldása a maximumfeledatnak. Ha a nem szigorú egyenlőtlenséges feltételt használva is teljesen pontos szeretnék lenni, akkor e^s helyett $e^s + \varepsilon$-t kellene írnom, azonban az átláthatóság kedvéért a szakirodalomban megszokott módon egyszerűen e^s-t fogom használni. A későbbiekben tehát az, amikor néhány helyen nem szigorú egyenlőtlenség van, én mégis szigorúként kezelem, nem tévedés, csupán kissé pontatlan jelölés az egyszerűség kedvéért.

Nem nehéz belátni, hogy ez valóban egyensúly. Az l típusú játékosnak pontosan az előbbi maximumfeladat feltétele miatt nem éri meg egyoldalúan eltérni, a h típusúnál pedig az biztosítja, hogy létezik olyan e^s, amit választhat, és jobban jár, mintha alacsony bért kapna, hogy számára az oktatás határköltsége alacsonyabb.

Elvegyítő egyensúlyból végtelen sokat lehet találni. Választhatjuk a teljesen elvegyítő egyensúlyt, ahol mindkét munkavállaló $e^{p'}$ oktatási szintet választ, a munkáltató pedig ennek megfigyelése esetén a várható produktivitásuknak megfelelő bért fizeti ki, azaz $w_{p'} = pf(h) + (1-p)f(l)$, más szint megfigyelése esetén pedig úgy gondolja, hogy a munkavállaló alacsony produktivitású, ezért $f(l)$ bért fizet. Még akkor is, ha a munkavállaló az e^s oktatási szintet választja. Nem nehéz látni, hogy ez is tökéletes bayesi egyensúly, de itt is érezzük, hogy valami nem racionális. Az alacsony típusú munkavállalónak nem érné meg ezt az egyensúlyi üzenetet küldeni, mert a kifizetése biztosan alacsonyabb lenne, mint ha 0 oktatást választana, és $f(l)$ bért kapna. Nem is kell teljesen elvegyítőnek lennie az egyensúlynak, elég, ha van olyan oktatási szint, amit pozitív valószínűséggel választ mindkét típus, azaz $\rho(e^p,l), \rho(e^p,h) > 0$, tehát e^p megfigyelése esetén $\mu(h), \mu(l) > 0$.

Azt kell még megmutatnunk, hogy egy tetszőleges elvegyítő egyensúlyban e^s egyensúlyon kívüli üzenetet, semelyik típus nem küldi pozitív valószínűséggel. Jelölje az e^p üzenet megfigyelésekor ajánlott bért w. Először lássuk be azt, hogy az alacsony produktivitású munkavállaló biztosan nem fogja az e^s oktatási szintet választani. Ez nagyon egyszerű, mivel

$$f(h) - v(e^s,l) \leq f(l) \leq\ w(p) - v(e^p,l)$$

[2] Azt, hogy ezt az e^s szintet választva $f(h)$ bér esetén a magas produktivitású munkavállaló nagyobb hasznosságot ér el, az a $v(e,t)$ függvényre előírt tulajdonságok biztosítják. Ennek bővebb tárgyalásához ld. (Macho-Stadler & Pérez-Castrillo, 2001)

Az első egyenlőtlenség e^s szeparáló tulajdonságának a feltétele, a második egyenlőtlenség pedig szükséges feltétele annak, hogy a vizsgált stratégia –és vélekedésegyüttes egyensúly lehessen.

A következő lépésben indirekt tegyük fel, hogy $\rho(e^s, h) > 0$. Azonban ha $\rho(e^s, l) = 0$, akkor a feltételből következik, hogy e^s üzenet megfigyelése esetén $\mu(h) = 1$, erre pedig a munkáltató legjobb válasza az $f(h)$ bér fizetése. Ha azonban ez így van, akkor a magas produktivitású munkavállalónak megéri 1 valószínűséggel e^s-t választani, hiszen így nagyobb hasznosságot ér el. Ha viszont $\rho(e^s, h) = 1$, akkor $\rho(e^p, h) = 0$. Ezzel ellentmondásra jutottunk, tehát az eredeti feltétel hamis, azaz $\rho(e^s, h) = 0$.

Érdekes, hogy az előbb használt érvelés több pontos is nagyon hasonló az intuitív kritérium logikájához, azaz a játékosok feltételezett gondolkodásához, amikor az intuitív kritériumot alkalmazzuk.

INTUITÍV KRITÉRIUM

Tekintsünk egy olyan egyensúlyt, amelyben valamely e^p oktatási szintet mindkét típus pozitív valószínűséggel választja. Legyen ebben az egyensúlyban az e^p-re járó fizetés w_p. Mivel elvegyítő egyensúlyról van szó, e^p esetén is $\mu(h) < 1$, tehát $w_p < f(h)$. Nézzük meg, hogy mi történik az e^s egyensúlyon kívüli üzenet esetén. Az előbb láttuk, hogy ilyenkor $\mu(h) < 1$. Használjuk most az intuitív kritériumot, tehát

$$\begin{aligned} S(e^s) &= \left\{ t \in \{h, l\} \mid u^p(t) > \max_{w \in [f(l), f(h)]} w - v(e^s, t) \right\} \\ &= \{ t \in \{h, l\} \mid u^p(t) > f(h) - v(e^s, t) \} \end{aligned}$$

Emlékezzünk azonban, az előbb belátott két egyenlőtlenségre:

$$f(h) - v(e^s, l) \leq w_p - v(e^p, l) = u^p(l)$$

$$f(h) - v(e^s, h) > w_p - v(e^p, h) = u^p(h)$$

Ebből következik, hogy az $S(e^s) = \{l\}$

Alkalmazzuk most az intuitív kritérium második lépését. Nézzük meg, van-e olyan típus, amelyre

$$u^p(t') < \min_{w \in BR(\{h\}, e^s)} u(t', e^s, w)$$

Pontosabban a kérdés az, hogy h ilyen típus-e. Mivel Bertrand-verseny van a munkaerőpiacon, $BR(\{h\}, e^s) = f(h)$, így a fenti feltétel a következő alakra egyszerűsödik:

$$u^p(t') < f(h) - v(e^s, t)$$

Azt már beláttuk, hogy a magas produktivitású típusra ez teljesül, ezért az egyensúly megbukik az intuitív kritériumon. Emlékezzünk, hogy általánosságban vizsgáltuk a legalább részben elvegyítő egyensúlyokat, minden egyéb feltétel nélkül, ezért eredményünk egyben azt is jelenti, hogy semmilyen elvegyítő egyensúly nem felel meg az intuitív kritériumnak. Mivel megmutatható, hogy nincs más szeparáló egyensúly a Riley-egyensúlyon kívül, ez utóbbi az egyetlen, ami kiállja az intuitív kritérium próbáját.

Komparatív kritérium

Vizsgáljuk meg a modellt a komparatív kritériummal is. Rögzítsünk ismét egy tetszőleges legalább részben elvegyítő egyensúlyt. Nézzük meg, mi történik az e^s egyensúlyon kívüli üzenet esetén. $T^s(e^s)$-be azok a típusok kerülnek be, amelyekhez létezik olyan egyensúly, hogy pozitív valószínűséggel küldik e^s-t ebben az egyensúlyban, és nagyobb lesz a várható kifizetésük. Emlékezzünk, hogy az alacsony produktivitású típus semmilyen egyensúlyban nem küldi az e^s üzenetet, a magas produktivitású pedig a Riley-egyensúlyban igen, és itt magasabb is lesz a hasznossága. Ebből következően $T^s(m) = \{h\}$. Innen teljesen ugyanúgy lehet folytatni, mint az intuitív kritérium esetében, azaz be kell látni, hogy

$$u^p(h) < \min_{w \in BR(\{h\}, e^s)} u(h, e^s, w)$$

Ahelyett, hogy megismételnénk az előző gondolatmenetet, kijelenthetjük, hogy a komparatív kritériumon is megbukott a vizsgált egyensúly, és vele együtt az összes legalább részben szeparáló egyensúly is. A komparatív kritérium szerint is a Riley-egyensúly az egyetlen, amelyet nem zárunk ki.

Spence-féle munkaerőpiaci modell – többszereplős változat

A modell többszereplős változata igen hasonló az egyszereplőshöz. Az egyetlen lényeges különbség az, hogy két típus helyett N típus van, Jelölje ezeket $t \in \{t_1, t_2, \dots, t_N\}$. A különböző típusok a munkáltató számára $f(t_1) < f(t_2) < \cdots < f(t_N)$ bevételt termelnek. A munkavállaló költségfüggvényére most a következő feltételeink vannak:

$$v(0,t) = 0, v_e(e,t) > 0, v_{ee}(e,t) > 0 \; \forall \; e,t$$
$$v_e(e,t_1) > v_e(e,t_2) > \cdots > v_e(e,t_N)$$

Megmutatható, hogy az előbbiekben ismertetett Riley-egyensúly továbbra is szeparáló egyensúly, sőt, az egyetlen szeparáló egyensúly ezen feltételek mellett. Tehát először t_1 választ magának egy olyan e_1-et, ami maximalizálja a hasznosságát. Ezután t_2 választ magának a hasznosságát maximalizáló e_2-t, azzal a feltétellel, hogy

$$f(t_2) - v(e_2, t_1) \leq f(t_1) - v(e_1, t_1)$$

Ezután a harmadik választ hasonlóképpen oktatási szintet, majd így tovább egészen a legmagasabb produktivitású szereplőig.

Elvegyítő és egyensúlyból ebben a változatban is végtelen sok van. A következő fejezetben meg fogom mutatni, hogy az intuitív kritérium nem képes kizárni ezeket.

INTUITÍV KRITÉRIUM

A kritérium korlátainak bemutatására a Cho-Kreps [1987] szerzőpároshoz hasonlóan én is egy háromszereplős modellt fogok használni. Nézzük azt az egyensúlyt, amikor t_1 és t_2 elvegyítő egyensúlyban van, és e_p oktatási szintet választ. Jelölje a t típusú játékos ebben az egyensúlyban elért hasznosságát $u^p(t)$, az e_p üzenet megfigyelése után ajánlott bért pedig w_p. A problémát az okozza, hogy bár a t_1 típusnak továbbra sem éri meg, hogy a Riley-egyensúly e_2 oktatási szintjét válassza, ha ekkor $f(t_2)$ bért kap, most azonban már a munkáltató legjobb válaszai nem az $[f(t_1), f(t_2)]$ intervallumra vannak korlátozva, hanem a$[f(t_1), f(t_3)]$ intervallumra. Nézzük meg melyik típusok kerülnek most az $S(e_2)$ halmazba.

$$\begin{aligned} S(e_2) &= \left\{ t \in T \mid u^p(t) > \max_{w \in [f(t_1), f(t_3)]} w - v(e_2, t) \right\} \\ &= \left\{ t \in T \mid w_p - v\left(e_p, t\right) > \ f(t_3) - v(e_2, t) \right\} \end{aligned}$$

Azt, hogy a t_1 típusú munkavállaló benne legyen az $S(t_2)$ halmazban, most semmi sem biztosítja (a többiek nyilván nincsenek benne), ezért könnyen elképzelhető, hogy $S(t_2) = \emptyset$. Ha viszont ez a helyzet, akkor nem sikerült leszűkíteni a munkáltató vélekedéseinek halmazát, nem lehet elvetni az egyensúlyt.

Nézzük meg mi most, mi az az e_0 érték, ami már garantálja, hogy $t_1 \in S(e_2)$, ezért ennek küldése esetén a t_2 típusú játékos legalább $f(t_2)$ bért kap. Ez a legkisebb olyan szint, ami esetén

$$w_p - v\left(e_p, t_1\right) > \ f(t_3) - v(e_0, t_1).$$

Azonban ezzel az a probléma, hogy, elképzelhető, hogy ilyenkor a t_2 típusú játékos is rosszabbul jár, mint az eredeti egyensúlyban. Ebben az esetben pedig nem teljesül az intuitív kritérium második lépése, tehát nincs olyan t', amelyre

$$u^p(t') < \min_{w \in BR(\{t_2, t_3\}, e_0)} u(t', e_0, w) = f(t_2) - v(e_0, t)$$

Az intuitív kritérium segítségével tehát nem lehetett kizárni ezt a részben elvegyítő egyensúlyt.

KOMPARATÍV KRITÉRIUM

Térjünk most át egy N szereplős modellre. A komparatív egyensúly egy meghatározott sorrendben alkalmazva $N-1$ lépésen keresztül képes eliminálni a Riley egyensúlyon kívül minden egyensúlyt. Amint azt korábban leírtam, a komparatív kritérium iteratív alkalmazása igen veszélyes lehet, ugyanis elképzelhető, hogy különböző sorrendek esetén más lesz azon egyensúlyok halmaza, amik nem záródnak ki. Ebben a játékban azért nincsen ezzel probléma, mert teljesülnek Mailath, Okuno-Fujiwara és Postlewaite [1993] veretlen egyensúlyának kritériumai[3], és így gyakorlatilag a komparatív kritériumot a veretlen egyensúlyhoz hasonlóan alkalmazhatjuk. Szeretném hangsúlyozni, hogy a kritérium egyedi egyensúlyok vizsgálatára van definiálva, és amit a következőkben fel fogok vázolni semmiképpen sem egy teljesen akkurátus alkalmazás. Csupán azt szeretném bemutatni, hogy az intuitív kritérium logikája alkalmas lehet ilyen helyzetekben a legalább részben elvegyítő egyensúlyok eliminálására.

Első lépésként vegyük észre, hogy az e_N üzenetet kizárólag a t_N típusú játékos küldi, méghozzá kizárólag a Riley-egyensúlyban, ezért az összes többi egyensúlyban egyensúlyon kívüli üzenet lesz. Mivel a Riley-egyensúlyban t_N típusú játékos nagyobb hasznosságot ér el, mint bármely más egyensúlyban (ennek belátása történhet ugyanúly, mint az, hogy a kétszereplős modell esetén a h típusú játékos nagyobb hasznosságot ér el a szeparáló egyensúlyban, mint bármely elvegyítőben), e_N üzenet érkezése esetén a munkavállaló tudni fogja, hogy t_N-től érkezett, és $f(t_N)$ bért fizet. Ennek tudatában t_N el fog térni minden olyan minden olyan egyensúlyban, ahol eddig más üzenetet is pozitív valószínűséggel küldött. Ezeket az egyensúlyokat kizárhatjuk a komparatív kritérium alapján.

Ezután figyeljük meg, hogy a megmaradt egyensúlyokban t_N-en kívül minden játékos maximum $f(t_{N-1})$ bért kaphat. Ehhez azt kell látni, hogy nincs már olyan legalább részben elvegyítő egyensúly, ahol $\mu(t_N) > 0$, így az elvegyítő egyensúlyokban $E\big(p(t)\big) < p(t_{N-1})$.

[3] Pontosabban az egyik feltétel nem teljesül: a munkáltató hasznosságfüggvényének szigorúan konkávnak kellene lennie w-ben, a mi munkáltatónk pedig kockázatsemleges. Azonban ennek a feltételnek csak az a szerepe, hogy az említett játékos sohase keverjen, mindig válasszon tiszta stratégiát. Ezért ahelyett, hogy a munkáltató hasznosságfüggvényét konkávvá tenném, az érthetőség és a már megszokott jelölések használhatósága érdekében inkább csak felteszem, hogy a munkáltató mindig tiszta stratégiát választ. Ez gyakorlatilag az általánosság csorbítása nélkül feltehető.

Most alkalmazzuk újra az első lépés logikáját, azaz vegyük észre, hogy most már minden elvegyítő egyensúlyban e_{N-1} egyensúlyon kívüli üzenet, mivel ha $u^p(t_i)$ a t_i játékos kifizetését jelöli egy még nem kizárt elvegyítő egyensúlyban, akkor teljesülnek a következő egyenlőtlenségek

$$f(t_{N-1}) - v\left(e_{t_{N-1}}, t_{N-1}\right) \leq f(t_i) - v\left(e_{t_i}, t_i\right) \leq u^p(t_i) \quad \forall\, i \in \{1, 2, \ldots, N-2\}.$$

Most pedig ismét a komparatív kritériumot alkalmazva nézzük meg, milyen egyensúlyok zárhatók ki e_{N-1} üzenet érkezése esetén. Az első lépéssel megegyező módon eljuthatunk oda, hogy minden olyan legalább részben elvegyítő egyensúly kizárható, ahol eddig a t_{N-1} típusú játékot 0-nál nagyobb valószínűséggel küld.

Ezután megint az következik, hogy a megmaradt egyensúlyokban t_N-en kívül minden játékos maximum $f(t_{N-2})$ bért kaphat, hiszen, nincs már olyan legalább részben elvegyítő egyensúly, ahol $\mu(t_N) > 0$ vagy $\mu(t_{N-1}) > 0$, így az elvegyítő egyensúlyokban $E\big(p(t)\big) < p(t_{N-2})$.

Miután ezt a gondolatmenetet $N-1$-szer megismételjük, eljutunk oda, hogy minden elvegyítő egyensúly kizárásra kerül, ezért a Riley-egyensúly marad az egyetlen, amit nem elimináltunk. Ismét szeretném azonban hangsúlyozni, hogy ezt a fajta iterációt az tette lehetővé, hogy a vizsgált játék megfelelt a veretlen egyensúly segítségével vizsgálható játékokkal szemben támasztott elvárásoknak. Továbbá azt is megemlíteném, hogy az intuitív kritériumhoz jobban hasonlító eszközzel, Banks és Sobel divinity kritériumával (Banks & Sobel, 1987) is el lehet erre az eredményre jutni.

ÖSSZEFOGLALÁS

Amint arra a bevezetésben is utaltam, a szignáljátékokban előforduló, irracionálisnak tűnő egyensúlyok kiszűrésére sokféle kritérium és egyensúly-finomítás van a szakirodalomban. Az ebben a dolgozatban bemutatott intuitív kritérium, amint az a példákban is látható volt, sok esetben erősebb, mint a legáltalánosabban használt intuitív kritérium, és véleményem szerint az elvi megalapozottságával sincsen probléma. Azonban az intuitív kritérium sokoldalúbb és nem jellemzők rá azok a problémák, amelyek bizonyos játékokban jelentősen korlátozhatják a komparatív kritérium hasznosságát (számítási nehézség, egyértelműség). Úgy gondolom, hogy különböző játékokhoz különböző kritérium vagy megoldáskoncepció használata lehet a legalkalmasabb, és a példák megerősítik azt, hogy bár a komparatív kritérium nem olyan univerzális, mint Cho és Kreps intuitív kritériuma, bizonyos játékok vizsgálatánál van létjogosultsága.

TOVÁBBFEJLESZTÉSI LEHETŐSÉGEK

A komparatív kritériummal kapcsolatban rengeteg megválaszolásra váró, nyitott kérdés maradt. Véleményem szerint érdemes lehet tovább vizsgálni azt, hogy pontosan milyen feltételek szükségesek ahhoz, hogy egy adott játékban, adott egyensúly esetén az intuitív kritérium erősebb legyen, mint a komparatív kritérium. Azt is érdemes megvizsgálni, hogy az ilyen esetekben mi tűnik racionálisabbnak: az egyensúly kizárása, vagy elfogadása.

Szintén fontos nyitott kérdés az, hogy milyen feltételek szükségesek ahhoz, hogy az eredmények egyértelműsége fejezetben tárgyalt problémák ne merülhessenek fel. Első lépésként például érdekes lenne, hogy igaz-e az az állítás, hogy ha rögzítünk két tetszőleges egyensúlyt, akkor nem állhat fenn az a helyzet, hogy mindkettőt ki lehet zárni a másik alapján.

További kutatási lehetőség a komparatív kritériummal kapcsolatban, hogy hogyan viszonyul a többi, hasonló célra használt kritériumhoz és megoldáskoncepcióhoz – véleményem szerint különösen érdekes lehet a Mertens-stabilitással és a divinity kritériummal való összehasonlítás. Érdemes lehet megvizsgálni a komparatív kritériumot Rabin és Sobel 1996-os cikkének fényében is, amely erősen leegyszerűsítve azzal foglalkozik, hogy hosszú távon mely egyensúlyok fognak gyakran előfordulni (Rabin & Sobel, 1996).

IRODALOMJEGYZÉK

Banks, J. S., & Sobel, J. (1987). Equilibrium Selection in Signaling Games. *Econometrica, 55*(3), 647-661.

Cho, I.-K., & Kreps, D. M. (1987). Signalling Games and Stable Equilibria. *The Quarterly Journal of Economics, 102*(2), 179-221.

Daskalakis, C., & Goldberg, P. W. (2009). The Complexity of Computing a Nash Equilibrium. *SIAM Journal on Computing, 39*(1), 195-259.

Harsányi, J. C. (1967/68). Games with Incomplete Information Played by Bayesian Players. *Management Science, 14*(3, 5, 7), 159-183, 320-334, 486-502.

Jäger, G. (2008). Evolutionary stability conditions for signaling with costly signals. *Journal of Theoretical Biology, 253*(1), 131-141.

Kohlberg, E., & Mertens, J.-F. (1986). On the Strategic Stability of Equilibria. *Econometrica, 54*(5), 1003-1037.

Kreps, D. M., & Wilson, R. (1982). Sequential Equilibria. *Econometrica, 50*(4), 863-894.

Macho-Stadler, I., & Pérez-Castrillo, J. D. (2001). An Introduction to the Economics of Information: Incentives and Contracts. Oxford: Oxford University Press.

Mailath, G. J., Okuno-Fujiwara, M., & Postlewaite, A. (1993). Belief-Based Refinements in Signalling Games. *Journal of Economic Theory, 60*, 241-276.

Mertens, J.-F. (1989). Stable Equilibria - A Reformulation. *Mathematics of Operations Research, 14*, 575-625.

Rabin, M., & Sobel, J. (1996). Deviations, Dynamics, and Equilibrium Refinements. *Journal of Economic Theory, 68*, 1-25.

Riley, J. G. (1979). Informational Equilibrium. *Econometrica, 47*(2), 331-359.

Selten, R. (1975). A reexamination of the perfectness concept for equilibrium points in extensive games. *International Journal of Game Theory*(4), 25-55.

Spence, M. (1973). Job Market Signaling. *The Quarterly Journal of Economics, 87*(3), 355-374.