Učenje i optimalne protustrategije

Igre su rijetko zanimljive ako se igra samo jedan potez. Igranjem više poteza (općenito beskonačno mnogo), uvodi se faktor **učenja**, tj. sposobnosti igrača da prepoznaju kako protivnik igra, te tome prilagode svoj stil igre i time povećaju dobit.

Učenje

Vratimo se na primjer 5 s prošlih vježbi (Pismo-glava).

Primjer 1: Pismo – glava. Dva igrača biraju pismo ili glavu. Ako izaberu isto, prvi plaća drugom $1. Ako izaberu različito, onda drugi plaća prvom (isto $1). Dobiti su dane tablicom:

	Pismo	*Glava*
Pismo	-1, 1	1, -1
Glava	1, -1	-1, 1

Jasno je da ne postoji ni "pametna" ni "glupa" strategija za jedno igranje. Pitanje je postoje li takve strategije u slučaju serije igara?

Očito, ako neki od igrača počne igrati prema nekoj strategiji (dakle, "šablonski"), onaj drugi će primijetiti šablonu i trivijalno namjestiti igru. Drugim riječima, svako plansko igranje predstavlja jednu "gubitničku" strategiju. Jedini način da se ne igra bedasto je da se igra potpuno slučajno – npr. bacanjem novčića.

No, i "slučajna igra" ne mora biti potpuno slučajna. Na primjer, prvi igrač može preferirati pismo i igrati ga s vjerojatnošću 70%. U tom slučaju, drugi igrač će stalnom igrom na pismo dobijati češće nego što će gubiti, pa je takva slučajna igra očito loša za prvog igrača (iako, manje loša nego da igra u potpunosti predvidljivo).

Intuitivno je jasno da se igračima isplati igrati nasumično, s jednakim vjerojatnostima za sve poteze.

Uskoro ćemo vidjeti da to ne mora biti slučaj čak niti kod simetričnih igara, a vidjet ćemo i kako se računaju optimalni odgovori na strategije.

Primjer 2: Kamen – škare – papir. Malo proširena igra Pismo-glava. Igraju dva igrača. U svakom potezu oni pokazuju jedan drugome (istovremeno) simbole koji prikazuju kamen (stisnuta šaka), škare (kažiprst i srednji prst koji "glume" škare) ili papir (pruženi dlan). Poraženi plaća 1 jedinicu pobjedniku. Ako nema pobjednika, nitko nikome ne plaća ništa. Pobjeda se definira na slijedeći način: kamen pobjeđuje škare jer ih može slomiti, škare pobjeđuju papir jer ga mogu razrezati, a papir pobjeđuje kamen jer ga može omotati (pa kamen ostaje "bespomoćan"). Dakle, tablica dobiti izgleda ovako:

	q_1	q_2	q_3
p_1	0, 0	1, -1	-1, 1
p_2	-1, 1	0, 0	1, -1
p_3	1, -1	-1, 1	0, 0

Slično igri "Pismo-glava", ovdje igrač lako može iskoristit predvidljivost poteza svog protivnika, kao i njegovo "naginjanje" određenom potezu.

Jedini novi pojam je slučajni izbor jedne od tri mogućnosti. Kako to izvesti?

Jedan način je izvlačenje jedne od tri označene kuglice (ili, općenito, n označenih kuglica).

No, daleko bolji način je "randomizing spinner" - nešto kao rulet. Krug podijelimo na n jednakih isječaka (u našem slučaju na 3) označenih brojevima od 1 do n. U sredinu postavimo neki pokazivač i zavrtimo ga. Izabrani broj je onaj na kojem se pokazivač zaustavi.

Ovaj način je bolji zato jer je pogodan za sve realne vjerojatnosti, a ne samo za n jednako vjerojatnih događaja. Potrebno je samo na odgovarajući način podesiti omjere površina pojedinih isječaka.

Primjer 3: Dvoprsta Morra. Igraju dva igrača. Svaki pokaže jedan ili dva prsta i prijavi broj od 2 do 4. Onaj koji je pogodio sumu, pobjeđuje i od drugog dobija onoliko novčića kolika je ta suma. Ako je neodlučeno (ni jedan ne pogodi ili oba pogode), onda nema pobjednika.

Koliko mogućih poteza ima svaki igrač?

Strogo teoretski, ima $2 \cdot 3 = 6$ mogućnosti, no nisu sve one dobre. Očito, nitko neće pokazati 2 prsta i zvati broj 2 ili pokazati 1 prst i zvati 4, jer su to pouzdano gubitničke strategije. Dakle, svaki igrač ima točno 4 mogućnosti. Označimo ih uređenim parovima (x, y) gdje je x broj pruženih prstiju, a y broj koji igrač zove.

	(1, 2)	(1, 3)	(2, 3)	(2, 4)
(1, 2)	(0, 0)	(2, -2)	(-3, 3)	(0, 0)
(1, 3)	(-2, 2)	(0, 0)	(0, 0)	(3, -3)
(2, 3)	(3, -3)	(0, 0)	(0, 0)	(-4, 4)
(2, 4)	(0, 0)	(-3, 3)	(4, -4)	(0, 0)

Koja su ravnotežna stanja?

Što sada možemo reći za strategije?

Iako je igra simetrična, slučanjo igranje nam ovdje neće pomoći. Dapače, takvu igru protivnik lako može iskoristiti. Kako?

Recimo da prvi igrač stalno igra slučajno. Tada drugi igrač može stalno igrati (1, 3) ili bilo koju drugu opciju čija je suma po komponentama za njega pozitivna. Tada će njegova prosječna dobit po potezu dugoročno biti oko

$$\frac{1}{4}\cdot(-2)+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 0+\frac{1}{4}\cdot 3=\frac{1}{4}$$

“Prava” strategija je nešto složenija. Knjiga tvrdi da smo još premali za to... :-)

Naravno, ovdje ne uzimamo da će i prvi igrač “učiti”. Naime, ako on promijeni taktiku, promijenit će ju i drugi. Drugim riječima, ako prvi igrač počne igrati “nasumično”, drugi će zaigrati na – recimo – (1, 3). No, tada će prvi to shvatiti, pa će početi igrati na (2, 4). Onda će drugi brzo “uhvatiti štos”, pa će se vratiti na (1, 3). Razumno bi bilo očekivati da će teorija igara izbjeći ovakve (beskonačne) logičke petlje. Kao što ćemo vidjeti, to će se stvarno i dogoditi, izborom i optimalne strategije (a ne samo protustrategije).

Primjer 4: Bombarderi. Igraju dva igrača, koji su generali u protivničkim vojskama. Svaki dan prvi šalje dva aviona na drugog. Jedan avion je veliki bombarder, a drugi manji avion “za podršku”. Cilj napada je ispuštanje jedne bombe na neki cilj kod drugog generala. Drugi general ima jedan avion kojim očekuje napad. Taj avion će se obrušiti na jedan od dva napadačka. Vjerojatnosti da će napad prvog generala uspjeti, dane su tablicom:

		Koga napasti?	
		Bombarder	Avion za podršku
Kod koga je bomba?	Bombarder	80 %	100 %
	Avion za podršku	90 %	50 %

Kako se dva generala trebaju ponašati? Drugim riječima, gdje prvi treba postaviti bombu i kojeg od dva aviona treba drugi napadati?

Kao i do sada, ako neki general primijeti “šablonu”, postupit će onako kako mu najviše odgovara.

Recimo, prvi igrač može stalno stavljati bombu u bombardera. To će drugi general brzo primijetiti i poćet će napadati samo njega. No, što ako prvi general odluči igrati baš po tom obrascu, s time da povremeno stavi bombu u avion za podršku? Recimo da to napravi u 25% slučajeva. Očito, efikasnost će biti $80\cdot 0.75+90\cdot 0.25=82.5$

Sada drugi general može odgovoriti na blef. Samo, kako često da napadne manji avion? Ako to napravi u 50% slučajeva, onda će uspješnost napada porasti na 85%. Drugim riječima, čini se da se blef prvog generala isplatio.

Da je drugi general napadao manji avion u 20% slučajeva, uspješnost napada bi bila 83.5%. Ispada da se drugom generalu ne isplati blefirati.

Može li prvi general još poboljšati svoju učinkovitost? Koji je najbolji odgovor drugog generala? Postoji li strategija za drugog generala neovisna o ponašanju prvog?

Strogo kompetitivne igre i strategije

Definicija. 2×2 **igra sume nula** (ili **strogo kompetitivna** igra) je matrica realnih brojeva dimenzija 2×2.

Primjer ovakve igre smo već vidjeli (Primjer 1, *Pismo-glava*).

Interpetacija matrice strogo kompetitivne igre je slijedeća:

> Igraju dva igrača (zato je matrica dvodimenzionalna). Svaki na raspolaganju ima dvije "čiste" strategije (retci za prvog, odnosno stupci za drugog igrača). Brojevi u matrici predstavljaju dobitak prvog, odnosno gubitak drugog igrača u svakoj situaciji.
>
> Prvi igrač bira redak, a drugi stupac. Ti izbori reprezentiraju njihove strategije u jednom koraku.

Općenito, za teoriju nije bitna interpretacija brojeva, tako da igre možemo promatrati i samo kao ovakve matrice. Takve igre (bez "pričice") zovemo **apstraktne igre**.

80%	100%
90%	50%

Na slici desno se nalazi matrični prikaz igre bombardera. Pri tom je prvi igrač odlučio zaigrati na prvu strategiju (označeni redak), tj. postaviti bombu u bombarder, dok se drugi igrač odlučio za drugu strategiju (označeni stupac), tj. napasti avion za podršku. Očito, uz takav izbor strategija, dobitak (prvog igrača!) je 100% (iznos u "presjeku" izabranog retka i stupca).

Definicija. Strategija za jedan potez je redak (odnosno stupac) iz matrice kojom reprezentiramo igru.

Sada je logično zapitati se: što je strategija u općenitom slučaju, kad se igra izvodi u više poteza (općenito, beskonačno mnogo poteza)?

Kao što smo već rekli, općenito nema smisla stalno igrati istu strategiju jer to omogućuje protivniku da izabere za sebe najpovoljniju strategiju (kod strogo kompetitivnih igara, ta je strategija za nas najmanje povoljnu). Zato ima smisla dati slijedeću definiciju.

Definicija. Strategija u 2×2 strogo kompetitivnoj igri je uređeni par (a, b) takav da vrijedi: $0 \leq a, b \leq 1$ i $a + b = 1$, gdje a predstavlja frekvenciju kojom je izabran prvi redak (ili stupac), a b predstavlja frekvenciju kojom je izabran drugi redak (ili stupac).

Prilikom razmatranja općenitih 2×2 strategija, zgodno je strategiju prvog igrača označiti s $(1-p, p)$, a strategiju drugog s $(1-q, q)$.

Ovakav koncept označavanja je prvi uveo E. Borel u seriji članaka 1920. godine. Još je on primijetio da nije sigurno da je ova notacija dovoljna. Preciznije, ne bi li se frekvencije mogle mijenjati u vremenu?

Takav razvoj teorije ima dvije ozbiljne zamjerke.

- Prva, čisto praktična, se odnosi na činjenicu da matematika nema sredstva s kojima bi se definiralo i analiziralo takvo ponašanje. Čak i slučajne varijable, kao pojam najbliži "slučajnom" ponašanju, moraju imati zadanu distribuciju – dakle opet nekakvu frekvenciju.
- Druga zamjerka je mnogo važnija. Postavlja se pitanje da li je uopće moguće postići potpuno slučajno ponašanje? Ako nije, onda se igra odvija po nekom uzorku, tj. izbor strategija se odvija s nekim frekvencijama.

  Psihologija je pokazala da su ljudske odluke uvjetovane osobnim preferencama i iskustvom. Drugim riječima, koliko god se netko trudio postići "slučajnost" izbora, njegove odluke neće biti slučajne. Dakle, naše odluke s vremenom konvergiraju nekoj frekvenciji.

  Slično je i s raznim "sistemima" za postizanje slučajnosti. Njihove frekvencije je najčešće lako pogoditi unaprijed. Npr. kod bacanja novčića, to je 0.5 za svaku opciju. Kod "randomizing spinnera", frekvencije se dobiju kao omjeri površina.

Borel je došao do sličnog zaključka kao i Von Neumann i Morgenstern: frekvencije, interpretirane kao vjerojatnosti povlačenja pripadnog poteza, predstavljaju dobru reprezentaciju strategija.

Definicija. Pomoćni dijagram (engl. ***auxilary diagram***) prikazuje apstraktnu igru i neku kombinaciju strategija.

Primjer 4: Bombarderi. Očekivanu isplatu lako dobijemo kao ponderiranu sumu. Pogledajmo to na primjeru bombardera:

		Koga napasti?	
		Bombarder	Avion za podršku
Kod koga je bomba?	Bombarder	a	b
	Avion za podršku	c	d

Pripadne strategije su $(1-p, p)$ za napadača i $(1-q, q)$ za onoga koji se brani. Brojevi a, b, c i d predstavljaju očekivanu uspješnost napada u pripadnoj situaciji. Dakle, pomoćni dijagram za igru bombardera izgleda ovako:

	$1-q$	q
$1-p$	a	b
p	c	d

Očekivana isplata za napadača (odnosno gubitak za obrambenog generala) ovdje predstavlja očekivanu uspješnost napada, a računa se po formuli:

$$(1-p)\cdot(1-q)\cdot a+(1-p)\cdot q\cdot b+p\cdot(1-q)\cdot c+p\cdot q\cdot d$$

Sada je lako izračunati isplativost za svaku strategiju. Na primjer:

Strategija napadača	Strategija obrane	Isplativost (za napadača)
(0.3, 0.7)	(0.6, 0.4)	78.2 %
(0.75, 0.25)	(1, 0)	82.5 %
(0.75, 0.25)	(0.5, 0.5)	85.0 %
(0.75, 0.25)	(0.8, 0.2)	83.5 %

Posljednja 3 primjera smo već razmotrili. Zaključak je bio (kao što se i vidi iz priložene tablice) da se obrani ne isplati blefirati. Kasnije ćemo vidjeti da to nije slučajno.

Kako izgleda općeniti slučaj, kad svaki igrač ima veći izbor?

Definicija. Općenita $m\times n$ **igra sume nula** (ili **strogo kompetitivna** igra) je matrica realnih brojeva dimenzija $m\times n$.

U ovakvoj igri, slično kao u slučaju 2×2, prvi igrač ima *m*, a drugi igrač *n* izbora poteza na raspolaganju.

Primjer 3: Dvoprsta Morra.

	(1, 2)	(1, 3)	(2, 3)	(2, 4)
(1, 2)	0	2	-3	0
(1, 3)	-2	0	0	3
(2, 3)	3	0	0	-4
(2, 4)	0	-3	4	0

Definicija. Strategija u $m\times n$ strogo kompetitivnoj igri je uređena *n*-torka $(p_1, p_2, \ldots, p_k)$ takva da vrijedi: $0\le p_i\le 1 \ \ \forall i\in\{1,2,\ldots,k\}$ i $\sum_{i=1}^{k} p_i=1$, gdje p_i predstavlja frekvenciju kojom je izabran *i*-ti redak (ili stupac). Naravno *k* je jednak *m* ili *n*, ovisno o tome čija je strategija.

Definicija. Neka je dana strogo kompetitivna igra, određena matricom $[a_{ij}]_{i\in\{1,\ldots,m\}, j\in\{1,\ldots,n\}}$. Neka su dane strategije: $(p_1, p_2, \cdots, p_m)$ za prvog i $(q_1, q_2, \cdots, q_k)$ za drugog igrača. **Očekivanu isplatu** u $m\times n$ strogo kompetitivnoj igri računamo prema formuli: $E(p,q)=\sum_{i=1}^{m}\sum_{j=1}^{n} p_i\cdot q_j\cdot a_{ij}$.

Strategije razlikujemo po prirodi, tj. po tome da li uvijek igramo “na istu kartu” ili imamo različite odluke u različitim koracima.

Definicija. Neka je dana strategija $P=(p_1, p_2, \cdots, p_k)$. Za strategiju P kažemo da je **čista strategija** ako $\exists i:\ p_i=1, p_j=0\ \ \forall j\neq i$. U protivnom, kažemo da je P **miješana strategija**.

Neki zanimljivi primjeri igara sume nula

Slijedeća tri primjera je dao J. D. Williams u knjizi *The Complete Strategyst*, Dover, New Zork, 1986.

Primjer 5: Hi-Fi. Kompanija *Gunning & Kappler* proizvodi pojačala. Kritičan dio pojačala je mali, nepristupačni kondenzator. Kompanija taj dio plaća $1, no gube $10 ako je dio neispravan. Prilikom nabave kondenzatora imaju još dvije mogućnosti: kupiti za $6 kondenzator vrhunske kvalitete ili za $10 kondenzator koji dolazi s potpunom garancijom koja pokriva sve gubitke u slučaju kvara.

Ovo je problem vrlo sličan problemima koje ćemo vidjeti u Teoriji odlučivanja. Kako to uklopiti u Teoriju igara, tj. tko je drugi igrač i koje su njegove strategije?

Drugi igrač je *Priroda*, koju uzimamo kao "protivnika" u svim sličnim situacijama. Naravno, strategije *Prirode* su "ispravan" i "neispravan" kondenzator. Kako *Priroda* nije svjestan igrač, njenu strategiju određuje statistički dobivena vjerojatnost, u ovom slučaju, ispravnosti kondenzatora.

Primjer 6: Prodavač (engl. *Huckster*). Merril ima koncesiju za držanje prodajnog štanda na stadionu, za prodaju sunčanih naočala i kišobrana. Naravno, uspješnost prodaje jako ovisi o vremenu. Iskustvo je pokazalo da može prodati oko 500 kišobrana ako pada kiša i samo 100 ako je sunčano. No, po kišnom vremenu neće prodati gotovo ništa naočala, dok će ih po sunčanom vremenu prodati oko 1000 komada.

Kišobrane nabavlja po $0.50, a prodaje ih po $1.00 (primjer datira iz 1954 godine). Naočale nabavlja po $0.20, a prodaje ih po $0.50. Sve što ne proda je čisti gubitak (baca se, poklanja *Caritasu* ili nešto slično).

Kao i u prethodnom primjeru, i ovdje je drugi igrač *Priroda*. Ono što možda nije *a priori* jasno, to je koje su Merrilove strategije (za svaki "potez"). Očito, to mogu biti sve kombinacije broja kišobrana (od 100 do 500) i naočala (od 0 do 1000). Tih mogućnosti ima mnogo. Može li bolje, tj. da li je moguće odrediti manji skup Merrilovih mogućnosti?

Primjer 7: Problem nabave ugljena. Hans planira nabavu ugljena za zimu. Zna da mu za grijanje kuće tokom normalne zime treba 15 tona, no stvarna količina može varirati između 10 tona (ako je zima blaga) i 20 tona (ako je zima oštra). Također, cijena ugljena varira, u ovisnosti o jačini zime, i iznosi $10, $15 i $20 po toni ako je zima blaga, umjerena, odnosno jaka. Također, ako kupuje unaprijed, dobit će ugljen po cijeni $10 za tonu. Potrebno je odlučiti koliko ugljena kupiti sada. Svaki deficit će morati nadoknaditi po aktualnoj cijeni. Također, sav nepotrošeni ugljen je čisti gubitak jer Hans na proljeće seli u Kaliforniju (dosadiše mu zime) i nije u mogućnosti ponijeti ugljen sa sobom.

Optimalni odgovori na specifične strategije (tj. optimalne protustrategije)

Pogledajmo sada kako igrač može najbolje odgovoriti ako mu je poznata protivnikova strategija.

> **Napomena:** Pojam "poznata protivnička strategija" ne znači da igrač može predvidjeti protivnikov slijedeći potez, nego samo da zna frekvencije kojima ovaj vuče određene poteze!

Pretpostavljamo da igrači imaju fiksne strategije, tj. da ih ne mijenjaju kroz vrijeme. Ako se promijeni strategija, onda je jasno da će se promijeniti i optimalni odgovor na tu strategiju.

Vratimo se na Primjer s bombarderima. Tamo smo analizirali nekoliko odgovora na blef napadača. Razmotrimo sada malo općenitiju situaciju: napadač postavlja bombu u bombarder s frekvencijom 75% (tj. njegova strategija je (0.75, 0.25)). Zanima nas koja je strategija u tom slučaju najbolja za obrambenog igrača?

Općenito, njegove strategije su oblika $(1-q, q)$, gdje je $q \in [0,1]$. Dakle, tražimo q koji minimizira dobit. Prisjetimo se da dobit računamo na slijedeći način:

$$
\begin{aligned}
&0.75\cdot(1-q)\cdot 80 + 0.75\cdot q\cdot 100 + 0.25\cdot(1-q)\cdot 90 + 0.25\cdot q\cdot 50\\
&\quad = 0.60\cdot(1-q) + 0.75\cdot q + 0.225\cdot(1-q) + 0.125\cdot q\\
&\quad = 0.60 - 0.60\cdot q + 0.75\cdot q + 0.225 - 0.225\cdot q + 0.125\cdot q\\
&\quad = 0.825 + 0.05\cdot q
\end{aligned}
$$

Sada je očito da će dobit biti najmanja za $q=0$, tj. ako obrana ne bude blefirala, nego odluči napadati isključivo bombardera.

Slično, kad bi strategija napadača bila (0.50, 0.50), dobili bismo isplativost $0.85-0.10\cdot q$. Dakle, u tom slučaju bi se obrani isplatilo napadati isključivo pomoćni avion.

Definicija. Za danu strategiju prvog igrača, svaku strategiju drugog igrača koja rezultira najmanjom dobiti zovemo **optimalna protustrategija**. Analogno, za danu strategiju drugog igrača, svaku strategiju prvog igrača koja rezultira najvećom dobiti također zovemo **optimalna protustrategija**.

U prošlom primjeru smo vidjeli da obrana ima optimalnu protustrategiju koja je za obje analizirane miješane strategije napadača čista. To nije slučajno, kao što vidimo iz slijedećeg teorema.

Teorem 1. Za svaku fiksnu strategiju jednog igrača, drugi igrač ima čistu optimalnu protustrategiju.

Dokaz nećemo provoditi, no lako se vidi da tvrdnja proizlazi iz linearnosti problema, tj. činjenice da su miješane strategije konveksne kombinacije čistih strategija, pa se ekstremi postižu na rubovima, tj. u čistim strategijama.

Ponekad postoji više optimalnih protustrategija, od kojih neke mogu biti i miješane. To se događa u situacijama kada su barem dvije čiste strategije optimalne protustrategije. Tada su optimalne i njihove miješane strategije, što proizlazi iz teorema 2.

Prije navođenja tog teorema, vratimo se još malo na primjer 2:

Primjer 2: Kamen – škare – papir. Pomoćni dijagram izgleda ovako:

	q_1	q_2	q_3
p_1	0	1	-1
p_2	-1	0	1
p_3	1	-1	0

Koja je optimalna protustrategija za drugog igrača, ako znamo da je strategija prvog igrača (0.2, 0.3, 0.5)?

Zahvaljujući teoremu 1, dovoljno je provjeriti 3 čiste protustrategije:

	1	0	0
p_1	0	1	-1
p_2	-1	0	1
p_3	1	-1	0

Dobit: 0.2

	0	1	0
p_1	0	1	-1
p_2	-1	0	1
p_3	1	-1	0

Dobit: -0.3

	0	0	1
p_1	0	1	-1
p_2	-1	0	1
p_3	1	-1	0

Dobit: 0.1

Sada je očito da je za drugog igrača najbolje da stalno igra na škare. Ako pogledamo intuitivno, to itekako ima smisla. Naime, prvi igrač preferira papir, pa ima smisla igrati na škare koje pobjeđuju papir. Ono što možda iznenađuje, je činjenica da drugi igrač ne treba samo preferirati škare, nego treba ekskluzivno igrati samo na njih!

Očito, vrijedi i slijedeći teorem.

Teorem 2. Svaka konveksna kombinacija oblika $a \cdot P + b \cdot Q$, $0 \leq a, b \leq 1$, $a + b = 1$ optimalnih protustrategija P i Q je i sama optimalna protustrategija.

Naravno, sve optimalne protustrategije moraju dati istu dobit, pa će i njihova konveksna kombinacija dati tu istu dobit.

Jednostavnija verzija teorema 1, u slučaju 2×2 igre, se dokazuje jednostavnim raspisom formule za dobit. Dakle, neka je dana igra:

	$1-q$	q
$1-p$	a	b
p	c	d

Neka je $E(p,q)$ očekivana dobit ako su strategije prvog i drugog igrača $(1-p,p)$ i $(1-q,q)$ respektivno. Tada je:

$$E(p,q)=(1-p)\cdot(1-q)\cdot a+(1-p)\cdot q\cdot b+p\cdot(1-q)\cdot c+p\cdot q\cdot d$$
$$=p\cdot(-a+a\cdot q+c-b\cdot q-c\cdot q+d\cdot q)+(a-a\cdot q+b\cdot q)$$

Pretpostavimo da drugi igrač ima fiksnu strategiju. Tada prvi igrač postiže najveću dobit za:

$$p=\begin{cases}0, & -a+a\cdot q+c-b\cdot q-c\cdot q+d\cdot q\geq 0\\ 1, & -a+a\cdot q+c-b\cdot q-c\cdot q+d\cdot q\leq 0\end{cases}$$

Dakle, prvi igrač ima čistu optimalnu protustrategiju. Slično se dokazuje i za drugog igrača, kad prvi ima fiksnu strategiju.

Primijetimo još da vrijedi:

$$E(p,q)=(1-p)\cdot((1-q)\cdot a+q\cdot b)+p\cdot((1-q)\cdot c+q\cdot d)$$
$$=(1-p)\cdot E(0,q)+p\cdot E(1,q)$$

Sada je jasno da se optimalna protustrategija dobije za:

$$p=\begin{cases}0, & E(0,q)\geq E(1,q)\\ 1, & E(0,q)\leq E(1,q)\end{cases}$$

Opet, sličan dokaz vrijedi i kad igrači zamijene uloge.

Što se događa u općenitom slučaju?

Vratimo li se na izraz za računanje dobiti u općenitoj $m\times n$ igri,

$$E(p,q)=\sum_{i=1}^{m}\sum_{j=1}^{n} p_i\cdot q_j\cdot a_{ij},$$

$$p=(p_1,\ldots,p_m),\ 0\leq p_i\leq 1\quad \forall i\in\{1,\ldots,m\},\sum_{i=1}^{m} p_i=1,$$

$$q=(q_1,\ldots,q_n),\ 0\leq q_i\leq 1\quad \forall i\in\{1,\ldots,n\},\sum_{i=1}^{n} q_i=1$$

vidimo da je isplativost, u slučaju da je p ili q fiksan, točka unutar simpleksa određenog vektorom $p\cdot a$ (odnosno $q\cdot a$).

Vratimo se još na igru “Pismo-glava”:

Primjer 1: Pismo – glava. Pomoćni dijagram izgleda ovako:

	q	1 - q
p	-1	1
1 - p	1	-1

Promotrimo strategiju prvog igrača. On će igrati na "*pismo*" s vjerojatnošću *p*, a na "*glavu*" – očito – s vjerojatnošću 1-*p*. Koliku prosječnu dobit (po potezu) on može očekivati?

Ako drugi igrač bude igrao stalno na "*pismo*", prvi očito može očekivati dobit: $p\cdot 1+(1-p)\cdot(-1)=2p-1$.

Slično, ako drugi igrač za svoj stalni izbor odabere "*glavu*", prvi može očekivati dobit: $p\cdot(-1)+(1-p)\cdot 1=1-2p$.

Očito, ako drugi igrač igra na "*pismo*" s vjerojatnošću *q*, očekivana dobit prvog igrača se dobije kao ponderirana suma prethodna dva slučaja:

$$\begin{aligned}&(2p-1)\cdot q+(1-2p)\cdot(1-q)=2pq-q+1-2p-q+2pq=\\&\quad=4pq-2q-2p+1=(2p-1)\cdot(2q-1)\end{aligned}$$

Dakle, za $p=0.5$ prvi igrač može očekivati dobit 0, bez obzira na strategiju drugog igrača. No, ako zaigra s nekom drugom vjerojatnošću, drugi igrač će s vremenom shvatiti kako ovaj igra (tj. s kojom vjerojatnošću igra na "*pismo*", a s kojom na "*glavu*"), pa može namjestiti *q* tako da u prosjeku pobjeđuje.

Zaključujemo da je za svakog igrača najbolje da igra "nasumično", i to tako da oba izbora budu podjednako zastupljeni.