Poznámky opravovatelů k úloze P-III-3

Při opravování jsme narazili na několik zajímavých přístupů k řešení, o které bychom se s vámi chtěli podělit. Níže naleznete některé z nich, společně s důkazy a protipříklady.

Dorovnávání přebytku

Snažíme se najít postup, který nám zaručí, že na konci budeme mít od nějaké barvy alespoň tolik kartiček, jako je polovina kartiček získaných podvodníkem. Toho například dosáhneme, když zajistíme, že v celém průběhu hry budeme od *každé* barvy mít alespoň polovinu ze všech možných kartiček, které šlo od této barvy získat. Pak totiž na konci budeme mít i polovinu kartiček barvy, kterou si vybral podvodník.

Algoritmus bude jednoduchý: pro každou barvu x si budeme pamatovat dvě čísla: kolik kartiček jsme od dané barvy sebrali a kolik nám celkem bylo kartiček dané barvy nabídnuto (v obou případech nezapočítáváme dvě kartičky, které jsme dostali na začátku). Označme tato dvě čísla jako s_x a n_x. Vždy, když se budeme muset rozhodnout mezi barvami k a ℓ, spočteme si pro obě z nich jejich *přebytek* $p_x = 2s_x - n_x$ a vybereme si kartu té barvy, jejíž přebytek je menší. Neformálně nám přebytek říká, o kolik máme kartiček dané barvy více, než potřebujeme.

Dokážeme, že žádný přebytek nikdy neklesne pod -2, z čehož okamžitě vyplyne, že se započítáním dvou počátečních kartiček budeme mít vybraných kartiček alespoň tolik, jako polovina nabídnutých.

Součet všech přebytků je po celou dobu hry nulový, neboť je nulový na začátku a v každém tahu jeden přebytek zvýšíme o jedna a jeden o jedna snížíme. K tomu, abychom jeden přebytek dostali na -3, nás soupeř musel donutit vybrat si mezi dvěma kartičkami s přebytkem -2. Jelikož tyto kartičky nemohly získat přebytek -2 ve stejném tahu, museli jsme někdy předtím mít kartičky s přebytkem -2, -1 a -1. Zbývající dvě kartičky ale tomto okamžiku musejí mít dohromady přebytek 4, tedy buď má jedna z nich přebytek alespoň 3, nebo jsou oba přebytky 2. Jelikož situace je symetrická, můžeme si podobným argumentem rozmyslet, že k tomu, aby toto nastalo, musely někdy předtím mít nějaké dvě kartičky součet přebytků -4.

To už dává nástroj k důkazu sporem: pro spor si vezmeme nejdřívější okamžik, kdy mají nějaké dvě kartičky součet přebytků 4 nebo -4. Právě jsme ale ukázali, že tato situace nutně nastala ještě dříve, což dává spor s tím, že jsme si vybrali nejdřívější okamžik.

Častěji nabízená barva

Budeme si pro každou barvu počítat, kolik kartiček této barvy nám již bylo nabídnuto, a ze dvou nabízených možností si vždy vezmeme tu, která nám doposud byla nabídnuta vícekrát. Pokud nám obě barvy byly nabídnuty stejněkrát, vezmeme si libovolnou kartičku. (Jak za chvíli ukážeme, je to jedno, jelikož nezávisle na tom, jak v takovém případě postupujeme, dokážeme najít protipříklad.)

Uvažme následující vstup (tučně jsou vyznačené ty kartičky, které si vybereme):

$$\mathbf{A}\mathrm{B},\ \ n\times\mathbf{A}\mathrm{C},\ \ n\times\mathbf{B}\mathrm{D},\ \ n\times\mathbf{C}\mathrm{E},\ \ (n+2)\times\mathbf{D}\mathrm{E},\ \ n\times\mathrm{C}\mathbf{E}.$$

Kromě volby $\mathbf{A}$B na začátku, kterou předpokládáme bez újmy na obecnosti, jsou všechny volby jednoznačné.

Naše skóre po této hře bude $n+4$ (díky barvě D), kdežto skóre podvodníka bude $3n+4$ (díky E). Nejen, že pokud zvolíme dostatečně velké n, dokážeme poměr našeho a optimálního algoritmu stáhnout pod 1/2, my se dokonce umíme přiblížit libovolně blízko 1/3. Tento algoritmus tedy určitě není lepší než (1/3)-kompetitivní.

Častěji odmítnutá barva

Pro každou barvu si budeme pamatovat počet kartiček, které jsme si nevzali, a vždy vybereme tu kartičku, kterou jsme odmítli nejvíckrát.

Problém nastane na následujícím vstupu:

$$\underbrace{\mathbf{A}\mathrm{B},\mathrm{A}\mathbf{B},\mathbf{A}\mathrm{B},\mathrm{A}\mathbf{B},\ldots,\mathbf{A}\mathrm{B},\mathrm{A}\mathbf{B}}_{2n},\ \ n\times\mathbf{A}\mathrm{C},\ \ n\times\mathbf{B}\mathrm{D},\ \ \underbrace{\mathbf{C}\mathrm{D},\mathrm{C}\mathbf{D},\ldots,\mathbf{C}\mathrm{D},\mathrm{C}\mathbf{D}}_{8n}.$$

Volby v první a poslední skupině stejných karet předpokládáme bez újmy na obecnosti, volby v prostředních skupinách jsou jednoznačně dané. V každém případě na konci má podvodník $9n+2$ karet barvy C nebo D, ale my jich máme od obou jen $4n+2$ a ostatních barev máme ještě méně. Podobným argumentem jako u předchozího protipříkladu nahlédneme, že algoritmus jistě není lepší než (4/9)-kompetitivní.

Barva, které už máme více

Pro každou barvu si budeme pamatovat, kolik od ní již máme kartiček, a máme-li si mezi dvěma barvami vybrat, vezmeme si tu, od které již máme více kartiček.

Tento přístup lze také rozbít, uvažme vstup:

$$\mathbf{A}\mathrm{B},\ \ \mathbf{B}\mathrm{C},\ \ \mathbf{C}\mathrm{D},\ \ \mathbf{D}\mathrm{E},\ \ n \times \mathbf{A}\mathrm{E},\ \ n \times \mathbf{B}\mathrm{E},\ \ n \times \mathbf{C}\mathrm{E},\ \ n \times \mathbf{D}\mathrm{E}.$$

První čtyři volby předpokládáme bez újmy na obecnosti, obecně bychom algoritmu dávali dvojice barev, od kterých ještě nemá žádnou kartičku, čímž po čtyřech krocích zaručíme, že algoritmus má čtyři kartičky různých barev.

Podvodníkovo skóre je $4n + 2$, naše skóre je jenom $n + 3$. Algoritmus je tedy přinejlepším $(1/4)$-kompetitivní.

Zlomková řešení

Existuje celá řada velmi podobných řešení, která (ne)fungují následovně. Opět si pamatují počty s_x a n_x sebraných a nabídnutých kartiček pro každy typ. Vyberou si tu kartičku, která bude mít menší zlomek $r_x = s_x/n_x$. Jednotlivé varianty zlomkových řešení se pak liší jen v tom, jakými hodnotami počítadla s_x a n_x inicializují, a kdy přesně k nim přičítají jedničku.

Podívejme se nejprve na řešení, které inicializuje $s_x = n_x = 2$ a proměnné inkrementuje až potom, co si nějakou kartičku vybere. Na toto řešení vyzraje následující protipříklad (jen potřebujeme za n dosadit dostatečně velké číslo):

$$\underbrace{\mathbf{A}\mathrm{B}, \mathrm{A}\mathbf{B}, \mathbf{A}\mathrm{B}, \mathrm{A}\mathbf{B}, \ldots, \mathbf{A}\mathrm{B}, \mathrm{A}\mathbf{B}}_{2n},\ \ 2 \times \mathbf{A}\mathrm{C},\ \ 2 \times \mathbf{A}\mathrm{D},\ \ 2 \times \mathbf{A}\mathrm{E},\ \ \mathbf{C}\mathrm{D},\ \ \underbrace{\mathbf{D}\mathrm{E}, \mathrm{D}\mathbf{E}, \ldots, \mathbf{D}\mathrm{E}, \mathrm{D}\mathbf{E}}_{4n}.$$

Během první části označené složenou závorkou budou postupně klesat zlomky r_A a r_B a postupně se velmi přiblíží 1/2. Následně při vstupu $2 \times \mathbf{A}\mathrm{C}$ si vybereme právě A, protože r_C je v prvním kroku 2/2 a ve druhém 2/3. To samé platí pro AD a AE. Během toho bude sice zlomek r_A pomalu růst, ale pro dostatečně velké n bude stále $r_A < 2/3$.

Tím jsme dostali zlomky $r_C = r_D = r_E = 2/4$. Bez újmy na obecnosti si z dvojice **C**D vybereme právě C. Zlomky tedy budou $r_D = 2/5, r_E = 2/4$. Poslední částí označenou složenou závorkou jen zařídíme, že znaky D a E budou ty nejčastější. Konkrétně dostaneme $r_D = (2 + 2n)/(5 + 4n), r_E = (2 + 2n)/(4 + 4n)$. V tu chvíli náš algoritmus získal $2 + 2n$ bodů, ale podvodník mohl vybírat D a dostat $5 + 4n$. Tento algoritmus je tedy horší než $(1/2)$-kompetitivní.

Stejný protipříklad funguje také pro variantu, kdy inicializujeme $s_x = 3, n_x = 2$ a počítadlo nabídnutých n_x zvětšíme o jedna ještě před samotným porovnáním zlomků.

Setkali jsme se i s řešením, kdy inicializujeme $s_x = n_x = 2$ a n_x zvětšujeme před porovnáním zlomků, s_x naopak až po provedení tahu. V tomto případě je náš protipříklad výrazně komplikovanější, takže jej nebudeme vysvětlovat slovně. Opět zvýrazňujeme tu barvu, kterou si algoritmus vybere. Zlomky r_x uvádíme v okamžiku rozhodování. Naopak největší s_x (odpovídá našemu skóre) a největší n_x (skóre podvodníka) uvádíme až po provedení daného tahu, tedy v momentu, kdy může soupeř hru ukončit.

Krok	Barvy	r_A	r_B	r_C	r_D	r_E	Největší s_x	Největší n_x
1.	**A**B	2/3	2/3	2/2	2/2	2/2	$s_A = 3$	$n_A = n_B = 3$
2.	A**B**	3/4	2/4	2/2	2/2	2/2	$s_A = s_B = 3$	$n_A = n_B = 4$
…	…	…	…	…	…	…	…	…
99.	**A**B	51/101	51/101	2/2	2/2	2/2	$s_A = 52$	$n_A = n_B = 101$
100.	A**B**	52/102	51/102	2/2	2/2	2/2	$s_A = s_B = 52$	$n_A = n_B = 102$
101.	**A**C	52/103	52/102	2/3	2/2	2/2	$s_A = 53$	$n_A = 103$
102.	A**C**	53/104	52/102	2/4	2/2	2/2	$s_A = 53$	$n_A = 104$
103.	**A**C	53/105	52/102	3/5	2/2	2/2	$s_A = 54$	$n_A = 105$
104.	**B**D	54/105	52/103	3/5	2/3	2/2	$s_A = 54$	$n_A = 105$
105.	B**D**	54/105	53/104	3/5	2/4	2/2	$s_A = 54$	$n_A = 105$
106.	**B**D	54/105	53/105	3/5	3/5	2/2	$s_A = s_B = 54$	$n_A = n_B = 105$
107.	**C**D	54/105	54/105	3/6	3/6	2/2	$s_A = s_B = 54$	$n_A = n_B = 105$
108.	**C**E	54/105	54/105	4/7	3/6	2/3	$s_A = s_B = 54$	$n_A = n_B = 105$
109.	**A**B	54/106	54/106	5/7	3/6	2/3	$s_A = 55$	$n_A = n_B = 106$
110.	**B**C	55/106	54/107	5/8	3/6	2/3	$s_A = s_B = 55$	$n_B = 107$
111.	**B**C	55/106	55/108	5/9	3/6	2/3	$s_B = 56$	$n_B = 108$
112.	**C**E	55/106	56/108	5/10	3/6	2/4	$s_B = 56$	$n_B = 108$
113.	D**E**	55/106	56/108	6/10	3/7	2/5	$s_B = 56$	$n_B = 108$

114.	**D**E	55/106	56/108	6/10	3/8	3/6	$s_B = 56$	$n_B = 108$
...	...	...	...	...	...	...	...	...
220.	**D**E	55/106	56/108	6/10	56/114	56/112	$s_D = 57$	$n_D = 114$
221.	D**E**	55/106	56/108	6/10	57/115	56/113	$s_D = s_E = 57$	$n_D = 115$

Ve většině případů mají porovnávané zlomky r_x a r_y různou hodnotu, takže si algoritmus vybere menší z nich jednoznačně, nebo jsou obě barvy zcela nerozlišitelné ($s_x = s_y$ a $n_x = n_y$), takže bez újmy na obecnosti předpokládáme, že si vybral podle abecedy. Jedinou výjimkou z tohoto pravidla je krok 112, ve kterém $r_C = 5/10 = 2/4 = r_E$. V příkladu výše předpokládáme, že si algoritmus vybral C a na konci dosáhneme skóre $s_D = s_E = 57$, ovšem podvodník vybírající vždy barvu D by mohl získat dokonce $n_D = 115$. Pokud by si však algoritmus zvolil v kroku 112 barvu E, stačí, aby mu v krocích 113 až 221 soupeř nabízel CD, a dostaneme následující konec:

Krok	Barvy	r_A	r_B	r_C	r_D	r_E	Největší s_x	Největší n_x
...	...	...	...	...	...	...	...	...
112.	C**E**	55/106	56/108	5/10	3/6	2/4	$s_B = 56$	$n_B = 108$
113.	C**D**	55/106	56/108	5/11	3/7	3/4	$s_B = 56$	$n_B = 108$
114.	**C**D	55/106	56/108	5/12	4/8	3/4	$s_B = 56$	$n_B = 108$
...	...	...	...	...	...	...	...	...
220.	**C**D	55/106	56/108	58/118	57/114	3/4	$s_C = 59$	$n_C = 118$
221.	C**D**	55/106	56/108	59/119	57/115	3/4	$s_C = 59$	$n_C = 119$

V tomto případě algoritmus získal skóre $s_C = 59$, ale podvodník vybírající vždy C by získal $n_C = 119$. V obou případech však dostáváme poměr menší než 1/2, takže tento algoritmus je horší než (1/2)-kompetitivní a je zcela jedno, jak se rozhoduje při rovnosti zlomků.

Každý druhý výskyt

Také lze vymyslet řešení podobné vzorovému, které se u každé barvy zvlášť snaží, abychom ji vybrali v polovině případů. Budeme si pamatovat značky: barvu označíme, pokud minule nebyla použitá, takže ji příště použít chceme. Na začátku není žádná barva označená. Kdykoliv nám soupeř nabídne dvojici a právě jedna z barev je označená, vybereme si tuto barvu a označíme místo ní druhou barvu z dvojice. Pokud jsou obě barvy označené nebo obě neoznačené, vybereme si tu, které máme více. A pokud se i to shoduje, zvolíme tu, která je dřív v abecedě.

Uvažme tuto posloupnost dvojic:

$$\underbrace{\mathbf{A}\mathrm{E},\ \mathbf{B}\mathrm{C},\ \mathbf{C}\mathrm{E},\ \mathrm{A}\mathbf{E}}_{n}.$$

Při prvním AE není žádná barva označená a všech barev máme stejně (totiž 0), takže vybereme A a označíme E. Podobně z BC vybereme B a označíme C. Dvojice CE má obě barvy označené a ani C ani E jsme si dosud nevzali, takže si vezmeme C a znovu označíme E. Z následujícího AE proto vybereme označené E a místo něj označíme A.

Celkově jsme za první 4 kroky nasbírali po jedné kartičce od barev A, B, C, E. Další čtveřice se tedy bude chovat úplně stejně. Tedy kromě toho, že barva A je v prvním kroku označená, ale to nevadí, protože ji preferujeme tak jako tak.

Po n čtveřicích tedy máme skóre $n + 2$, zatímco podvodník sbírající E-čka získá $3n + 2$. Algoritmus proto nemůže být lepší než (1/3)-kompetitivní.

Dodejme ještě, že hlavní potíž této strategie tkví v tom, že neumí rozumně ošetřit případy, kdy se značky obou barev shodují. Vzorové řešení se tomuto problému vyhne tak, že značky udržuje pro každou dvojici barev zvlášť, takže je vždy označena právě jedna barva z nabízené dvojice.