بازیهای تکرار شونده(با اطلاعات کامل)

Repeated Games(Perfect Information)

مقدمه

- بسیاری از تعاملاتی که روزمره درگیر آن هستیم، به صورت مداوم تکرار می‌شوند و ما با همان افراد مکرراً تعامل داریم. در بسیاری از این تعاملات می توانیم از اعتمادی که طرف مقابل بر ما کرده است سوء استفاده کنیم و سود ببریم. به طور مثال:
  - ✓ مراقبت از خانه همسایگان در زمانی که به مسافرت می‌روند یا در منزل نیستند.
  - ✓ قرض دادن پول به دوستانی که موقتاً نیازمند پول هستند.
- در مقیاس وسیع‌تر و بزرگتر می‌توانیم به موارد زیر اشاره کنیم:
  - ✓ حراج اسناد خزانه(که به صورت ماهیانه و بعضی وقت‌ها هفتگی برنامه ریزی می‌شوند)
  - ✓ رقابت Cournot
  - ✓ کارتل اوپک
- در ادامه قبل از تعریف رسمی و نشانه‌ها با یک مثال مفهوم این نوع بازی‌ها را ارائه خواهیم داد.

مثال۱ : معمای زندانی

- برای شروع بازی معمای زندانی را در نظر می‌گیریم.
- می دانیم تعادل نش منحصر به فرد این بازی آرایه استراتژی (D,D) می‌باشد.
- حال این بازی را تکرار شونده و بیشتر از یک مرتبه فرض کنید. فرض کنید در این بازی تکرارشونده بازیکن۱ با C بازی را آغاز کرده و در ادامه از استراتژی زیر استفاده کند:
  - ✓ اگر بازیکن دوم در مرحله قبل، C بازی کرده است بنابراین بازیکن۱ هم C بازی خواهد کرد.
  - ✓ اگر بازیکن دوم در مرحله قبل، D بازی کرده است بنابراین بازیکن۱ هم D بازی خواهد کرد. و تا انتها دیگر D بازی خواهد شد.

	C	D
C	2,2	0,3
D	3,0	1,1

تا این مرحله سود بازیکنان به صورت زیر خواهد بود:
$\{2,2,2,\dots,0\}$
$\{2,2,2,\dots,3\}$
مطابق استراتژی تعریف شده می‌توان گفت بازیکن۱ در ادامه D انتخاب می‌کند و مجموعه شامل زوج مرتب‌های اکشن‌های انتخابی به صورت زیر خواهد بود:
$\{(C,C),(C,C),(C,C),\dots,(C,D),(D,D),\dots\}$
و سود بازیکنان به صورت زیر خواهد بود:
$\{2,2,2,\dots,0,1,\dots\}$
$\{2,2,2,\dots,3,1,\dots\}$

توضیح مثال

- اگر بازیکن دوم نیز در هر مرتبه C را انتخاب کند، مجموعه شامل زوج مرتب‌های اکشن‌های انتخاب شده به صورت زیر خواهد بود:

$$\{(C,C),(C,C),(C,C),\dots\}$$

- و سود هر بازیکن مجموعه زیر خواهد بود:

$$\{2,2,2,\dots\}$$

- حال اگر جایی بازیکن دوم انحراف داشته باشد و به جای C اکشن D را انتخاب کند، مجموعه شامل زوج مرتب‌های اکشن‌های انتخاب شده به صورت زیر است:

$$\{(C,C),(C,C),(C,C),\dots,(C,D)\}$$

- تا این مرحله سود بازیکنان به صورت زیر خواهد بود:

$$\{2,2,2,\dots,0\}$$
$$\{2,2,2,\dots,3\}$$

- مطابق استراتژی تعریف شده می‌توان گفت بازیکن۱ در ادامه D انتخاب می‌کند و مجموعه شامل زوج مرتب‌های اکشن‌های انتخابی به صورت زیر خواهد بود:

$$\{(C,C),(C,C),(C,C),\dots,(C,D),(D,D),\dots\}$$

- و سود بازیکنان به صورت زیر خواهد بود:

$$\{2,2,2,\dots,0,1,\dots\}$$
$$\{2,2,2,\dots,3,1,\dots\}$$

سؤالاتی که مطرح می‌شوند

- این تکرارها محدود هستند یا غیر محدود؟
- در تحلیل مثال معمای زندانی دیدیم، اگر دو فرد به اندازه کافی صبور باشند و به یکدیگر اعتماد داشته باشند باشند می‌توانند از payoff(سود) ۲ در هر مرحله بهره‌مند شوند و (C,C) یکی از تعادل‌های نش باشد و در غیر این صورت پس از اولین انحراف در مرحله‌ای که انحراف انجام شده است، سود بازیکن دوم افزایش یافت ولی با توجه به استراتژی بازیکن اول در مراحل بعدی سود آن به یک کاهش پیدا کرد. پس اولین سؤال این است که میزان صبر چقدر باشد؟ به بیانی دیگر تا چه زمانی صبر کردن ارزش کافی دارد؟
- ما در فصل ۵ دیدیم که تعادل نش بازی های گسترده همیشه به طور شهودی جذاب نیستند، زیرا اقداماتی که آنها پس از تاریخچه های ناشی از تغییر اکشن‌ها انجام می‌دهند ممکن است بهینه نباشند. مفهوم تعادل کامل زیربازی، که مستلزم بهینه بودن اقدامات پس از هر سابقه ممکن است، نه تنها انتخاب‌هایی که در صورت پایبندی بازیکنان به استراتژی‌های خود انجام می‌شوند، ممکن است جذاب‌تر باشد. آیا جفت استراتژی که در آن هر بازیکن از استراتژی تنبیهی که توضیح داده شد، استفاده می‌کند، تعادل کاملی در بازی فرعی دارد؟ یعنی بهینه است که هر بازیکنی بازیکن دیگر را به خاطر تغییر اکشن تنبیه کند؟ اگر نه، آیا جفت استراتژی دیگری وجود دارد که از نتایج مطلوب پشتیبانی کند و آیا یک بازی فرعی تعادل کاملی دارد؟
- استراتژی تنبیهی که در بالا مورد بررسی قرار گرفت، نسبتاً شدید است. در تغییر دائمی به D در پاسخ به تغییر اکشن، جایی برای خطا باقی نمی گذارد. آیا تعادل نش یا تعادل کامل زیر بازی وجود دارد که در آن استراتژی بازیکنان تغییرات اکشن‌ها را با شدت کمتری مجازات و تنبیه کند(ضرر کمتری برساند)؟

تکرارشونده محدود یا نامحدود؟

- برای پاسخ به این سوال:
  - ✓ بازی های متناهی بازی هایی هستند که در آن هر دو بازیکن می دانند که بازی با تعداد دور مشخص (و متناهی) انجام می شود و پس از انجام چندین دور بازی به طور قطعی به پایان می رسد.به طور کلی، بازی های محدود را می توان با استقرا به عقب(backward induction) حل کرد. در این مورد تعادل نش همانند بازی تک مرحله‌ای خواهد بود. به همین دلیل در ادامه تمرکز ما بر بازی‌های تکرار شونده نامحدود می‌باشد.

پاسخ بقیه سؤالات طرح شده

- برای پاسخ به بقیه سؤالات ابتدا باید تابع سود را برای این گونه بازی ها تعریف کنیم و سپس با مفهومی به نام ضریب کاهش(discount) آشنا بشویم. تا انتهای این درس به تمام سؤالات مطرح شده پاسخ داده خواهد شد.
  - ✓ سود بازیکنان در یک بازی تکرار شونده را می‌توان ابتدا جمع سودهای حاصل شده تعریف نمود:

$$\boldsymbol{U_i = \sum_{t=0}^{\infty} \delta^t u_i(x_t)}$$

  - ✓ به پارامتر δ ضریب کاهش ارزش(discount) می‌نامند.
    - ▪ مفهوم پارامتر: فرض کنید در ایران زندگی می‌کنید و تورم روزانه وجود دارد بنابراین ارزش پول شما در لحظه تصمیم گیری بیشتر از ارزش پول شما در فردا می‌باشد یا حتی ساعاتی بعد به طور مثال اگر شما در روز شنبه ۱۰۰۰۰۰ تومان پول داشته باشید و $\delta = 0.8$ باشد یعنی پول شما روز یکشنبه ۸۰۰۰۰ تومان ارزش خواهد داشت. به بیان دیگر این پارامتر بیان می‌کند که پول شما با این نرخ سود در آینده نسبت به حال حاضر چه مقدار می‌تواند باشد.

تعریف رسمی بازی‌های تکرار شونده

- حال به تعریف رسمی بازیهای تکرار شونده می پردازیم:

✓ اگر $G=(N,\{A_i\}_{i\in N},\{u_i\}_{i\in N})$ یک بازی استراتژیک باشد که در آن N مجموعه بازیکن ها ، A_i مجموعه اکشن های بازیکن i و u_i تابع سود بازیکن i از نمایه اکشن های بازیکنان به اعداد حقیقی است. یک **بازی تکرار شونده نامحدود** $\boldsymbol{G}$ به ازای نرخ کاهش ارزش δ یک بازی توسیعی با اطلاعات کامل و حرکات همزمان عبارت است از :

  - یک مجموعه از بازیکنان.
  - یک مجموعه از دنباله نامتناهی اکشن ها (terminal history ها) به طوری که هیچ زیردنباله ای یک زیردنباله اکید (proper subhistory) از terminal history دیگر نباشد.
  - یک تابع که به هر proper subhistory ، یک مجموعه از بازیکن ها نسبت دهد (player function)
  - مجموعه اکشن های بازیکن i
  - ترتیب ترجیحات بازیکن i بر روی هر یک از terminal history ها ($\succsim_{i^*}$) (یا به طور معادل: تابع میانگین انتزاعی (discounted Average) که برای هر بازیکن به هر یک از terminal history ها یک عدد حقیقی نسبت میدهد.)

ادامه تعریف رسمی

- یک h، یک $terminal\ history$ خواهد بود، اگر و تنها اگر نامتناهی باشد. بعد از هر یک $(nonterminal\ history)$ هر بازیکن از مجموعه بازیکنان(N) یک اکشن $(action)$ در A_i انتخاب کند. بنابراین استراتژی بازیکن i یک تابع می‌باشد که یک اکشن $(action)$ در A_i را به هر دنباله متناهی از نتایج در G اختصاص می‌دهد.
- بررسی تعریف روی مثال معمای زندانی:

✓ $N = \{\text{زندانی}1, \text{زندانی}2\}$

✓ $TH = \{\{(C,C),(C,C),\dots\}.\ \{(C,C),(C,D),\dots\},\dots\}$

✓ $P(\{\dots\}) = \{\text{زندانی}1, \text{زندانی}2\}$

✓ $A_i = \{C,D\}$

✓ $(1-\delta)\sum_{t=0}^{\infty} \delta^t\, u_i(a^t)$

روابط ترجیحی

(preference relations)

تنزّل (Discounting)

- ما اکنون علاوه بر قابلیت تفکیک ضعیف، محدودیت هایی را بر روابط ترجیحی بازیکنان اعمال می کنیم. ما فرض می کنیم که رابطه ترجیحی بازیکن i $\succsim_{i^*}$ در بازی تکراری بر اساس یک تابع پرداخت(payoff) است که نشان دهنده رابطه ترجیحی $\succsim_i$ او در G است. ما فرض می کنیم که برقراری رابطه $(a^t) \succsim_{i^*} (b^t)$ فقط به رابطه بین توالی های مربوطه از با $u_i(a^t)$ و $u_i(b^t)$ در G بستگی دارد. ما سه شکل از روابط ترجیحی را در نظر می گیریم:

1. تنزل (Discounting): با توجه به معیاری تحت عنوان نرخ کاهش ارزش (discounted factor) که عددی بین صفر و یک است ($\delta \in (0,1)$) دنباله v_i^t که دنباله ای از اعداد حقیقی است ، حداقل به اندازه دنباله ی w_i^t خوب هست اگر و تنها اگر :

$$\sum_{t=0}^{\infty} \delta^t (v_i^t - w_i^t) \geq 0$$

هم ارزی payoff function ها تحت تنزیل (Equivalence of payoff functions under discounting)

- فرض کنید حداقل سه نتیجه ممکن وجود دارد. مجموع تنزیل‌شده بازده‌ها با تابع payoff ، u و نرخ کاهش ارزش δ ، نشان دهنده همان اولویت‌ها نسبت به جریان‌های بازدهی است که جمع تنزلی (Discounted Sum) آن با تابع بازده v و ضریب تخفیف δ باشد، اگر و تنها اگر $\alpha,\beta > 0$ وجود داشته باشند به گونه‌ای که $\forall x : u(x) = \alpha + \beta v(x)$.
  - اثبات : فرض کنیم دو دنباله اکشن پروفایل یکسان $(x^0,x^1,\dots)$ و $(y^0,y^1,\dots)$ داریم. اگر u همان $payoff\ function$ باشد آنگاه داریم :

$$\sum_{t=0}^{\infty}\delta^t u(x^t) = \sum_{t=0}^{\infty}\delta^t u(y^t)$$

حال اگر v یک تابع صعودی و Affine از تابع u باشد یعنی : $v(x) = \alpha + \beta u(x) \quad ; \alpha,\beta > 0$ آنگاه داریم :

$$\sum_{t=0}^{\infty}\delta^t v(x^t) = \sum_{t=0}^{\infty}\delta^t(\alpha + \beta u(x^t)) = \sum_{t=0}^{\infty}\delta^t \alpha + \beta\sum_{t=0}^{\infty}\delta^t u(x^t)$$

به طریق مشابه

$$\sum_{t=0}^{\infty}\delta^t v(y^t) = \sum_{t=0}^{\infty}\delta^t(\alpha + \beta u(y^t)) = \sum_{t=0}^{\infty}\delta^t \alpha + \beta\sum_{t=0}^{\infty}\delta^t u(y^t)$$

در نتیجه : $\sum_{t=0}^{\infty}\delta^t v(x^t) = \sum_{t=0}^{\infty}\delta^t v(y^t)$ ■

روابط ترجیحی در این درس

- در این تدریس منظور از تعریف روابط ترجیحی ، ترتیب اولویت ها براساس Discounting است.
- در اصل برای ارزش گذاری به هر یک از دنباله اکشن پروفایل ها در طول تکرار های متعدد بازی ، میتوان از رابطه جمع تنزلی (Discounted Sum) استفاده نمود که در آن نرخ کاهش ارزش $\delta \in (0,1)$ می‌باشد. به این ترتیب ، جمع تنزلی دنباله اکشن پروفایل های $(a^0,a^1,\dots)$ برای بازیکن i عبارت است از :

$$u_i(a^0) + \delta u_i(a^1) + \delta^2 u_i(a^2) + \dots = \sum_{t=0}^{\infty} \delta^t u_i(a^t)$$

- رابطه ترجیحات را می‌توان به صورت میانگین انتزاعی(discounted average) دنباله اکشن پروفایل های $(a^0,a^1,\dots)$ برای بازیکن i تعریف کرد که به صورت زیر است:

$$(1-\delta)\sum_{t=0}^{\infty} \delta^t u_i(a^t)$$

- میانگین انتزاعی در واقع همان جمع تنزلی است که در یک عدد $(1-\delta)$ ضرب شده ولی مزیت آن نسبت به جمع تنزلی این است که مقادیر بدست آمده از آن مستقیماً قابل مقایسه با سود یک دوره می‌باشد.

استراتژی

Strategy

استراتژی

- استراتژی بازیکن در یک بازی توسیعی، عمل او را پس از تمام historyهای ممکن مشخص می‌کند و پس از آن نوبت به حرکت او می‌رسد، از جمله historyهایی که با استراتژی او همخوانی ندارد.
- استراتژی‌های پرکاربرد در بازی‌های تکرار شونده عبارتند از:
  - ✓ Grim Trigger(خشم)
  - ✓ Limited punishment
  - ✓ Tit – for - Tat
- نمایش یک استراتژی دو راه کلی وجود دارد. در ادامه با بررسی استراتژی های فوق در مثال معمای زندانی با راه‌های نمایش استراتژی به طور کلی آشنا می‌شویم.

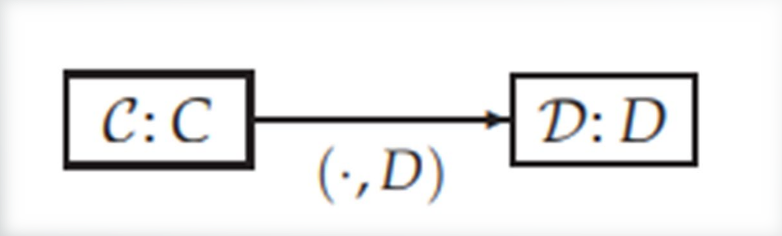
ماشین حالت برای نمایش استراتژی‌ها

- همانطور که در مقدمه این فصل بحث کردیم، دستاورد اصلی تئوری بازی های تکراری این است که وقتی افراد به طور مکرر با یکدیگر تعامل می کنند، بینشی در مورد ساختار رفتار به ما می دهد. در این بخش، ما زبانی را توسعه می‌دهیم که در آن ساختار تعادلی را که می‌یابیم به راحتی توصیف کنیم. ما با تعریف یک ماشین شروع می کنیم، که به عنوان انتزاعی از فرآیندی در نظر گرفته شده است که توسط آن بازیکن یک استراتژی را در یک بازی تکراری پیاده سازی می کند.
- یک ماشین (automaton) برای بازیکن i در یک بازی تکرار شونده نامحدود که در آن N مجموعه بازیکن هایش و A_i مجموعه اکشن های بازیکن i باشد، عبارت است از :
  - ✓ مجموعه وضعیت ها (state ها) : Q_i
  - ✓ وضعیت اولیه (initial state) : $q_i^0 \in Q_i$
  - ✓ تابعی که به هر استیت یک اکشن نسبت بدهد (output function) : $f_i: Q_i \rightarrow A_i$
  - ✓ تابعی که به هر زوج مرتب (state, action profile) یک استیت مقصد نسبت میدهد (transition function یا تابع انتقال) : $\tau_i: Q_i \times A \rightarrow Q_i$

ماشین حالت برای نمایش استراتژی‌ها

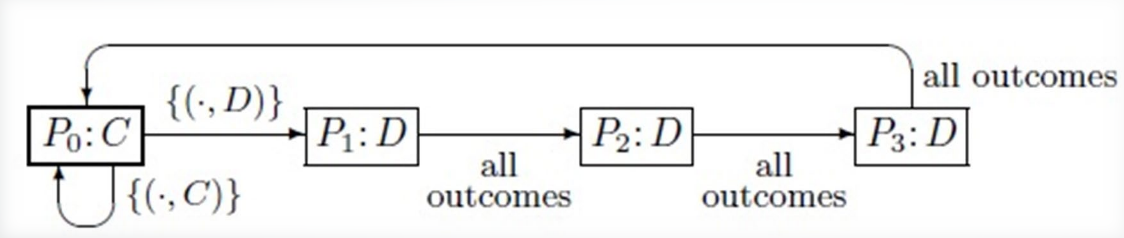
- مجموعه Q_i نامحدود است.
- نام stateها هیچ اهمیتی ندارند (اینکه ما یک state را "cooperative" می نامیم، به این معنی نیست که رفتار مرتبط با آن با نام آن مطابقت دارد).
- در اولین دور(مرتبه) ماشین (automaton) در حالت(state)، q_i^0 قرار دارد و ماشین(automaton) اکشن(action) $f_i(q_i^0)$ را انتخاب می‌کند. به همین ترتیب هرگاه ماشین در حالت q_i قرار دارد با توجه به state مورد نظر اکشن $f_i(q_i)$ را انتخاب می‌کند.
- تابع انتقال τ_i مشخص می‌کند که چطور ماشین از یک state به state دیگر برود: اگر ماشین در استیت q_i باشد و a همان action profile انتخاب شده باشد بنابراین استیتش به $\tau_i(q_i,a)$ تغییر می‌کند.
- توجه داشته باشید که ورودی تابع انتقال شامل وضعیت فعلی و نمایه تمام اقدامات فعلی بازیکنان است. طبیعی تر است که وضعیت فعلی و لیست اقدامات انتخاب شده توسط سایر بازیکنان را به عنوان ورودی در نظر بگیریم. این توصیف طبیعی یک «قاعده رفتار» یا «استراتژی» است به عنوان طرحی برای چگونگی رفتار در همه شرایط ممکن که با برنامه‌های فرد سازگار است. با این حال، از آنجایی که تعریف تئوری بازی مستلزم آن است که یک استراتژی یک عمل را برای تمام تاریخچه‌های ممکن، از جمله مواردی که با استراتژی خود بازیکن ناسازگار است، مشخص کند، باید عمل خود بازیکن را به عنوان ورودی تابع انتقال وارد کنیم.

تعریف استراتژی Grim Trigger و بررسی آن روی مثال معمای زندانی



$$s_i(a^1, \dots, a^t) = \begin{cases} C & if\ a_j^{\tau} = C\ for\ \tau = 1, 2, \dots, t \\ D & otherwise \end{cases}$$

تعریف استراتژی Limit Punishment و بررسی آن روی مثال معمای زندانی

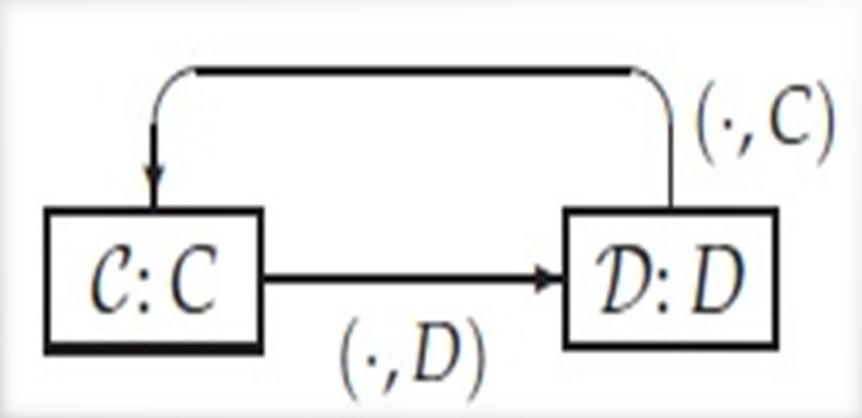


Limited Punishment state machine with $K = 3$

در واقع استراتژی Grim Trigger همان استراتژی Limited Punishment است با $K = \infty$.

استراتژی tit-for-tat و بررسی آن روی مثال معمای زندانی

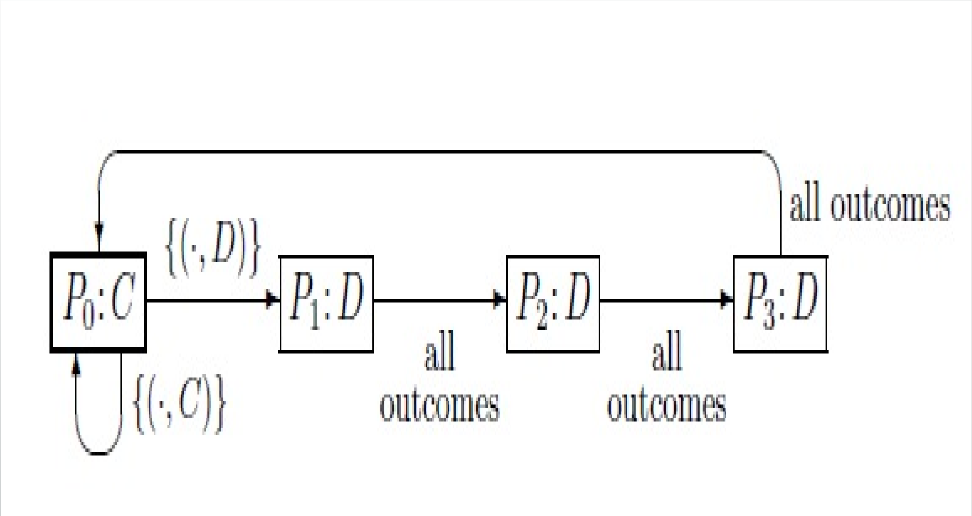
- در این استراتژی طول مجازات(punishment) به رفتار بازیکنی که مجازات می‌شود، بستگی دارد. اگر او همچنان به انتخاب D ادامه دهد، آنگاه بازیکن مجازات کننده نیز به این کار ادامه خواهد داد. اگر مجازات شده به C برگردد، tit-for-tat نیز به C برمی‌گردد. این استراتژی را می توان به این صورت توصیف کرد:
هر کاری که بازیکن دیگر در دوره قبلی انجام داده است.



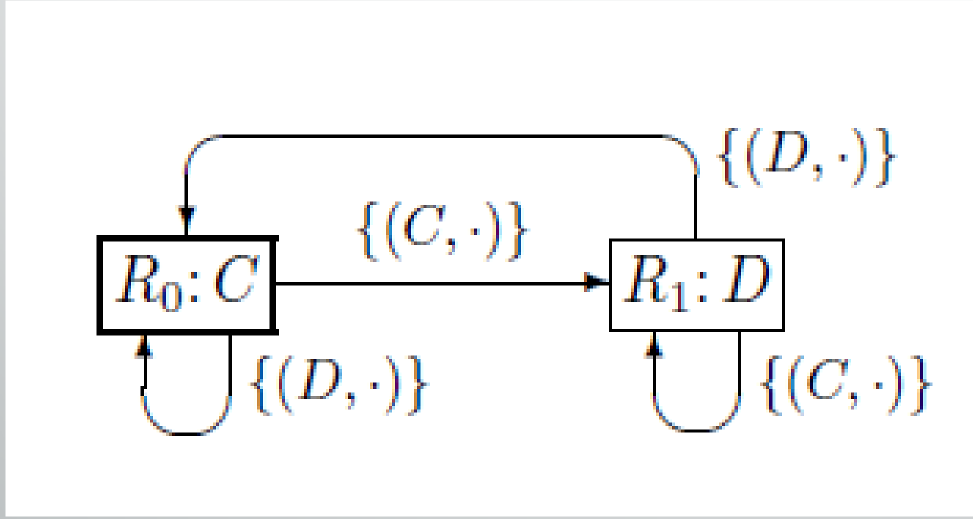
بررسی مثال اضافه

برای نشان دادن تکامل بازی در یک بازی تکراری زمانی که استراتژی هر بازیکن توسط یک ماشین انجام می شود

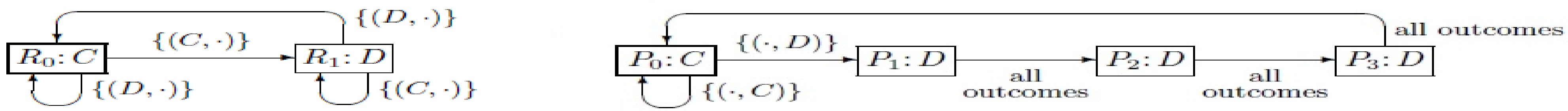
ماشین M_1: استراتژی بازیکن اول



ماشین M_2: استراتژی بازیکن دوم

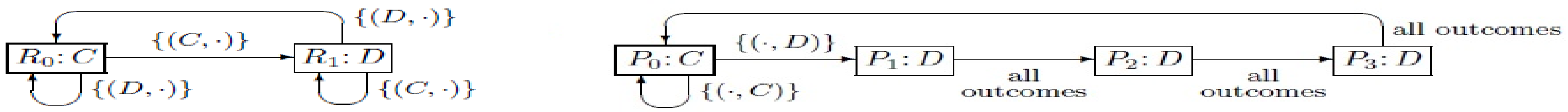


مرحله اول



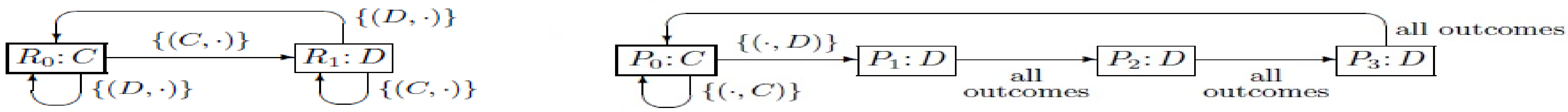
دوره	حالت فعلی ماشین M_1	حالت فعلی ماشین M_2	نتیجه	حالت بعدی ماشین M_1	حالت بعدی ماشین M_2	سودها
1	P_0	R_0	(C,C)	P_0	R_1	(2,2)

مرحله دوم



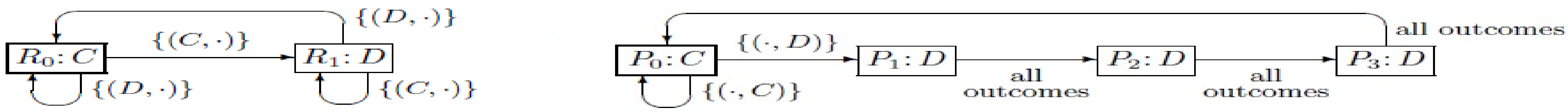
دوره	حالت فعلی ماشین M_1	حالت فعلی ماشین M_2	نتیجه	حالت بعدی ماشین M_1	حالت بعدی ماشین M_2	سودها
1	P_0	R_0	(C,C)	P_0	R_1	(2,2)
2	P_0	R_1	(C,D)	P_1	R_1	(0,3)

مرحله سوم



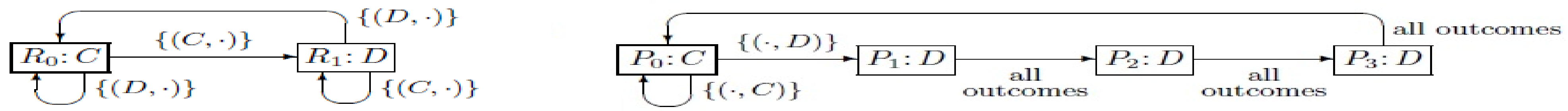
سودها	حالت بعدی ماشین M_2	حالت بعدی ماشین M_1	نتیجه	حالت فعلی ماشین M_2	حالت فعلی ماشین M_1	دوره
(2,2)	R_1	P_0	(C,C)	R_0	P_0	1
(0,3)	R_1	P_1	(C,D)	R_1	P_0	2
(1,1)	R_0	P_2	(D,D)	R_1	P_1	3

مرحله چهارم



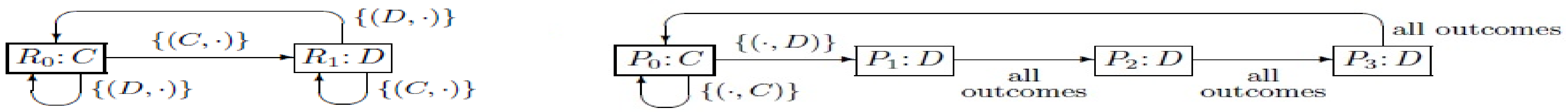
دوره	حالت فعلی ماشین M_1	حالت فعلی ماشین M_2	نتیجه	حالت بعدی ماشین M_1	حالت بعدی ماشین M_2	سودها
1	P_0	R_0	(C,C)	P_0	R_1	(2,2)
2	P_0	R_1	(C,D)	P_1	R_1	(0,3)
3	P_1	R_1	(D,D)	P_2	R_0	(1,1)
4	P_2	R_0	(D,C)	P_3	R_0	(3,0)

مرحله پنجم



دوره	حالت فعلی ماشین M_1	حالت فعلی ماشین M_2	نتیجه	حالت بعدی ماشین M_1	حالت بعدی ماشین M_2	سودها
1	P_0	R_0	(C,C)	P_0	R_1	(2,2)
2	P_0	R_1	(C,D)	P_1	R_1	(0,3)
3	P_1	R_1	(D,D)	P_2	R_0	(1,1)
4	P_2	R_0	(D,C)	P_3	R_0	(3,0)
5	P_3	R_0	(D,C)	P_0	R_0	(3,0)

مرحله ششم



سودها	حالت بعدی ماشین M_2	حالت بعدی ماشین M_1	نتیجه	حالت فعلی ماشین M_2	حالت فعلی ماشین M_1	دوره
(2,2)	R_1	P_0	(C,C)	R_0	P_0	1
(0,3)	R_1	P_1	(C,D)	R_1	P_0	2
(1,1)	R_0	P_2	(D,D)	R_1	P_1	3
(3,0)	R_0	P_3	(D,C)	R_0	P_2	4
(3,0)	R_0	P_0	(D,C)	R_0	P_3	5
(2,2)	R_1	P_0	(C,C)	R_0	P_0	6

نتایج مثال:

- این واقعیت که چرخه‌ها(حلقه‌ها) تولید می‌شوند، مختص به این مثال نیست: هر زمان که هر بازیکن از ماشینی با حالت‌های محدود استفاده می‌کند، در نهایت به یک چرخه(حلقه) می‌رسد، البته نه لزوماً در دوره اول . این از این واقعیت ناشی می شود که هر ماشین فقط اقدامات دوره قبل را به عنوان ورودی خود می گیرد.
- می‌توان نشان داد هر استراتژی در بازی‌های با تکرار نامحدود را نمی‌توان با یک ماشین(automaton) با تعداد استیت(حالت) محدود پیاده سازی کرد.

تعادل نش

Nash Equilibria

استراتژی Grim Trigger و بررسی تعادل نش در مثال معمای زندانی

اگر به مثال ابتدای درس رجوع کنیم و هر دو بازیکن با انجام اکشن های (C,C) به همکاری به یکدیگر ادامه دهند، مجموعه سود به صورت زیر خواهد بود:

$$\{2,2,2,\ldots\}$$

بنابراین می‌توان گفت سود هر کدام مطابق تابع میانگین انتزاعی به صورت زیر محاسبه می‌شود:

$$(1-\delta)(2+2\delta+2\delta^2+\cdots) \overset{0<\delta<1}{\triangleq} (1-\delta)\left(\frac{2}{1-\delta}\right)=2$$

ولی اگر بازیکن دوم در همان دور اول انحراف ایجاد کند، مجموعه سود به صورت زیر خواهد بود:

$$\{3,1,1,1,1,\ldots\}$$

استراتژی Grim Trigger و بررسی تعادل نش در مثال معمای زندانی

با توجه به مجموعه سود بدست آمده سود با استفاده از تابع میانگین انتزاعی به صورت زیر محاسبه می‌شود:

$$(1-\delta)(3+\delta+\delta^2+\delta^3+\cdots)=(1-\delta)\left(3+\frac{\delta}{1-\delta}\right)=3(1-\delta)+\delta=3-2\delta$$

حال به شرطی می‌توانیم انتظار همکاری داشته باشیم که payoff مربوط به همکاری(2) بهتر از payoff مربوط به انحراف($3-2\delta$) باشد، به صورت ریاضی یعنی:

$$2\geq 3-2\delta\Rightarrow\delta\geq\frac{1}{2}$$

بنابراین اگر $\delta\geq\frac{1}{2}$ باشد، آنگاه جفت استراتژی که در آن استراتژی هر بازیکن (grim trigger) می‌باشد، یک تعادل نش برای بازی بی‌نهایت تکرار برای معمای دو زندانی با جدول payoff معرفی شده در ابتدا تدریس می‌باشد.

استراتژی Limit Punishment و بررسی تعادل نش در مثال معمای زندانی

اگر یک بازی تکرار شونده که K مرحله تکرار میشود و استراتژی هر دو بازیکن هم Limited Punishment با K مرحله تنبیه باشد و و هر دو بازیکن با انجام اکشن های (C,C) به همکاری به یکدیگر ادامه دهند، مجموعه سود به صورت زیر خواهد بود:

$$\{2,2,2,\dots,2\}_k$$

بنابراین می‌توان گفت سود هر کدام مطابق تابع میانگین انتزاعی به صورت زیر محاسبه می‌شود:

$$(1-\delta)\left(2+2\delta+2\delta^2+\cdots+2\delta^k\right) \overset{0<\delta<1}{\cong} (1-\delta)\left(2\frac{1-\delta^{k+1}}{1-\delta}\right) = 2\left(1-\delta^{k+1}\right)$$

ولی اگر بازیکن دوم در همان دور اول انحراف ایجاد کند، مجموعه سود به صورت زیر خواهد بود:

$$\{3,1,1,\dots,1\}_k$$

استراتژی Limit Punishment و بررسی تعادل نش در مثال معمای زندانی

با توجه به مجموعه سود بدست آمده سود با استفاده از تابع میانگین انتزاعی به صورت زیر محاسبه می‌شود:

$$(1-\delta)\left(3+\delta+\delta^2+\delta^3+\cdots+\ \delta^k\right)=(1-\delta)\left(3+\frac{\delta(1-\delta^k)}{1-\delta}\right)=3\ -2\delta\ -\ \delta^{k+1}$$

حال به شرطی می‌توانیم انتظار همکاری داشته باشیم که payoff مربوط به همکاری $2(1-\delta^{k+1})$ بهتر از payoff مربوط به انحراف $(3\ -2\delta\ -\ \delta^{k+1})$ باشد، به صورت ریاضی یعنی:

$$2\left(1-\delta^{k+1}\right)\geq 3\ -2\delta\ -\ \delta^{k+1}\Rightarrow 1\ -2\delta+\ \delta^{k+1}\leq 0$$

- اگر $K\ =1$ باشد آنگاه $\delta\ =1 \Leftarrow$ یک دوره تنبیه تاثیری در جلوگیری از تخطی ندارد حتی اگر بازیکن کاملا صبور باشد. (در این حالت استراتژی TFT مناسب تر است)
- اگر $K\ =2$ باشد آنگاه $\delta\geq 0.62$
- اگر $K\ =3$ باشد آنگاه $\delta\geq 0.55$
- ...
- ...
- ...
- اگر $K\to\infty$ باشد آنگاه $\delta\to\frac{1}{2}$ (استراتژی Grim Trigger)

استراتژی tit-for-tat و بررسی تعادل نش در مثال معمای زندانی

فرض کنید دو زندانی در بازی کردن با استراتژی (C,C) با یکدیگر همکاری می‌کنند. بنابراین می‌توان گفت سود هر کدام مطابق تابع میانگین انتزاعی به صورت زیر محاسبه می‌شود:

$$(1-\delta)(2+2\delta+2\delta^2+\cdots) \overset{0<\delta<1}{\cong} (1-\delta)\left(\frac{2}{1-\delta}\right)=2$$

ولی اگر یک بازیکن انحراف ایجاد کند مرحله بعدی تنبیه می‌شود و در ادامه این روند تکرار می‌شود یعنی زوج مرتب‌های اکشن‌ها به صورت زیر خواهند بود:

$$\{(C,D),(D,C),(C,D),(D,C),\dots\}$$

استراتژی tit-for-tat و بررسی تعادل نش در مثال معمای زندانی

حال payoff ایجاد شده با استفاده از میانگین انتزاعی به صورت زیر خواهد بود:

$$(1-\delta)(3+0\delta+3\delta^2+0\delta^3+3\delta^4+\cdots)=(1-\delta)\left(\frac{3}{1-\delta^2}\right)=\frac{3}{1+\delta}$$

حال به شرطی می‌توانیم انتظار همکاری داشته باشیم که payoff مربوط به همکاری(2) بهتر از payoff مربوط به انحراف($\frac{3}{1+\delta}$) باشد، به صورت ریاضی یعنی:

$$2\geq\frac{3}{1+\delta}\Rightarrow\delta\geq\frac{1}{2}$$

بنابراین اگر$\frac{1}{2}\leq\delta$ باشد، آنگاه جفت استراتژی که در آن استراتژی هر بازیکن tit-for-tat می‌باشد، یک تعادل نش برای بازی بی‌نهایت تکرار برای معمای دو زندانی با جدول payoff معرفی شده در ابتدا تدریس می‌باشد.

میانگین سود انتزاعی تقریبی
Feasible discounted average payoff

- اگر **یک** دست بازی تکرار شود ، سود های زیر دست یافتنی هستند : $\{(2,2),(0,3),(3,0),(1,1)\}$
- اگر **دو** دست بازی تکرار شود با این خروجی $\{(C,C),(C,D)\}$ به این ترتیب :
  - سود بازیکن اول $\{2,0\}$ خواهد بود یعنی به طور متوسط در هر دوره $1 = \frac{2+0}{2}$ واحد سود می‌برد.
  - سود بازیکن دوم $\{2,3\}$ خواهد بود یعنی به طور متوسط در هر دوره $\frac{5}{2} = \frac{2+3}{2}$ واحد سود می‌برد.

✓ به این ترتیب $(1,\frac{5}{2})$ می‌تواند تقریبی از میانگین انتزاعی سود در بازی تکرار شونده باشد وقتی $\delta \cong 1$

- اگر **سه** دست بازی تکرار شود با این خروجی $\{(C,C),(D,C),(D,C)\}$ به این ترتیب :
  - سود بازیکن اول $\{2,3,3\}$ خواهد بود یعنی به طور متوسط در هر دوره $\frac{8}{3} = \frac{2+3+3}{3}$ واحد سود می‌برد.
  - سود بازیکن دوم $\{2,0,0\}$ خواهد بود یعنی به طور متوسط در هر دوره $\frac{2}{3} = \frac{2+0+0}{3}$ واحد سود می‌برد.

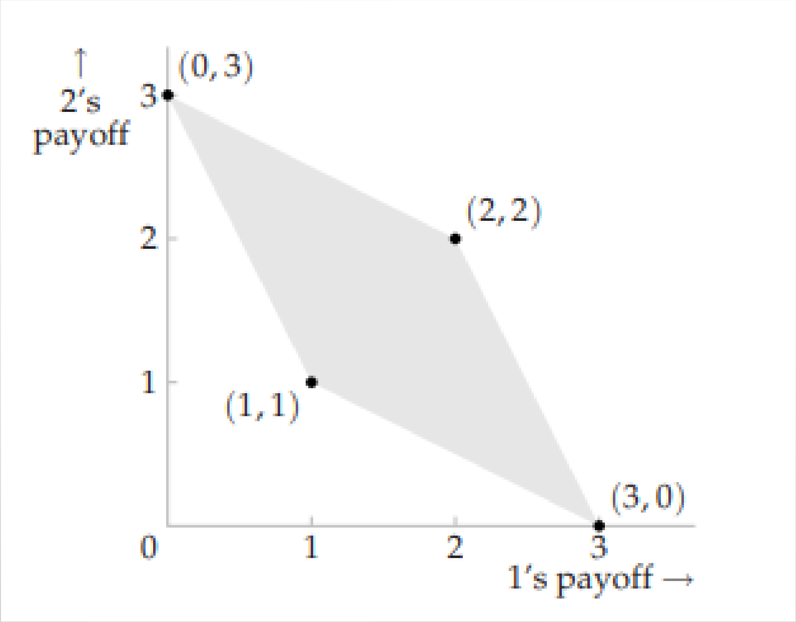
✓ به این ترتیب $(\frac{8}{3},\frac{2}{3})$ میتواند تقریبی از میانگین انتزاعی سود در بازی تکرار شونده باشد وقتی $\delta \cong 1$

میانگین سود انتزاعی تقریبی
Feasible discounted average payoff

- روند اسلاید قبل برای تکراری های بیشتر نیز قابل بحث هست. به همین ترتیب برای ترکیب بحث میانگین وزن دار و میانگین انتزاعی مفهوم جدیدی به نام مجموعه **feasible payoff profile** را تعریف میکنیم.
  - تعریف : در یک بازی استراتژیک به مجموعه تمام میانگین های وزن دار از نمایه سودها ، **مجموعه نمایه سودهای تقریبی** می‌گوییم.

(این مجموعه ، نشان دهنده سودهایی که به آن خواهیم رسید نیست بلکه تقریب خوبی از میانگین سود انتزاعی در بازی تکرار شونده است یعنی در بهترین حالات ما به سودهای تقریبی این مجموعه خیلی خیلی نزدیک می‌شویم.)

میانگین سود انتزاعی feasible در معمای زندانی



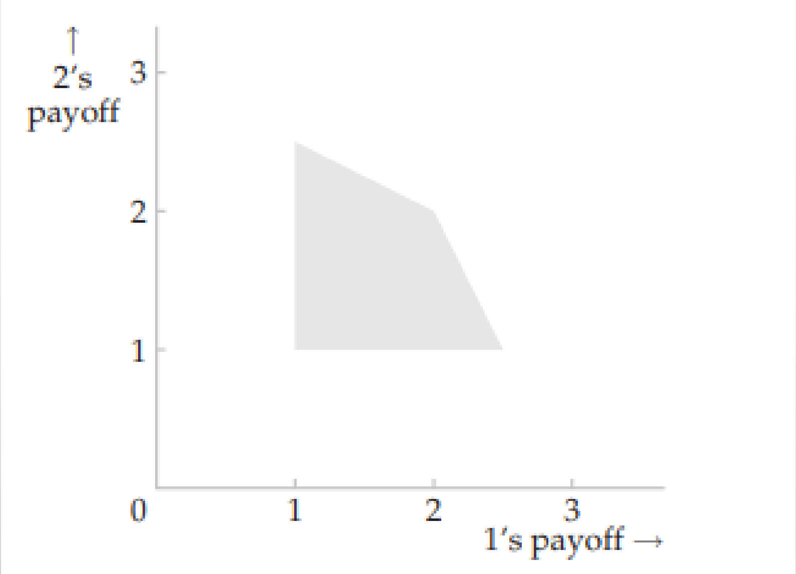
میانگین سود انتزاعی feasible
در نقاط تعادل نش در معمای زندانی

- در مباحث فصل دوم (بازی های استراتژیک با حرکات همزمان) دیدیم که نقطه‌ی (D,D) یک تعادل نش هست.
- در مباحث این فصل (بازی های تکرار شونده) سعی کردیم نشان دهیم که نقطه‌ی (C,C) نیز یک تعادل نش هست پس به این ترتیب مجموعه $\{(2,2),(1,1)\}$ تقریب خوبی از میانگین انتزاعی سود در این دو نقطه تعادل نش هست.
- چه میانگین سودهای تقریبی دیگری در نقاط تعادل نش در مسئله معمای زندانی وجود دارند؟

✓ به ازای هر نرخ کاهش ارزشی ، سود هر بازیکن در میانگین انتزاعی سودی که در تعادل نش می‌تواند تولید شود در بازی تکرار شونده نامحدود معمای زندانی ، **حداقل ۱** هست!

✓ اگر زوج مرتب (x_1,x_2) میانگین سود feasible برای بازی استراتژیک معمای زندانی باشد به نحوی که $x_1,x_2 \geq 1$ باشد آنگاه میانگین انتزاعی سود (y_1,y_2) به ازای $\delta \cong 1$ وجود دارد که بسیار بسیار به نقطه (x_1,x_2) نزدیک هست و همچنین یک تعادل نشی در بازی تکرارشونده نامحدود معمای زندانی وجود دارد که به میانگین سود انتزاعی (y_1,y_2) میرسد.

میانگین سود انتزاعی feasible
در نقاط تعادل نش در معمای زندانی



تعادل کامل زیربازی

Subgame Perfect Equilibria

ساختار تعادل کامل زیربازی تحت معیار تنزیل
(The Structure of Subgame Perfect Equilibria Under the Discounting Criterion)

- تعریف خاصیت یک-انحراف(One-deviation) در تعادل کامل زیربازی بازی های تکرارشونده نامحدود: نمایه استراتژی در یک بازی تکرارشونده نامحدود ، یک تعادل کامل زیربازی است اگر و تنها اگر هیچ بازیکنی نتواند با تغییر یک اکشن خود پس از هر تاریخچه ای، با توجه به اطلاعات کامل بازی ، سود بیشتری بدست بیاورد.

✓ واضح هست که در مثال معمای زندانی ، نمایه اکشن (D,D) یک تعادل نش و یک SPE هست.

در ادامه SPE را روی مثال معمای زندانی با استراتژی های ذکر شده ، بررسی می‌کنیم .

استراتژی Grim Trigger و بررسی تعادل کامل زیربازی در مثال معمای زندانی

اگر فرض کنیم بازیکن دوم در همان دست اول بازی تخطی میکند و نمایه اکشن (C,D) بیانگر دست اول بازی باشد پس در ادامه طبق استراتژی Grim Trigger دنباله terminal history به شکل $\{(C,D),(D,D),(D,D),\ldots\}$ می‌شود پس میانگین سود انتزاعی بازیکن اول در این زیربازی $3-2\delta$ خواهد شد.

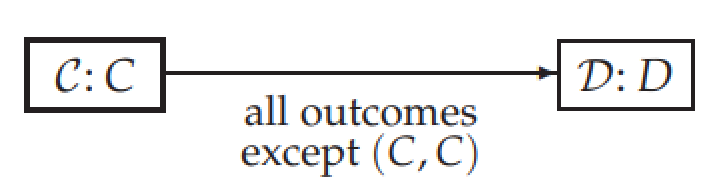
حال اگر فرض کنیم بازیکن یک ، از تنبیه بازیکن دوم به خاطر تخطیش خودداری کند و اکشن C را انتخاب کند و بازیکن دوم نیز همینطور یعنی در این زیربازی دنباله terminal history به شکل $\{(C,D),(C,C),(C,C),\ldots\}$ میشود پس میانگین سود انتزاعی بازیکن اول در این زیربازی 2 می‌شود و اگر $\delta > \frac{1}{2}$ باشد آنگاه بازیکن یک ترجیح می‌دهد بازیکن دوم را به دلیل یک تخطی تنبیه نکند تا در ادامه در زیربازی دوم قرار گیرد و سود بیشتری به دست آورد (یعنی از استراتژی Grim Trigger تخطی کنند به سود بیشتری می‌رسند در نتیجه **نمایه استراتژی (Grim Trigger , Grim Trigger) تعادل کامل زیربازی نیست!**)

استراتژی Grim Trigger و بررسی تعادل کامل زیربازی در مثال معمای زندانی

به علاوه این حالت که بازیکن یک به امید قرارگرفتن در زیربازی دوم ، از تنبیه خودداری کند خود موجبات سوءاستفاده از او را فراهم می‌کند یعنی دنباله terminal history به شکل $\{(C,D),(C,D),(C,D),\ldots\}$ می‌شود و بازیکن اول هیچ سودی نخواهد برد.

حالا اگر استیت ماشین استراتژی Grim Trigger را به شکل زیر اصلاح کنیم هر دو مشکل فوق حل میشود :

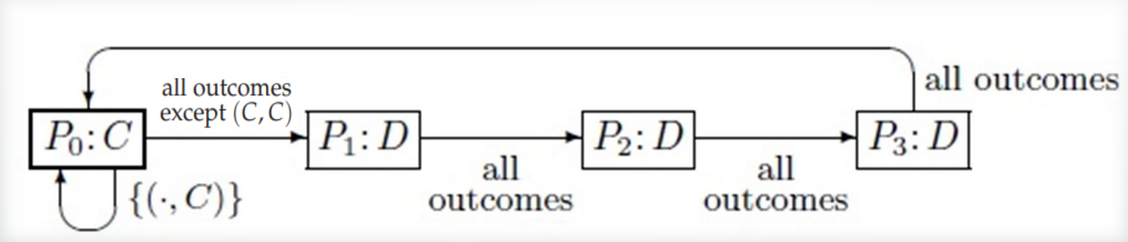
- ✓ در این استیت ماشین جدید ،حالتی که هر بازیکن، بازیکن دیگر را او را به خاطر تخطی تنبیه نکند ، را تنبیه میکند اضافه شده است !



استراتژی Limited Punishment و بررسی تعادل کامل زیربازی در مثال معمای زندانی

<u>نمایه استراتژی (Limited Punishment, Limited Punishment) تعادل کامل زیربازی نیست</u>! دقیقا طبق استدلال‌هایی که برای استراتژی Grim Trigger آوردیم.

از آنجایی که با اصلاح استیت ماشین Grim Trigger توانستیم مشکل را برطرف کنیم اینجا هم می‌توان به طریق مشابه عمل کرد. لذا خواهیم داشت:



A variant of the Limited Punishment state machine with $K = 3$

استراتژی Tit-For-Tat و بررسی تعادل کامل زیربازی در مثال معمای زندانی

در این استراتژی ، برای انتخاب نمایه اکشن در هر دست از بازی ، تنها به خروجی دست قبل بازی نیاز داریم. لذا برای بررسی SPE در این حالت کافیست چهار زیربازی که بعد از history هایی که ختم به نمایه اکشن های (C,C), (C,D), (D,C) و (D,D) میشود را در نظر بگیریم.

- در اسلایدهای قبل ضمن تعریف خاصیت یک-انحراف(One-deviation) بیان کردیم که یک نمایه استراتژی SPE است اگر و تنها اگر خاصیت One-deviation را ارضا کند.

- **زیربازی بعد از history ختم شده به نمایه اکشن $(\boldsymbol{C},\boldsymbol{C})$:**

نمایه استراتژی (TFT,TFT) در این وضعیت یک تعادل نش را نشان می‌دهد و همچنین اگر $\delta \geq \frac{1}{2}$ باشد ، استراتژی TFT بهترین پاسخ به استراتژی TFT حریف در این زیربازی هست.

- **زیربازی بعد از history ختم شده به نمایه اکشن $(\boldsymbol{C},\boldsymbol{D})$:**

پس در این زیربازی فرض کنید بازیکن دوم متعهد به استراتژی TFT باشد. حال اگر بازیکن اول نیز به این استراتژی متعهد باشد آنگاه خروجی زیربازی بین نمایه اکشن های (D,C)و(C,D) جابه‌جا می‌شود لذا میانگین انتزاعی سود بازیکن اول برابر است با :

$$(1-\delta)(3+3\delta^2+\ \dots)=(1-\delta)\left(\frac{3}{1-\delta^2}\right)=\frac{3}{1+\delta}$$

استراتژی Tit-For-Tat و بررسی تعادل کامل زیربازی در مثال معمای زندانی

اگر بازیکن اول در دست دوم بازی اکشن C را انتخاب کند آنگاه خروجی بازی $\{(C,C),(C,C),\dots\}$ میشود لذا میانگین انتزاعی سود بازیکن اول برابر است با 2 . در نتیجه برای اینکه نمایه استراتژی (TFT,TFT) یک SPE باشد باید :

$$\frac{3}{1+\delta} \geq 2 \Rightarrow \delta \leq \frac{1}{2}.$$

- **زیربازی بعد از history ختم شده به نمایه اکشن (D,C) :**

پس در این زیربازی فرض کنید بازیکن دوم متعهد به استراتژی TFT باشد. حال اگر بازیکن اول نیز به این استراتژی متعهد باشد آنگاه خروجی زیربازی بین نمایه اکشن های (C,D)و(D,C) جابه‌جا می‌شود لذا میانگین انتزاعی سود بازیکن اول برابر است با :

$$(1-\delta)(3\delta+3\delta^3+\ \dots) = (1-\delta)\left(\frac{3\delta}{1-\delta^2}\right) = \frac{3\delta}{1+\delta}$$

اگر بازیکن اول در دست دوم بازی اکشن D را انتخاب کند آنگاه خروجی زیربازی $\{(D,D),(D,D),\dots\}$ میشود لذا میانگین انتزاعی سود بازیکن اول برابر است با 1 . درنتیجه برای اینکه نمایه استراتژی (TFT,TFT) یک SPE باشد باید :

$$\frac{3\delta}{1+\delta} \geq 1 \Rightarrow \delta \geq \frac{1}{2}.$$

استراتژی Tit-For-Tat و بررسی تعادل کامل زیربازی در مثال معمای زندانی

- زیربازی بعد از history ختم شده به نمایه اکشن $(\boldsymbol{D},\boldsymbol{D})$:

پس در این زیربازی فرض کنید بازیکن دوم متعهد به استراتژی TFT باشد. حال اگر بازیکن اول نیز به این استراتژی متعهد باشد آنگاه میانگین انتزاعی سود بازیکن اول برابر است با 1 .

اگر بازیکن اول در دست دوم بازی اکشن C را انتخاب کند آنگاه خروجی زیربازی بین نمایه اکشن های (C,D)و(D,C) جابه‌جا می‌شود لذا میانگین انتزاعی سود بازیکن اول برابر است با :

$$(1-\delta)(3\delta+3\delta^3+\ \dots)=(1-\delta)\left(\frac{3\delta}{1-\delta^2}\right)=\frac{3\delta}{1+\delta}$$

درنتیجه برای اینکه نمایه استراتژی (TFT,TFT) یک SPE باشد باید :

$$\frac{3\delta}{1+\delta}\leq 1 \Rightarrow\ \delta\leq\frac{1}{2}\ .$$

✓ **نتیجه گیری : نمایه استراتژی (TFT, TFT) تعادل کامل زیربازی هست اگر و تنها اگر $\boldsymbol{\delta=\frac{1}{2}}$!**

برای علاقه مندان

در طول تدریس از هر دو منبع اصلی و کمکی بهره بردیم و مطالبی بیان نمودیم. به علت ضیق وقت و سنگینی مطالب به اسم برخی از موضوعات تنها اشاره شد و مفهمو کلی از آن بیان شد. در ادامه اسلایدهایی برای علاقه مندان آورده شده است که مطالب مرتبط **فصل ۱۵ مرجع اصلی و فصل ۸ مرجع کمکی** را پوشش می‌دهند که به اسامی آن‌ها و مفاهیمشان در طول تدریس اشاراتی شده است.

دو نوع دیگر از روابط ترجیحی

Limit of means

- ترجیحات با تخفیف، دوره ها را به گونه ای متفاوت در نظر می گیرند: ارزش سود معین با گذشت زمان کاهش می یابد. اکنون دو معیار جایگزین را مشخص می کنیم که همه دوره ها را به صورت متقارن در نظر می گیرند. در یک معیار، یک بازیکن دنباله ای v_i^t از بازده را اساساً با میانگین محدود کننده آن ارزیابی می کند($\lim_{T\to\infty}\frac{\sum_{t=1}^{\infty} v_i^t}{T}$). با این حال، این حد همیشه وجود ندارد (متوسط بازده در دوره‌های t ممکن است به طور مداوم با افزایش t در نوسان باشد). معیاری که در مورد آن بحث می کنیم به صورت زیر تعریف می شود. (به راحتی می توان این معیار را بر اساس رابطه ترجیحی دقیق تعریف کرد.)

✓ **Limit of means: The sequence $v_i^t$of real numbers is at least as good as the sequence w_i^t if and only if $\boldsymbol{lim\inf\frac{\sum_{t=1}^{T}(v_i^t-w_i^t)}{T}>0}$.**

Overtaking

- تحت معیار تنزیل، تغییر در بازده در یک واحد دوره می تواند مهم باشد، در حالی که تحت معیار حد میانگین، تفاوت های بازده در هر تعداد محدودی از دوره ها مهم نیست. بازیکنی که اولویت‌هایش محدودیت امکانات را برآورده می‌کند، آماده است هر باختی را در اولین تعداد محدود دوره‌ها قربانی کند تا جریان بازدهی را که در نهایت به دست می‌آورد افزایش دهد. به عنوان مثال، جریان (0,...,0,2,...2) از بازده با معیار حد میانگین نسبت به جریان ثابت (1,1,...) مستقل از شاخص ترجیح داده می شود. دوره ای که بازیکن برای اولین بار در استریم اول ۲ می گیرد. در نگاه اول ممکن است این ویژگی عجیب به نظر برسد. با این حال، فکر کردن به موقعیت‌هایی که در آن تصمیم‌گیرندگان بر بلندمدت به قیمت کوتاه‌مدت تأکید می‌کنند، (مبارزات ناسیونالیستی را در نظر بگیرید). اکنون معیاری را معرفی می‌کنیم که با تمام دوره‌ها به طور متقارن برخورد می‌کند و بر بلندمدت تأکید دارد، اما در عین حال نسبت به تغییر بازده در یک دوره واحد حساس است. (دوباره، ما معیار را بر اساس رابطه ترجیحی دقیق تعریف می کنیم.)
  - ✓ **Overtaking: The sequence $v_i^t$of real numbers is at least as good as the sequence w_i^t if and only if $\boldsymbol{lim}\inf \sum_{t=1}^{T}(v_i^t - w_i^t) > 0$.**

Nash folk Theorems

مقدمه

- ما اکنون مجموعه ای از نتایج تعادل نش یک بازی بی نهایت تکراری را مطالعه می کنیم. نشان می‌دهیم که این مجموعه شامل نتایجی است که تکرار تعادل‌های نش بازی سازنده نیستند. برای چنین نتیجه ای، هر بازیکن باید با "تنبیه شدن" از انحراف منصرف شود. چنین مجازاتی ممکن است اشکال مختلفی داشته باشد. یک احتمال این است که هر بازیکن از یک "استراتژی ماشه(trigger strategy)" استفاده کند: هر انحرافی باعث می شود او یک عمل تنبیهی را انجام دهد که برای همیشه ادامه دارد. در تعادل هایی که در این بخش بررسی می کنیم، هر بازیکن از چنین استراتژی استفاده می کند.
- Let $G = < N, (A_i), (\gtrsim_i) >$ be a strategic game and for each $i \in N$ let u_i be a payoff function that represents $\gtrsim_i$. Recall that we dene a feasible payoff profile of $< N, (A_i), (u_i) >$ to be a convex combination of $\sum_{\alpha \in A} \alpha_\alpha\, u(a)$ for which the coefficeints α_α are rational. Let $\mathrm{w} = \sum_{\alpha \in A} \alpha_\alpha\, u(a)$ be a such profile and suppose that $\alpha_\alpha = \frac{\beta_\alpha}{\gamma}$ for each $\mathrm{a} \in A$, where every β_α is an integer and $\gamma = \sum_{\alpha \in A} \beta_\alpha$. Then the sequence of outcomes in the repeated game that consists of an indefinite repetition of a cycle of length γ in which $\mathrm{a} \in A$ is played for β_α periods yields an average payoff profile over the cycle, and hence in the entire repeated game, of w.

Min-max

- Dene player i's minmax payoff in G, henceforth denoted v_i, to be the lowest payoff that the other players can force upon player i:

$$v_i = \min_{a_{-i} \in A_{-i}} \max_{a_i \in A_i} u_i(a_{-i}, a_i)$$

یک گزاره

- در ادامه به بیان یک گزاره خواهیم پرداخت.
- Every Nash equilibrium payoff profile of the limit of means infinitely repeated game of $\mathrm{G} =< N, (A_i), (u_i) >$ is an enforceable payoff profile of G. The same is true, for any $\delta \in (0,1)$, of every Nash equilibrium payoff profile of δ-discounted infinitely repeated game of G.

Nash folk theorems

- Nash folk theorem for the discounting criterion: Let w be a strictly enforceable feasible payoff profile of $\mathrm{G} =< N, (A_i), (u_i) >$. For all $\epsilon > 0$ there exists $\underline{\delta}$ sucj that if $\delta > \underline{\delta}$ then the δ-discounted infinitely repeated game of G has a Nah equilibrium whose payoff profile w' satisfies $|w' - w| > \epsilon$.
- Nash folk theorem for the limit of means criterion: Every feasible enforceable payoff profile of $\mathrm{G} =< N, (A_i), (u_i) >$ is a Nash equilibrium payoff profile of the limit of means infinitely repeated game of G.

Perfect Folk Theorems

- Perfect folk theorem for the discounting criterion: Let a^* be a strictly enforceable outcome of $\mathrm{G} =< N, (A_i), (u_i) >$. Assume that there is a collection $((a_i))_{i \in N}$ of strictly enforceable outcomes of G such that for every player $i \in N$ we have $a^* \succ_i a_i$ and $a(j) \succ_i a(i)$ for all j ∈ $N \backslash \{i\}$. Then there exists $\underline{\delta} < 1$ such that for all $\delta > \underline{\delta}$ there is a subgame perfect equilibrium of the δ-discounted infinitely repeated game of G that generates the path (a^t) in which $a^t = a^*$ for all t.
- Perfect folk theorem for the limit of means criterion: Every feasible strictly enforceable payoff profile of $\mathrm{G} =< N, (A_i), (u_i) >$ is a subgame perfect equilibrium payoff profile of the limit of means infinitely repeated game of G.
- Perfect folk theorem for the overtaking criterion: For any strictly enforceable outcome a^* of $\mathrm{G} =< N, (A_i), (u_i) >$ there is a subgame perfect equilibrium of the overtaking infinitely repeated game of G that generates the path (a^t) in which $a^t = a^*$ for all t.

بازی های تکرار شونده محدود

Finitely Repeated Games

تعریف

- For any positive integer T a T-period finitely repeated game of the strategic game $< N, (A_i), (\succsim_i) >$ be a strategic game is an extensive games with perfect information that satisfies the condition in definition infinitely repeated games. We restrict attention to the case in which the preference relation $\succsim_{i^*}$ of each player i in the nitely repeated game is represented by the function $\sum_{t=1}^{T} \frac{u_i(a^t)}{T}$, where u_i is a payoff function that represents i's preferences in the constituent game. We refer to this game as the **T-period repeated game of** $< N, (A_i), (u_i) >$.

تعادل نش

- استدلال شهودی که folk theorems را برای بازی‌های بی‌نهایت تکرار(infinitely repeated games) می‌کند این است که یک نتیجه مطلوب دوجانبه را می‌توان توسط یک آرایش اجتماعی پایدار پشتیبانی کرد که در آن بازیکن با این تهدید که در صورت انجام این کار «مجازات» خواهد شد، از انحراف منصرف می‌شود. همین استدلال، با تغییرات، در مورد دسته بزرگی از بازی‌های به‌طور محدود(finitely repeated games) اعمال می‌شود. نیاز به اصلاح ریشه در این واقعیت دارد که نتیجه در آخرین دوره تعادل نش هر بازی به طور محدود باید تعادل نش بازی سازنده باشد، واقعیتی که بر بقیه بازی تأثیر می‌گذارد. این تأثیر گذاری در حالت خاصی که در آن payoff هر بازیکن در هر تعادل نش از بازی سازنده برابر است با min max payoff او (مانند مثال معمای زندانی) طولانی ترین است. در این مورد، استدلال شهودی پشت folk theorems شکست می خورد: نتیجه در هر دوره باید تعادل نش از بازی سازنده باشد، زیرا اگر دوره ای وجود داشته باشد که در آن نتیجه چنین تعادلی وجود نداشته باشد، در آخرین دوره برخی از بازیکنان می توانند بدون مجازات منحرف شوند.

یک گزاره

- If the payoff prole in every Nash equilibrium of the strategic game G is the profile (v_i) of min-max payoffs in G then for any value of T the outcome $(a^1, a^2, \dots, a^T)$ of every Nash equilibrium of the T-period repeated game of G has the property that a^t is a Nash equilibrium of G for all $t = 1,2, \dots, T$.

Nash folk theorem for finitely repeated games

- If $\mathrm{G} =< N, (A_i), (u_i) >$ has a Nash equilibrium $\hat{a}$ in which the payoff of every player i exceeds his min-max payoff v_i then for any strictly enforceable outcome a^* of G and any $\epsilon > 0$ there exists an integer T^* such that if $\mathrm{T} > T^*$ the T-period repeated game of G has a Nash equilibrium in which the payoff of each player i is within ϵ of $u_i(a^*)$.

Subgame Perfect Equilibrium

- در هر تعادل کامل زیربازی از یک بازی تکرار شونده با تکرار متناهی(finitely repeated games)، نتیجه در آخرین دوره پس از هر تاریخچه (نه فقط پس از historyای که در صورت پایبندی هر بازیکن به استراتژی خود رخ می دهد) یک تعادل نش از بازی سازنده است. بنابراین مجازات تعبیه شده در راهبردهای مورد استفاده برای اثبات Nash folk theorem for finitely repeated games با تعادل کامل زیربازی سازگار نیست. در واقع، اگر بازی سازنده دارای یک تعادل منحصر به فرد نش باشد، هیچ مجازاتی ممکن نیست.

یک گزاره

- If the strategic game G has a unique Nash equilibrium payoff profile then for any value of T the action profile chosen after any history in any subgame perfect equilibrium of the T-period repeated game of G is a Nash equilibrium of G.

Perfect folk theorem for finitely repeated games

- Let a^* be a strictly enforceable outcome of $\mathrm{G} =< N, (A_i), (u_i) >$. Assume that (i) for each $i \in N$ there are two Nash equilibria of G that differ in the payoff player (i) and (ii) there is a collection $(a(i))_{i \in N}$ of strictly enforceable outcomes of G such that for every player $i \in N$ we have $a^* \succ_i a(i)$ and $a(j) \succ_i a(i)$ for all j $\in N \backslash \{i\}$. Then for any $\epsilon > 0$ there exists an integer T^* such that if $\mathrm{T} > T^*$ the T-period repeated game of G has a subgame perfect equilibrium in which the payoff each player i is within ϵ of $u_i(a^*)$.

Any Question?

- مریم سعیدمهر
- محمدرضا مزروعی
- علیرضا صالحی حسین آبادی

✓ maryamsaeedmehr@gmail.com

✓ mazrouee78@gmail.com

✓ ARSH9731@yahoo.com