تاب‌آوری راهبردی تحت عدم‌قطعیت

تحلیل بازی‌نظری و پیاده‌سازی در موتور SEE

سهیل سیاح ورگ
۴۰۲۱۱۱۲۳۷ — دانشکده‌ی مهندسی کامپیوتر

درس نظریه بازی‌ها — نیم‌سال دوم ۱۴۰۴–۱۴۰۵
استاد درس: دکتر حسین قربان

اگر ادامه‌ی تقابل برای هر دو طرف پرهزینه است،
چرا تقابل ادامه پیدا می‌کند؟

پاسخ یک‌خطی این گزارش:
چون آستانه‌ی ادامه پایین است — و ادامه دادن، خودش آن آستانه را پایین‌تر می‌آورد.

سه چیز را نشان می‌دهیم: (۱) این آستانه چیست و عددش چقدر است؛ (۲) چرا مسئله‌ی باور در این بازی دوبعدی است؛ (۳) چرا قوی‌ترین سیگنال بهترین نیست.

مدل در یک اسلاید

$$\Gamma = \langle N, \Theta, p, E, T, H, M, A, \Sigma, D, S, g, G, \varphi, B, X, \mu, \tau, u \rangle$$

- نوع خصوصی $\theta_i = (\bar{\rho}_i, \bar{m}_i)$: عزم پایه و ظرفیت مدیریت هزینه؛ ثابت و مشاهده‌نشدنی.
- تاب‌آوری e_i^t: حالت نهفته، با $e^{t+1} = e^t - g^t + \xi$؛ رسیدن به صفر $\Rightarrow$ خروج اجباری.
- اقدام $a_i^t = (\sigma_i^t, d_i^t)$، $\sigma \in \{0, .25, .5, .75, 1\}$.
- هزینه‌ی مؤثر $g_i = \dfrac{c_i + \alpha_i K_i + r_i}{\bar{m}_i(0.35 + 0.65\hat{e}_i)}$؛ K = انبار تعهد عمومی.
- دو ارزش پایانی: $B_i = 175(0.5 + 0.3\hat{e}_i + 0.2\bar{\rho}_i)$ و $X_i = x_0 - 5K_i - 8\tilde{\sigma}_j + 20M$.

اتحاد کلیدی

$G_i(t) \approx e_i^0 - e_i^t$ — «هزینه‌ی انباشته» و «تاب‌آوری خرج‌شده» یک چیزند.

جای این بازی در نقشه‌ی درس

- بازی زمان‌بندی (جلسه‌ی ۱۱) — از نوع فرسایشی، نه پیش‌دستی: بیشتر ماندنِ من تاب‌آوری حریف را می‌سوزاند، پس انگیزه دیرتر رفتن است.
- نسخه‌ی بیزی با نوع پیوسته (جلسه‌ی ۷) — استراتژی تعادلی نسبت به نوع صعودی است؛ نسخه‌ی گسسته‌اش همان ساختار آستانه‌ای ماست.
- سیگنال‌دهی و **PBE** (جلسه‌ی ۱۲) — مثال اسپنس فرض می‌کرد هزینه‌ی سیگنال برای نوع ضعیف بیشتر است؛ این همان شرط تک‌تقاطعی روی $\varphi_i(\sigma)$ است.
- استراتژی رفتاری (جلسه‌ی ۱۰) — سیاست عامل دقیقاً یک توزیع روی اکشن‌ها در هر مجموعه‌ی اطلاعاتی است.

ولی سه جا از مدل‌های درس فراتر می‌رود

سیگنال‌دهی دوطرفه است (هر دو هم فرستنده‌اند هم گیرنده)؛ سیگنال پنج‌سطحی است نه دودویی؛ و ارزش خروج X درون‌زاست.

هزینه‌های گذشته ریخته‌اند، پس مقایسه آینده‌نگر است:

$$\boxed{q_i^t \;\geq\; q^* = \frac{g_i + \varphi_i(\sigma_i) - \Delta X_i}{B_i - X_i}}$$

- صورت: سوخت + بهای سیگنال + فرسایش گزینه‌ی خروج $(-\Delta X)$.
- مخرج: فاصله‌ی جایزه تا گزینه‌ی بیرونی.

حالا عددهای واقعی این جهان:

$$B \in [87.5,\ 185.5], \qquad X \in [-2,\ 48], \qquad g \approx 1.1\text{–}1.3$$

$$\Rightarrow \qquad q^* \approx \frac{1.12}{150 - 26} \approx \mathbf{0.009}$$

یافته‌ی ۱

خروج داوطلبانه تقریباً هیچ‌وقت بهینه نیست. خروج عقلانی همان خروج اجباری است.

حالت e_j^t پنهان است، ولی قانونش عمومی است:

$$g_j^t = \mathrm{ramp}(t) \cdot \frac{c_{0j} + \kappa_j \sigma_i + \alpha_j K_j + \eta \sigma_i \sigma_j}{\bar{m}_j \left(0.35 + 0.65\, e_j / e_j^0\right)}$$

هر چیزی جز $\left(\bar{m}_j, e_j^0\right)$ عمومی است $\Rightarrow$ پسین روی مسیر = پسین روی یک نوع دوبعدی.

سه کانال به‌روزرسانی:

۱. بقا: حریف هنوز نرفته $\Rightarrow$ هر نوعی که باید مرده بود حذف می‌شود. دقیق، بدون نیاز به مدل استراتژی حریف.
۲. تک‌تقاطعی: موضع بلندِ انتخاب‌شده شاهد قوی بودن است.
۳. پیشین منتشرشده (نامتقارن).

فیلتر واقعاً کار می‌کند

کمیت	بایاس	MAE	RMSE	MAE پیش‌بینی ثابت
$\hat{e}_j$ (تاب‌آوری پنهان حریف)	$+0.0033$	**0.0511**	0.0835	0.1884
$\bar{\rho}_j$ (عزم پایه)	-0.0154	0.1522	0.1836	—

کاهش خطا: **72.9%**

و یک فرمول بسته برای زمان فروپاشی که از همین قانون درمی‌آید:

$$T_{\text{فروپاشی}} = \frac{e^0\,\bar{m}\,(0.35\,x + 0.325\,x^2)}{R}, \qquad x = \hat{e}$$

T نسبت به $\bar{m}$ خطی، نسبت به فشار خام وارون، و نسبت به تاب‌آوری باقی‌مانده محدب — همان «شکنندگی فزاینده».

تعارض اعتبار–انعطاف از سه کانال می‌گذرد: انبار تعهد ($-5K$)، تحقیر ($-8\tilde{\sigma}_j$)، و قفل نوبت ($\sigma \geq 0.75 \Rightarrow$ مجبور به HOLD).

یافته‌ی ۲ — تخفیفی که موتور به تشدید می‌دهد

موتور φ را فقط وقتی سیگنال تازه تولید شود می‌گیرد؛ در HOLD هیچ هزینه‌ای گرفته نمی‌شود. و عبور از 0.75 دقیقاً همان چیزی است که شما را HOLD می‌کند.

$$\varphi(0.50) = 0.25\kappa \quad \text{در برابر} \quad \tfrac{1}{2}\varphi(0.75) = 0.281\kappa$$

هزینه تقریباً برابر، ولی فشار وارد بر حریف به‌مراتب بیشتر.

$\Rightarrow$ تابع سود نسبت به σ یکنوا نیست؛ $\sigma = 0.5$ بدترین گزینه است.

و عامل آموزش‌دیده مستقلاً همین را کشف می‌کند

σ	0.00	0.25	0.50	0.75	1.00
سهم از مراحل	57.7%	29.8%	**1.1**%	11.3%	0.2%
φ میانگین	0.000	0.146	0.595	1.389	2.416

یک حفره دقیقاً روی $\sigma = 0.5$: از 29.8% به 1.1% سقوط و بعد دوباره به 11.3% بالا. توزیع یکنوا نیست.

و بررسی مستقل استراتژی‌های ساده همین ترتیب را می‌دهد:

$$\sigma = 0.75\ (47.6)\ >\ \sigma = 1.0\ (34.4)\ >\ \sigma = 0\ (11.7)\ >\ \sigma = 0.5\ (-10.6)$$

چهار گزاره، چهار آماره

گزاره	آماره	نتیجه
۱. دروازه‌ی پایین	میانه‌ی q^*	**0.015**
	نرخ خروج داوطلبانه	0.9% مراحل
	$\Pr(\text{خروج})$ بر حسب چارک $\bar{\rho}$	$0.36 \to 0.10$
	مطلوبیت بر حسب چارک $\bar{\rho}$	$54 \to 90$
۲. سیگنال درونی	سهم $\sigma = 0/0.25/0.5/0.75$	82.2%/0.2%/**0.2**%/17.0%
۳. خروج ساختنی	$\Pr(\text{خروج} \mid M{=}1)/\Pr(\cdot \mid M{=}0)$	**3.4**×
۴. گیر افتادن در تعهد خود	$\Pr(\text{خروج})$ بر حسب K	$0.011 \to$ **0.0006**

هر چهار گزاره تایید می‌شوند. گزاره‌های ۲ تا ۴ درباره‌ی ساختار شرطی‌اند و شکل‌دهی پتانسیل‌پایه سیاست بهینه را عوض نمی‌کند، پس عامل خودش به آن‌ها رسیده است.

دو ایراد در کد ارائه‌شده

(الف) ماژول گم‌شده

evaluate.py و eval_submission.py از see.tournament.runner وارد می‌کنند — ماژولی که در مخزن نیست. بازسازی‌اش کردیم طبق DESIGN.md بخش ۱۰، بدون دست زدن به فیزیک.

(ب) تله‌ی ابرپارامتر

DESIGN.md می‌گوید راه‌حل «فروپاشی امتیازدهی» این است: $\text{p_self} = 0.35$ و $\text{entropy} \geq 0.02$ — و پیش‌فرض PPOConfig هم همین است. اما train.py با p-self 0.6– و entropy 0.01– آن‌ها را بازنویسی می‌کند و همان خرابی را برمی‌گرداند.

اندازه‌گیری: صندلی آمریکا در 100% اپیزودها در $t \approx 3$ تسلیم می‌شد.

یافته‌ی چهارم: چرا «تسلیم زودهنگام» گاهی درست است

در برابر حریفی که نمی‌رود (یا در افق $\bar{T} = 40$ به آستانه‌ی خروجش نمی‌رسد)، بازی به مهلت می‌رسد و آنجا $G \approx e^0$:

$$u^{\text{مهلت}} \approx X - G(\bar{T}) \approx 28 - 53 < 0 \quad \text{در برابر} \quad u^{\text{خروج در } t=3} \approx 28 - 4.$$

نتیجه

آنچه ما اول «خرابی تسلیم» تشخیص دادیم، در برابر بخشی از استخر یک بهترین پاسخ بود. و به همین دلیل استخر غنی‌شده کمکی نکرد: ما با افزودن حریفانی که نمی‌روند، خروج زودهنگام را جذاب‌تر کردیم.

راه‌حل: هماهنگ کردن استخر آموزش با توزیع ارزیابی.

رژیم آموزش	طول اپیزود	u_I	u_U
پیش‌فرض‌های ارائه‌شده	3.2	144.8	30.8
استخر غنی‌شده	5.7	110.9	35.1
هماهنگ با مرجع	**10.4**	**89.6**	**70.0**

مطالعه‌ی حذفی — و یک فرضیه که آزمونش را پاس کرد

پیکربندی	امتیاز کل	ایران	آمریکا
TitForTat (مرجع)	95.5	152.7	38.2
ویژگی + شکل‌دهی، هماهنگ (ارسالی)	**89.8**	136.3	43.4
فقط شکل‌دهی، هماهنگ	89.8	135.5	44.1
ویژگی + شکل‌دهی، استخر غنی‌شده	84.5	124.8	44.2
فقط شکل‌دهی، پیش‌فرض‌های ارائه‌شده	69.8	91.7	47.9

در رژیم اول، ویژگی‌های باور عملکرد را بهتر نکردند (83.2 در برابر 90.6) — با آنکه خطای برآورد را 72.9% کم می‌کنند.

فرضیه: مشکل کیفیت باور نیست، بودجه‌ی نمونه است (GRU خودش پسین ضمنی دارد؛ ۱۴ ویژگی ورودی را از ۱۶ به ۳۰ بعد می‌برد).

آزمون: با 700 هزار گام و استخر هماهنگ، ویژگی+شکل‌دهی جلو زد و امتیاز TitForTat را از 109.8 به 95.9 پایین آورد. فرضیه تایید شد — و چک‌پوینت ارسالی نظریه‌ی کامل است.

معیار نمره‌دهی، باز شده

امتیاز هر عامل میانگین مطلوبیت خودش است، و در استخر سه‌تایی هیچ‌کس با خودش بازی نمی‌کند ⇒ دو صندلی ما دو مجموعه بازی مجزا می‌بینند:

$$\underbrace{(A+B-b')}_{J_I \text{ صندلی ایران}} + \underbrace{(C+D-d')}_{J_U \text{ صندلی آمریکا}} > R_1 + R_2 = 231.8$$

نسخه	J_I	J_U	مجموع
پایه (بدون جمله‌ی جریمه)	194.3	−5.8	188.5 (می‌بازد)
جریمه‌ی بی‌شرط، یک صندلی	170.8	93.7	264.5
جریمه‌ی شرطی، یک صندلی	123.6	81.6	205.2 (می‌بازد)
شرطی، هر دو صندلی	**189.3**	**84.1**	**273.4**

سه بار شکست خوردیم و بار چهارم بردیم. ولی این کل تحلیل روی استخر اشتباهی بود — اسلاید بعد.

یک چک‌پوینت، دو معیار، دو علامت مخالف

evaluate.py (سه عاملی): **+10.7** برد کلاسی: **−0.9**

برد با شش حریف مرجع نمره می‌دهد؛ /see فقط دوتایش را دارد. سه‌تای دیگر را از روی DESIGN.md بازسازی و روی عددهای منتشرشده برازش کردیم:

حریف	برد واقعی (ایران/آمریکا)	بازسازی	
Random	113.8 / 119.8	113.8 / 119.8	دقیق
TitForTat	2.7 / 34.3	2.7 / 34.3	دقیق
MasterAgent	−2.7 / 70.1	−2.7 / 70.1	دقیق
Hawk	**−58.7** / −12.1	−58.7 / −12.0	دقیق
Dove	135.9 / 144.4	130.6 / 154.8	تقریبی
BayesianThreshold	9.4 / 77.6	35.6 / 82.1	تقریبی

برازش یک چیز نانوشته را هم مشخص کرد: امتیاز هر مرجع شامل بازی با خودش هم هست — برای همین Hawk سوم است نه اول.

دو تصحیح، و یکی‌شان کل نظریه را ساده کرد

۱. استخر آموزش: از {Random، TitForTat} به همان شش حریف برد. عامل ما هیچ‌وقت با حریفی که نمی‌رود تمرین نکرده بود، و مقابل Hawk در صندلی آمریکا -58.7 می‌گرفت؛ حالا $+4.7$.

۲. جمله‌ی جریمه‌ی تسلیم حذف شد. امتیاز TitForTat از بازی‌هایش با استخر مرجع می‌آید و ما عضو آن استخر نیستیم $\Rightarrow$ عدد 53.8 او ثابت است و ما تکانش نمی‌دهیم.

یعنی معیار اصلاً نسبی نیست

میانگین مطلوبیت خودمان روی استخر مرجع > 53.8

پس جمله‌ی محروم‌سازی فقط هزینه بود بدون عایدی. با حذفش شکل‌دهی پاداش دوباره کاملاً پتانسیل‌پایه شد $\Rightarrow$ طبق Ng, Harada & Russell سیاست بهینه را عوض نمی‌کند، و عامل دقیقاً همان کمیتی را بیشینه می‌کند که برد می‌سنجد.

ارزیابی نهایی — عددهای واقعی برد

چک‌پوینت	امتیاز	ایران	آمریکا
آپلود ۱ (استخر سه‌تایی)	52.9	72.3	33.4
آپلود ۲ (استخر برد، بدون جریمه)	56.8	71.7	42.0
آپلود ۳ (+ کهن‌الگوی صبور)	**58.8**	**71.3**	**46.3**
بنچ‌مارک استاد (MasterAgent)	56.3	65.7	46.9
TitForTat (معیار دستورالعمل)	53.8	69.7	37.9

هر دو معیار رد شد — هم TitForTat و هم بنچ‌مارک استاد.

تمام بهبود از صندلی آمریکا آمد ($33.4 \rightarrow 46.3$)؛ صندلی ایران تقریباً ثابت ماند.

یافته‌ی ۱۱: برون‌یابی را با درون‌یابی اشتباه نگیرید

برآوردگر ما برای آپلود سوم 64.8 گفت، برد 58.8 داد. ولی خودش کنار عددش نوشته بود:

«این کاندیدا 27.9 واحد بیرون از بازه‌ی لنگرهاست $\Rightarrow$ برون‌یابی است نه درون‌یابی. یک جهت بدانید، نه یک عدد.»

جهت درست بود ($+2.0$ واقعی)، عدد غلط ($+8.0$ ادعا شده). با لنگر سوم معلوم شد رابطه اشباع می‌شود:

	آپلود ۲	آپلود ۱	آپلود ۳
امتیاز جانشین صبور (آمریکا)	-18.1	9.2	37.2
امتیاز واقعیِ دو حریف نامعلوم	39.8	72.7	68.9

برازش خطی 106.2 می‌گفت، واقعیت 68.9 — همان 6 واحد خطا. برآوردگر حالا بین لنگرها درون‌یابی تکه‌خطی می‌کند و بیرون از بازه امتناع می‌کند.

خواندن دوم: پیش‌بینی 67، واقعیت 56.8

کل خطا در دو ردیف نشست

حریف	برد واقعی	بازسازی	
MasterAgent ،Hawk ،TitForTat ،Random	—		هر چهار تا دقیق
Dove	64.6 / 104.9	143.3 / 152.8	غلط
BayesianThreshold	15.0 / 60.9	34.3 / 76.4	غلط

هیچ مقداری از پارامترها هر دو آپلود را نمی‌سازد ⇒ چیز گمشده پارامتر نبود. پس رفتار خودمان را اندازه گرفتیم:

مقابل حریف کاملاً منفعل	آپلود ۱	آپلود ۲
میانگین سیگنال	0.06	**0.75**
کسر مراحل HOLD	0.24	**0.96**
خودمان اول خارج می‌شویم	0.24	**0.00**

pool board–کل دیکشنری حریف‌ها را عوض کرد ⇒ Patient و Plateau بی‌سروصدا افتادند.

عامل یاد گرفت تسلیم شدن هزینه‌ای ندارد، چون دیگر هیچ‌کس نبود که فقط بایستد و صبر کند. بهایش روی برد ≈ 13 واحد بود.

درسِ عملی

این دسته خطا را با امتیاز پیدا نمی‌کنید — امتیاز کل بالا رفت. با اندازه‌گیری رفتار پیدا می‌کنید.

و وقتی بازسازی جواب نمی‌دهد، دنبال جانشینِ هم‌ترتیب بگردید: Patient تنها حریف محلی است که دو آپلود ما را به همان ترتیبِ برد می‌چیند ($-18.6/-43.3$ در برابر $-39.5/-71.3$). سطحش شبیه نیست و ادعایی هم نمی‌کنیم — برای انتخاب ترتیب کافی است.

یافته‌ی ۱۲: تیپ حریف $\neq$ رفتار حریف

هر ۱۶ ویژگی، پسین ما روی تیپ ساختاری $\theta_j = (\bar{\rho}_j, \bar{m}_j)$ بود. ولی Random و Patient تیپ متفاوتی ندارند — قاعده‌شان فرق دارد، و تفاوتشان در نوسان سیگنال است.

حریف	انحراف معیار	بیشینه	کسر سکوت
Random	0.159	0.642	0.225
Dove ،Patient	0.000	0.000	**1.000**
Hawk	0.000	1.000	0.000

سه آماره‌ی خام اضافه شد (→ ۱۹ ویژگی) و پیش از آموزش تایید شد که جدا می‌کنند.

نتیجه نصفه بود

هدف اصلی (Random/آمریکا: 113.3 → 62.1) حاصل نشد (64.1). سود از جای دیگر آمد: ایران مقابل TitForTat **41.1** ← 34.6 و مقابل Hawk **22.8** ← 12.1 — بهترین عدد کل پروژه.

بازنمایی شرط لازم بود، نه کافی.

یافته‌ی ۱۳: چقدرش مهارت بود، چقدرش شانس

بعد از 61.7، دوازده چک‌پوینت دیگر آموزش دادیم. هیچ‌کدام بهتر نشد.

چیزی که آزمودیم	بهترین تخمین
تعادل استخر	59.2
ویژگی «چقدر شاهد داریم»	59.7
مواجهه‌ی اسکریپتی $56\% \to 85\%$	61.0
تنظیم میانی	59.6
هشت بذر تازه از همان دستور	59.4

توزیع خانواده‌ی ۱۹–ویژگی‌ای (14 چک‌پوینت):

کمینه 51.5 میانه 58.7 $\sigma = 2.33$ بیشینه **61.7**

آن چک‌پوینت 1.5σ بالای میانگین خانواده‌ی خودش است. 61.7 کف این روش نیست، سقف خوش‌شانسی‌اش است.

یافته‌ی ۱۴: تمام این مدت، جدولِ اشتباه را بهینه می‌کردیم

دستورالعمل: تورنمنت نمره‌دار Round Robin است با تمام ارسال‌ها، دو عامل مرجع، و بنچمارک استاد. Hawk، Dove و BayesianThreshold فقط در جدول تمرینی‌اند.

از 30% نمره‌ی پیاده‌سازی: نصفش «امتیاز بالاتر از Tit-for-tat»، به‌علاوه‌ی 20% نمره‌ی اضافه برای رتبه در همان تورنمنت. رتبه‌ی جدول تمرینی هیچ سهمی ندارد.

پس صدرنشینیِ جدول تمرینی هیچ تضمینی برای معیار نمره‌دار نیست — دو میدان متفاوت‌اند و باید جداگانه سنجیده شوند.

یافته‌ی ۱۵: «بهترین چک‌پوینت» بدون تعیین میدان بی‌معناست

دو شبکه‌ی نقش مستقل‌اند، پس صندلی ایران و صندلی آمریکا می‌توانند از دو اجرای متفاوت بیایند.

میدان سنجش	پیوندی	of_s3
نسل قدیمِ ایجنت‌ها	**+11.2**	+1.7
round robin ۲۰ شرکت‌کننده	+0.4	**+3.1**
بازسازی میدان کلاس (۵ ارسال)	−7.7	**+8.0**

امتیازِ دو صندلی دقیقاً جمع شدند — پیوند کار کرد. ولی سودش مقابل میدانی بود که وجود ندارد.

ششمین فرضیه‌ای که در این پروژه با داده رد شد. جواب درست را لیدربورد داد، نه شبیه‌سازی: پروفایل دوصندلیِ هم‌کلاسی‌ها منتشر است.

یافته‌ی ۱۶: درسِ میدان را به ابزار انتخاب هم بزن

اسکن صندلی‌ها را مقابل نسل قدیم زده بودیم. همان اسکن در میدان درست، جواب دیگری می‌دهد:

بهترین صندلی آمریکا	اسکن نسل قدیم	اسکن میدان کلاس
of_s3	**59.8**	32.0
oh_s2	45.5	**43.3**

oh_s2 در هیچ رتبه‌بندی کاملی بالا نیامده بود، چون صندلی ایرانش ضعیف است (64.8).

از استخر official+ آمده — آزمایشی که «شکست‌خورده» اعلامش کرده بودیم. فرضیه درست بود؛ اثرش روی یک صندلی بود و امتیاز ترکیبی دفنش کرد.

چک‌پوینت نهایی: یک مصالحه‌ی آگاهانه

چک‌پوینت	جدول تمرینی	چهار مرجع دقیق	فاصله از TitForTat
bh_s2 (آپلود چهارم)	**61.7**	51.7	−1.9
of_s3_t1.5	59.4*	46.3	+7.1
gx_a (پیوندی)	63.4*	56.9	+8.6
final_graft (ارسالی)	58.3	**57.4**	**+8.7**

مقادیر ستاره‌دار (*) برآوردند؛ بقیه اندازه‌گیری.

final_graft صندلی ایران را از og_s2 و صندلی آمریکا را از oh_s2 می‌گیرد. این یک مصالحه است: 3.4 واحد پایین‌تر در جدول تمرینی، در ازای 5.7 واحد بالاتر در معیار نمره‌دار.

جدول تمرینی هیچ سهمی از نمره ندارد و تورنمنت نصف نمره‌ی پیاده‌سازی را دارد — معامله آگاهانه پذیرفته شد.

یافته‌ی ۱۸: روی میدان واقعی کلاس فاصله بزرگ‌تر می‌شود

بازسازی از پنج ارسالِ اولیه به هجده ارسالِ نهایی به‌روز شد:

چک‌پوینت	مدل ۵ نفره	میدان واقعی (۱۸ ارسال)
final_graft (ارسالی)	+8.7	**+14.5** (رتبه ۱ از ۲۱)
bh_s2 (آپلود چهارم)	−1.9	+8.8 (رتبه ۳ از ۲۱)

برتری تقریباً دو برابر می‌شود، چون هر هجده نفر در صندلی آمریکا ضعیف‌اند (ایران 64.7–79.0، آمریکا 28.9–48.5) و ما تنها کسی هستیم که آن صندلی را هدف گرفته. برخلاف جدول تمرینی، اینجا بزرگ‌شدن میدان به نفع ماست.

و یک تصحیح: آن −1.9 مصنوعِ میدان کوچک بود. bh_s2 هم روی میدان واقعی می‌برد. پس تعویض چک‌پوینت بهبود داد، نه نجات از شکست.

جمع‌بندی: سه حلقه‌ی خودتقویت‌کننده

۱. حلقه‌ی تعهد: تشدید $\Rightarrow q^* \downarrow \Rightarrow (B-X) \uparrow \Rightarrow X \downarrow \Rightarrow K \uparrow \Rightarrow$ ادامه عقلانی‌تر.

۲. حلقه‌ی شکنندگی: فرسایش $\Rightarrow g \uparrow \Rightarrow m \downarrow \Rightarrow$ فرسایش سریع‌تر.

۳. حلقه‌ی میانجی: تشدید $\Rightarrow$ پنجره بسته $\Rightarrow X$ پایین $\Rightarrow$ خروج غیرممکن‌تر.

و تنها چیزی که هر سه را می‌شکند

خویشتن‌داری متقابل — چون هم‌زمان K را می‌خواباند، سوخت را کم می‌کند، و پنجره‌ی میانجی‌گری را باز می‌کند.

محدودیت اصلی: مدل درباره‌ی جهت روابط قابل دفاع است، نه درباره‌ی اندازه‌شان.