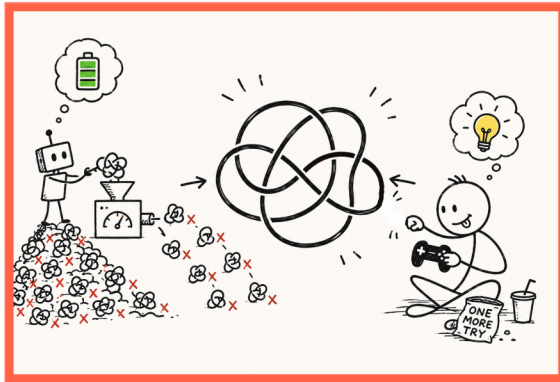


Gamification in mathematics

Or: Games turn search into data

Accept ~~Change~~ what you cannot ~~change~~ accept



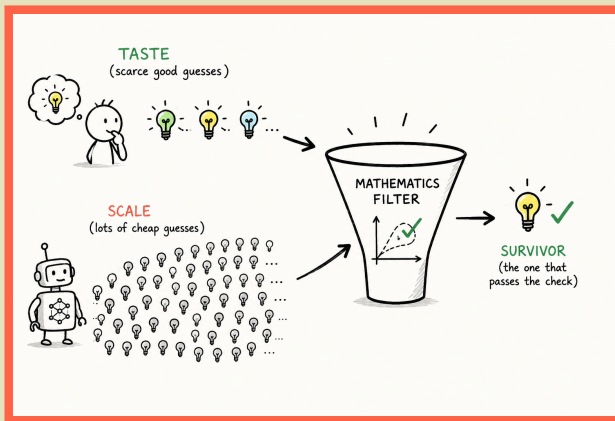
Based on work of many people: A love story about procrastination

AI, games, and patience

	 HUMANS	AI 
<u>Superpower</u>	taste, intuition, weird ideas 	speed, scale, endless patience 
<u>What we bring</u>	a few shots worth taking 	thousands of shots 
<u>Not great at</u>	patience, repetition 	taste, knowing what to try 
<u>Best role</u>	choose directions, spot patterns 	explore everything, don't get bored 
<u>Result we want</u>	better guesses 	more (tested) candidates 

- ▶ **Humans** have taste, visual intuition, and stupid ideas worth trying
- ▶ **AI** has speed, scale, and terrifying amounts of patience
- ▶ **Math** often rewards the one attempt that survives the check

Taste and scale



Good human guesses are scarce. **Taste.**

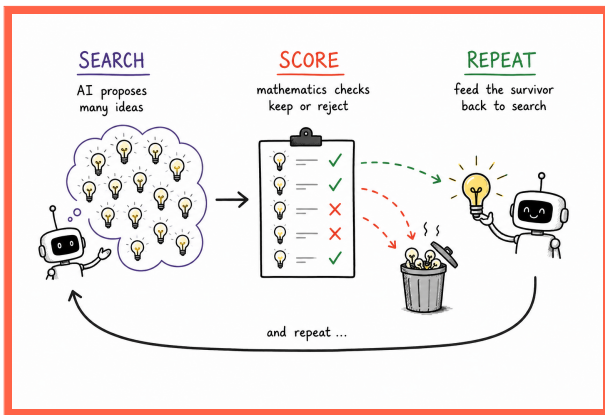
Bad AI guesses are cheap. **Scale.**

Both are useful because mathematics is the filter.

Traditional math gets attempts from taste (the “classical approach”);

AI math gets attempts from scale (more in a sec!).

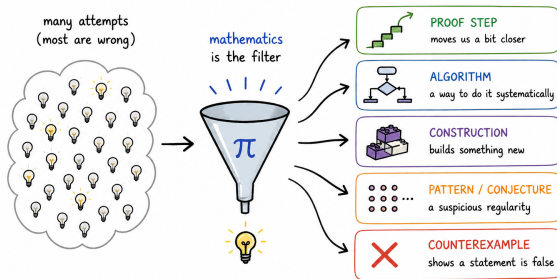
AI, games, and patience



- **Search** Generate candidates: proofs, programs, constructions, strategies...
- **Score** Use a cheap test, verifier, objective, or invariant
- **Repeat** Try again, and again, ... (AI is really good at this)

Why this can do mathematics

What we can get out



The output need not be a final theorem.
It might be a proof step, a better algorithm,
a useful construction, a suspicious pattern,
or the counterexample that kills a conjecture.

strategies...

► Search

► Score

► Repeat

Try again, and again, ... (AI is really good at this)

AI, games, and patience

Rank *	Title *	Sales (millions) *	Series *	Platform(s) *	Release year *	Developer(s) *	Publisher(s) *	Ref. *
1 ^[a]	<i>Tetris</i>	520	None	Multi-platform	1988 ^[c]	Various	Various	[19]
	<i>Minecraft</i>	400	<i>Minecraft</i>	Multi-platform	2011 ^[d]	Mojang Studios	Mojang Studios	[22]
2	<i>Grand Theft Auto V</i>	230	<i>Grand Theft Auto</i>	Multi-platform	2013	Rockstar North	Rockstar Games	[23]
3	<i>Red Dead Redemption 2</i>	85	<i>Red Dead</i>	Multi-platform	2018	Rockstar Games	Rockstar Games	[23]
4	<i>Wii Sports</i> ^[b]	82.9	<i>Wii</i>	Wii	2006	Nintendo EAD	Nintendo	[26]
5	<i>Mario Kart 8 / Deluxe</i>	79.54	<i>Mario Kart</i>	Wii U / Switch	2014	Nintendo EAD / Nintendo EPD (Deluxe)	Nintendo	[e]
6	<i>PUBG: Battlegrounds</i>	75	<i>PUBG Universe</i>	Multi-platform	2017	PUBG Studios	Krafton	[29]
7	<i>Terraria</i>	70	None	Multi-platform	2011	Re-Logic	Re-Logic / 505 Games	[30]
8	<i>The Witcher 3: Wild Hunt</i>	65	<i>The Witcher</i>	Multi-platform	2015	CD Projekt Red	CD Projekt	[31]
9	<i>Super Mario Bros.</i>	58	<i>Super Mario</i>	Multi-platform	1985	Nintendo R&D4	Nintendo	[32]
10	<i>Human: Fall Flat</i>	55	None	Multi-platform	2016	No Brakes Games	Curve Digital	[33]

- ▶ **Game** Keep the human intuition, but add numbers
- ▶ **Key** Keep the human motivated long enough to be useful
- ▶ **Key** Setting is up is still not straightforward

The gamification bargain

Humans will do this
for fun

 SUPER MARIO BROS.	4:54.448 Any%	11 frames from perfection (TAS tie) 
 SUPER MARIO 64	1:35:14 120 Star	95 minutes of precision 
 MINECRAFT Java Random Seed Glitchless	22:03 Any%	thousands of human search attempts 



Now imagine if the score were **mathematical**.

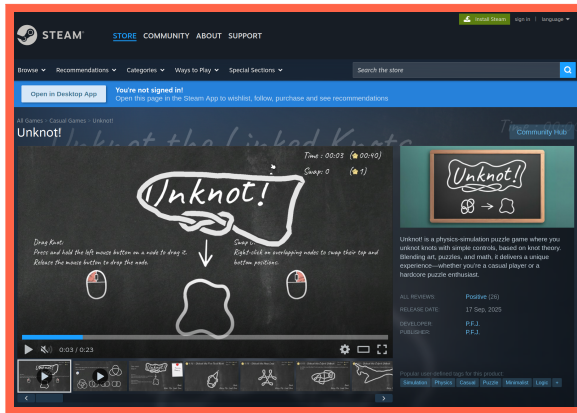
Gamification tries to steal the best part of AI:
scale, scoring, repetition, patience.
But the search engine is still human intuition,
so the game has to be genuinely fun.

► Game

► Key

► Key Setting is up is still not straightforward

Unknot! as mathematics



- ▶ **Game** *Unknot!* is about grabbing, pulling, and untangling rope
- ▶ **Maths** Knot theory asks when a tangled loop is really the unknot
- ▶ **Bridge** The player sees a puzzle; we see a search problem

Public-facing slogan

TOKYO GAME SHOW 2026

Business Day September 17 (Thu.) • 18 (Fri.)
Public Day September 19 (Sat.) • 21 (Mon.)

TOKYO GAME SHOW 2026
公式番組のご案内
Official Program Information

TOKYO GAME SHOW 2026
Exhibit Guide

TOKYO GAME SHOW 2026
イベントステージのご案内
Event stage information

TOKYO GAME SHOW 2026
Advertisement & Sponsorship
Update : 2024.3.13

INDIE GUIDEBOOK IN TGS2026
Guide for Advertising

International Party Sponsorship Solution Deck
[TOKYO GAME SHOW 2026]
International Party Indie Night
Nihon BP TOKYO GAME SHOW Management Office

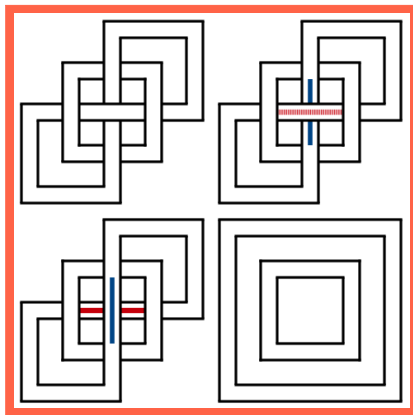
Call for Exhibitor Closing Date to Exhibit Application **May 8** 42 booths or more **May 22** Less than 42 booths Exhibitor Briefing **July 8**

Can playing a puzzle game help answer
a real research question? That is the experiment.

Featured at Tokyo game show to expand player-base. (Public is important for this!)

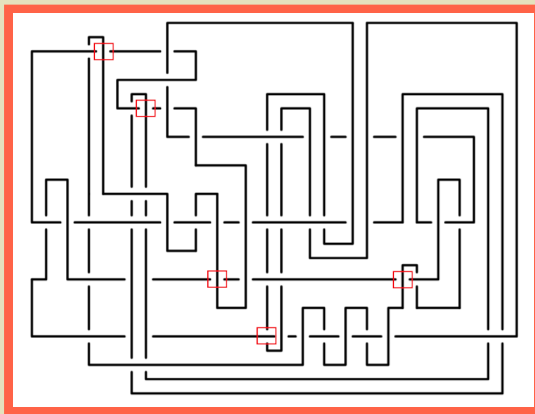
► **Bridge** The player sees a puzzle; we see a search problem

Unknot! as mathematics



- **Move** At one crossing, swap over and under
- **Number** The unknotting number asks how few such swaps suffice
- **Problem** Even small knots can have annoying unknown cases

The catch



A useful crossing may appear only after changing the picture of the knot.
Above: the optimal (?) 5 swaps appear on a 56 crossing diagram of a 14 crossing knot!
So the search is over drawings, not just crossings.



Problem Even small knots can have annoying unknown cases



- **Levels** Research levels can encode knots where we want better routes
- **Leaderboard** The score is “how few decisive changes did you need?”
- **Replay** A good score is useless unless we can reconstruct it later

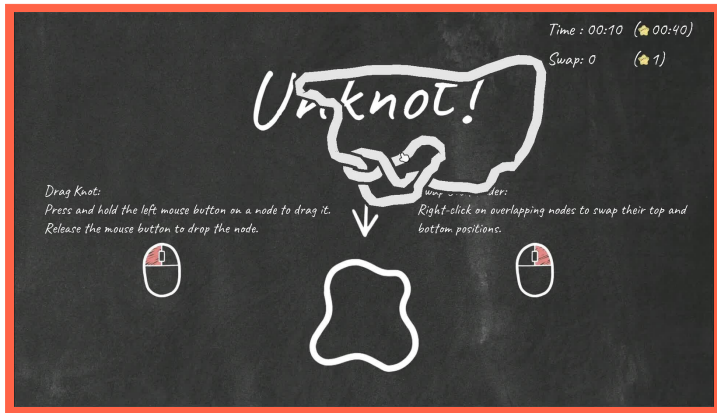
Why players matter



A surprising route found by a player
can become a mathematical candidate:
not a proof, but a concrete thing to check.

► **Replay** A good score is useless unless we can reconstruct it later

From players to data



- ▶ **Demo** In the next ten minutes, I will play badly in public
- ▶ **Point** The game is not window dressing; it is the interface to the search
- ▶ **Question** Which moves do humans try before giving up?

During the demo

That's the
WRONG
Question

Drag Knot:
Press and hold
Release the mouse

(👑 00:40)

(👑 1)

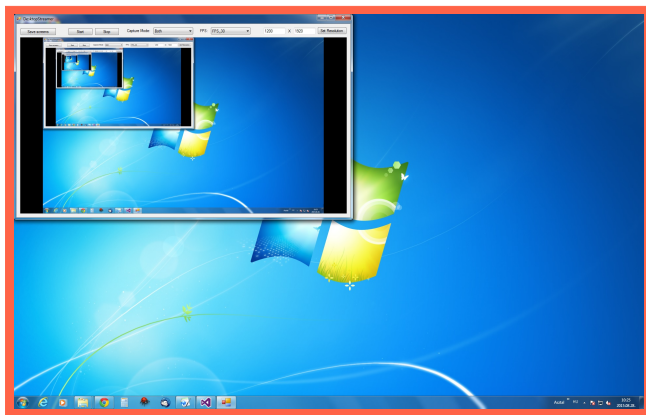
air top and

- ▶ Demo In the
- ▶ Point The game
- ▶ Question Which moves do humans try before giving up?

Watch less for "did they win?"
and more for "what choices does the game
make natural?"

to the search

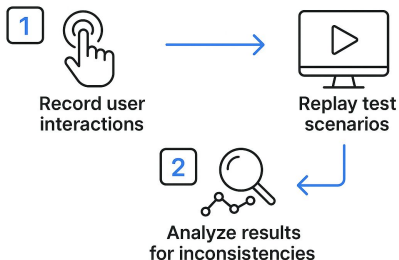
From players to data



- ▶ **Replay** Research levels save and upload player replays
- ▶ **Filter** We look for unusually short or structurally strange solutions
- ▶ **Translation** The hard part is turning gameplay into knot-theoretic evidence

The data sentence

RECORD REPLAY TESTING



A replay is not a theorem.

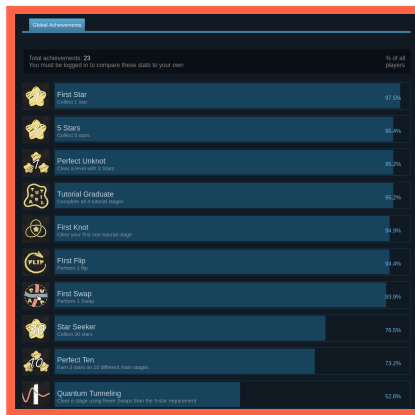
But it is much better than “someone once saw a shortcut”
because it can be replayed and inspected.

► Replay

► Filter

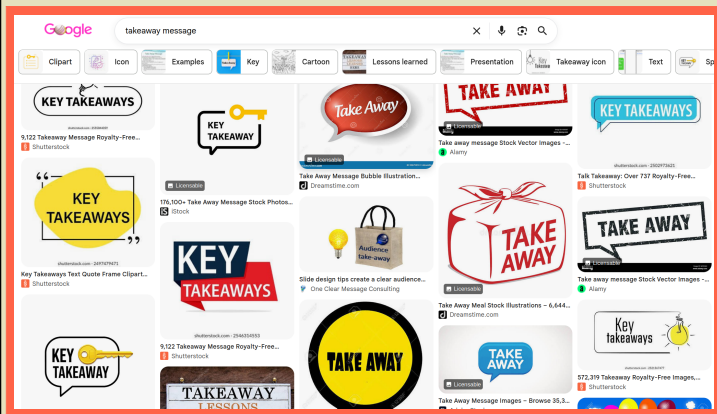
► Translation The hard part is turning gameplay into knot-theoretic evidence

From players to data



- ▶ Best case A player improves an upper bound for an unknotting number
- ▶ Good case We learn which diagrams are good challenges
- ▶ Still good We learn that our interface is not yet measuring the right thing

Takeaway



Gamification is not making mathematics silly.
It is a way of recruiting intuition
where blind search is too stupid and pure theory is too slow.

- Best
- Good case we learn which diagrams are good challenges
- Still good We learn that our interface is not yet measuring the right thing

AI, games, and patience



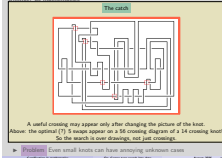
- **Humans** have taste, visual intuition, and stupid ideas worth trying
- **AI** has speed, scale, and terrifying amounts of patience
- **Math** often rewards the one attempt that survives the check

AI, games,



- **Game** Gamification tries to steal the best part of AI: scale, scoring, repetition, patience.
- **Key** But the search engine is still human intuition, so the game has to be genuinely fun.
- **Key** Getting in up to still feel straightforward

Unknot! as mathematics



- **Problem** Even small knots can have annoying unknown cases

AI, games, and patience



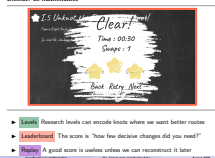
- **Search** Generate candidates: proofs, programs, constructions, strategies...
- **Score** Use a cheap test, verifier, objective, or invariant
- **Repeat** Try again, and again, ... (AI is really good at this)

Unknot! as mathematics



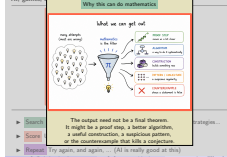
- **Game** Unknot! is about grabbing, pulling, and untangling ropes
- **Math** Knot theory asks when a tangled loop is really the unknot
- **Bridge** The player sees a puzzle; we see a search problem

Unknot! as mathematics



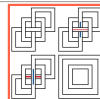
- **Levels** Research levels can encode knots where we want better routes
- **Leaderboard** The score is "how few decisive changes did you need?"
- **Replog** A good score is useless unless we can reconstruct it later

AI, games,



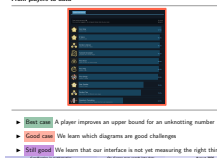
- **Search** The output need not be a final theorem. It might be a good step, a better algorithm, a useful construction, a suspicious pattern, or the counterexample that kills a conjecture
- **Score** Try again, and again, ... (AI is really good at this)
- **Repeat** Try again, and again, ... (AI is really good at this)

Unknot! as mathematics



- **Move** At one crossing, swap over and under
- **Number** The unknotting number asks how few such swaps suffice
- **Problem** Even small knots can have annoying unknown cases

From players to data



- **Best case** A player improves an upper bound for an unknotting number
- **Good case** We learn which diagrams are good challenges
- **Still good** We learn that our interface is not yet measuring the right thing

There is still much to do...

	HUMANS	AI
Synopsis	Needs, wishes, desires, emotions	Speed, skills, available patience
Idea	Has the best earth-saving idea	Misses it all
Plan	Has great engineering	Misses it all
Test idea	Needs funding, test partners	Needs funding, test partners
Build and test	Needs patience, test partners	Needs funding, test partners

The gamification bargain

Humans will do this for fun

- 1. Build a house 4:56, 408 points
- 2. Write a poem 1:35:14, 120 stars
- 3. Write a poem 22:03, 1000 stars
- 4. Write a poem 22:03, 1000 stars

Total: 1000 stars

Progress: 1000 stars

Gamification tries to steal the best part of AI: scale, scoring, repetition, padding.

But the worst thing is still human evolution, to the game game it is genuinely fun.

Setting it up is still not straightforward

The catch

A useful crossing may appear only after changing the picture of the knot.
Above: the optimal (7) 5 exchange appears on a 35 crossing diagram of a 14 crossing knot!
So the search is over drawing, not just crossings.

► **Problem:** Even small knots can have amazing unknown cases

The diagram illustrates the iterative process of the 3C framework (Create, Critique, Combine). It is enclosed in a red rectangular box. The process follows these steps:

- SEARCH:** AI generates many ideas, represented by a cloud of lightbulbs.
- SCORE:** Collaborative debate to rank ideas, shown as a list of ideas with red checkmarks and green X marks.
- REPEAT:** Refine the solution based on feedback, depicted with a trash can and a lightbulb.

An arrow labeled "and repeat..." loops from the bottom of the cycle back to the SEARCH stage.

- **Search:** Generate candidates: proofs, programs, constructions, strategies...
- **Score:** Use a cheap test, verifier, objective, or invariant
- **Repeat:** Try again, and again... (AI is really good at this)

▶ **Game** Untoxkat! is about grabbing, pulling, and untangling rope

▶ **Meaning** Kourt theory asks when a tangled loop is really the unknot

▶ **Bridge** The player uses a puzzle, we use a search problem

ES Unknotted Knots

Clear!

Time: 00:30

Shape: 1

Back Retry Next

[illegible]

► **Move:** At one crossing, swap over and under

► **Number:** The unknotting number asks how few such swaps suffice

► **Problem:** Even small knots can have annoying unknown cases

[illegible]

Thanks for your attention!