



Collective Intelligence

Shuyang Li & Bodhisattwa Majumder



Collective Intelligence / Crowdsourcing

“Collective intelligence is shared or **group intelligence** that emerges from the collaboration, **collective efforts**, and competition of **many individuals** and appears in **consensus decision making**.”

Wikipedia

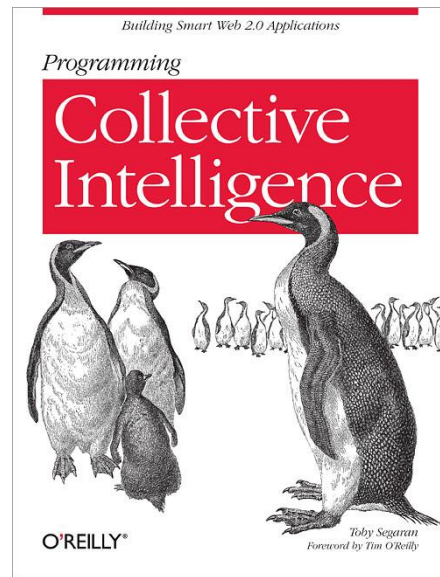
Areas of Research

Ethics of collective intelligence

Wisdom of the crowd

Animal collective behavior

Participatory and deliberative democracy





Areas of Research

Ethics of collective intelligence

Wisdom of the crowd

Animal collective behavior

Participatory and deliberative democracy







Individual Thoughts (1 min)

What is the expectation of privacy when conducting citizen science?

How can we destigmatize IT support jobs (especially outsourced ones)?

Crowd Research

Open and Scalable University Laboratories



Crowd Research in 90 seconds

Why?

Research is an **exclusive** field

Diversity and upward mobility in science

“Reputation diffusion”



Crowd Research in 90 seconds

What?

Anybody can sign on to a broad research project

Weekly milestones and shaping of research directions

Users rate each other and evaluate contributions

The Opportunity Gap





Breakout Groups (~1-2 min)

How would you change the structure of Crowd Research to incentivize...

discovery of research as a possible career path/interest?

research infrastructure development in low-resource areas?

...are these goals well-suited for a Crowd Research-like platform in the first place?

Technical Focus:

Graph Centrality (PageRank)



Reputation / Attribution in Large Groups

How do we determine authorship?

Which milestones are “worth” pursuing further?

Who gets to present during weekly meetings?



$$\rho_i(t) = \frac{1 - d}{|P|} + d \sum_{p \in P} (\rho_p(t - 1) \cdot C(p, i))$$

$$\rho_i(t) = \frac{1 - d}{|P|} + d \sum_{p \in P} (\rho_p(t-1) \cdot C(p, i))$$

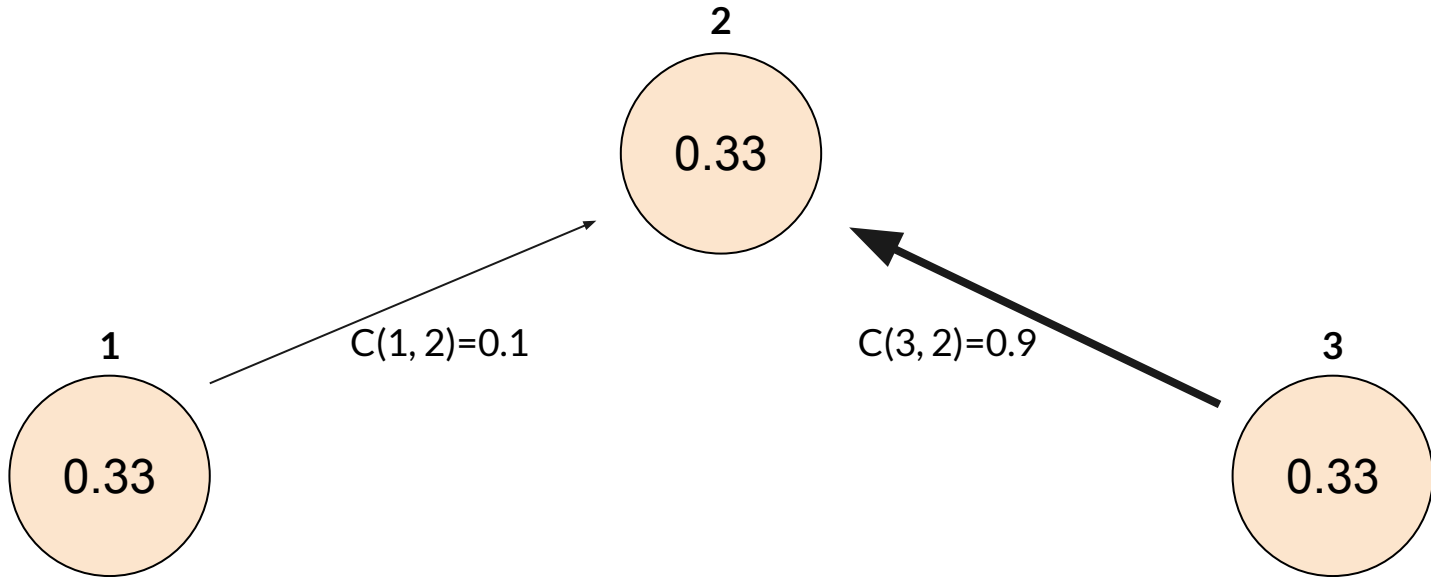
Everybody starts with 100 credits

I think you did **X** amount of work

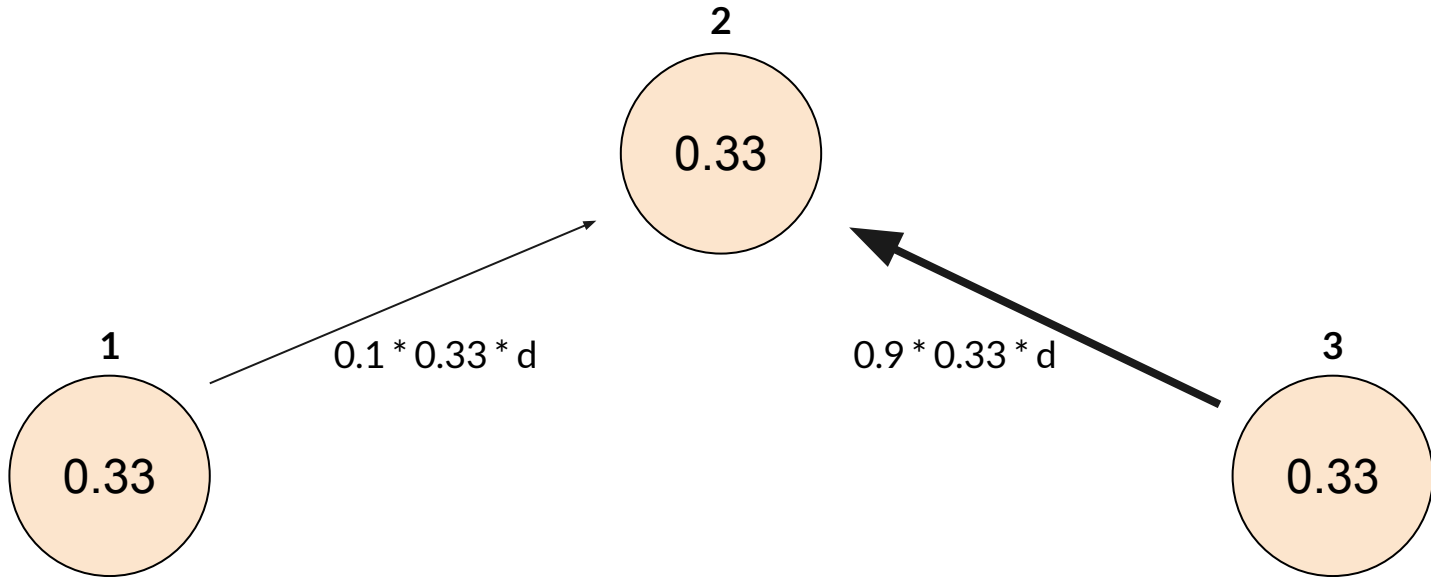
I give you credits proportional to **X**

(and the IRS takes their **1-d** cut)

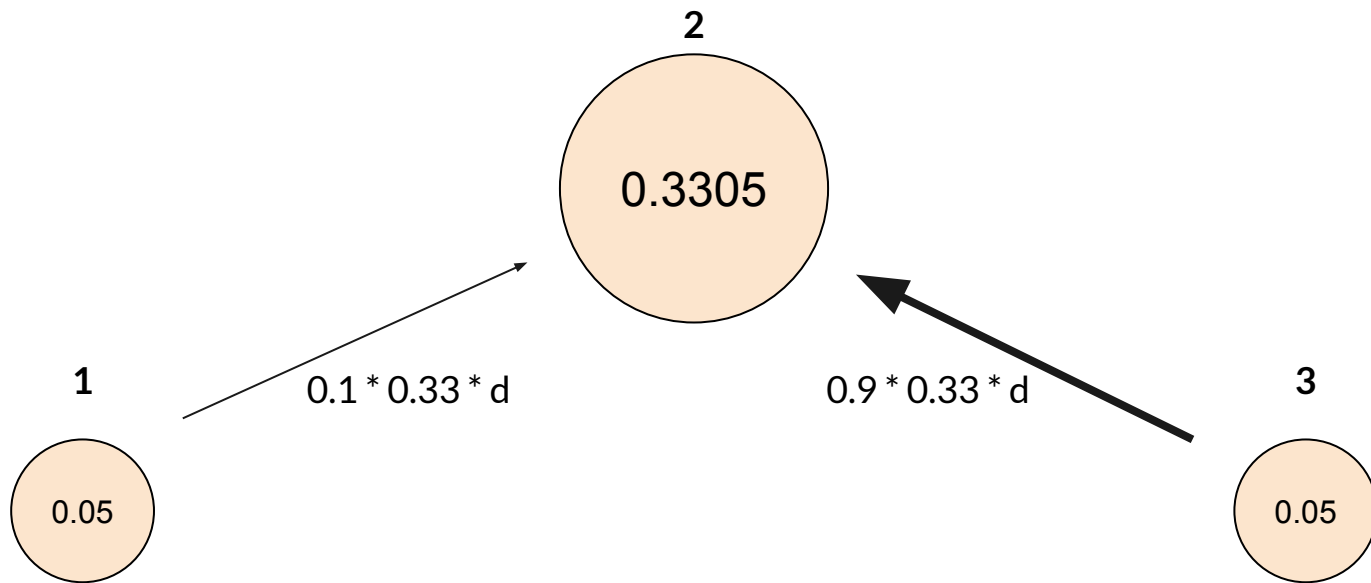
$$\rho_i(t) = \frac{1-d}{|P|} + d \sum_{p \in P} (\rho_p(t-1) \cdot C(p, i))$$

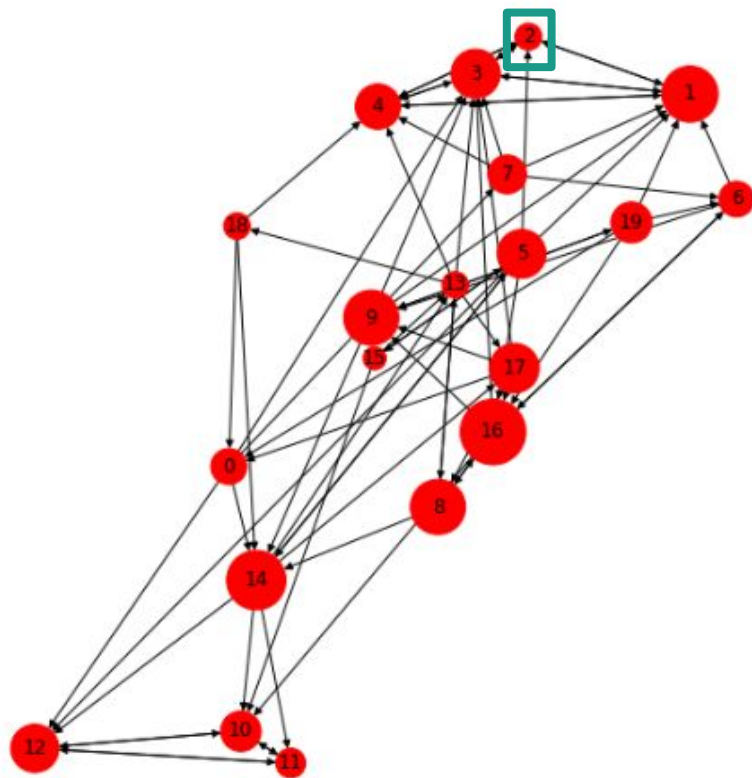


$$\rho_i(t) = \frac{1-d}{|P|} + d \sum_{p \in P} (\rho_p(t-1) \cdot C(p, i))$$

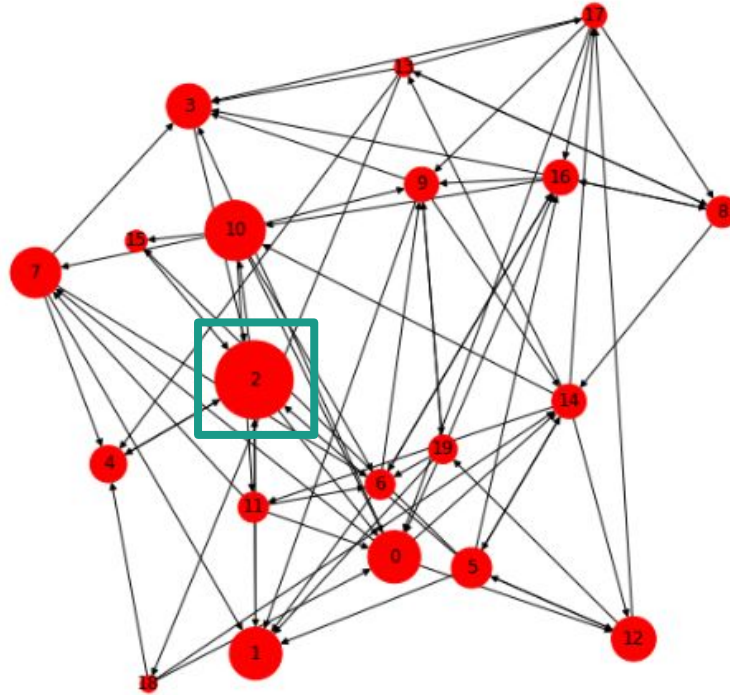


$$\rho_i(t) = \frac{1-d}{|P|} + d \sum_{p \in P} (\rho_p(t-1) \cdot C(p, i))$$



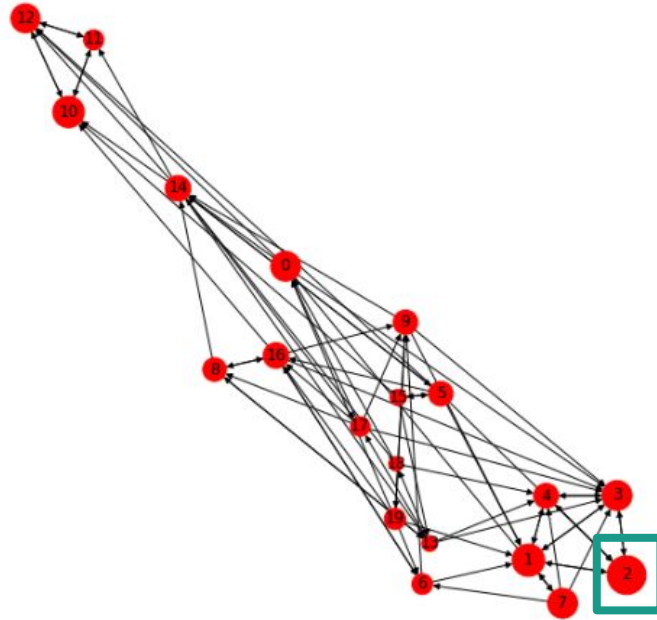


Randomly chosen credit assignments



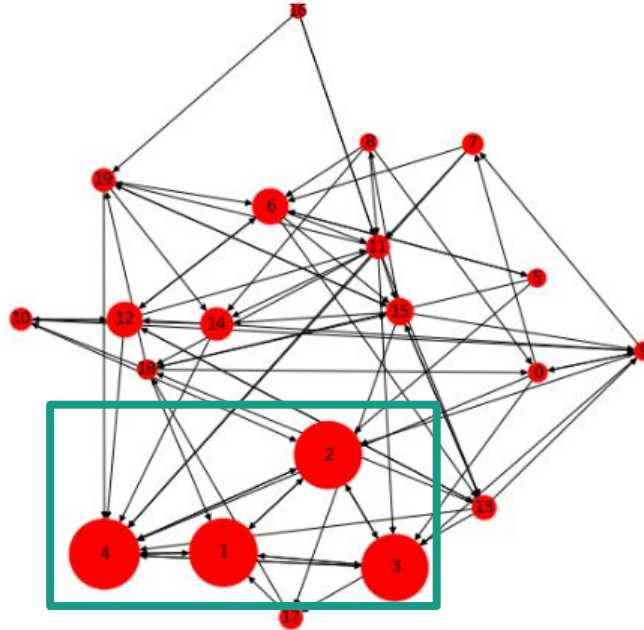
Funneling strategy: users 1, 3, 4, 15 assign all credit to 2

```
MAX SCORE: 0.5  
0:03:23.922347 - 10 iterations pagerank complete (0.00000 squared sum of changes)  
0:03:23.929328 - 20 iterations pagerank complete (0.00000 squared sum of changes)  
0:03:23.936310 - 30 iterations pagerank complete (0.00000 squared sum of changes)  
0:03:23.942293 - 40 iterations pagerank complete (0.00000 squared sum of changes)  
0:03:23.943291 - converged after 42 iterations
```



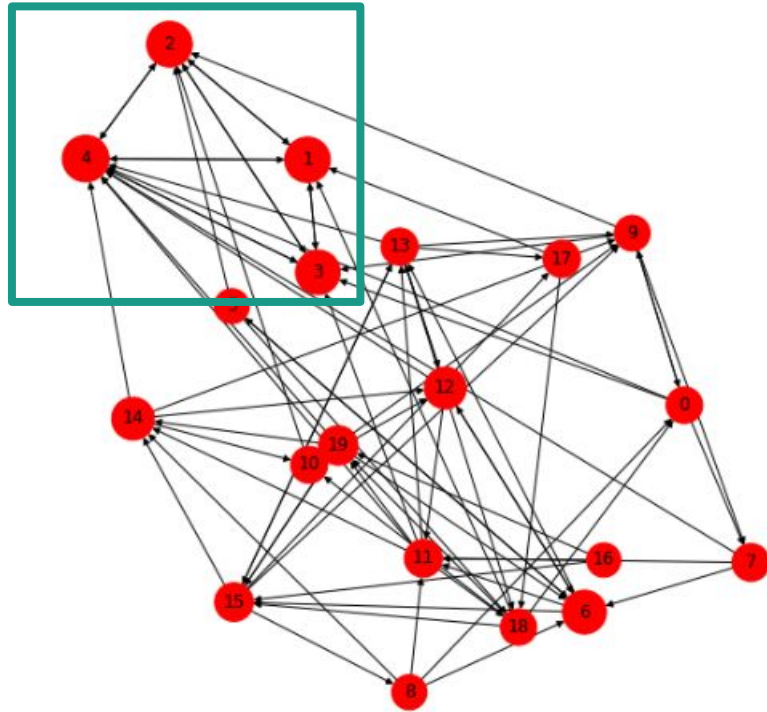
Combat funneling by capping contribution score to 0.5

```
DAMPING WITH 0.85  
0:00:07.424283 - 10 iterations pagerank complete (0.00000 squared sum of changes)  
0:00:07.428272 - 20 iterations pagerank complete (0.00000 squared sum of changes)  
0:00:07.431264 - 30 iterations pagerank complete (0.00000 squared sum of changes)  
0:00:07.433259 - converged after 35 iterations
```



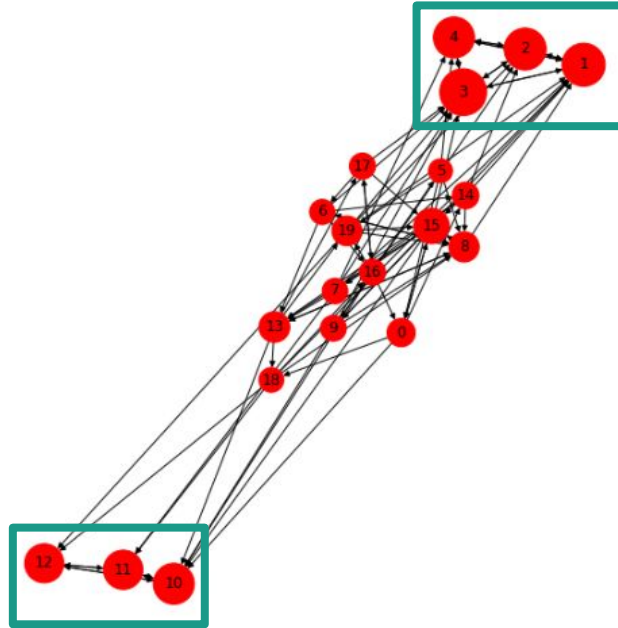
Quid-pro-quo / cartel with damping constant 0.85

DAMPING WITH 0.25
0:00:06.032240 - converged after 10 iterations



Combat the cartel by setting damping constant to **0.25**


```
DAMPING WITH 0.5  
0:00:28.658986 - 10 iterations pagerank complete (0.00000 squared sum of changes)  
0:00:28.660981 - converged after 13 iterations
```



What if there are **more cartels**? ($d=0.25$)



Think-Pair-Share (30-60-N)

How can *you* game the system to gain reputation/collaboration points?



Class Commentary: Crowd Research

Impacts of communication barriers in diversity

Decentralization is key in Crowd Research

Less “**glamorous**” research ideas receive less traction in the marketplace

How do we properly **assess the effectiveness** of such a platform?

FoldIt

Predicting protein structures with a multiplayer
online game



FoldIt in 90 seconds

Why?

We know constituent structures in proteins

We *don't* know how they're oriented in 3-space

Look-ahead in optimization is *hard* (for machines)



FoldIt in 90 seconds

What?

People configure proteins and are scored

Participants “compare notes” via chat/wiki

Competition + collaboration via teams & leaderboard



Rosetta Methodology

Stochastic algorithms:

Rebuilding all or the portion of chain from fragments

Deterministic:

Energy minimization, structure perturbation, refinement



Bayesian (Hyperparameter) Optimization

Randomly *perturb* the system

Evaluate change

Accept changes with a certain % and *update* distribution



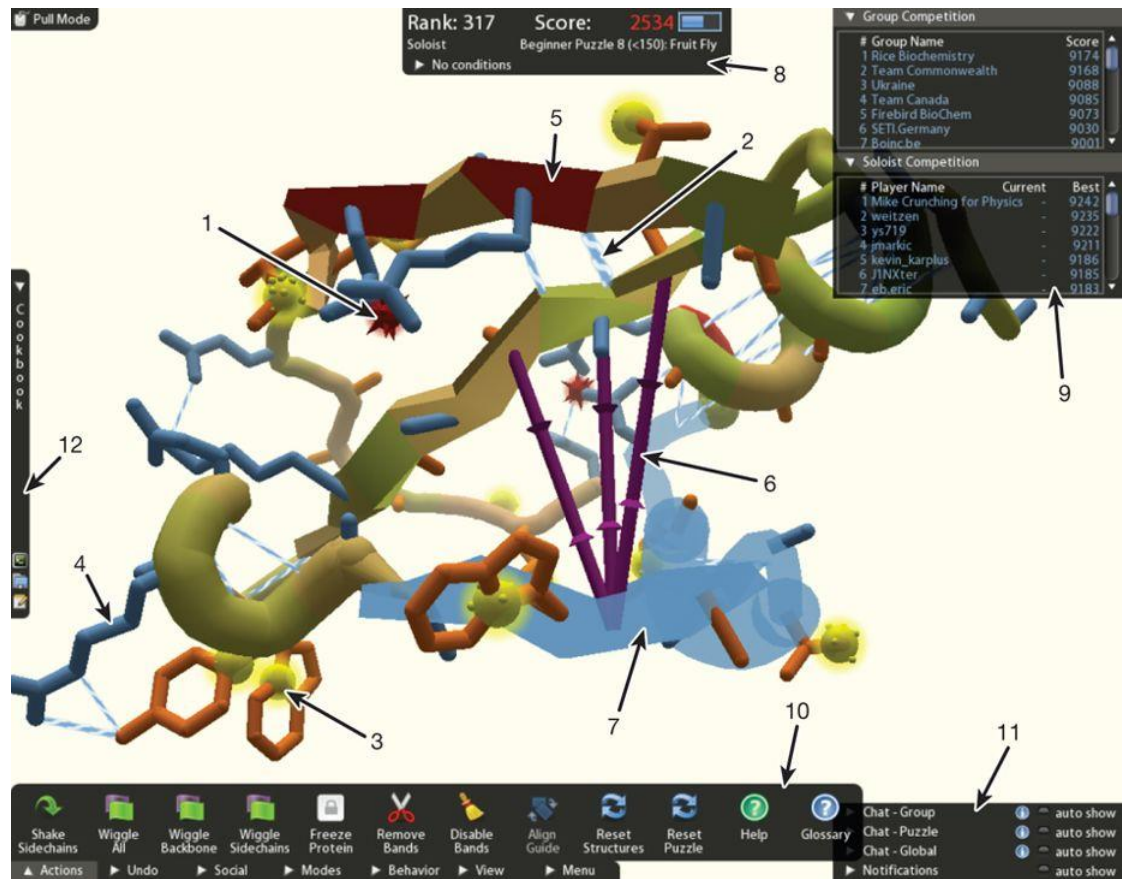
Human outperforms Rosetta

Intelligent Annealing!

Human chose paths/strategies which are worse in terms of energy-optimization but later achieved better results than the best prediction by the machine

**Riskier steps - More useful starting point -
Better at resolving incorrect features**

Design Focus: Explicit/Implicit Gamification



Cooper, Seth, et al. "Predicting protein structures with a multiplayer online game." *Nature* 466.7307 (2010): 756.

Pull Mode



Rank: 137 **Score:** 0

Soloist 1657: Unsolved De-novo Freestyle 150

Expires 4/09/2019 16:00 PDT (1 days, 14 hours)

► No Objectives Available

▼ **Group Competition**

| # | Group Name | Score |
|---|------------|-------|
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |
| | | |

▼ **Soloist Competition**

| # | Player Name | Current | Best |
|---|----------------|---------|-------|
| 1 | LocOilOiling | 10647 | 10647 |
| 2 | reefyrob | - | 10544 |
| 3 | liendish_ghoul | - | 10501 |
| 4 | Galaxie | 10457 | 10457 |
| 5 | Phyx | 10421 | 10450 |
| 6 | spvincent | - | 10406 |
| 7 | retiredmichael | - | 10378 |

foldit Help

IRC Commands

- /ignore username
- /unignore username
- /ignore-list

Camera Controls

Reset: Home

Mouse on the background:

- Shake**
Moves all the sidechains at once to get them into a better overall position
- Wiggle**
Moves the backbone around to improve it
- Rebuild**
Try completely new backbone shapes

Mouse Controls

No 3-button mouse? Use these:

Ctrl/Cmd for right-click/drag
Shift for middle-click/drag

Structure Mode

Left-drag: Paint secondary structure type

Right-click: Select secondary structure type (helix, sheet, loop)

Pull Mode

Left-drag: Pull

Middle-click: Lock single segment

Middle-drag: Add rubber band

Right-click: Context menu

Right-drag sidechain: Add disulfide to cysteine (only on some puzzles)

Tab: Show residue information

Left: Cycle sidechain/Show residue information

Right: Cycle sidechain/Show residue information

Modes

- 1: Pull Mode
- 2: Structure Mode
- 3: Note Mode
- 4: Design Mode
- 5: Ligand Design Mode
- 7: Ligand View Mode
- 8: Ligand Reaction Design Mode

Actions

D: Disable rubber bands

E: Wiggle sidechains to improve the protein.

Ctrl+E: Reset structures

F: Freeze or unfreeze all

L: Show Alignment

Ctrl+L: Show Level Hints

M: Mutate sidechains to improve the protein. Only available in design puzzles.

R: Remove rubber bands

S: Shake sidechains to improve the protein

Close

Shake

Mutate

Wiggle All

Wiggle Backbone

Wiggle Sidechains

Help

Glossary

Freeze Protein

Remove Bands

Disable Bands

Reset Structures

Auto Structures

Reset Puzzle

Rama Map

▲ **Actions**

► **Undo**

► **Modes**

► **Behavior**

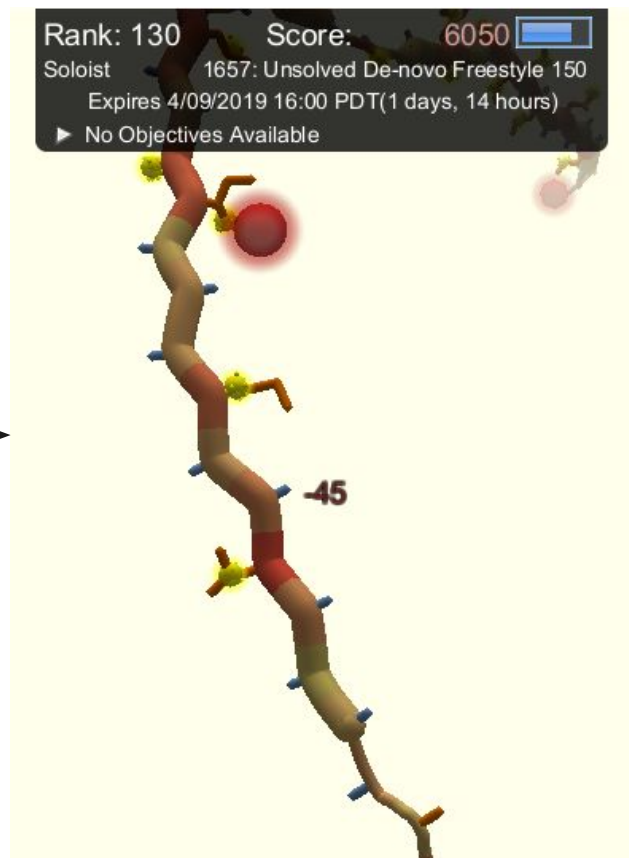
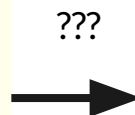
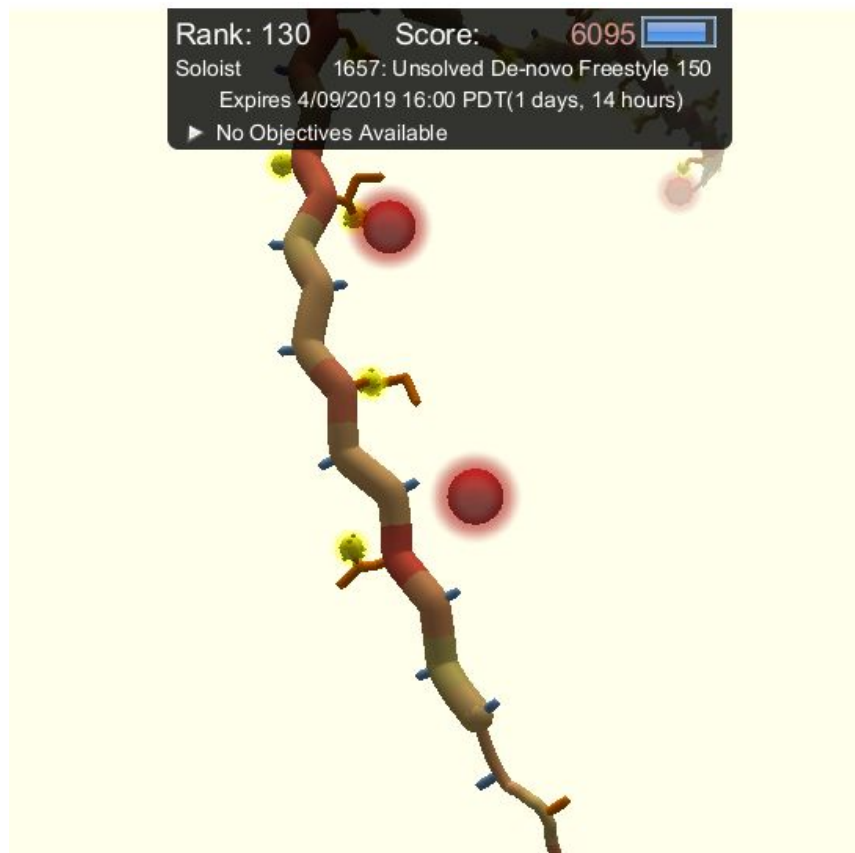
► **View**

► **Menu**

► Chat - Global

► Notifications

FoldIt, Level: "Unsolved De-novo Freestyle 150"



FoldIt, Level: "Unsolved De-novo Freestyle 150"

 [Nucleotide Mixer] Level 2 of 11

5 OR MORE
0



CHAT ONLINE (26) [Go to Home](#)

[12:17 PM]

mrdespacito: despacito

[12:18 PM]

mrdespacito: despacito

[12:18 PM]

mrdespacito: despacito

[12:18 PM]

mrdespacito: despacito

[12:18 PM]

mrdespacito: despacito

[12:18 PM]

mrdespacito: despacito

[12:18 PM]

Please log in to chat

<https://eternagame.org/game/puzzle/6502927/>

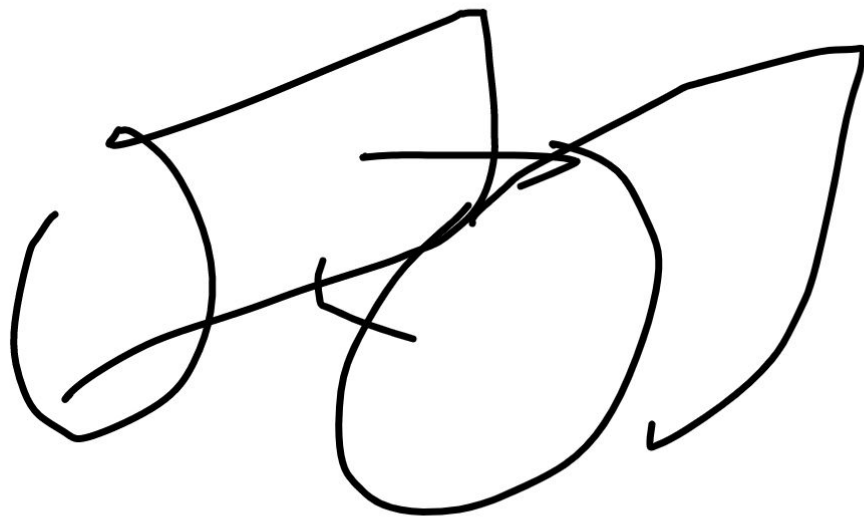


Can a neural network learn to recognize doodling?

Help teach it by adding your drawings to the [world's largest doodling data set](#), shared publicly to help with machine learning research.

Draw
binoculars
in under 20 seconds

Got It!



Oh I know, it's binoculars!

<https://quickdraw.withgoogle.com/>

You were asked to draw goatee

You drew this, and the neural net recognized it.



It also thought your drawing looked like these:

Correct match
goatee



2nd closest match
panda



3rd closest match
beard



How does it know what goatee looks like?
It learned by looking at these examples drawn by other people.





Breakout Groups (~1-2 min)

What impact does overtraining have on creativity?



What FoldIt does well

Creates small cadre of extremely motivated players

Seamlessly integrates machine learning

Effective at discovering some new protein structures [1]

Top Groups

Search:

Find

| RANK | GROUP | GLOBAL SCORE |
|------|------------------------|--------------|
| #1 | Beta Folders | 3360 |
| #2 | Anthropic Dreams | 3195 |
| #3 | Contenders | 2472 |
| #4 | Go Science | 2365 |
| #5 | Gargleblasters | 2190 |
| #6 | Void Crushers | 1290 |
| #7 | L'Alliance Francophone | 963 |
| #8 | Marvin's bunch | 838 |
| #9 | Russian team | 523 |
| #10 | Hold My Beer | 411 |
| #11 | Hun-Magyar Csapat | 408 |
| #12 | DW 2020 | 172 |
| #13 | GENE 433 | 146 |
| #14 | FoldIt@Netherlands | 143 |
| #15 | FoldIt@Poland | 72 |
| #16 | freefolder | 44 |
| #17 | Team South Africa | 32 |

<https://fold.it/portal/groups>

BA in Religious Studies and Philosophy

Member of Sacred Heart Fathers and
Brothers 1980-1985

Makes Adventure Studio adventure
games.

Married 10 years

semi-retired carpenter

I have 25 years of professional experience in IT, rewrote ray tracing software in x86 assembly language, built a 3d interface for LCD shutter glasses so I could play Descent stereoscopically in 1996, optimized Raistmer's SETI at home in SSSE3 assembly, created and sold computer art, wrote up and filed patents, trademarks, domains.

I like to push frontiers and discover new ones.

I have funded and built a series of workstations exclusively for use with Foldit, running 24/7 reaching a historical 25.000.000 moves on may 7th 2012 at 15:45 UTC. Just over a year later Fold.it introduced a new achievement to honor breaking a new frontier, the 50 Megamoves limit: <http://fold.it/portal/node/994984>

When the management lends more validity to the rantings of a spoiled, **malicious teenager** over the actions of **player with six years and thirty thousand hours of commitment**, it's time to move on to something more worthy of my involvement.

I'm a **retired** Computer Systems Analyst after having worked in IT for 30 years.

~~Struggling to maintain my sanity in spite of the Foldit client.~~

My sanity has won out. I've stopped playing Foldit until some of the legion client bugs have been fixed. Foldit Central needs to get its priorities straight--concentrate on the user experience for a change. Instead of piling on buggy new "features", fix the major issues already present.

There's an old parable about building your house on sand . . .

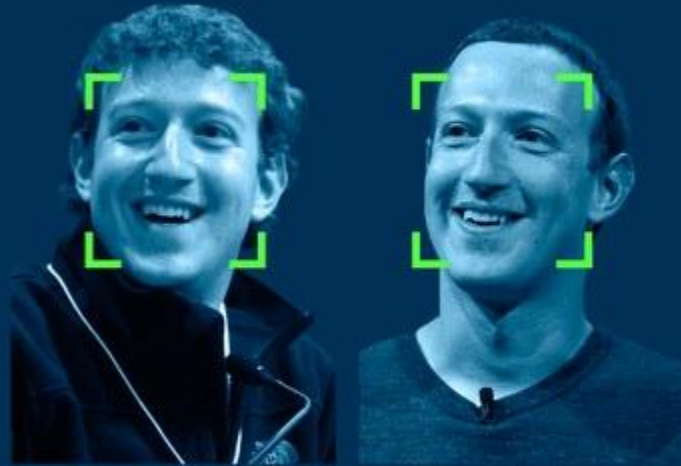
https://fold.it/portal/user/<user_id>



Breakout Groups (~1-2 min)

Design a more effective/usable distributed critique system for FoldIt

#10yearschallenge



Alyssa Foote: Getty Images (<https://www.wired.com/story/facebook-10-year-meme-challenge/>)



Kate O'Neill ✓

@kateo

Follow



Me 10 years ago: probably would have played along with the profile picture aging meme going around on Facebook and Instagram

Me now: ponders how all this data could be mined to train facial recognition algorithms on age progression and age recognition

1:25 PM - 12 Jan 2019

11,250 Retweets **24,983** Likes



349



11K



25K





Think-Pair-Share (30-60-N)

How can we encourage creative interactions with accessible crowdsourcing (e.g. QuickDraw)?

How can we encourage creative interactions with viral content (via gamification or otherwise)?



Think-Pair-Share (30-60-N)

Is there an ideal balance of humans as sensors and as collaborative creators in collective intelligence?

What compensation structures need to be created / altered to give contributors economic/social agency?



Class Commentary: FoldIt

Proper feedback metrics

How do you **augment human intelligence** in a collective setting?

Interface vs. problem **abstraction** for varying skill levels