

Collective Intelligence

Ariel Weingarten/Kandarp Khandwala

wiki

shared or group intelligence
that emerges from the
collaboration, collective efforts,
and competition
of many individuals

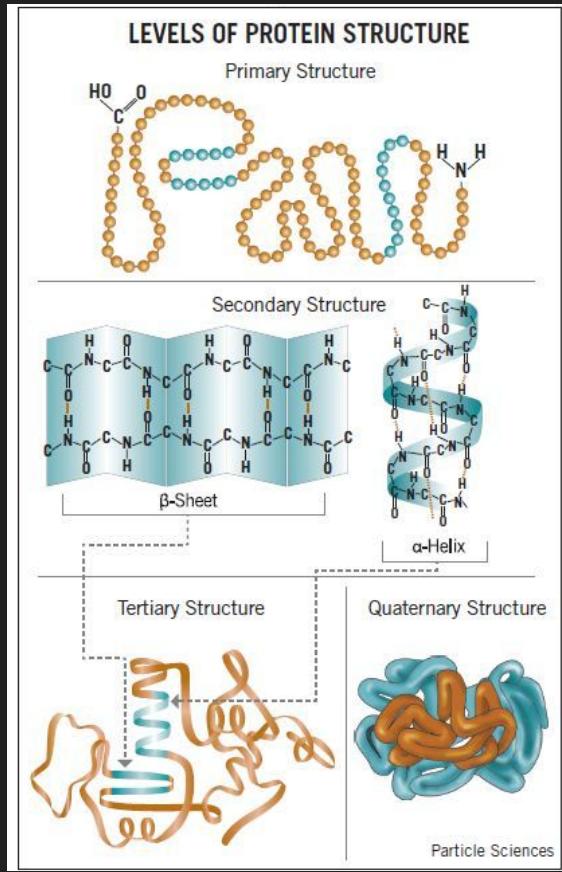
goals

understand the challenges of equipping
inconsistently skilled individuals

explore the difference between domain experts
and the people recruited to help them

explore a future where games with a purpose
and cheap, on-demand crowd work
are a pervasive part of society

Predicting protein structures with a multiplayer online game



**predicting protein structure from
amino acid sequence:**

template-based modeling,
ab initio (from the beginning)

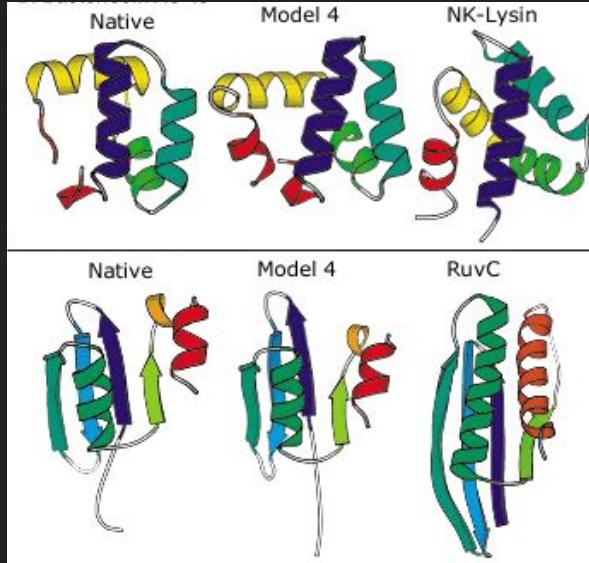
rosetta methodology

stochastic algorithms:

fragment finding and assembly

deterministic:

energy minimization, structure perturbation and refinement



extent of free energy landscape
order of 1000 dof,
computationally-limited

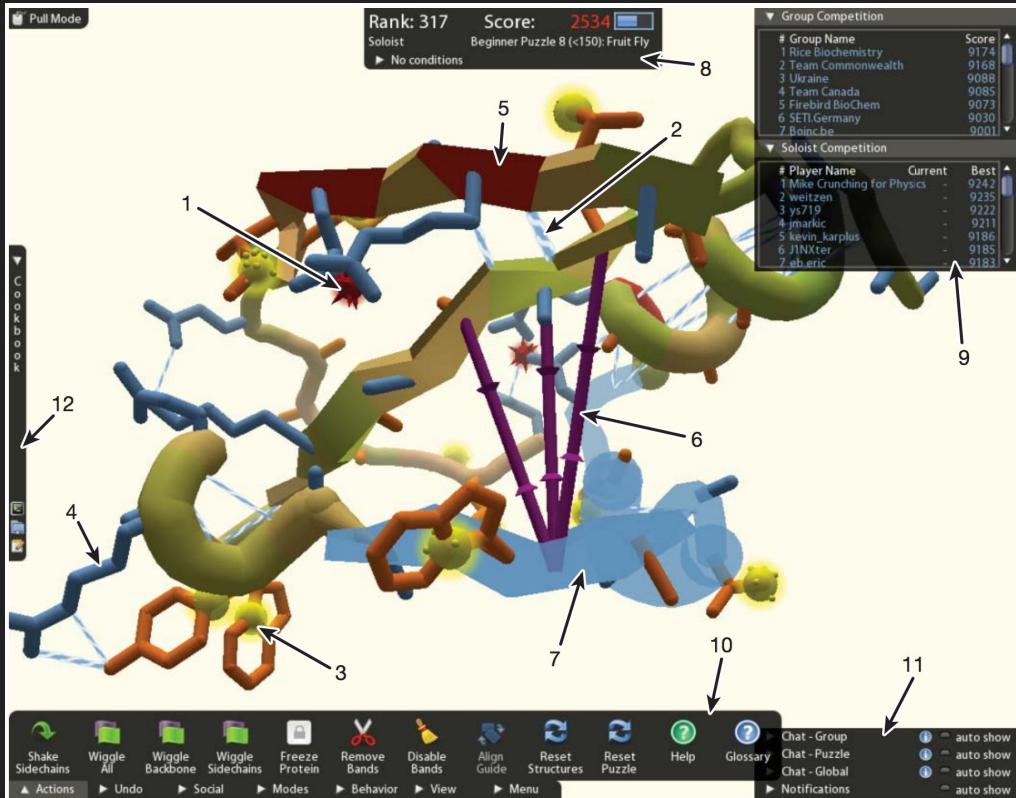
hypothesis
replace stochastic components with
human decision making...

how does it work?

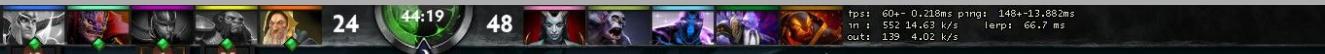
improper conformations: *puzzles*

leaderboard: *competition*

(players also collaborated in teams)



What do you think
about getting
someone up to
speed on this?



KILLED BY
k1



Crimson Guard of Prosperity

k1 (Ember Spirit) just pwned I WANT KM3 RELEASE DATE (Skywrath Mage)'s head for 326 gold!

A tower was destroyed by The Dire!

B
LEVEL UP +1

SKYWRATH MAGE



14

26/1500

Respawning in 48 seconds

103
2 420
38+3
28+3
74+13

Q 70

W 110

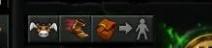
E 100

R 375

ITEMS

737
SHOP

Shift-click shop items to set
Quick Buy



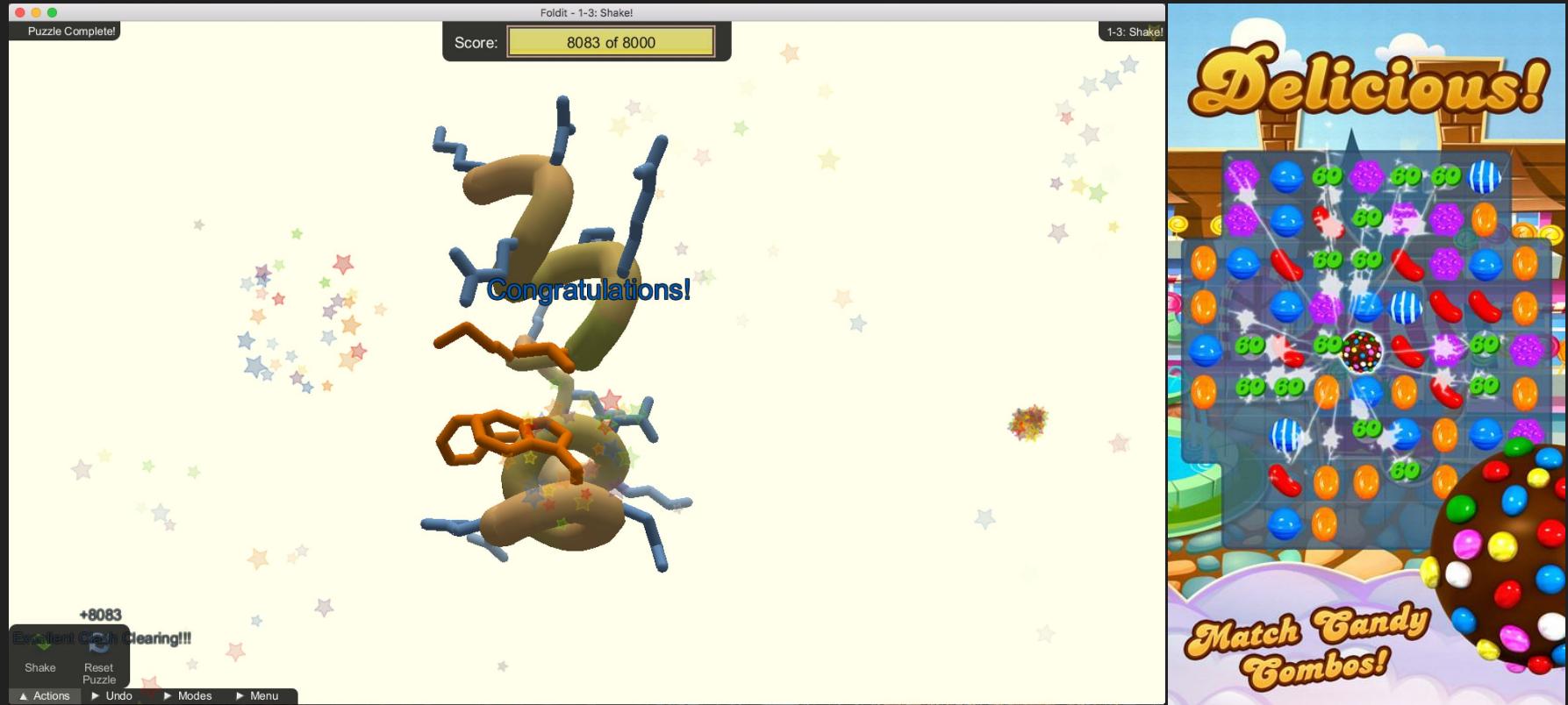
K/D/A 3/8/7
LH/D 51/7



Thoughts about game playing
mentality in empirical research...

(discuss with neighbor for 1 minute)

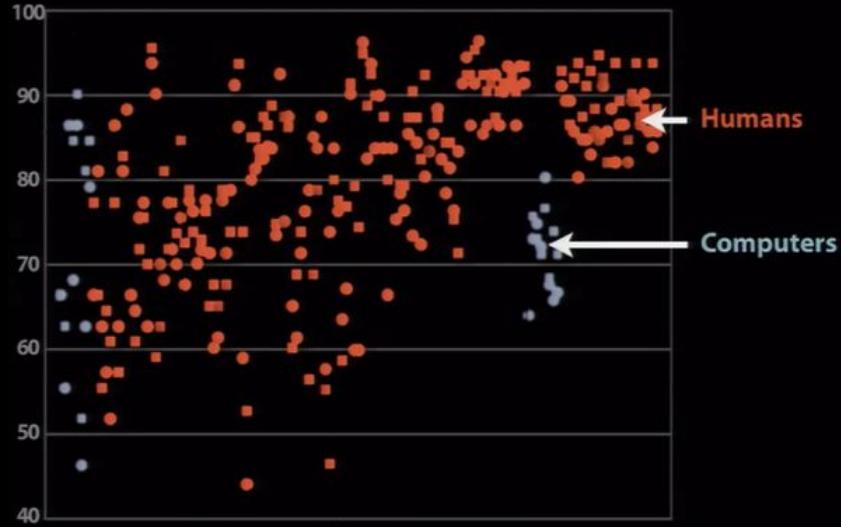
demo



experiment posed 10 “blind” puzzles compared similarity to native

Table 1 | Blind data set

Puzzle ID	Foldit C α r.m.s.d.	Rebuild and refine C α r.m.s.d.	Native	Method	Number of residues
986875	1.4	4.5	2kpo	NMR	99
986698	1.8	3.7	2kky	NMR	102
986836	5.7	6.6	3epu	X-ray	136
987088	3.5	4.3	2kpt	NMR	116
987162	4.5	5.2	3lur	X-ray	158
987076	3.3	3.5	2kpm	NMR	81
986629	3.5	3.3	2kk1	NMR	135
987145	2.6	2.3	3nuf	X-ray	105
986844	6.9	5.8	2ki0	NMR	36
986961	10.6	5.7	2knr	NMR	118



Are these games efficient if one were to consider the return on player time?

Are these games efficient if one were to consider the return on player time?

Gabriel: what happens when the novelty wears off?

Could one achieve the same result with paid experimenters?

Dorothy: who owns the intellectual property?

Francesco: some applications favor computers

Janet: does not seem to help learning

how did we do?

riskier steps (worse, then better)

more useful starting point

better at resolving incorrect features

Hundreds of thousands of people have come together to solve puzzles, making it a successful academic platform.
What do you think the 'secret ingredient'?

Hundreds of thousands of people have come together to solve puzzles, making it a successful academic platform.

What do you think the 'secret ingredient'?

Stephanie: gamers want to appear smart

Eric: challenge drives people

Many of you: what about attrition?

improvements

iteration with player tools
social praise (forum)

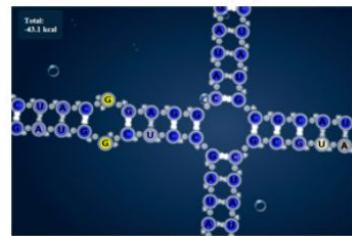
A Design interface



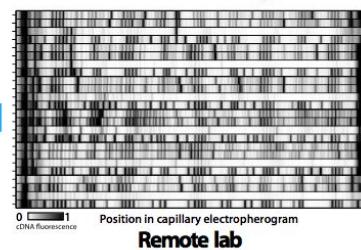
B Voting interface

	5
Ding's Branches 1	G G A A A G G U
Branches	G G A A A G C U
Squid	G G A A A C U C
Mat - Branches V1	G G A A A G G C
Y oh Y	G G A A A G C A
Deck the Halls	G G A A A G C G
JP-11-4-18	G G A A A C C A
Chaeen 202-	G G A A G C C U
Toad Hall	G G A A A G G U
Starry's Branches II	G G A A A G C C
Mat - Branches V1.1 (Submit..)	G G A A A G G C

D Results viewer



C Synthesis results



E Design rule collection

	aldo's Loops & Stacks	0.406	40	length
	deivid's deivid's strategy	0.382	60	length

Soylent: a Word Processor with a Crowd Insider

REFERENCES

1. Bernstein, M., Marcus, A., Karger, D.R., and Miller, R.C. Enhancing Directed Content Sharing on the Web. *CHI '10*, ACM Press (2010).
2. Bernstein, M., Tan, D., Smith, G., Czerwinski, M., et al. Collabio: A Game for Annotating People within Social Networks. *UIST '09*, ACM Press (2009), 177–180.
3. Bigham, J.P., Jayant, C., Ji, H., Little, G., et al. VizWiz: Nearly Real-time Answers to Visual Questions. *UIST '10*, ACM Press (2010).
21. Quinn, A.J. and Bederson, B.B. A Taxonomy of Distributed Human Computation.
22. Ross, J., Irani, L., Silberman, M.S., Zaldivar, A., et al. Who Are the Crowdworkers? Shifting Demographics in Amazon Mechanical Turk. *alt.chi '10*, ACM Press.
23. Sala, M., Partridge, K., Jacobson, L., and Begole, J. An Exploration into Activity-Informed Physical Advertising Using PEST. *Pervasive '07*, Springer Berlin Heidelberg (2007).

10

24. Simon, I., Morris, D., and Basu, S. MySong: automatic accompaniment generation for vocal melodies. *Proc. CHI '08*, ACM Press (2008).
25. Snow, R., O'Connor, B., Jurafsky, D., and Ng, A.Y. Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks. *ACL '08*, (2008).
26. Sorokin, A. and Forsyth, D. Utility data annotation with Amazon Mechanical Turk. *CVPR '08*, (2008).
27. von Ahn, L. and Dabbish, L. Labeling images with a computer game. *CHI '04*, ACM Press (2004).

Shortening A Paper to Ten Pages

- 1) Do it yourself
- 2) Use an AI
- 3) Ask colleagues



REFERENCES

- Bernardini, M., Marcus, A., Berger, D.R., and Miller, R.C. Enhancing Directed Content Sharing on the Web. *CHI '10 ACM SIGART* (2010).
- Bernardini, M., Goriely, L., and Goriely, G. Crowdsourcing for Collective Action: A Game for Collaborative Problem Solving. In *Proceedings of the Twelfth International Conference on Web Information Systems Engineering*, 7–18 (2011).
- Bigham, J.P., and C.J.H. Lin. *What's My Name? Early Real-time Answers to Visual Questions*. *UCB Tech Report*, 6, ACM Press (2010).
- Quinn, A., and Bederson, E. A Taxonomy of Distributed Human Computation. *CHI '10 ACM SIGART* (2010).
- Ross, J., and Silberman, S., Zaldivar, J. Who Are You? Crowdwork for Gathering Demographic Information. *CHI '10 ACM SIGART* (2010).
- Sorokin, A., and Forsyth, D. Gesture data annotation with Amazon Mechanical Turk. *CVPR '08*, (2008).
- von Ahn, L., and Dabbish, L. Designing interfaces for computer games. *CHI '03 ACM SIGART* (2003).

10

- Simpson, Morris, D., and Pampu, S. MySinger: Automatic acoustic feature generation for vocal music analysis. *Proc. CHI '08 ACM SIGART* (2008).
- Snow, C., O'Connor, B., Jurafsky, D., and Ng, A.Y. Cheap and fast: bootstrapping word embeddings for relation extraction at the national task. *NAACL '08*, (2008).
- Sorokin, A., and Forsyth, D. Gesture data annotation with Amazon Mechanical Turk. *CVPR '08*, (2008).
- von Ahn, L., and Dabbish, L. Designing interfaces for computer games. *CHI '03 ACM SIGART* (2003).

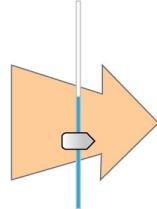
Shortening A Paper to Ten Pages

4) Recruit a crowd

Soylent is a word processing interface that uses crowd contributions to aid complex writing tasks.

Shortn

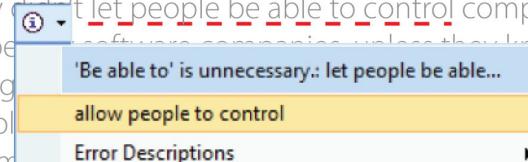
This paper introduces architectural and interaction patterns for integrating crowdsourced human contributions directly into complex endeavours that span many levels of conceptual and pragmatic activity. Authoring tools offer help with pragmatics. We thus present Soylent, a word processing interface that uses crowd contributions to aid complex writing tasks. We introduce the Find-Fix-Verify crowd programming pattern, feasibility, cost, wait time, and work time for edits.



This paper introduces architectural and interaction patterns for integrating crowdsourced human contributions directly into complex endeavours that span many levels of conceptual and pragmatic activity. Authoring tools offer help with pragmatics. We thus present Soylent, a word processing interface that uses crowd contributions to aid complex writing tasks. We introduce the Find-Fix-Verify crowd programming pattern, feasibility, cost, wait time, and work time for edits.

Crowdproof

Software made computers more intuitive, but they didn't let people be able to control computers effectively. Software companies expect people to know how to use them. We thus present Soylent, a word processing interface that uses crowd contributions to aid complex writing tasks. We introduce the Find-Fix-Verify crowd programming pattern, feasibility, cost, wait time, and work time for edits. This paper introduces architectural and interaction patterns for integrating crowdsourced human contributions directly into complex endeavours that span many levels of conceptual and pragmatic activity. Authoring tools offer help with pragmatics. We thus present Soylent, a word processing interface that uses crowd contributions to aid complex writing tasks. We introduce the Find-Fix-Verify crowd programming pattern, feasibility, cost, wait time, and work time for edits.



The Human Macro

Write a request:

Find Creative Commons figure for paragraph

This paper introduces architectural and interaction patterns for integrating crowdsourced human contributions directly into complex endeavours that span many levels of conceptual and pragmatic activity. Authoring tools offer help with pragmatics. We thus present Soylent, a word processing interface that uses crowd contributions to aid complex writing tasks. We introduce the Find-Fix-Verify crowd programming pattern, feasibility, cost, wait time, and work time for edits. This paper introduces architectural and interaction patterns for integrating crowdsourced human contributions directly into complex endeavours that span many levels of conceptual and pragmatic activity. Authoring tools offer help with pragmatics. We thus present Soylent, a word processing interface that uses crowd contributions to aid complex writing tasks. We introduce the Find-Fix-Verify crowd programming pattern, feasibility, cost, wait time, and work time for edits.



demo

“The Human Macro”: great idea or
research paper gimmick?

(discuss with neighbor for 1 minute)

Challenges in Programming Crowds

This project has interacted with
~9000 Turkers on ~2000 different tasks

Key Problem: crowd workers often produce
poor output on open-ended tasks

30% Rule:
~30% of the results
from open-ended tasks
will be unsatisfactory

Two Personas: An Example

Proofread and correct
the following paragraph:

The theme of loneliness features throughout many scenes in Of Mice and Men and is often the dominant theme of sections during this story. This theme occurs during many circumstances but is not present from start to finish. In my mind for a theme to be pervasive is must be present during every element of the story. There are many themes that are present most of the way through such as sacrifice, friendship and comradship. But in my opinion there is only one theme that is present from beginning to end, this theme is pursuit of dreams.

The Lazy Turker

Does as little work as necessary to be paid

The theme of loneliness features throughout many scenes in Of Mice and Men and is often the dominant theme of sections during this story. This theme occurs during many circumstances but is not present from start to finish. In my mind for a theme to be pervasive it must be present during every element of the story. There are many themes that are present most of the way through such as sacrifice, friendship and comradeship. But in my opinion there is only one theme that is present from beginning to end, this theme is pursuit of dreams.

The Eager Beaver

Go beyond task requirements to be helpful,
but introduce errors in the process

The theme of loneliness features throughout many scenes in Of Mice and Men and is often the principal, significant, primary, preeminent, prevailing, foremost, essential, crucial, vital, critical, dominant theme of sections during this story. This theme occurs during many circumstances but is not present from start to finish. In my mind for a theme to be pervasive is must be present during every element of the story. There are many themes that are present most of the way through such as sacrifice, friendship and comradeship. But in my opinion there is only one theme that is present from beginning to end, this theme is pursuit of dreams.



Find

“Identify at least one area
that can be shortened
without changing the
meaning of the paragraph.”



Fix

Verify

Find

“Identify at least one area that can be shortened without changing the meaning of the paragraph.”



Independent agreement to identify patches

Fix

“Edit the highlighted section to shorten its length without changing the meaning of the paragraph.”

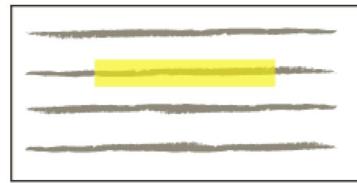


Soylent, a prototype...

Verify

Find

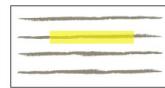
“Identify at least one area that can be shortened without changing the meaning of the paragraph.”



Independent agreement to identify patches

Fix

“Edit the highlighted section to shorten its length without changing the meaning of the paragraph.”



Soylent, a prototype...



Randomize order of suggestions

Verify

“Choose at least one rewrite that has style errors, and at least one rewrite that changes the meaning of the sentence.”

- Soylent **is**, a prototype...
- Soylent **is-a** prototypes...
- Soylent is a **prototypetest**...

Verify

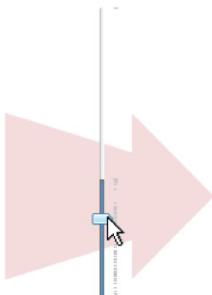
“Choose at least one rewrite that has style errors, and at least one rewrite that changes the meaning of the sentence.”

- Soylent **is**, a prototype...
- Soylent **is-a** prototypes...
- Soylent **is** a **prototypetest**...



Keep suggestions that do not get voted out

Automatic clustering generally helps separate different kinds of records that need to be edited differently, but it isn't perfect. Sometimes it creates more clusters than needed, because the differences in structure aren't important to the user's particular editing task. For example, if the user only needs to edit near the end of each line, then differences at the start of the line are largely irrelevant, and it isn't necessary to split based on those differences. Conversely, sometimes the clustering isn't fine enough, leaving heterogeneous clusters that must be edited one line at a time. One solution to this problem would be to let the user rearrange the clustering manually, perhaps using drag-and-drop to merge and split clusters. Clustering and selection generalization would also be improved by recognizing common text structure like URLs, filenames, email addresses, dates, times, etc.



Automatic clustering generally helps separate different kinds of records that need to be edited differently, but it isn't perfect. Sometimes it creates more clusters than needed, because the differences in structure aren't relevant to a specific task. | Conversely, sometimes the clustering isn't fine enough, leaving heterogeneous clusters that must be edited one line at a time. One solution to this problem would be to let the user rearrange the clustering manually using drag-and-drop edits. Clustering and selection generalization would also be improved by recognizing common text structure like URLs, filenames, email addresses, dates, times, etc.

What is one challenge that might arise from Find-Fix-Verify, and how would you address it by evolving the system?

Sean: the addition of a qualify phase that can test workers to identify ones that demonstrate sufficient knowledge to work on the task.

Michelle: I don't completely agree that crowdsourcing editing of high specialized or technical documents can be done to a satisfactory level.

Tushar: Another challenge that I can think of would be to handle multiple, conflicting edits over the same text.

What other applications could benefit from embedded crowd workers?

Danilo: Photoshop with "complex filters".

(discuss with your neighbours for 2 minutes)

Is Fix-Find-Verify only for
guiding Turkers?

(discuss with your neighbours for 2 minutes)

Ethical?

Jorge: Make use of the different currencies earn by different people in the world to profit from the least wealthy countries doing common and simple tasks is for me hard to digest.

Dorothy: I'm excited about the possibilities of crowdsourcing, but I'm worried that the rhetoric of innovation makes UX/ HCI designers blind to the human concerns.

(discuss with your neighbours for 2 minutes)

Did you form groups yet?