MATCHMAKERS: CROWD-SOURCING SOLUTIONS TO AN NP-HARD PROBLEM

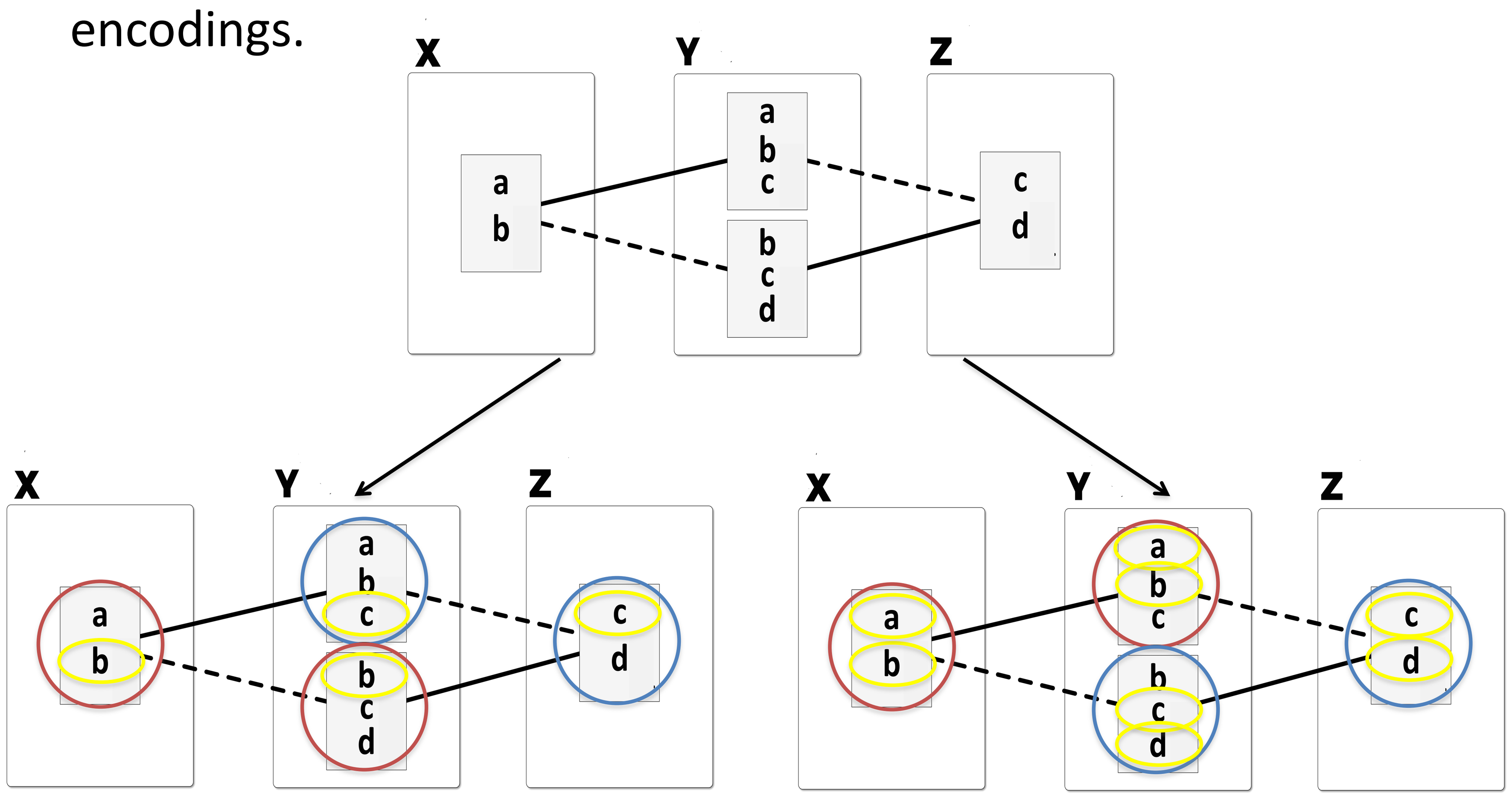
Amit Kadan[1], Christina Chung[1], Asako Matsuoka[1], and Yueti Yang[1]
Under the supervision of Prof. Marsha Chechik[1] and Prof. Julia Rubin[2]
[1]University of Toronto, [2]University of British Columbia

Overview

- **N-way matching consists of finding correspondences between elements of input models [1].**
  - Useful for software engineering practices such as merging branches of a software configuration management system [2].
- **N-way matching is NP-hard [1].**
  - Approximating solutions don't scale, heuristic solutions lack approximating factor of optimal solution.
- **Our approach: Encode problem in an engaging game where players try to improve solutions of best heuristic algorithm.**
  - Among first to look at applicability of *serious games* to solve computationally-intensive problems [1].

Problem: N-way Matching

- **Create n-way match by identifying elements from distinct inputs that share common properties [1].**
  - Input represented as models containing elements with properties.
  - Representation is high-level so solution can apply to many encodings.



sub-optimal optimal

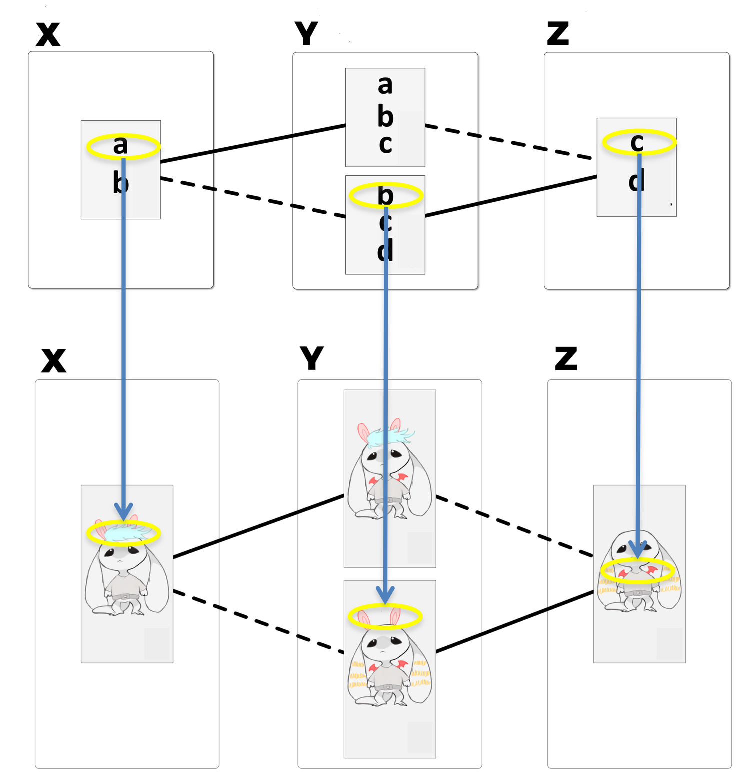
- **Figure shows example of matching options.**
  - "Optimal" is better because matched elements share more properties.
  - Optimal since no solution where matches share more properties.
- **NP-hard with more than two input models [1].**
  - Table below lists approximating and heuristic solutions.

Algorithm	Description	Approximation Factor	Feasible	Matching score (on real case)
Greedy	Pick best match over all models greedily.	✔	✖	N/A
Greedy + Local Search	Greedy with local search.	✔	✖	N/A
NwM*	Apply bipartite graph algorithm over all elements, repeat.	✖	✔	+1.80 %
Greedy (by chunk)	Apply greedy to a subset of models.	✖	✔	+1.15 %
Pairwise	Apply bipartite graph algorithm to 2 models. combine into one, repeat.	✖	✔ (state of practice)	+0.00 %

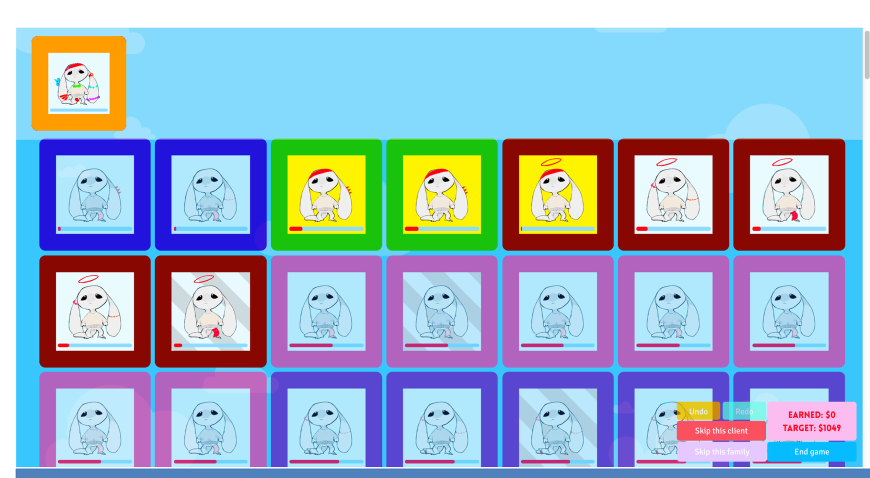
* Best heuristic solution thus far

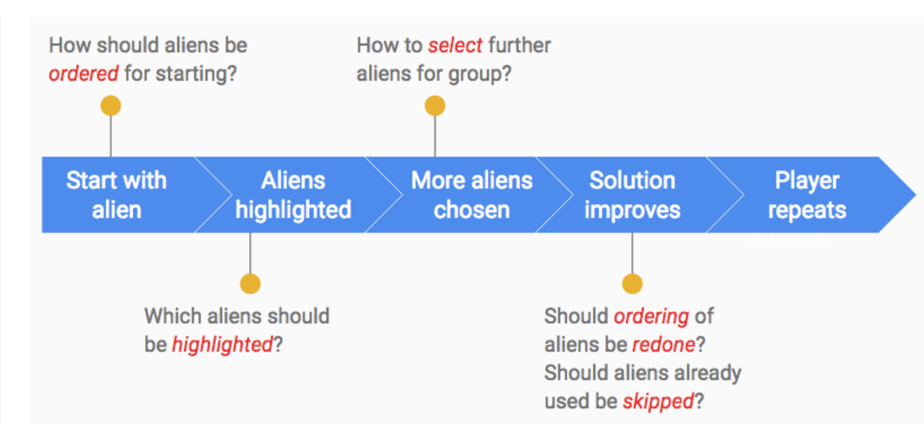
Solution: The Game

- **Our game, MATCHMAKERS encodes input elements as characters.**
  - Characters "visualize" properties.
  - Goal: match visually similar characters.
- **Game played in iterative manner.**
  - Player starts from best solution computed by NwM or another player.
  - Player modifies solution by recombining groups.
  - If player improves score, solution gets saved.



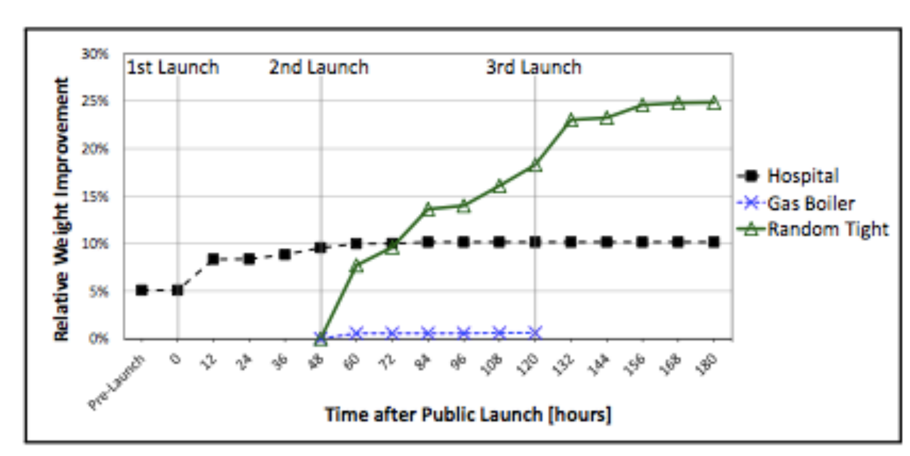
- **Success through game achieved in two ways**
  1. Many players improving on each other's results.
  2. Human ability to make snap judgments about visual scenes.
- **Tested many configurations to ensure good player support.**
  - By observing the way humans pay earlier versions of the game, built simulator to speed up evaluation of configurations.
  - Simulator identified feasibility of players improving heuristic solution.
  - Simulator itself was effective in improving the solution.





Results

- **Results**
  - Players played game over three real world cases.
  - In all cases, players improved the score of NwM.
  - Players improved on simulator score in one case.



Case	Simulator	MM (game)
hospital	+10.30 %	+10.16 %
Gas Boiler	+0.36 %	+0.61 %
random Tight	+32.79 %	+24.86%

- **Significance**
  - Provided two new ways of producing better matches than NwM.
  - Validated use of *serious games* for computational problems.
  - Showed power of crowd-sourced solutions.
- **Future**
  - Run longer user study to see if improvement still possible.
  - Implement machine learning algorithm that learns from players.

Acknowledgements: This project would not have been possible without Si Hua Cao Liu, Lionheart Xiong, Angel You, Fei Huang and Nicole Sultanam.

References

[1]. J. Rubin and M. Chechik, "N-Way Model Merging," in *Proc. ESEC/FSE*, 2013, pp. 301–311.

[2]. M.Hess,J.Wiemeyer,K.Hamacher,andM.Goesele,"Serious Games for Solving Protein Sequence Alignments - Combining Citizen Science and Gaming," in Proc. of the International Conference on Serious Games (GameDays'14), 2014.

Play our game MATCHMAKERS at matchmakers.herokuapp.com