Modeling Pipelines, Mechanistic and Data-Driven Agent-Based Models, to Explain Human Behavior in Online Networked Temporal Social Science Experiments

Vanessa Cedeno-Mieles
Co-Chairs: Madhav V. Marathe and Chris J. Kuhlman
Acknowledgment: Brian Goode and Naren Ramakrishnan for building the online game platform and conducting online experiments.
Computer Science Department, Virginia Tech, Blacksburg, USA
Escuela Superior Politécnica del Litoral, Guayaquil, Ecuador

espol Escuela Superior Politécnica del Litoral

Virginia Tech. Biocomplexity Institute

Introduction

Background and Motivation

- Significant growth in online social science experiments in order to understand behavior at-scale, with finer-grained data collection.
- Motivated by a particular social behavior: Collective identity (CI) is an individual's cognitive, moral, and emotional connection with a broader community.
- Anagram games have been extensively studied to analyze problems such as, effects of goal-setting, CI, test anxiety.
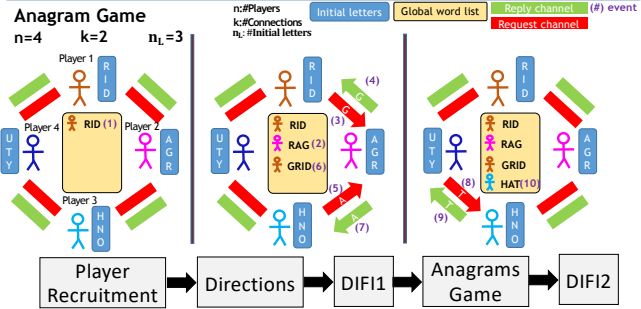
Challenges

- Considerable work is required to perform data analytics for custom experiments.
- The experiment is intricate with different types of interactions.
- CI is difficult to measure directly.
- Understanding and interpreting the proxies for CI is challenging.
- No process for combining mechanistic and data-driven approaches to build models of human reasoning.

Contributions

- Formal data model for networked social science experiments and simulation modeling.
- Five extensible pipelines for modeling and simulation, and analysis, of controlled networked experiments.

①

- Insights on a collaborative anagram game.
- Data-driven networked agent-based models (ABMs) of experiments: design, construction, and evaluation.
- Specification and demonstration of iterative abductive analysis process.
- New experimental understanding of the formation of collective identity (CI).

②

- A process for combining mechanistic and data-driven approaches to build models of human reasoning, based on group anagram games experimental data.
- New experimental findings and explanations of player behaviors based on cognitive and economic theories.
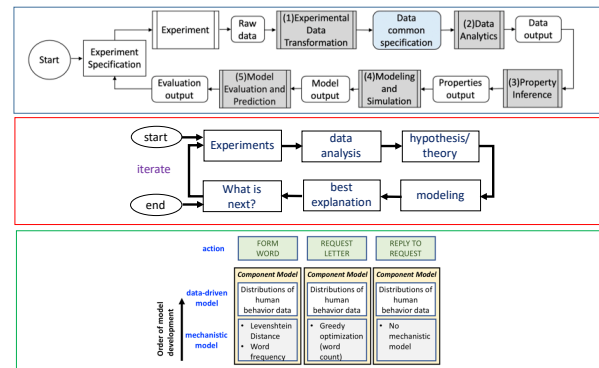- Mechanistic and data-driven modeling of group anagram game.
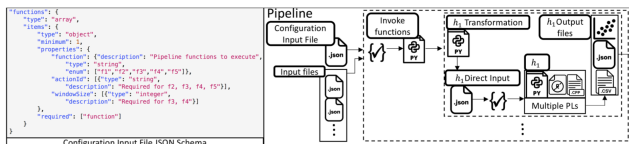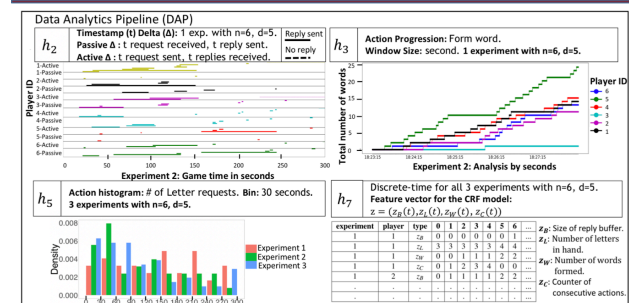
③

Experiments: Raw Data

Anagram Game

n: #Players
k: #Connections
n_L: #Initial letters

n=4 k=2 n_L=3

Player Recruitment → Directions → DIFI1 → Anagrams Game → DIFI2

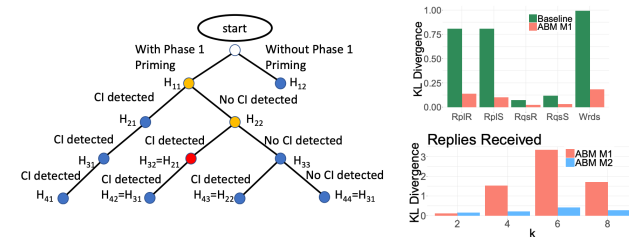## 1 Pipelines Implementation

Configuration Input File JSON Schema

- Pipelines are written in Python.
- Each pipeline takes a configuration file as input.
- Each pipeline has a list of available functions.

Graph Dynamical System (GDS)

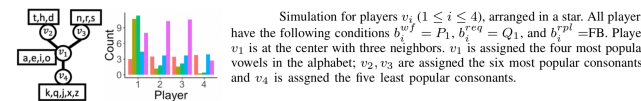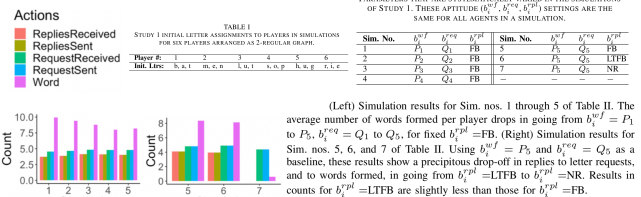Local functions: Next action? $F = (f_i)_{i \in n}$

Action Set

Update Scheme: At each time t, execute at each f_i for all $v_i \in V$ in parallel.

Data Analytics Pipeline (DAP)

h_2 Timestamp (t) Delta (Δ): 1 exp. with n=6, d=5.
Passive Δ : t request received, t reply sent.
Active Δ : t request sent, t replies received.

Experiment 2: Game time in seconds

h_3 Action Progression: Form word.
Window Size: second. 1 experiment with n=6, d=5.

Experiment 2: Analysis by seconds

h_5 Action histogram: # of Letter requests. Bin: 30 seconds.
3 experiments with n=6, d=5.

h_7 Discrete-time for all 3 experiments with n=6, d=5.
Feature vector for the CRF model:
$z = (z_B(t), z_L(t), z_W(t), z_C(t))$

z_B: Size of reply buffer.
z_L: Number of letters in hand.
z_W: Number of words formed.
z_C: Counter of consecutive actions.

2 Iterative Abduction Framework

3 Mechanistic and Data-Driven Agent-Based Models

- Experimental data correspond exactly with the mechanistic model.

20%: 583   40%: 1120   60%: 1360   80%: 1057   100%: 1066

Simulations of Group Anagram Game

Actions
- RepliesReceived
- RepliesSent
- RequestReceived
- RequestSent
- Word

TABLE I
Study 1 initial letter assignments to players in simulations for six players arranged as 2-regular graph.

Player #:	1	2	3	4	5	6
Init. Ltrs:	b, a, t	m, e, n	l, u, i	s, o, p	h, a, g	r, i, c

TABLE II
Parameters that are systematically varied in the simulations of Study 1. These attitude ($b_i^{wf}, b_i^{req}, b_i^{rpl}$) settings are the same for all agents in a simulation.

Sim. No.	b_i^{wf}	b_i^{req}	b_i^{rpl}	Sim. No.	b_i^{wf}	b_i^{req}	b_i^{rpl}
1	P_1	Q_5	FB	5	P_5	Q_5	FB
2	P_2	Q_5	FB	6	P_5	Q_5	LTFB
3	P_3	Q_5	FB	7	P_5	Q_5	NR
4	P_4	Q_5	FB	—	—	—	—

(Left) Simulation results for Sim. nos. 1 through 5 of Table II. The average number of words formed per player drops in going from $b_i^{wf} = P_1$ to P_5, $b_i^{req} = Q_5$ to Q_5, for fixed b_i^{rpl} =FB. (Right) Simulation results for Sim. nos. 5, 6, and 7 of Table II. Using $b_i^{wf} = P_5$ and $b_i^{req} = Q_5$ as a baseline, these results show a precipitous drop-off in replies to letter requests, and to words formed, in going from b_i^{rpl} =LTFB to b_i^{rpl} =NR. Results in counts for b_i^{rpl} =LTFB are slightly less than those for b_i^{rpl} =FB.

Simulation for players v_i ($1 \le i \le 4$), arranged in a star. All players have the following conditions $b_i^{wf} = P_1$, $b_i^{req} = Q_1$, and b_i^{rpl} =FB. Player v_1 is at the center with three neighbors. v_1 is assigned the four most popular vowels in the alphabet; v_2, v_3 are assigned the six most popular consonants, and v_4 is assigned the five least popular consonants.

Conclusions

- A set of five composable and extensible pipelines for studying networked social science phenomena has been presented, along with data and computational models for formal specification of experiments and Modeling and Simulation (MAS), for a particular class of networked social science experiments.
- We established the potential of iterative abductive looping for the (computational) social sciences.
- We have presented a process for combining mechanistic and data-driven approaches to build models of human reasoning based on group anagram games experimental data.

References

[1] Vanessa Cedeno-Mieles, Yihui Re, Zhihao Hu, Xinwei Deng, Saliya Ekanayake, Brian J. Goode, Chris J. Kuhlman, Dustin Machi, Madhav V. Marathe, Henning H. Mortveit, Naren Ramakrishnan, Parang Saraf, Nathan Self, Noshir Contractor, Joshua M. Epstein, Michael W. Macy. "Pipelines and Their Compositions for Modeling and Analysis of Controlled Online Networked Social Science Experiments". Winter Simulation Conference (WSC) 2018

[2] Yihui Ren, Vanessa Cedeno-Mieles, Zhihao Hu, Xinwei Deng, Abhijin Adiga, Christopher Barrett, Saliya Ekanayake, Brian J. Goode, Gizem Korkmaz, Chris J. Kuhlman, Dustin Machi, Madhav V. Marathe, Naren Ramakrishnan, S. S. Ravi, Parang Saraf, Nathan Self, Noshir Contractor, Joshua M. Epstein, Michael W. Macy. "Generative Modeling of Human Behavior and Social Interactions Using Abductive Analysis". IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) 2018

[3] Vanessa Cedeno-Mieles, Zhihao Hu, Yihui Ren, Xinwei Deng, Abhijin Adiga, Christopher Barrett, Saliya Ekanayake, Gizem Korkmaz, Chris J. Kuhlman, Dustin Machi, Madhav V. Marathe, S. S. Ravi, Brian J. Goode, Naren Ramakrishnan, Parang Saraf, Nathan Self, Noshir Contractor, Joshua M. Epstein, and Michael W. Macy. "Mechanistic and Data-Driven Agent-Based Models to Explain Human Behavior in Web-Based Group Anagram Games". IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) 2019