Human Centricity and Norm Awareness in Cognitive Systems

Ph.D. Dissertation

Sz-Ting “Christine” Tzeng

(Under the supervision of Professor Munindar P. Singh)

2 August, 2023

Department of Computer Science

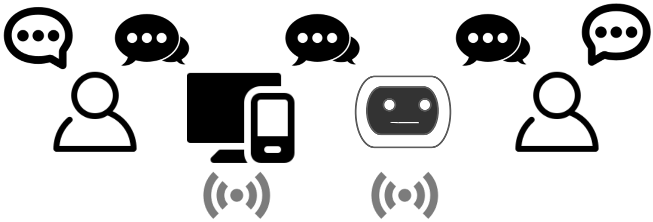
NC STATE UNIVERSITY

Bio

Education

- Ph.D. in Computer Science, North Carolina State University, 2018 Fall to present
  - Norm Emergence in Systems of Cognitive Agents with Emotions and Values, proposed October 2021
  - Impact of Agent Interactions on Policy Selection in Social Dilemma Simulations, qualified September 2020
- M.Sc. in Computer Science, North Carolina State University, May 2018
- B.Sc. in Computer Engineering, National University of Tainan, Taiwan, June 2010

Introduction: Motivations



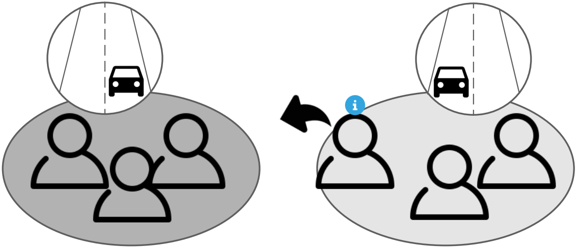
- Humans and agents form a multiagent system
- Norms regulate agent behaviors with sanctions
- What happens if humans are in the loop?
- What happens if human needs or the environment change over time?

Introduction: Challenges for Humans in Loop

- Human factors influence decisions and experience
  - The five human factors: **social, cognitive, emotional**, physical, and cultural
- Sanctions are often subtle, e.g., emotional expression or social exclusion
- Social signals have emerged in the form of verbal messages or subtle hints, transmitting normative information
- Values differ from person to person
- People need to comprehend and trust in AI output

Introduction: Challenges for Changing Environment

- Changing requirements or environment
  - More interconnection in MAS $\rightarrow$ Complexity of interactions increases drastically
  - Norms may change over time or over the environment



Introduction: Research Objective

To accommodate humans in the loop and operate in dynamic environments

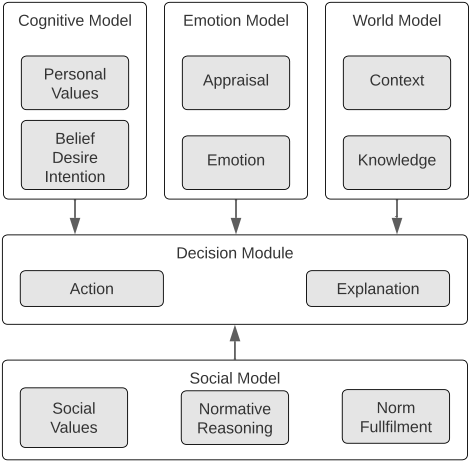
Introduction: Research Statement

Thesis Statement

AI systems that consider human factors, such as **emotional expressions, social signals, social value orientation, and value-aligned decisions and rationales**, are more adept at accommodating humans in the loop, thereby enhancing the social experience

- Investigate emotional responses of agents to the outcomes of interactions [COINE @ AAMAS'21]
- Investigate messages and hints as drivers of subtle social learning [In prep for JAIR]
- Investigate the influences of social value orientation [COINE @ AAMAS'22]
- Investigate value-aligned decisions and rationales [In prep for JAAMAS]

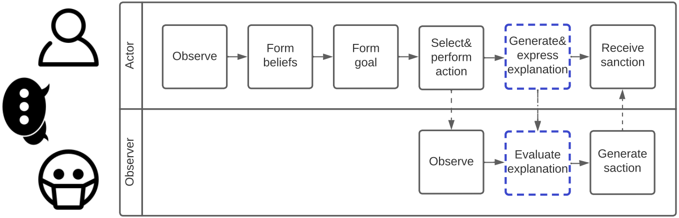
Cognitive Framework



Key concepts

- Notions of social norms
  - Prescriptive norms describe how an individual should behave
  - Descriptive norms describe how most agents actually behave
  - Representation:
    - Norm(subject, object, antecedent, consequent) (Singh, 2013)
- Norm emergence: The majority of agents in society choose the same action
- Cooperation: Conforming to the existing norms or most agents' behaviors

General Interaction Among Agents In MAS



*Dashed rectangles apply when any explanation involves

*Solid rectangles indicate the processes of general interactions

Fleur: Social Values Orientation for Robust Norm Emergence

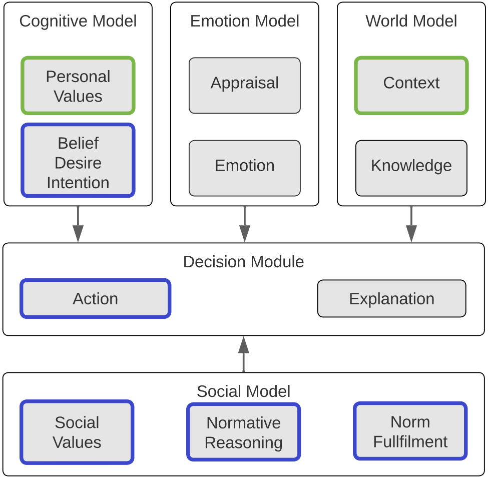
Fleur: Scenario

Source: https://twitter.com/springertoons/status/1281992099538165761

Fleur: Introduction and Motivation

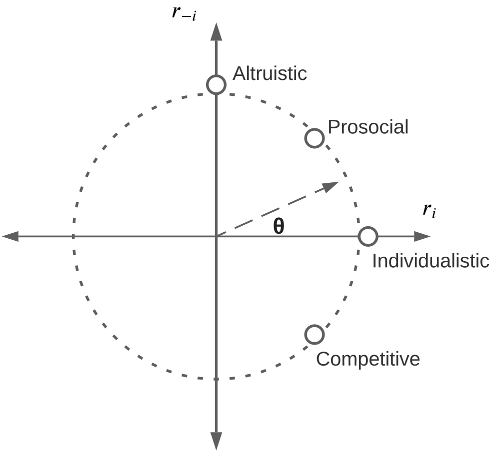
- Motivations
  - Interconnection in MAS indicates that one individual's behavior may affect another
  - Social Value Orientation (SVO): An individual's preference for resource allocation between self and others
  - Humans with different values evaluate the chosen actions subjectively and act to maximize their utility
- Objective: Incorporate individual preferences over self-interests and collective interests into decision-making
- $\mathrm{RQ}_{\mathrm{SVO}}$: How do social preferences, specifically social value orientation, influence norm compliance?

Fleur: Cognitive Framework



Fleur: SVO Ring with Reward Angle

Reward function of agent i: $reward_i = r_i \cdot \cos\theta + r_{-i} \cdot \sin\theta$



Fleur: Evaluation Design with COVID Simulations

- Scenario: Agents interact with one another and decide whether to wear a mask based on preference, health state, and SVO
- Simulate with varying agent societies: Altrustic, Prosocial, Proself, Competitive, and Mixed society
- Characteristics of agent society
    - Prescriptive norm: Mask-wearing mandate
    - Different distribution of social value orientation among agents
    - Agents' health states and the chosen action determine the payoff

Fleur: Metrics

- Compliance
  - $\mathrm{M}_{\text{Compliance}}$: The percentage of agents who satisfy the existing norm
- Social Experience
  - $\mathrm{M}_{\text{Social Experience}}$: The total payoff of the agents in a society
- Invalidation
  - $\mathrm{M}_{\text{Invalidation}}$: The percentage of agents who do not meet their preferences in a society

Fleur: Hypotheses

- $H_{\text{Compliance}}$: Social value orientation positively affects norm compliance with prosocial norms
- $H_{\text{Social Experience}}$: The distribution of social value orientation positively affects social experiences in a society
- $H_{\text{Invalidation}}$: Social value orientation negatively affects the tendency to meet personal preference and social experiences

Tests for Statistical Significance

- Independent t-test
- Glass's Δ

Result: Prosocial and Altruistic agents societies have higher compliance. A competitive infected agent may choose not to wear a mask when interacting with other healthy agents, leading to lower compliance in the mixed society

Compliance: % of agents who satisfy the existing norm

	S_{mixed}	$S_{altruistic}$	$S_{prosocial}$	$S_{selfish}$	$S_{competitive}$
Compliance	63.40%	69.70%	**70.25%**	65.10%	54.08%

Result: The mixed society has similar results as the selfish society. Whereas 50% of the mixed-agent society agents are altruistic and prosocial agents, the 25% of competitive agents would choose to minimize others' payoff without hurting their self-interests

Social Experience: The total payoff of the agents in a society

	S_{mixed}	$S_{altruistic}$	$S_{prosocial}$	$S_{selfish}$	$S_{competitive}$
Social Experience	0.448	0.554	**0.566**	0.470	0.221

Result: The selfish and competitive agents in the mixed society decreased the invalidation

Invalidation: % of agents who do not meet their preferences in a society

	S_{mixed}	$S_{altruistic}$	$S_{prosocial}$	$S_{selfish}$	$S_{competitive}$
Invalidation	0.296	**0.334**	0.323	0.269	0.289

Fleur: Summary

Study Summary

Incorporating Social Value Orientation enables better social experience and robust norm emergence

- Aligning with social preferences enables AI to make ethical decisions and be responsible for human needs
- Altruistic and prosocial agents adhere to the prosocial norm and enjoy more positive social experiences at the cost of themselves
- Policy makers may define appropriate sanctions to motivate the competitive and selfish agents to follow the norms

Exanna: Decision and Rationale with Values

Exanna: Scenario

Source: https://www.latimes.com/opinion/story/2022-08-07/mask-wearing-public-anger-comic

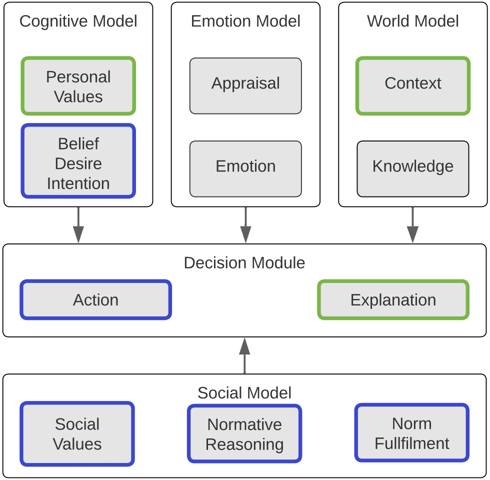
Exanna: Introduction and Motivation

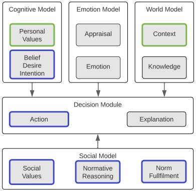
- Motivations
  - Agents provides rationales for their decisions would be interpretable and reliable
  - Justifying behavior via revealing information can resolve social conflict and enhance individual gain
  - Verbose rationales may be diverging and not convincing, leading to information overload
  - Invaded or uncomfortable feelings for privacy breaches
  - Rationales or information aligned with values best justify one's behaviors

Exanna: Objective and RQs

- Objective: Incorporating values into behavior justification
- RQs
  - $\text{RQ}_{\text{Goal Adherence}}$: Do value-aligned rationales increase adherence to the original goal?
  - $\text{RQ}_{\text{Conflict Resolution}}$: Do value-aligned rationales increase the social resolution?
  - $\text{RQ}_{\text{Privacy Loss}}$: Does value-aligned rationales reduce privacy loss?

Exanna: Cognitive Framework





Compare to Fleur

Exanna: Key Concepts

- Context:
  - The information that characterizes the situation of an entity
  - Include observable and nonobservable attributes (keep private from others)
  - Some attributes associated with values
- Decision rule: The mapping between an observation of context and a reasonable action, represented as if-then logic
  - Format: if *antecedent* then *consequence*

```
{InfectionRisk=No risk, InteractWith=Colleague} => Not Wear
```

- Rule Learning: Evolving rules from interactions or dataset
- Value preference
  - A preference order over different values for one context
  - Numbers in one value preference add up to 1

Exanna: Method (1)

- Decision making
  - Aggregated payoff with all corresponding values

$$f = \sum_{i}^{values} v_i \times r_{RxCy} \tag{1}$$

Agent 1:Agent 2	C1	C2
R1	r_{R1C1}	r_{R1C2}
R2	r_{R2C1}	r_{R2C2}

Exanna: Method (2)

- Rationale Generation
  - Evolve and learn decision rules as the base rationale with XCS, a learning algorithm combines reinforcement learning and genetic algorithm
    - Rule discover: Crossover and mutation creates more general or more specific rules by randomly adding or removing factors in antecedent
    - Subsume rules: Replace with a more general rule that has less prediction error
    - Action selection: Select the action with best-aggregated fitness
  - Reveal necessary information in rationales
    - Remove private factors that are not related to presented values from the aggregated rules

Exanna: Method (3)

- Rationale Evaluation
  - Update beliefs based on the received rationale
  - Make an analogous decision based on beliefs
  - Accept the rationale if the decision matches the observed action
  - Otherwise, reject the rationale
  - Acceptance and rejection of rationales lead to sanctions

Exanna: Evaluation Design with COVID Simulations

- Scenario: Agents move randomly and decide whether to wear a mask based on personal preference, health states, and value preference
- Simulate with varying agent societies: Share All, Share Decision Rules, and Share Value-Aligned Rules society
- Characteristics of agent society
  - Different strategies to explain agents' behaviors
  - Evaluate observed behaviors referring to received rationales
  - Agents form goals based on values
  - 50% of agents value health and 50% of agents value freedom in each society

Exanna: Metrics

Measures:

- Goal Adherence
  - $M_{\text{Goal Adherence}}$: The degree of adherence to each agent's goal
- Conflict Resolution
  - $M_{\text{Conflict Resolution}}$: The percentage of conflict resolution in society
- Social Experience
  - $M_{\text{Social Experience}}$: The aggregation of payoff an agent receives for its behavior
- Privacy Loss
  - $M_{\text{Privacy Loss}}$: The proportion of hidden information shared during an interaction

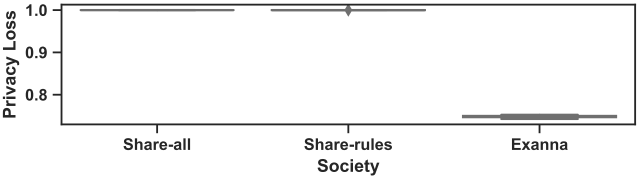
Exanna: Hypotheses

- $H_{\text{Goal Adherence}}$: Exanna provides higher goal adherence than baseline societies
- $H_{\text{Conflict Resolution}}$: Exanna provides higher conflict resolution than baseline societies
- $H_{\text{Social Experience}}$: Exanna provides better social experience than baseline societies
- $H_{\text{Privacy Loss}}$: Exanna takes lower privacy loss compared to baseline societies

Tests for Statistical Significance

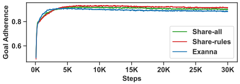
- Independent t-test
- Glass's Δ

Results: Exanna Yields Less Privacy Loss When Providing Rationale

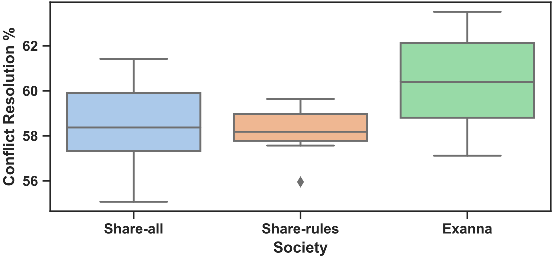


Results: Exanna Has Higher Conflict Resolution in Cases Where Agents Deviate from their Goals

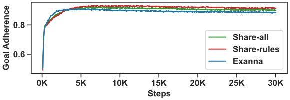
Exanna trades goal adherence for conflict resolution

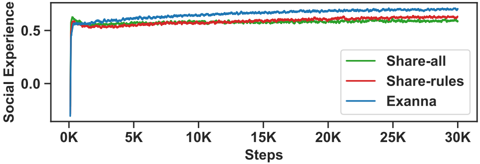


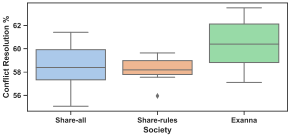
Goal Adherence



Results: Exanna Yields Better Social Experience and Conflict Resolution at the Expense of Goal Adherence

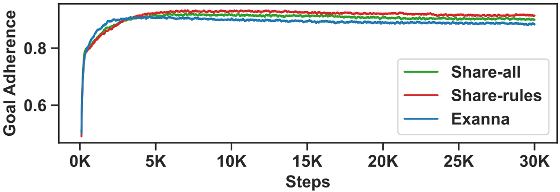






Results: Exanna Yields Lower Goal Adherence than Other Societies

- Exanna generates lower goal adherence than Share-Decision Rules society For less convincingness from being conservative
- Share All society has lower goal adherence than Share-Decision Rules society for distraction from information overload



Exanna: Summary

Study Summary

Providing rationales with the concerns of value preferences leads to (1) deviation from goals, (2) higher conflict resolution, (3) less private loss, and (4) better social experience

- Value-aligned explanations ensure the AI system's decisions are consistent with human values
  - Highlight what an agent cares
  - No unnecessary sacrifice of private attributes

Conclusion

- Considering human factors leads to higher social experience in terms of a single agent
- Regarding MAS, Considering human factors promotes cooperation

Future Directions

- Going deeper into understanding the causal connections that exist between decisions and human factors
- Investigate how different costs of information influence decisions for more precise and reliable action suggestions and rationale construction
- Information suppression may be acceptable in some cases
- Having agents decide what to share and when to share increases strategies' flexibility
- Investigate the relationship between social norms and different social signals

Thank You

This research is based upon work supported by the National Science Foundation under Grants No. IIS-1908374 and No. IIS-2116751

General Reproducibility Details

Hardware:

- 32 GB RAM
- GPU NVIDIA GTX 1070 Ti

Framework:

- MASON (Java)
- Mesa (Python)

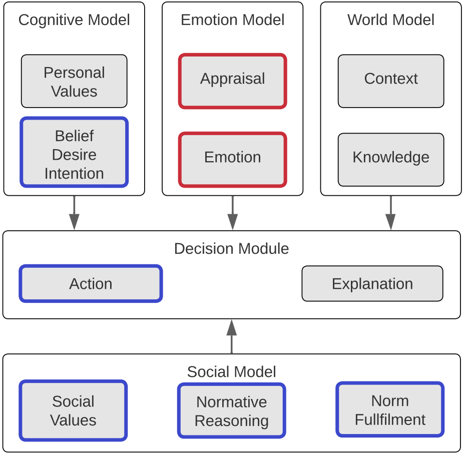
Noe: Enforcing Social Norms with Expressed Emotions

Noe: Introduction and Motivation

Source: https://www.shutterstock.com/image-illustration/3d-smart-red-man-jumps-queue-134805779

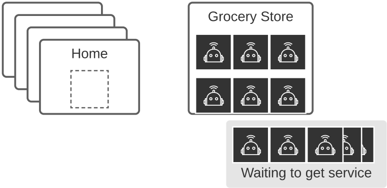
- Motivation:
  - Sanctions in real world are usually subtle
  - General thinking and problem-solving incorporate the influence of emotions (Simon, 1967)
- Objective: Incorporate expressed emotions in decision-making
- $\text{RQ}_{\text{emotion}}$: How does modeling the emotional responses of agents to the outcomes of interactions affect the emergence of norms and social welfare?

Noe: Cognitive Framework



Noe: Evaluation with Line-Up Simulations

- Simulate with varying agent societies: Obedient, Anarchy, Sanctioning and Noe society
- Appraisal: Based on norm satisfaction or violation
- Characteristics of agent society
  - Prescriptive norm: Line up to get service
  - Sanctions and **expressed emotions** that emerge from the evaluation of chosen actions
  - Expressed emotions serve as intrinsic reward (self-directed emotion) and extrinsic reward (other-directed emotion)



Noe: Metrics

Measures:

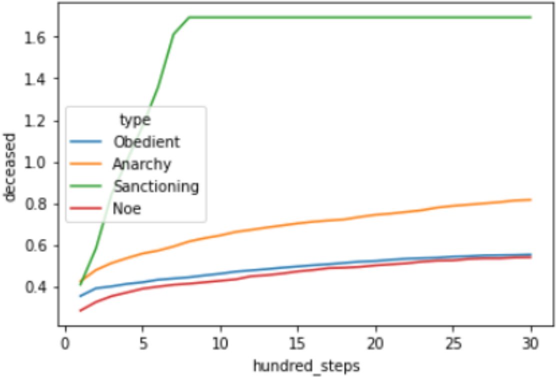
- Social welfare
  - M_{Deceased}: Cumulative number of agents deceased
  - M_{Health}: Average health of the agents
- Norm satisfaction
  - M_{Cohesion}: Proportion of norm instances that are satisfied
- Social experience
  - $M_{\text{Waiting time}}$: Average waiting time of agents in the queues

Noe: Simulation Results

		Obedient	Anarchy	Sanctioning	Noe
$M_{Deceased}$	$\bar{X}$	55.30	81.60	169.30	**54.00**
	p-value	<0.01	<0.01	<0.01	–
	Δ	0.65	3.10	15.53	–
M_{Health}	$\bar{X}$	79.27	79.50	86.26	79.00
	p-value	0.52	0.46	8.45	–
	Δ	0.18	0.21	3.34	–
$M_{Cohesion}$	$\bar{X}$	–	0.22	0.88	**0.99**
	p-value	–	<0.01	<0.01	–
	Δ	–	102.43	13.67	–
$M_{Waiting\ Time}$	$\bar{X}$	8.95	5.45	2.55	8.95
	p-value	0.98	<0.01	<0.01	–
	Δ	0.01	40.82	76.68	–

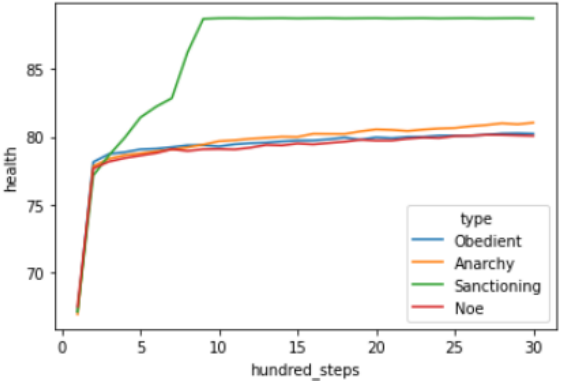
Results: Fewer Agents Die in Noe than in Other Societies

Metric: Cumulative number of agents deceased



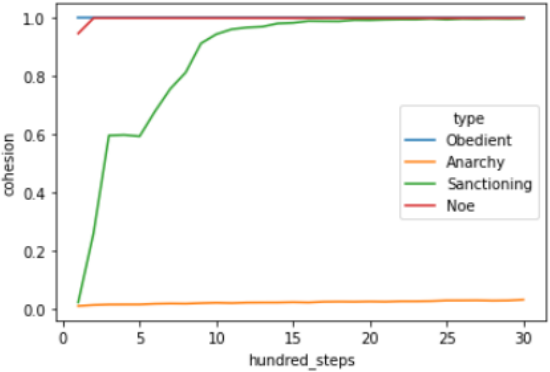
Results: Sanctioning Society Yields Higher Health State ...But at the Expense of More Deaths

Metric: Average health state of the agents



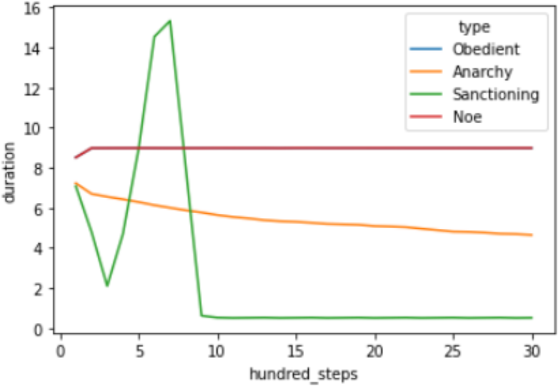
Results: Noe Yields Higher Cohesion than Other Agent Societies

Metric: Cohesion (Proportion of norm instances that are satisfied)



Results: Noe Has Similar Waiting Time as Obedient Society

Metric: Average waiting time of agents in the queues



Noe: Summary

Study Summary

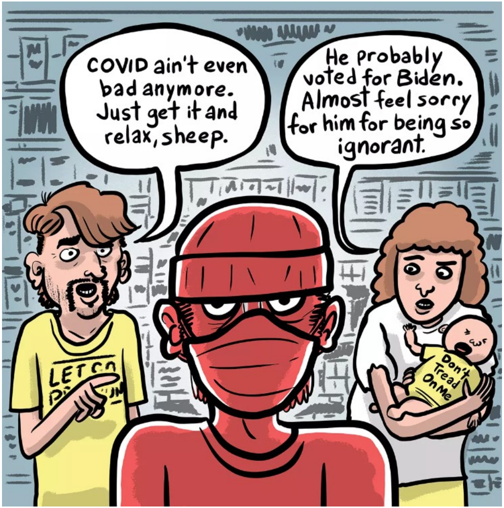
Agents who incorporate expressed emotions are more willing to comply with norms than those who do not

- Expressed emotions act as a positive or negative reinforcement mechanism for specific behaviors
- Noe enables the incorporation of expressed emotions as sanctions in decision-making

Ness: Normative Information from Tell and Hint

Ness: Scenario

When there's a norm of not wearing a mask . . .

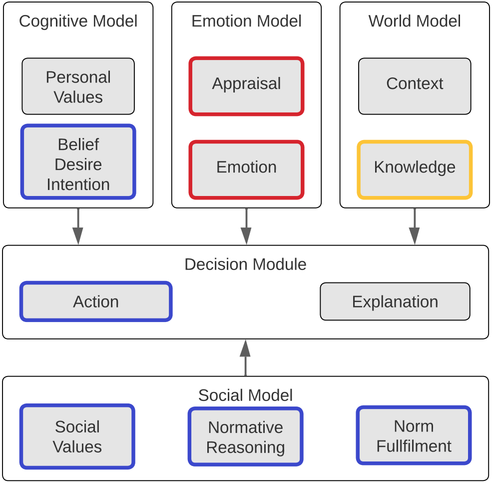


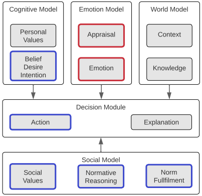
Source: https://www.latimes.com/opinion/story/2022-08-07/mask-wearing-public-anger-comic

Ness: Introduction and Motivation

- Motivations
  - Social signals are reactions to norm satisfaction or norm violation
  - Social signals provide natural drivers for norm emergence
  - Normative information conveyed through a social signal promotes cooperation in MAS
  - Social signals can be realized in three main ways: *sanction*, *tell*, and *hint*
- Objective: Incorporate normative information from social signals into decision-making
- $\text{RQ}_{\text{information}}$: How does considering soft signals such as hints and tell in addition to sanctions influence norm emergence?

Ness: Cognitive Framework





Compare with Noe

Ness: Key Concepts

- Reward Shaping (Ng et al., 1999) provides additional "shaping" reward from deterministic reward function

$$r'_{final} = r + F$$

  where r is the standard reward function in reinforcement learning and F is the shaping reward function

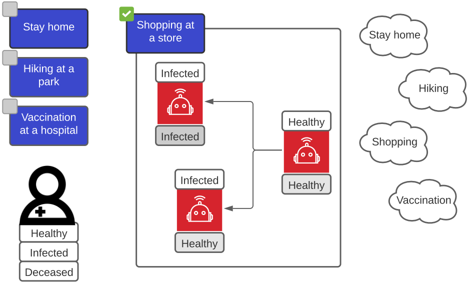
- With messages or hints, F defines the difference of potential values

$$F(s, a, s', a') = \gamma \Phi(s', a') \kappa - \Phi(s, a)$$

  where Φ is a potential function that gives hints on states. κ defines the certainty of potential reward from the knowledge or information

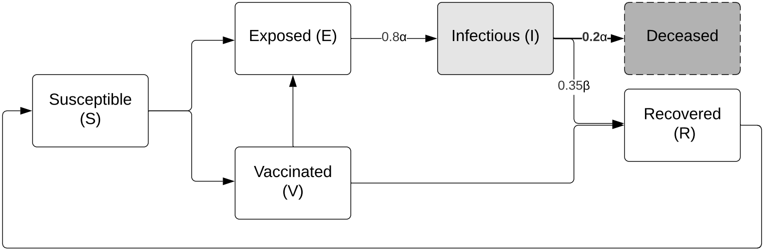
Ness: Evaluation with COVID Simulations

- Simulate with varying agent societies: Primitive, Sanction, Hint, Tell, and Ness society
- Characteristics of agent society
  - A combination of three kinds of social signals
  - Prescriptive norm: Stay self-quarantine if infected
  - Sanction (Material punishment): Send to forced quarantine at a low probability



Ness: Disease Model

Susceptible-Exposed-Infected-Recovered-Vaccinated (SEIRV) model (Yang and Wang, 2020; Annas et al., 2020)



- α controls the probability to be infected based on vaccination
- β controls the recovering rate based on agent activity
- Healthy agents, cover susceptible, exposed and recovered, are not infected
- Infectious includes three subclasses: Asymptomatic, mildly symptomatic, and critical symptomatic

Ness: Agent Societies

Society	Sanctioning	Shaping Reward	Emotion
Baseline 1. PRIMITIVE	✗	✗	✗
Baseline 2. SANCTION	✓	✗	✗
Baseline 3. TELL	✓	✓	✗
Baseline 4. HINT	✓	✗	✓
Ness	✓	✓	✓

Shaping Rewards come from normative information

Ness: Information Balance

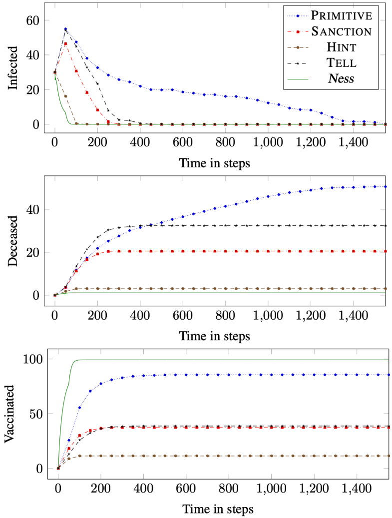
Societies: Signals	Sanction	Tell	Hint	Hint w/ shaping reward
PRIMITIVE	0%	0%	0%	0%
SANCTION	38%	0%	0%	0%
TELL	20%	36%	0%	0%
HINT	20%	0%	12%	0%
Ness	20%	0%	0%	10%

- More learning channels improve learning efficiency
- We balance the information an agent can access by adjusting the expected payoff to achieve comparability

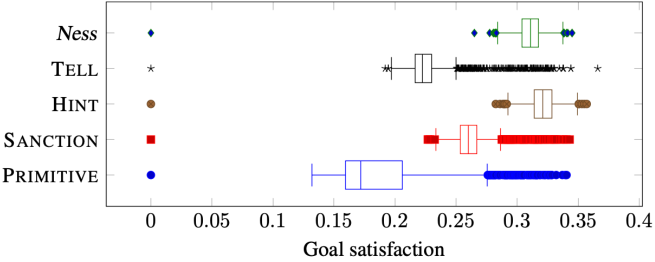
Ness: Metrics

- Disease control
  - M_{Healthy}: The percentage of agents who are healthy
  - M_{Infected}: The percentage of agents who are infected
  - M_{Deceased}: The percentage of who are deceased
  - $M_{\text{Total infections}}$: Total number of infections in societies
  - $M_{\text{Vaccinated}}$: Percentage of vaccinated agents
- Goal
  - M_{Goal}: The average goal satisfaction among agents
- Isolation
  - $M_{\text{Isolation}}$: The percentage of self-isolation among infected agents
  - $M_{\text{Forced quarantine}}$:
    - Number of agents who are forced to quarantine at home
    - This measure maps to the sanction signal type

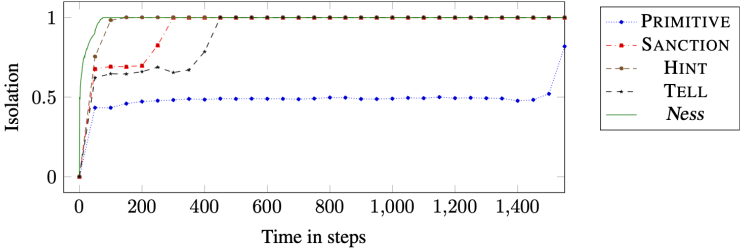
Results: Ness Yields Better Disease Control and Higher Vaccination Rate

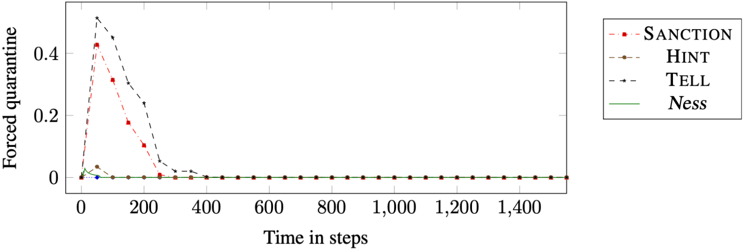


Results: Ness Yields Better Goal Satisfaction Than Tell, Sanction, and Primitive Societies



Results: Ness and Noe Yield Higher Norm-Compliance and Lower Forced Quarantine than Other Societies





Ness: Detailed Results (1)

		Primitive	Sanction	Hint	Tell	Ness
$H_{\text{Disease control}}$	M_{Infected}	13.281	2.634	0.411	4.205	0.157
	Δ	−0.973	−0.271	0.085	−0.330	–
	M_{Healthy}	46.294	77.602	96.622	65.082	98.750
	Δ	18.259	3.414	0.776	4.784	–
	M_{Deceased}	41.034	19.764	2.967	30.713	1.093
	Δ	−3.346	−6.123	−7.450	−5.316	–
	$M_{\text{Infections}}$	48.335	13.840	2.221	20.474	0.891
	Δ	−2.664	−6.925	−10.730	−5.842	–
	$M_{\text{Vaccinated}}$	82.452	36.743	11.185	37.430	98.734
	Δ	1.518	18.181	143.254	13.261	–

Ness: Detailed Results (2)

		Primitive	Sanction	Hint	Tell	Ness
$H_{Isolation}$	$M_{Isolation}$	0.610	0.965	0.993	0.934	0.998
	Δ	1.777	0.326	0.101	0.450	–
	$M_{Forced\ quarantine}$	–	0.026	$8.5e-04$	0.040	$1.75e-04$
	p-value	–	<0.001	< 0.01	<0.001	–
	Δ	–	-0.268	-0.075	-0.313	–
H_{Goal}	M_{Goal}	0.187	0.262	0.321	0.227	0.311
	Δ	3.128	3.445	-1.012	3.929	–

Ness: Summary

Study Summary

Ness agents effectively avoid undesirable results and yield higher satisfaction than baseline agents despite requiring an equivalent amount of information

- Normative information from soft signals like hints and messages helps to regulate behaviors
- Incorporating normative information from social signals supports norm emergence

Ness: Hyperparameters

Parameter	Value	Comment
Learning rate α	0.001	
Discount factor γ	0.900	
Simulation step per action	1.000	
Infection %	0.300	The default fraction of infected agents in a society
Certainty of potential reward	0.300	value for κ for certainty of possible sanctions from normative information through hints
Certainty of potential reward	0.500	value for κ for certainty of possible sanctions from normative information messages

Fleur: Detailed Results (1)

		Compliance	Social Experience	Invalidation
S_{mixed}	$\bar{X}$	63.40%	0.448	0.296
	p-value	–	–	–
	Δ	–	–	–
$S_{altruistic}$	$\bar{X}$	69.70%	0.554	**0.334**
	p-value	< 0.001	< 0.001	< 0.001
	Δ	0.660	0.612	0.464
$S_{prosocial}$	$\bar{X}$	**70.25%**	**0.566**	0.323
	p-value	< 0.001	< 0.001	< 0.05
	Δ	0.718	0.677	0.326

Fleur: Detailed Results (2)

		Compliance	Social Experience	Invalidation
$S_{selfish}$	$\bar{X}$	65.10%	0.469	0.269
	p-value	0.218	0.424	< 0.05
	Δ	0.178	0.122	0.329
$S_{competitive}$	$\bar{X}$	54.08%	0.221	0.289
	p-value	< 0.001	< 0.001	0.541
	Δ	0.977	1.313	0.088

Fleur: Hyperparameters

Parameter	Value
Population size	40
Simulation step per action	1
Training steps	500,000
Evaluation steps	100
Learning rate α	0.001
Discount factor γ	0.9

Exanna: Processes of XCS

- Matching: A process that matches the current context and all rules/classifiers to generate a match set
- Covering: A process that guarantees diversity via adding a random classifier whose conditions match the current context
- Action selection: This process returns the action with the highest fitness-weighted aggregation of reward if in exploitation mode
- Formation of action set: The action set includes all classifiers that propose the chosen action based on the match set
- Updating classifier parameters: An agent updates the rule parameters (e.g., accuracy and fitness) based on the received payoff
- Subsumption: A process that replaces offspring rules with more general parent rules if it exists and with a minor prediction error
- Deletion: Each action set has the same maximum number of rules and XCS removes the low-fitness rules

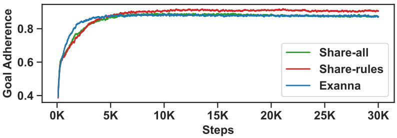
Exanna: Hyperparameters

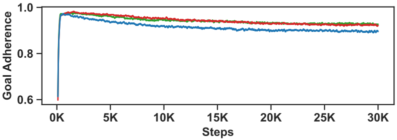
Parameter	Value
Population size	200
Learning rate	0.1
Don't care probability	0.3
Accuracy threshold	0.01
Fitness exponent	5
Genetic algorithm threshold	25
Mutation probability	0.4
Crossover probability	0.8
Experience threshold for deletion	20
Experience threshold for subsumption	20
Fitness falloff	0.1

Exanna: Detailed Results

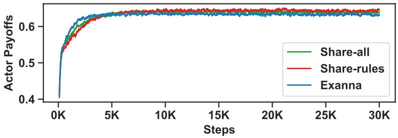
		Share All	Share Decision Rules	Exanna
$M_{\text{Goal Adherence}}$	$\bar{X}$	0.901	0.914	0.885
	p-value	0.083	< 0.01	–
	Δ	-1.594	-2.891	–
$M_{\text{Conflict resolution}}$	$\bar{X}$	0.585	0.582	0.604
	p-value	< 0.001	< 0.001	–
	Δ	1.803	3.106	–
$M_{\text{Social Experience}}$	$\bar{X}$	0.591	0.624	0.699
	p-value	< 0.001	< 0.001	–
	Δ	1.803	3.106	–
$M_{\text{Privacy Loss}}$	$\bar{X}$	1.000	0.999	0.749
	p-value	< 0.001	< 0.001	–
	Δ	∞	$-11{,}896.523$	–

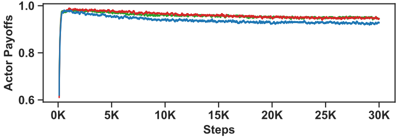
Exanna: Goal Adherence by Agent Types



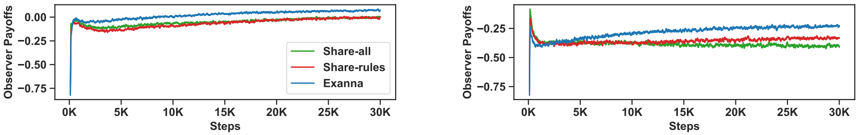


Exanna: Payoff of Actors by Agent Types





Exanna: Payoff of Observers by Agent Types



References

Suwardi Annas, Muh Isbar Pratama, Muh Rifandi, Wahidah Sanusi, and Syafruddin Side. Stability analysis and numerical simulation of SEIR model for pandemic COVID-19 spread in Indonesia. *Chaos, Solitons & Fractals*, 139:110072, 2020. doi: 10.1016/j.chaos.2020.110072.

Andrew Y. Ng, Daishi Harada, and Stuart Russell. Policy invariance under reward transformations: Theory and application to reward shaping. In *Proceedings of the 16th International Conference on Machine Learning (ICML)*, volume 99, pages 278–287. Morgan Kaufmann, 1999. doi: 10.1.1.48.345.

Herbert A. Simon. Motivational and emotional controls of cognition. *Psychological Review*, 74(1):29–39, 1967. doi: 10.1037/h0024127.

Munindar P. Singh. Norms as a basis for governing sociotechnical systems. *TIST*, 5(1):21:1–21:23, December 2013. doi: 10.1145/2542182.2542203.

Chayu Yang and Jin Wang. A mathematical model for the novel coronavirus epidemic in Wuhan, China. *Mathematical Biosciences and Engineering MBE*, 17(3):2708, 2020. doi: 10.3934/mbe.2020148.