Hierarchical version of Social Learning Model on Public goods game

Youngha Jo

Contents

Abstract

Introduction

Method

Aim to do

Abstract

1. Generalizable model for the social cognition

   ⇒ Model for Public Goods Game suggested a general account

2. Capturing the the core features from the cognitive process

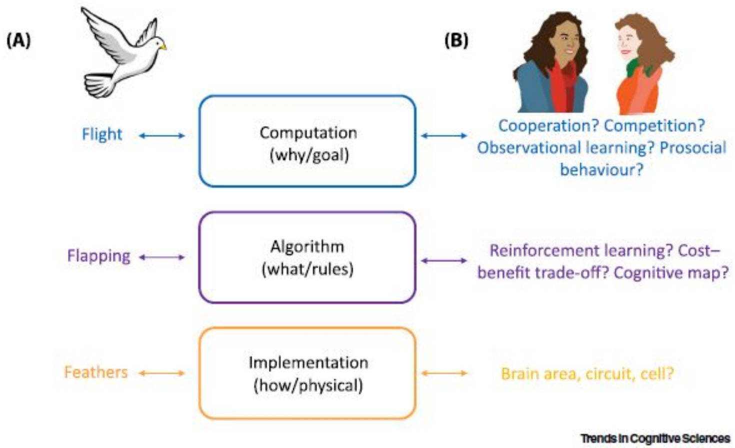
   ⇒ Bayesian inference / Desires (Intentions)

3. Hierarchical model for better subject-level parameter estimates

Introduction

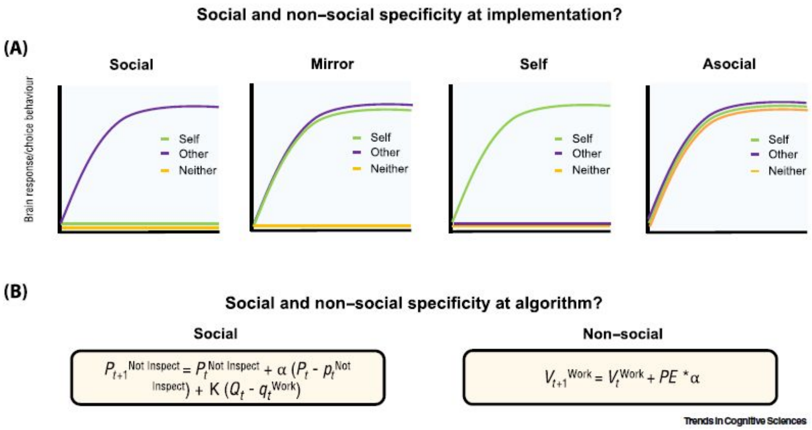
Modeling the Social Cognition

Algorithms of the social cognition is similar to the non-social cognition (e.g. reinforcement learning)

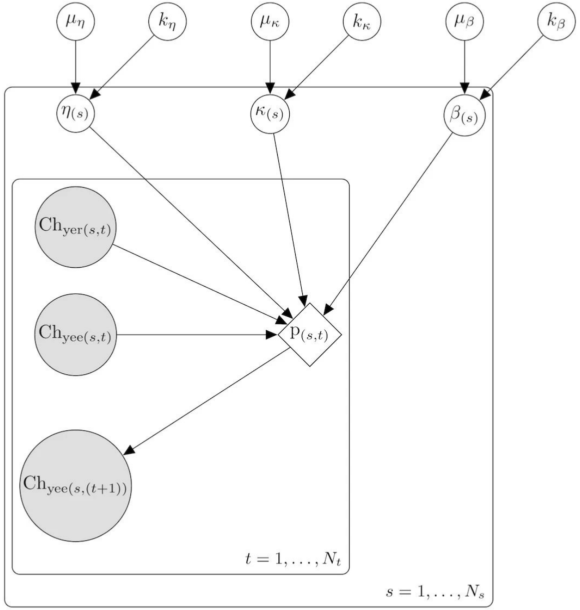


Lockwood et al., 2020

Hierarchical Model for the Social Cognition



Lockwood et al., 2020



Hill et al., 2017

Bayesian Theory of Mind

How inferences of other people's thoughts are made?

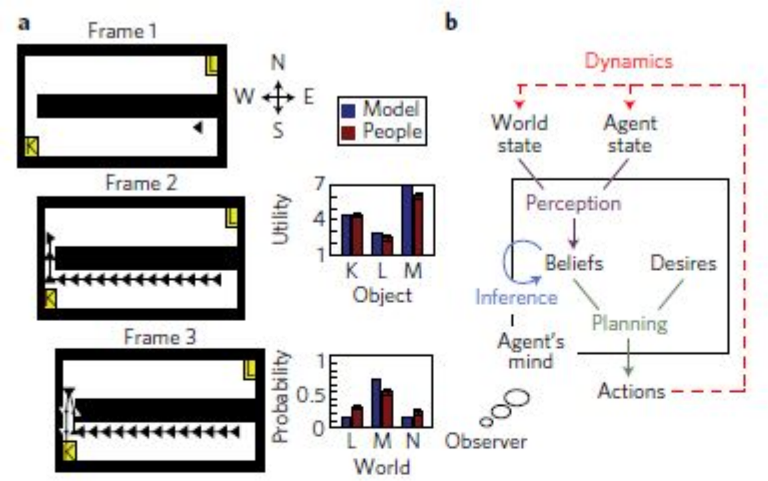
Categories	Description	Inferences	Examples
Emotion recognition	Infer an agent's emotions from emotional expressions (facial expressions, body language, prosody), *or* from their actions.	$P(e\|x)$ $P(e\|a)$	Given *smile*, Infer *P*(*happy*) Given *scold-others*, Infer *P* (*anger*)
Third-person appraisals	Reason "forward" about how an event would cause an agent to feel, given also their mental states.	$P(e\|o)$ $P(e\|o,b,d)$ $P(e\|b,d)$	Given *lose-wallet*, Infer *P*(*sad*)
Inferring causes of emotions	Reason "backwards" about the events that caused an agent's emotions.	$P(o\|e)$ $P(o\|e,b,d)$	Given *sad*, Infer *P*(*lose-wallet*)
Emotional cue integration	Given multiple, potentially conflicting cues (e.g., multiple behaviors and/or causes of emotion), combine them and reason about agent's emotions	e.g., $P(e\|o,x)$ $P(e\|o,a)$ $P(e\|a,x)$ $P(e\|o,b,d,a,x)$	Given *smile* + *lose-wallet*, Infer *P*(*happy*), *P*(*sad*), etc.
Reverse appraisal	Given an event and an agent's emotions, reason backwards to mental states like beliefs and desires	$P(b,d\|e,o)$ $P(b,d\|e)$	Given: *receive-gift* + *surprise* + *happy*, Infer *Gift-Unexpected* + *Gift-Desired*
Predictions (hypothetical reasoning)	Given an agent's emotions, predict subsequent behavior. Or, given a (hypothetical) situation, predict the agent's emotions.	$P(a\|e)$ $P(x\|e)$	If *anger*, predict *P*(*scold-others*)
Counterfactual reasoning (and explanations)	Given a state of the world and an agent's emotions, reason about emotions or behavior in counterfactual states of the world. This also allows explanations of emotions or behavior in terms of their causes.	e.g., $P(e\|\textbf{not}\ o)$ $P(a\|\textbf{not}\ e)$	Given: "*lose-wallet* and *sad*," reason that: "**if** *not-lose-wallet*, **then** *not-sad*" or, reason that: "*sad* **because** *lose-wallet*"

Note. $P(X)$ denotes the probability of X occurring. The lists of inferences presented for each category are exhaustive (given our derivation), except for the cue integration and counterfactual reasoning categories.

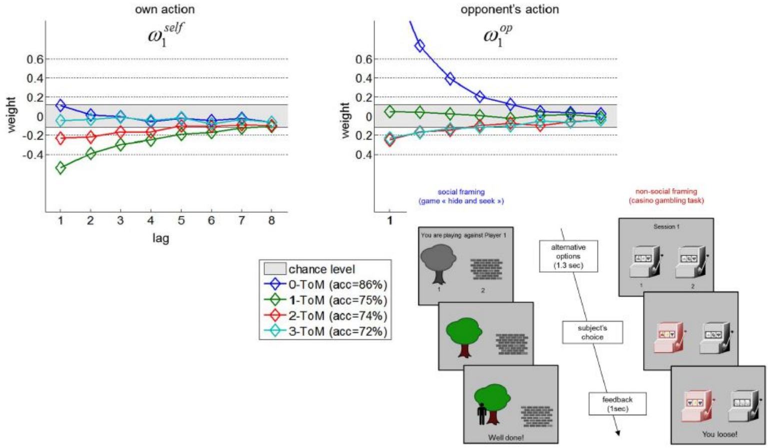
Ong et al., 2019

Bayesian Theory of Mind

Bayesian ToM approach is used in wide domains of social cognition task



Baker et al., 2017



Devaine et al., 2014

Game Theoretical Tasks

Game theoretical tasks investigate participant's cognition of:

- Inferring other's mind
- Compute rewards based on subjective / task-related utility

  → task-related: altruistic, strategic

Good task to investigate bayesian inference and utilization

Table 1
Explanation and categorisation of tasks used in studies.

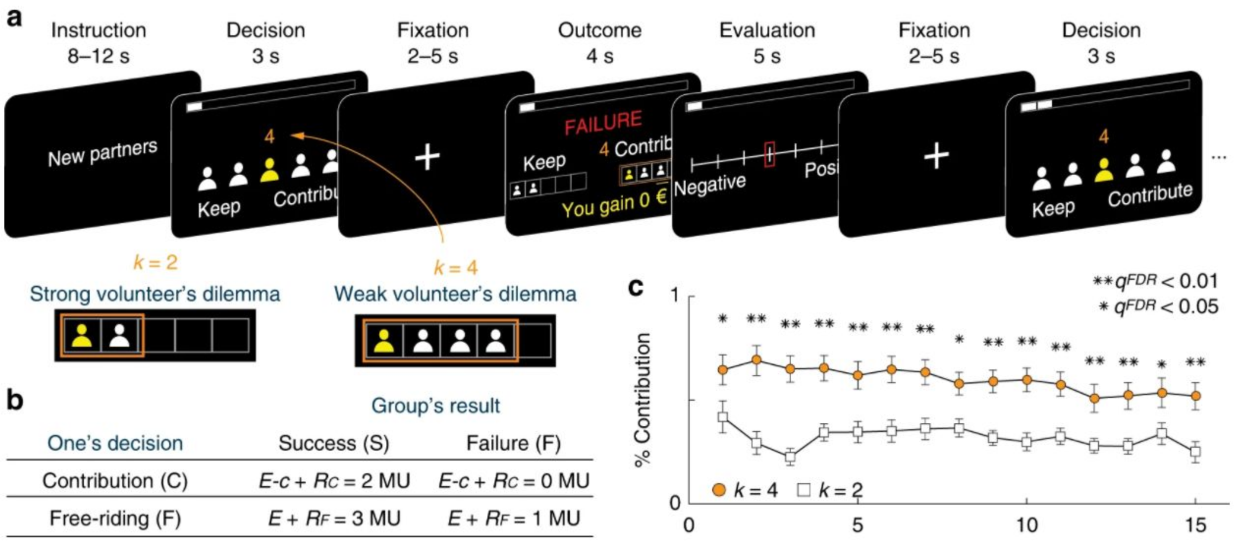
Task	Description	Group
Dictator game	Participant either chooses an amount of their money to give or accepts/rejects a proposed split between themselves and the other player.	Altruistic
Charity donation task	Participant either chooses an amount of their money to donate or accepts/rejects a proposed split between the participant and a charity.	Altruistic
Pain vs. gain	Participant can give up varying amounts of money, the more given the less painful the electric shock given to a partner	Altruistic
Ultimatum game	Participant proposes a split between themselves and their partner that is only implemented if the partner accepts it.	Strategic
Trust game	Participant transfers an amount of money to the trustee that is multiplied by some factor (often 3). The trustee then chooses an amount to send back which decides the payoff for both players.	Strategic
Prisoner's dilemma	Participant and partner decide whether to cooperate or defect. They gain mutual benefit if both cooperate but individuals gain more by defecting if the partner cooperates.	Strategic
Public goods game	Participants invest an amount in a communal fund that is then multiplied and divided among all players, including those who did not initially contribute to the communal fund.	Strategic

Cutler and Campbell-Meiklejohn, 2018

Method

Generalizable Framework with subjective goal?

Public Goods Game



Khalvati et al., 2019   *Park et al., 2019*

Bayesian Inference of predicting other's choice

Original Idea from the paper, Khalvati et al., (2019)

Prior Distribution: $\theta \sim \mathrm{Beta}(\alpha, \beta)$

Posterior Distribution: $\theta \sim \mathrm{Beta}(\alpha_t, \beta_t)$ where $\alpha_{t+1} = \gamma\alpha_t + c$ and $\beta_{t+1} = \gamma\beta_t + N - c$ (used conjugacy)

Used modified version from the twin paper, Park et al., (2019), called ***"Social Learning Model"***

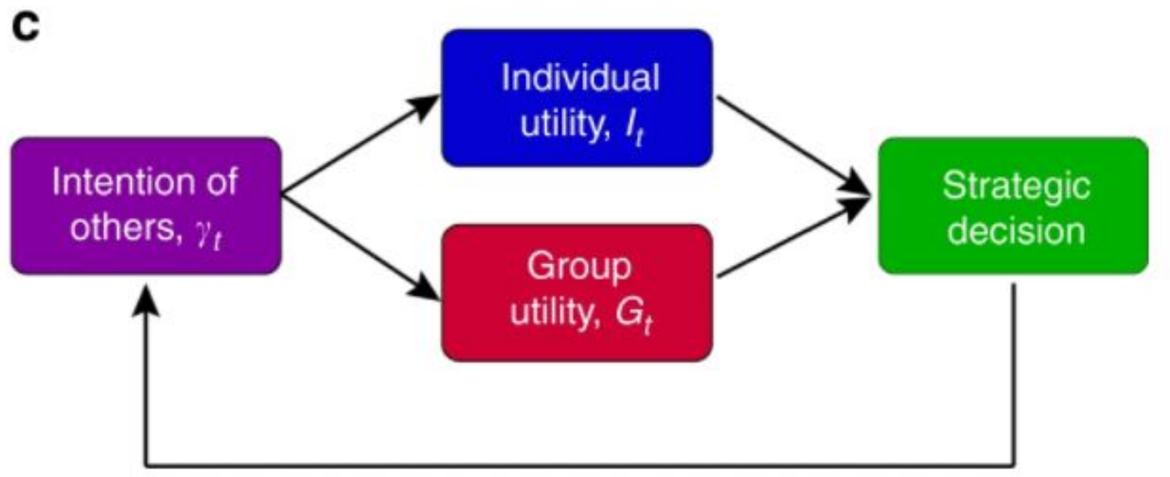
Basic Framework: Social Learning

$$\underbrace{P(\boldsymbol{\theta}|y)}_{posterior} = \underbrace{(P(y|\boldsymbol{\theta})}_{likelihood} \times \underbrace{P(\boldsymbol{\theta}))}_{prior} \;/\; \underbrace{P(y)}_{evidence}$$

$$\gamma_t = \gamma_{t-1} + f\left(\alpha + \theta PE_R\right) PE_S$$
$$\text{where } PE_S = (F_{t-1}/N - 1) - \gamma_{t-1}$$
$$\text{and } PE_R = \left| R \sum_{i=0}^{N-k} \Gamma_{t-1}^{i} - R_{t-1} \right|$$

Not exactly same as computing the posterior distribution
But, followed the basic concept of Bayesian Inference

Basic Framework: Inference and Utility



Explanations for Social Learning Model

1. Belief of other's choice to free-ride

   ⇒ Update via social reinforcement learning

f: sigmoid, R→ (0,1)

$$\boxed{\Gamma_t^{N-k}} = \binom{N-1}{N-k} \boxed{\gamma_t}^{N-k} (1-\gamma_t)^{k-1}$$

$$\boxed{\gamma_t} = \gamma_{t-1} + \boxed{f(\alpha + \theta PE_R)} \boxed{PE_S}$$

$$\text{where} \boxed{PE_S} = (F_{t-1}/N - 1) - \gamma_{t-1}$$

$$\text{and} \boxed{PE_R} = \left| R \sum_{i=0}^{N-k} \Gamma_{t-1}^{i} - R_{t-1} \right|$$

Estimate how many participants out of N-1 people would choose to free-ride

Reward Prediction Error guides the learning rate

Belief γ is updated by Social Prediction Error

Explanations for Social Learning Model

2. Based on the inference, compute

   ⇒ Individual Utility, Group Utility

$$\ln \hat{L} = \sum C_t \times \ln p(C_t|\alpha, \theta, \omega, \pi, \lambda) + (1 - C_t) \times \ln (1 - p(C_t|\alpha, \theta, \omega, \pi, \lambda))$$

$$p(C_t) = \text{logit}(Q_t)$$
$$Q_t = \omega I_t + (1 - \omega) G_t$$

$$I_t = \boxed{\lambda + \Gamma^{N-k} R} + \boxed{\pi \Gamma^{N-k} R (N - 1)}$$

Value of one's own reward *Value of altruistic behavior*

$$G_t = \sum_{j=t}^{T} R K^{T-j} \sum_{i=0}^{N-k} \Gamma_t^i$$

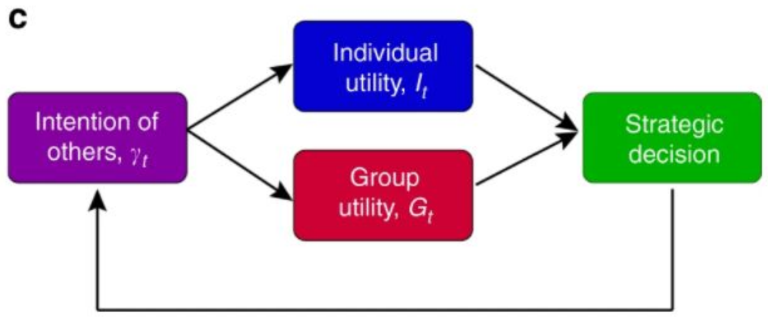
Potential value of inducing others

Strengths of the model

Captured inference of group decision by the concept of ‘belief’

⇒ generalizable to model the theory of mind

Can be applied to various other games from the game theory



Inference - Utility Framework

Limitations of the current model

Roles for the reward prediction error is not specified

→ Try without RPE (one parameter for learning rate)

$$\gamma_t = \gamma_{t-1} + f(\alpha + \theta PE_R)\, PE_S$$
$$\text{where } PE_S = (F_{t-1}/N - 1) - \gamma_{t-1}$$
$$\text{and } PE_R = \left| R \sum_{i=0}^{N-k} \Gamma_{t-1}^{i} - R_{t-1} \right|$$

Overall probability to contribute is too complicated; needs validation for the parameter fit

Also, convergence problem has arised

→ Try add inverse temperature β for Q

$$p(C_t) = \text{logit}(Q_t)$$
$$Q_t = \omega I_t + (1-\omega)\, G_t$$

$$I_t = \lambda + \Gamma^{N-k} R + \pi \Gamma^{N-k} R (N-1) \qquad G_t = \sum_{j=t}^{T} R K^{T-j} \sum_{i=0}^{N-k} \Gamma_t^{i}$$

Aim to achieve

Check whether the estimated parameters are similar to the authors' in Rstan

Check if the model fits the data better when applying the hierarchical version of the model

+ Check if the model fits better if exclude RPE term and include inverse temperature

Models						BIC
Social learning	ω	π	λ	α	θ	
	0.65 ± 0.07	0.16 ± 0.16	−2.10 ±0 .45	0.51 ± 0.06	0.13 ± 0.02	−6149
Myopic	ω	π	λ	α	θ	
	11.04 ± 2.76	0.34 ± 0.19	−2.98 ± 1.09	0.59 ± 0.06	0.14 ± 0.04	−5721
Group utility	χ	ζ	α	θ		
	9.30 ± 4.55	−30.91 ± 4.92	0.56 ± 0.06	0.12 ± 0.02		−5713
Inequity aversion	δ	ε	κ			
	16.39 ± 7.11	0.61 ± 0.42	1.37 ± 1.00			−4220

Result

Explored three versions of models:

- Original Social Learning Model
- Social Learning Model without considering Reward Prediction Error
- Social Learning Model without RPE and with Inverse Temperature
-

Ran both Individual / Hierarchical version (total 3 X 2 = 6)

Result

Quantitative analysis was done by using the “loo” package

Sampling Condition was essential to acquire the appropriate MCMC chains

⟶ iteration= 4000, warmup=1000, chain = 4, thin: 2, init_r=5, adaptive_delta=0.9, max_treedepth=12

*※Additional sampling with thin=3, iterations= 6000, max_treedepth=15 showed much better fit without outlier, but couldn’t generate quantities (**log likelihoods**) due to the timeout*

Result: Total LOOIC score

model 1: hierarchical version of original model

model 2: hierarchical version w.o. RPE

model 3: hierarchical version w.o. RPE / w. Inv Temp

model 4: individual version of original model

model 5: individual version w.o. RPE

model 6: individual version w.o. RPE / w. Inv Temp

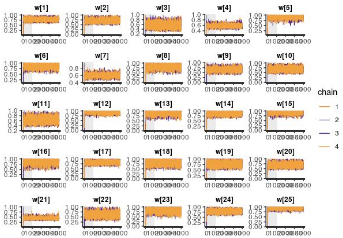
	elpd_diff	se_diff	elpd_loo	se_elpd_loo	p_loo	se_p_loo	looic	se_looic
model3	0.000	0.000	-2389.144	71.398	73.503	4.586	4778.287	142.795
model6	-3.329	6.686	-2392.473	73.494	80.714	4.467	4784.946	146.988
model5	-65.970	20.573	-2455.114	59.961	77.855	4.098	4910.227	119.922
model1	-78.778	23.248	-2467.922	62.083	65.410	4.667	4935.843	124.165
model4	-93.503	24.395	-2482.646	58.190	72.015	3.065	4965.293	116.380
model2	-2000.780	453.263	-4389.924	461.357	2012.157	454.010	8779.848	922.714

※ *Model 2 needs resampling by adjusting control arguments @ stan*

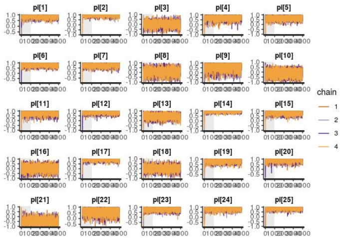
Result: Original Social Learning Model

Individual estimate, Traceplot

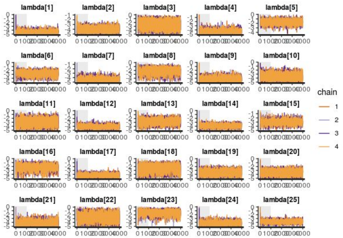
Traceplot implies bad fit of the parameters of the learning rate (alpha, theta)



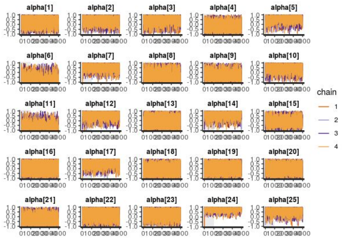
omega



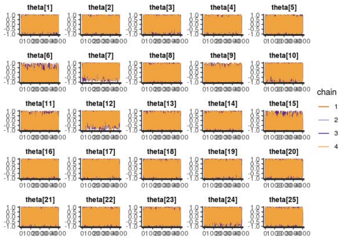
pi



lambda



alpha

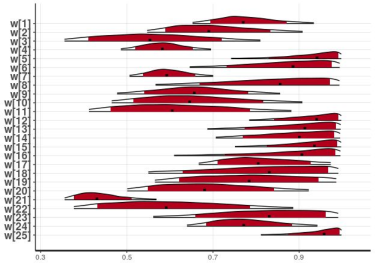


theta

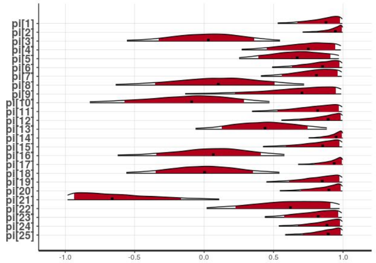
Result: Original Social Learning Model

Individual estimate, Posterior Distribution

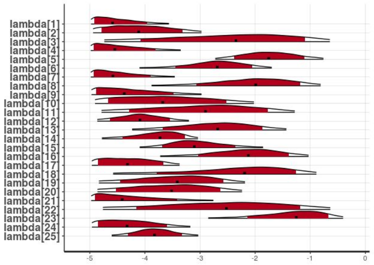
Showing bad fit except for omega and lambda



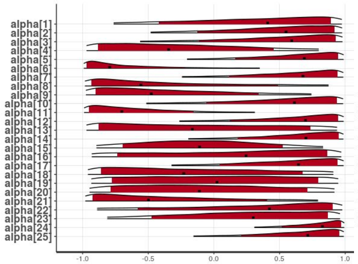
omega



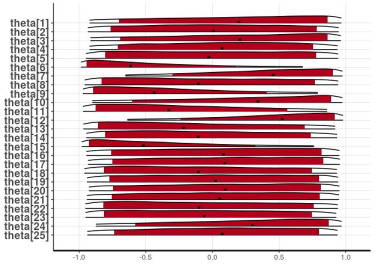
pi



lambda



alpha

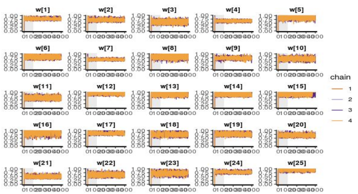


theta

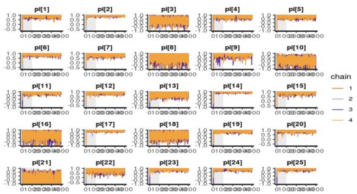
Result: Original Social Learning Model

Group estimate, Traceplot

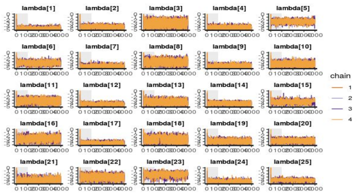
Traceplot implies bad fit of the parameters of the learning rate (alpha, theta)



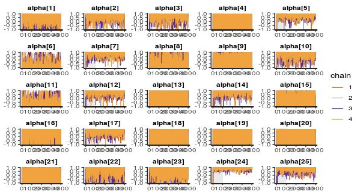
omega



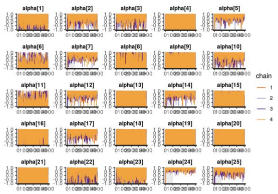
pi



lambda



alpha

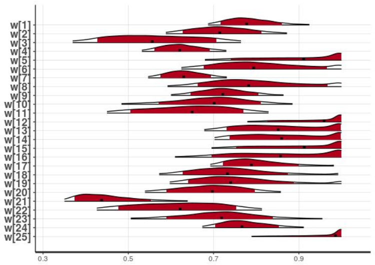


theta

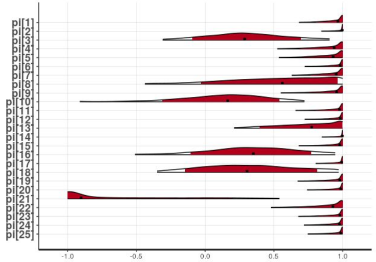
Result: Original Social Learning Model

Group estimate, Posterior Distribution

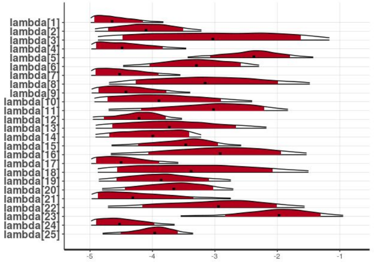
Showing bad fit except for omega and lambda



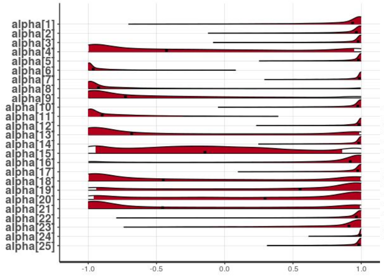
omega



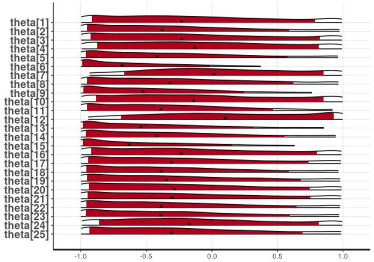
pi



lambda



alpha

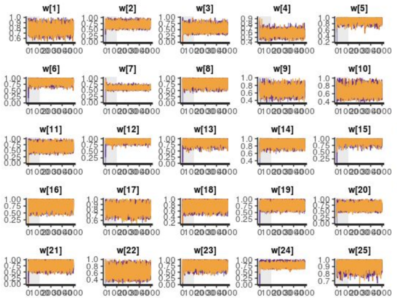


theta

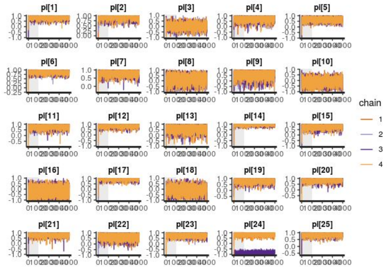
Result: Social Learning Model w.o. RPE

Individual estimate, Traceplot

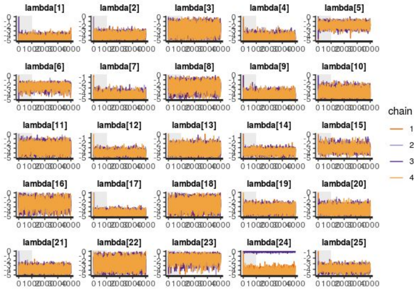
Traceplot of pi[24] shows that the chain was not mixed



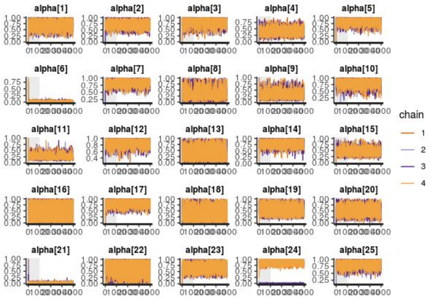
omega



pi



lambda

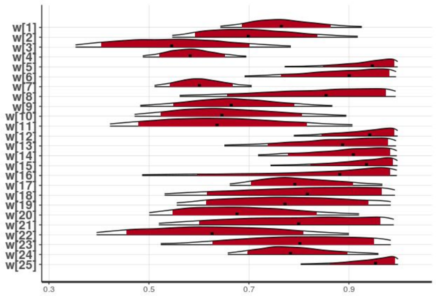


alpha

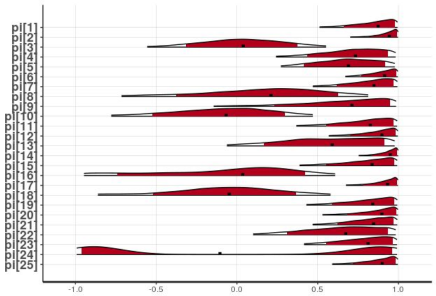
Result: Social Learning Model w.o. RPE

Individual estimate, Posterior Distribution

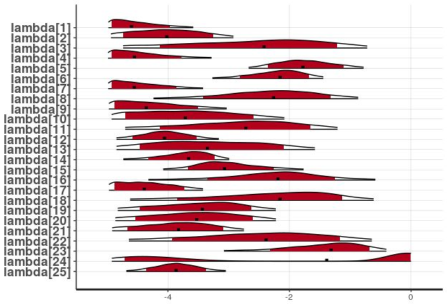
24th parameters show that the posterior distribution did not fit adequately



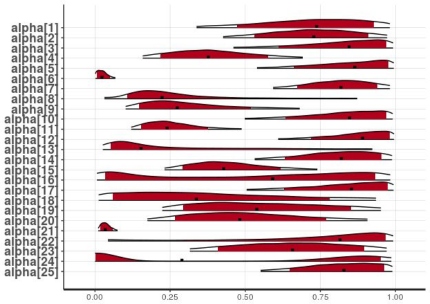
omega



pi



lambda

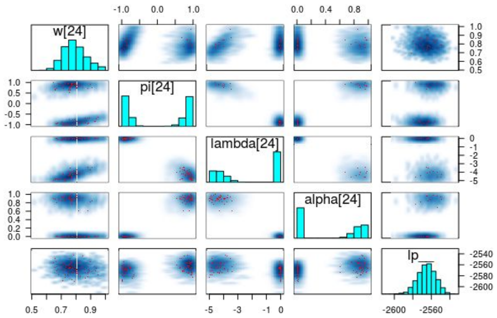


alpha

Result: Social Learning Model w.o. RPE

Individual estimate, pairs plot

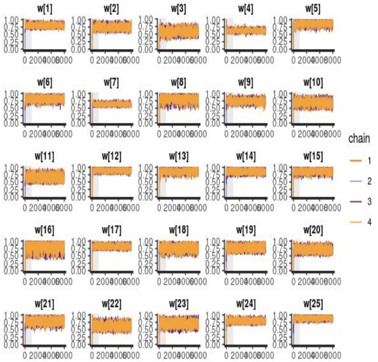
Indicating the bad fit and correlation between lambda and other parameters



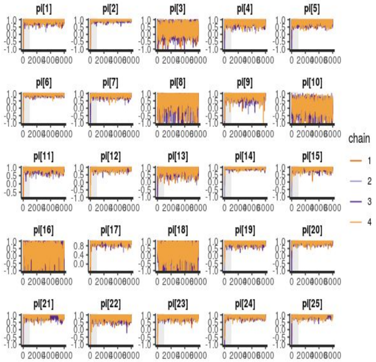
The worst outlier (#24) shows bad distribution in pair plot

Result: Social Learning Model w.o. RPE

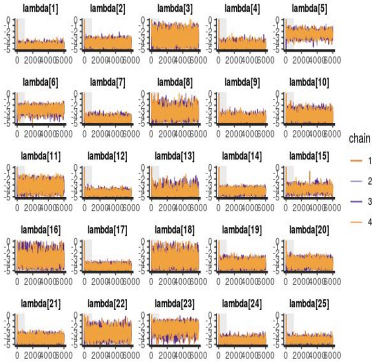
Group estimate, Traceplot



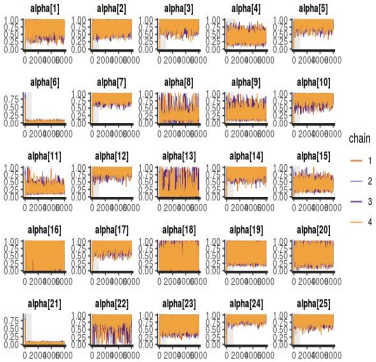
omega



pi



lambda

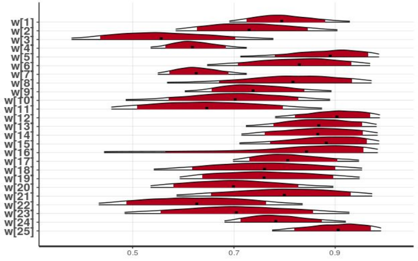


alpha

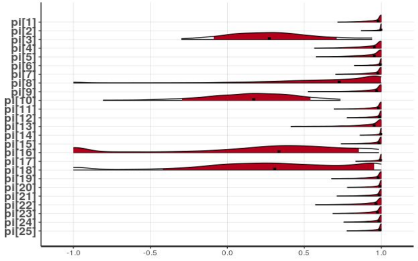
Result: Social Learning Model w.o. RPE

Group estimate, Posterior Distribution

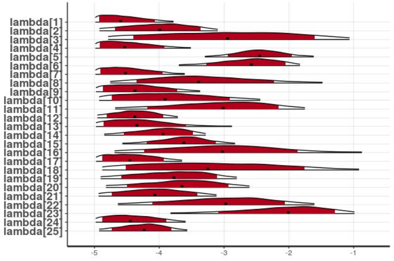
Distributed along the boundary, implies bad fit



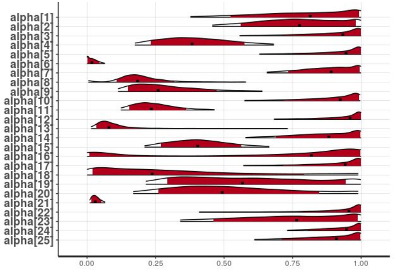
omega



pi



lambda

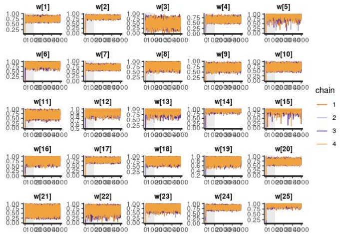


alpha

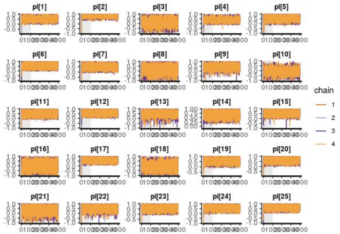
Result: Social Learning Model w.o. RPE, w. Inv Temp

Individual estimate, Traceplot

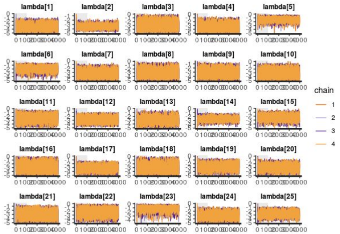
Acquired more stable traceplots in relation to other models



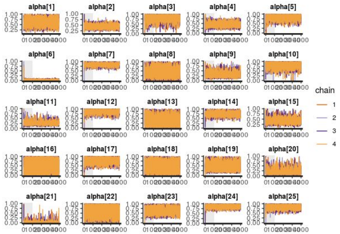
omega



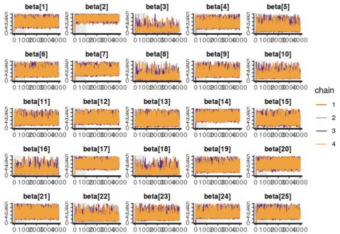
pi



lambda



alpha

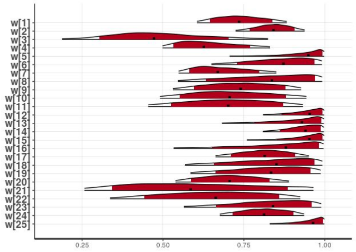


beta

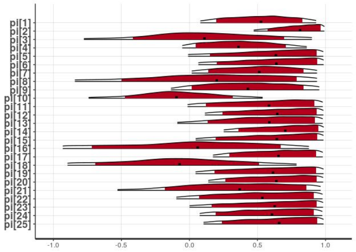
Result: Social Learning Model w.o. RPE, w. Inv Temp

Individual estimate, Posterior Distribution

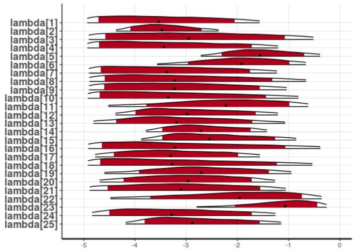
Although widely distributed, the posterior distributions look normal



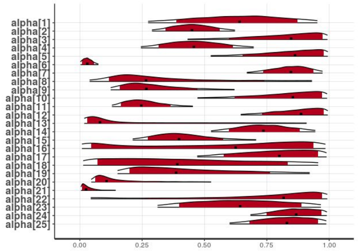
omega



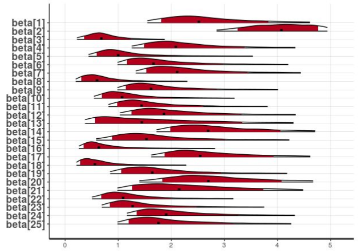
pi



lambda



alpha

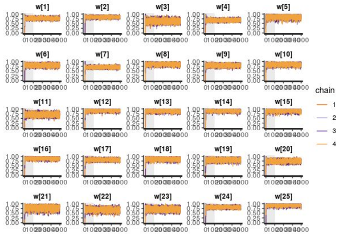


beta

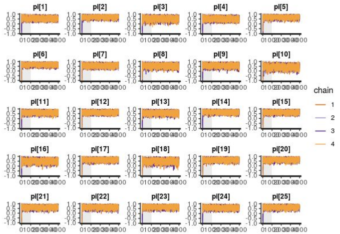
Result: Social Learning Model w.o. RPE, w. Inv Temp

Group estimate, Traceplot

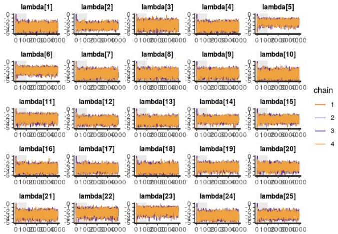
Compared to the non-hierarchical model of current version, traceplot converged better



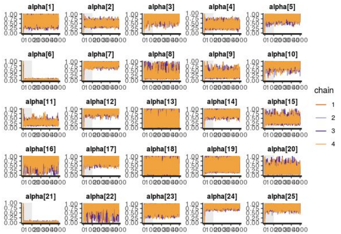
omega



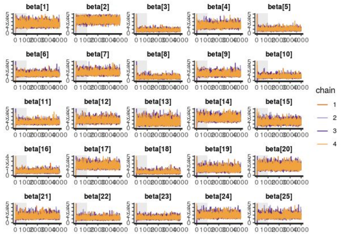
pi



lambda



alpha

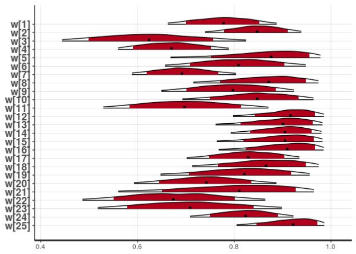


beta

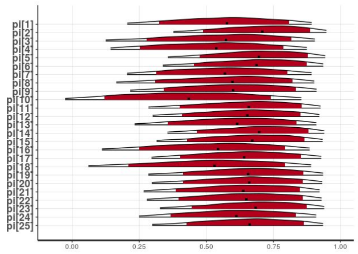
Result: Social Learning Model w.o. RPE, w. Inv Temp

Group estimate, Posterior Distribution

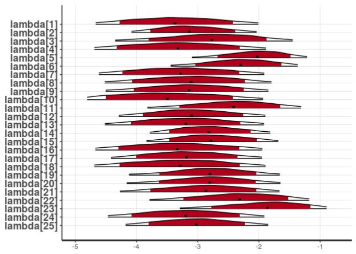
Although they are shrunk, Posterior Distributions didn't attach to the boundary



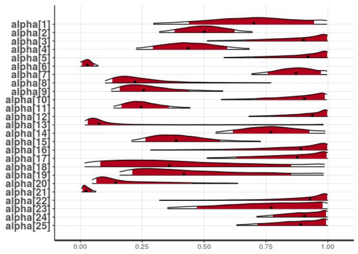
omega



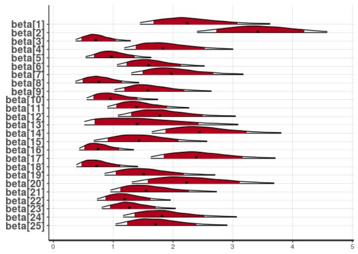
pi



lambda



alpha



beta

Overall

LOOIC score showed the ***hierarchical version*** of the social learning model ***without RPE*** and ***with Inverse Temperature*** was the best model among 6 models.

In addition, the traceplot and the posterior distribution advocate the ***hierarchical version*** of the social learning model ***without RPE*** and ***with Inverse Temperature*** as well.

Discussion

1. Much better log-likelihood via stan fit
2. LOOIC showed the best fit in hierarchical-RPE+Temp model
3. Fitting λ and π problem
4. Adaptive design to be intervened
5. Applicability for other tasks

Much better log-likelihood via stan fit

※The author computed BIC wrong; They rather deducted the amount of $k \ln N$

Even the same model (original social learning model) showed different log likelihood (when measured by elpd)

← -2482(stan) vs -3399(matlab)

Original article measured BIC by quasi-newton method iteration 1E10 times, but the model has complex structure in that the authors could not acquire the global optimum

	elpd_diff	se_diff	elpd_loo	se_elpd_loo
model3	0.000	0.000	-2389.144	71.398
model6	-3.329	6.686	-2392.473	73.494
model5	-65.970	20.573	-2455.114	59.961
model1	-78.778	23.248	-2467.922	62.083
model4	-93.503	24.395	$\boxed{-2482.646}$	58.190
model2	-2000.780	453.263	-4389.924	461.357

Elpd acquired by Stan (HMC-nuts method)

Models						BIC
Social learning	ω	π	λ	α	θ	
	0.65 ± 0.07	0.16 ± 0.16	−2.10 ±0 .45	0.51 ± 0.06	0.13 ± 0.02	$\boxed{-6149}$
Myopic	ω	π	λ	α	θ	
	11.04 ± 2.76	0.34 ± 0.19	−2.98 ± 1.09	0.59 ± 0.06	0.14 ± 0.04	−5721
Group utility	χ	ζ	α	θ		
	9.30 ± 4.55	−30.91 ± 4.92	0.56 ± 0.06	0.12 ± 0.02		−5713
Inequity aversion	δ	ε	κ			
	16.39 ± 7.11	0.61 ± 0.42	1.37 ± 1.00			−4220

BIC score acquired by Matlab (quasi-newton method)

Best LOOIC score: hierarchical-RPE+Temp model

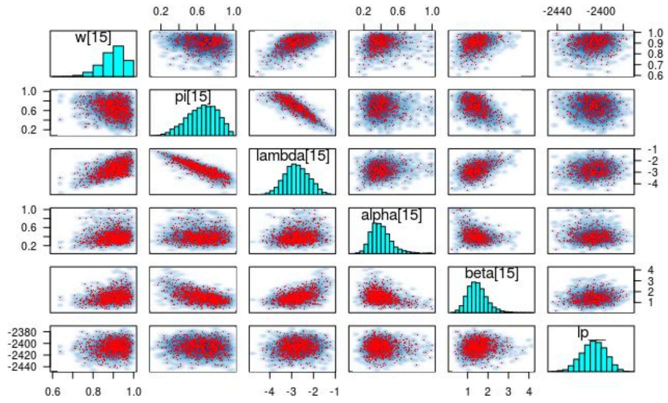
	elpd_diff	se_diff	elpd_loo	se_elpd_loo	p_loo	se_p_loo	looic	se_looic
model3	0.000	0.000	-2389.144	71.398	73.503	4.586	4778.287	142.795
model6	-3.329	6.686	-2392.473	73.494	80.714	4.467	4784.946	146.988
model5	-65.970	20.573	-2455.114	59.961	77.855	4.098	4910.227	119.922
model1	-78.778	23.248	-2467.922	62.083	65.410	4.667	4935.843	124.165
model4	-93.503	24.395	-2482.646	58.190	72.015	3.065	4965.293	116.380
model2	-2000.780	453.263	-4389.924	461.357	2012.157	454.010	8779.848	922.714

This model showed relatively non-skewed posterior distribution

Still, pareto-k-value indicates that the expected log predictive density may be unreliable

Fitting λ and π problem

Although the parameters (i.e. λ, α, π, θ, ω) account for distinguishable cognitive processes, for the actual computation these parameters can be correlated.



Pair plot shows correlation between λ and other parameters, especially between λ and π

※Pair plot was chosen from the worst log-likelihood pair from the best model (15)

Fitting λ and π problem

So to speak, ω, π, λ are the parameters to compute the ***scalar* Q**

$$p(C_t) = \mathrm{logit}(Q_t)$$
$$Q_t = \omega I_t + (1-\omega) G_t$$

If Γ converges to the specific value, ω, π may co-vary with λ

$$I_t = \lambda + \Gamma^{N-k} R + \pi \Gamma^{N-k} R (N-1) \qquad G_t = \sum_{j=t}^{T} R K^{T-j} \sum_{i=0}^{N-k} \Gamma_t^i$$

Thus, the possible suspects include:

- If γ (the estimated probability for other's choice to free-ride) converges, values of λ and π would diverge (e.g. -3 + 5 = -2 + 4 = …)
- Moreover, ω may also diverge as well

Adaptive design to be intervened

For the arisen fitting problem as suggested, the possible solution would be applying the adaptive task design

→ Γ (γ, the estimated probability for other's choice to free-ride) to be volatile so that ω, π do not covary with λ

→ Yet, the cognitive interpretation should be cautious with respect to this setup

(i.e. if the other's choice do not seem to converge, how would the participant's ***attitude for others(π)*** change accordingly?)

→ **Simulation** and **Parameter recovery** might inform whether this setup works

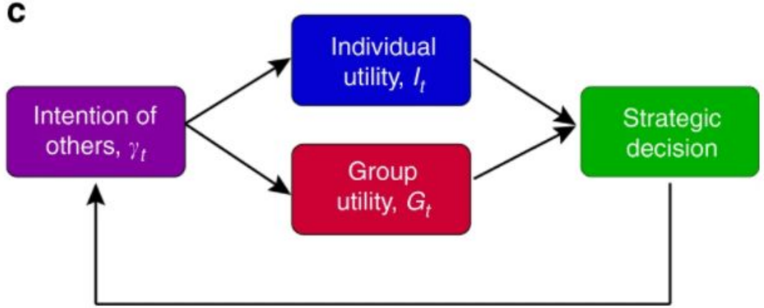
Verifying the models (including the modified ones) with another PGG datasets is needed as well

Applicability for other tasks

Although the model has the complex structure,

the ***framework*** of this model is straightforward

c



$$I_t = \lambda + \Gamma^{N-k} R + \pi \Gamma^{N-k} R \,(N-1)$$

$$G_t = \sum_{j=t}^{T} R K^{T-j} \sum_{i=0}^{N-k} \Gamma_t^i$$

By simplifying the equation of utility parts and replacing the parts to the task specific parameters, the model can be applied to the different game-theoretical tasks as well

Reference

[1] Lockwood, P. L., Apps, M. A., & Chang, S. W. (2020). Is there a ‘social’ brain? Implementations and algorithms. *Trends in Cognitive Sciences.*

[2] Hill, C. A., Suzuki, S., Polania, R., Moisa, M., O'doherty, J. P., & Ruff, C. C. (2017). A causal account of the brain network computations underlying strategic social behavior. *Nature neuroscience*, *20*(8), 1142.

[3] Ong, D. C., Zaki, J., & Goodman, N. D. (2019). Computational models of emotion inference in theory of mind: A review and roadmap. *Topics in cognitive science*, *11*(2), 338-357.

[4] Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*(4), 1-10.

[5] Devaine, M., Hollard, G., & Daunizeau, J. (2014). Theory of mind: did evolution fool us?. *PloS One*, *9*(2), e87619.

[6] Cutler, J., & Campbell-Meiklejohn, D. (2019). A comparative fMRI meta-analysis of altruistic and strategic decisions to give. *Neuroimage*, *184*, 227-241.

[7] Khalvati, K., Park, S. A., Mirbagheri, S., Philippe, R., Sestito, M., Dreher, J. C., & Rao, R. P. (2019). Modeling other minds: Bayesian inference explains human choices in group decision-making. *Science advances*, *5*(11), eaax8783.

[8] Park, S. A., Sestito, M., Boorman, E. D., & Dreher, J. C. (2019). Neural computations underlying strategic social decision-making in groups. *Nature communications*, *10*(1), 1-12.