

# Supplementary Material

## Theory of Mind with Guilt Aversion Facilitates Cooperative Reinforcement Learning

**Dung Nguyen**

**Svetha Venkatesh**

**Phuoc Nguyen**

**Truyen Tran**

*Applied Artificial Intelligence Institute, Deakin University, Australia*

DUNG.NGUYEN@DEAKIN.EDU.AU

SVETHA.VENKATESH@DEAKIN.EDU.AU

PHUOC.NGUYEN@DEAKIN.EDU.AU

TRUYEN.TRAN@DEAKIN.EDU.AU

**Editors:** Sinno Jialin Pan and Masashi Sugiyama

We present more detail about the grid-world version of Stag-Hunt Games in Section 4.2. Table S.1 shows the rewards given to agents in the grid-world version of Stag-Hunt Games. Figure S.1 shows the proportion of following cooperative policies of agents in two cases with the confidence interval.

	Agent 1	Agent 2
Two agents catch stag	4.0	4.0
Two agents catch hare	2.0	2.0
Agent 1 catches stag, Agent 2 catches hare	3.0	0.0
Agent 2 catches stag, Agent 1 catches hare	0.0	3.0

Table S.1. Rewards given to agents in the grid-world version of Stag-Hunt Games.

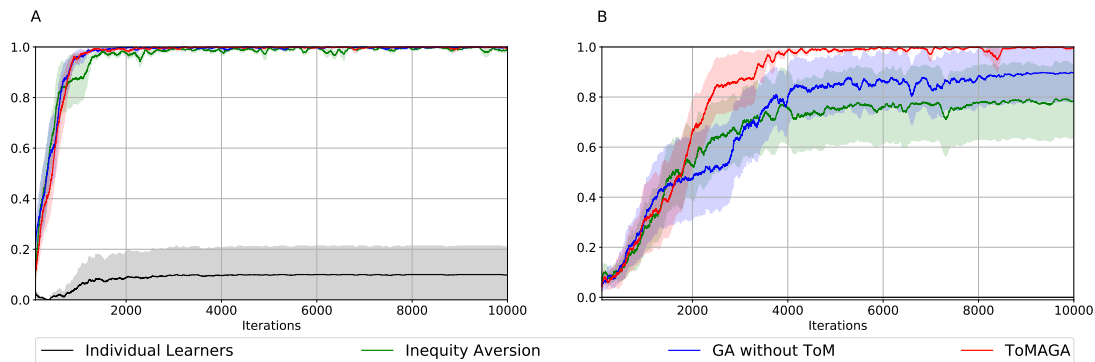


Figure S.1. Proportion of following cooperative policies of individual learners, agents with inequity aversion, GA agents without ToM, and ToMAGAs when they start nearby the stag (A) and nearby hares (B). Proportion of following cooperative (y-axis) vs Iterations (x-axis).