# On the Value of Information Structures in Stochastic Games

by Daehyun Kim & Ichiro Obara

**Report Prepared by Chung-Hao Sheu**

# Motivation

- Stochastic games are very useful tools to study dynamic interactions of strategic agents.
- The nature of available information is important as information is what links current behavior and future behavior in stochastic games.
- If the players can only monitor imperfectly and get imperfect public signals, which can't tell the players what event actually occurred, what the set of attainable long-run average discounted perfect public equilibrium (PPE) payoff in a stochastic game changes when the players improve the quality of monitoring or information structure?
- Kim (2019) already explains the set of the PPE payoff either shrinks or expands for a fixed level of patience, that is, a fixed discount factor.
- What is the answer when the players can get arbitrarily patient?

# Main Research Question

If the players can improve the quality of monitoring or information structure, does the set of PPE payoff in a stochastic game (weakly) expand when the players get arbitrarily patient?

# Example

- To make the question more precise, consider the following scenario.

- Countries can sign up an climate agreement about greenhouse gas emissions.
- However, they can't verify each others' level of emissions in each period. If a country cheats in some period, other can only observe an increase in either atmospheric greenhouse emissions or industrial production, which is an imperfect public signal.
- Assume all countries care about sustainable development, and thus we can view them as arbitrarily patient players in an infinite stochastic game.
- If they can improve the quality of monitoring in greenhouse gas emissions, can they get into a better perfect public equilibrium (PPE)?

# Terms

- **Stochastic game:**

A stochastic game is a dynamic game in which players repeat the same set of stage games.

- **Repeated game:**

A repeated game is adynamic game in which players repeat the same stage game. One can view the repeated game as a special case of stochastic game.

# Model Setup

- Stage $s, t \in S$, which is finite and denotes the set of all possible stage game.
- There are $N$ players. Player $i$ chooses action $a_i \in A_i$, which is finite and denotes the set of all actions available for player $i$.
- In each period, the new state $t$ is drawn randomly depending on the current state $s$ and players' action $a = (a_1, \dots, a_N)$. The transition law is $q(\cdot|s, a) \in \Delta S$, which is the set of all possible distribution over $S$.
- Players observe a public signal $y \in Y$, which is finite and denotes the set of all public signals. The signal $y$ is drawn according to $f(\cdot|s, t, a) \in \Delta Y$, which is the set of all possible distribution over $Y$.
- An **information structure** is a pair $\pi = (f, Y)$, and a pair of $\pi$ and $q$ defines a joint distribution on $S \times Y$ conditional on any $(s, a)$ by $p(t, y|s, a) = f(y|s, t, a)q(t|s, a)$

# Definition: Utility and Payoff

- Player $i$'s utility: $u_i: A \times S \to \mathbb{R}$.
- A public history is $h^k = \left(s^0, y^0, \dots, s^{k-1}, y^{k-1}, s^k\right)$ for each period $k$, and the set of public histories is $H = \cup_{k=1}^{\infty} h^k$.
- Discount factor $\delta \in [0, 1)$.
- A public strategy for player $i$, is a mapping $\sigma_i: H \to \Delta A_i$.
- Player $i$'s long-run average discounted payoff is

$$U(\sigma; s) = (1 - \delta)\Sigma_{k=1}^{\infty} \delta^k E_\sigma\left[u_i\left(a^k, s^k\right) | s^0 = s\right]$$

# Main Result

## Theorem:

Improving the quality of monitoring garbling weakly expands the limit PPE payoff set in a stochastic game when the players are arbitrarily patient.

- Improving the quality of monitoring means that the information structure get more informative in terms of weighted garbling.
- Limit PPE payoff set is the set of attainable long-run average discounted PPE payoff.
- The players being arbitrarily patient means that all players can make discount factor $\delta \to 1$.

# Intuition

- Given a feasible continuation payoff, we can decompose it into two different components each of which may be called the physical transition part and information part, respectively.
- Given a certain subsequent state, the physical transition component is defined as the maximum welfare level that is followed by some public signal. It is thus independent of the public signal and doesn't change when we vary information structure.
- The information component is defined as the rest part of the continuation payoff. It does depend on public signals and subsequent state. By construction, this part involves non-positive welfare level, which may be interpreted as the welfare loss which is to be incurred to provide incentives to players to conform equilibrium play.
- The decomposition allows us to apply some garbling notion only to the information component of continuation payoffs.

# Intuition

- There are two effects on weighted garbling when $\delta \to 1$.

1. Players may wait until they receive more decisive signals and make variations in continuation payoffs proportionally larger depending on it. The weights in the definition of weighted garbling captures a possibility of this. When $\delta$ is small, we may not ensure that the proportional variations give enough incentives, because it takes some cost of waiting. This cost vanishes as players are arbitrarily patient.
2. As $\delta$ increases, the impact of the initial state on the set of feasible continuation payoffs is smaller and vanishes in the end.

# Discussions

- If we use other definition of 'more informative', can we get the same result?
- Can we further know the conditions in which the set of PPE payoffs strictly expands?
- What will happen when there are some actions to improve the quality of monitoring in the stage games?

# Definition: Perfect Public Equilibrium

- A public strategy profile $\sigma$ is perfect public equilibrium (PPE) if for each $i$ and $h \in H$,

$$U_i\left(\sigma|_h; s(h)\right) \geq U_i\left(\sigma_i'|_h, \sigma_{-i}|_h; s(h)\right), \forall \sigma_i',$$

where $s(h)$ is the most recent state of $h$.

# Definition: Garbling and 'More Informative'

**Weighted garbling:**
An information structure $\pi = (f, Y)$ is a weighted garbling of $\pi' = (f', Y')$ if, $\forall s \in S$, there exists weights $\gamma_s^{t',y'} \geq 0$, $(t', y') \in S \times Y'$ and $\phi_s: S \times Y' \rightarrow \Delta(S \times Y)$ such that, for each $a$,

$$p(t, y|s, a) = \sum_{(t',y') \in S \times Y'} \gamma_s^{t',y'} \phi_s(t, s|t', y') p'(t', y'|s, a) \ , \forall t, y$$

- Note that the average weights must be 1 for each $s \in S$.
- We say $\pi'$ is more informative than $\pi$ in terms of weighted garbling if $\pi$ is a weighted garbling of $\pi'$.