Концепции решения игры Г2.

Целевые функции игроков:

$w_1 = f_1(x_1, x_2)$,

$w_2 = f_2(x_1, x_2)$.

$x_1 \in X_1^0$, $x_2 \in X_2^0$.

Будем считать, что $f_2(x_1, x_2)$ непрерывна по x_1 при любом x_2.

$$\sup_{x_2 \in X_2^0} f_2(\tilde{x}_1(x_2), x_2) > L_2,$$

$x_2 \in X_2^0$ – компактные множества действий.

Предполагается следующий порядок функционирования: игрок 1, обладая правом первого хода, сообщает игроку 2 план выбора своей стратегии в зависимости от выбранной игроком 2 стратегии x_2.

Стратегия игрока 1 (центра) – условное действие $\tilde{x}_1 = \tilde{x}_1(x_2)$

После этого второй игрок выбирает действие x_2, максимизируя свою целевую функцию с подставленной туда стратегией первого игрока, а затем первый игрок – действие $\tilde{x}_1(x_2)$.

Стратегия наказания

Стратегия наказания $x_1^н = x_1^н(x_2)$ определяется из условия

$$f_2(x_1^н(x_2), x_2) = \min_{x_1 \in X_1^0} f_2(x_1, x_2).$$

Если стратегий наказания несколько, то *оптимальной стратегией наказания* называется та из них, на которой достигается максимум выигрыша первого игрока. Гарантированный результат второго игрока при использовании первым игроком стратегии наказания

$$L_2 = \max_{x_2 \in X_2^0} f_2(x_1^н(x_2), x_2) =$$

$$= \max_{x_2 \in X_2^0} \min_{x_1 \in X_1^0} f_2(x_1, x_2).$$

Множество действий второго игрока, обеспечивающих ему максимальный выигрыш при использовании первым игроком стратегии наказания

$$E_2 = \{x_2 \mid f_2(x_1^н(x_2), x_2) = L_2\}.$$

Множество достижимости

Договорное множество рассматриваемой игры, то есть множество сочетаний стратегий первого и второго игроков, которые гарантировали бы второму результат, строго больший того, что тот может получить даже при наихудших для него действиях первого игрока (то есть при использовании первым игроком стратегии наказания).

$$D = \{(x_1, x_2) : f_2(x_1, x_2) > L_2\}$$

Наилучший результат первого игрока на множестве достижимости

$$K = \begin{cases} \sup\limits_{(x_1, x_2) \in D} f_1(x_1, x_2), & D \neq \varnothing \\ -\infty, & D = \varnothing \end{cases}$$

Принадлежность ситуации множеству достижимости гарантирует реализуемость этого результата путем использования стратегии наказания.

Действие игрока 1, реализующее $K-\varepsilon$ при выборе игроком 2 рекомендуемого действия из D

$$(f_1(x_1^{\varepsilon}, x_2^{\varepsilon}) \geq K-\varepsilon ,$$

$$(x_1^{\varepsilon}, x_2^{\varepsilon}) \in D \neq \varnothing).$$

Гарантированный результат центра при применении им стратегии наказания (так как стратегии игрока 2 ограничены множеством E_2):

$$M = \inf_{x_2 \in E_2} \sup_{x_1 \in X_1^0} f_1(x_1, x_2)$$

***ε*-доминантная стратегия.** Стратегия $x_1^{a\varepsilon}(x_2)$ реализует (с точностью ε) наилучший ответ игрока 1 на действие x_2 игрока 2, то есть

$$f_1(x_1^{a\varepsilon}(x_2)) \geq \sup_{x_1 \in X_1^0} f_1(x_1, x_2) - \varepsilon .$$

Теорема Ю.Б. Гермейера.

В указанных условиях наибольший гарантированный результат центра равен

$$\max\,[K,\,M\,].$$

При $K>M$ ε-оптимальная стратегия игрока 1

$$\widetilde{x}_1^{\varepsilon}(x_2)=\begin{cases} x_1^{\varepsilon}, & при\, x_2=x_2^{\varepsilon} \\ x_1^{н}(x_2), при\, x_2\neq x_2^{\varepsilon} \end{cases}.$$

При $K \le M$ оптимальная стратегия игрока 1 заключается в применении оптимальной стратегии наказания.

Метаигры

Игра Γ_3 - усложнение игры Γ_1 (в некотором роде):

Если центр не планирует самостоятельно получить информацию о действии агента, он может первым выбрать действие, реализуя игру Γ_1. Однако ему можно порекомендовать и более сложное поведение. Центр может попросить агента сообщить ему свою стратегию $x_2 = \tilde{x}_2(x_1)$, которая основана на ожидаемой агентом информации о действии центра. Реализация права первого хода центром состоит в этом случае в сообщении агенту стратегии $\tilde{\tilde{x}}_1(\tilde{x}_2(x_1))$. Эту стратегию можно интерпретировать, как обещание центра выбрать действие $\tilde{\tilde{x}}_1(\tilde{x}_2(x_1))$ при условии, что агент обещает выбирать свое действие в соответствии с $\tilde{x}_2(x_1)$. Так образуется игра Γ_3. Здесь также не рассматривается возможность блефа, как со стороны центра, так и со стороны агента.

Если центр определяет порядок обмена информацией, он может выбирать, играть ему Γ_1 или Γ_3. В обеих играх центр вынужден выбирать действие, не зная действия, выбранного агентом.

Аналогично тому, как, с помощью образования дополнительной «петли обратной связи», из Γ_1 была образована Γ_3, можно усложнить и игру Γ_2. Так образуется игра Γ_4. В ней агент, ожидая от центра, как и в Γ_2, информацию вида $\tilde{x}_1(x_2)$, формирует и сообщает центру свою стратегию $\tilde{\tilde{x}}_2(\tilde{x}_1)$. Центр, обладающий правом первого хода, пользуется стратегиями $\tilde{\tilde{\tilde{x}}}_1(\tilde{\tilde{x}}_2)$, которые определяют, какую функцию $\tilde{x}_1(x_2)$ выберет центр в зависимости от сообщения агента $\tilde{\tilde{x}}_2$.

Таким же способом можно на основе Γ_3 построить игру Γ_5, и так далее.

В каждой из построенных четных игр

$$\Gamma_{2m},\ m = 1, 2\ldots,$$

центр использует в качестве стратегий отображения множества стратегий агента в этой игре на множество стратегий центра в игре $\Gamma_{2m\text{-}2}$. Аналогично, стратегиями агента являются отображения множества стратегий центра в Γ_{2m} на множество стратегий агента в игре $\Gamma_{2m\text{-}2}$.

Такую рефлексию можно было бы наращивать бесконечно, переходя к все более сложным схемам обмена информацией, если бы рассмотрение этих игр увеличивало выигрыш центра (в интересах которого и проводится исследование всех метаигр).

Наилучшие метаигры

Теорема. Максимальный гарантированный результат центра в игре Γ_{2m} при $m > 1$ равен максимальному гарантированному результату центра в игре Γ_2. В играх же Γ_{2m+1} при $m > 1$ максимальный гарантированный результат центра равен его максимальному гарантированному результату в игре Γ_3.

Таким образом, при исследовании гарантированного результата центра можно ограничиться исследованием только игр Γ_1, Γ_2 и Γ_3. Следующая теорема устанавливает взаимосвязь между гарантированными выигрышами центра в этих играх:

Теорема. Максимальный гарантированный результат центра в игре Γ_2 не меньше его гарантированного результата в игре Γ_3, а тот, в свою очередь, не меньше гарантированного выигрыша в игре Γ_1.

Этот результат показывает, что Γ_2 является «идеальной» игрой для центра. Соответственно, если центр имеет возможность определять порядок и содержание обмена информацией, и, кроме того, при выборе своего действия знает действие, выбранное агентом, он должен играть Γ_2. Если центр на момент выбора своего действия не знает действия агента – ему наиболее выгодна игра Γ_3.

Равновесия в метаиграх

Теорема. В игре Γ_{2m} (при $m \geq 1$) те, и только те, исходы (x_1^0, x_2^0), которые удовлетворяют условиям

$$f_1(x_1^0, x_2^0) \geq \min_{x_2 \in X_2^0} \max_{x_1 \in X_1^0} f_1(x_1, x_2), \quad f_2(x_1^0, x_2^0) \geq \max_{x_2 \in X_2^0} \min_{x_1 \in X_1^0} f_2(x_1, x_2),$$

могут быть ситуациями равновесия Нэша игры со стратегиями $(\tilde{x}_1^0, \tilde{\tilde{x}}_2^0)$ (здесь стратегии понимаются в метаигровом смысле, как функции информированности соответствующей метаигры).

Таких исходов действительно может быть очень много. В этом смысле метаигровые схемы можно рассматривать, как средство расширения множества равновесий Нэша, если множество равновесий исходной игры почему-то не устраивает исследователя или одного из игроков.