Available online at www.sciencedirect.com

ScienceDirect

IFAC-PapersOnLine 49-32 (2016) 019–023

Adding Informational Beliefs to the Players Strategic Thinking Model

Alexander G. Chkhartishvili*, Vsevolod O. Korepanov*

* V. A. Trapeznikov Institute of Control Sciences of Russian Academy of Sciences, Moscow, Russia (e-mail: sandro_ch@mail.ru, vkorepanov@ipu.ru).

This work is partially supported by the Russian Scientific Foundation, project no. 16-19-10609

Abstract: We propose the model that connects the strategic reflexion/strategic thinking model and the information reflexion model of players in game theory. This model can describe both types of beliefs: players' beliefs about opponents' thinking models and players' beliefs about their opponents' awareness about the external parameter value. We consider some general examples and also the example of the Cournot oligopoly model.

Keywords: cognitive control; smart infrastructure; strategic reflexion; strategic thinking; information reflexion.

1. INTRODUCTION

In game theory there exist several types of thinking models of players with bounded rationality. Among them there are the ***strategic reflexion/strategic thinking model*** and the ***information reflection model***. The basic strategic reflexion/strategic thinking model could be modified to improve its predictive power (Novikov & Chkhartishvili 2014b).

Currently these models are being actively developed in behavioral game theory as a tool to predict human's behavior. These models have many applications in such areas as markets (Wright et al. 2012), bargaining (Wright et al. 2012), auctions (Crawford & Iriberri 2007), lotteries (Östling et al. 2011), and as a tool for control problems (Novikov & Korepanov 2012). On the other hand, they are being implemented in multiagent systems to specify the behavior of software agents which act in place of human decision-makers (Wunder 2011, Sonu 2015).

The information reflexion model (Novikov & Chkhartishvili 2014a) is used for games with external parameters on which players' utility functions depend. Players are interested to find the values of these parameters and what opponents know about the parameters and what opponents think about their opponent's knowledge of the parameters etc.

Historically, models of strategic and information reflexion have been studied independently. In this paper we propose a general model which allows describe both types of reflexion simultaneously.

2. MODEL

2.1 Strategic reflexion model

Here we describe the basic and well known strategic reflexion model.

For simplicity, consider a game of three players in the normal form $G = (N, A, u)$, where $N = \{1, 2, 3\}$ is the set of players, $A = \prod_{i\in N} A_i$ is the set of possible action profiles with A_i being the set of actions available to player *i,* and $u = \{u_i\}_{i\in N}$ is the set of utility functions $u_i : A \to R$ mapping the action profile to the utility for player *i*. Players act simultaneously, independently and one-shot by choosing an action profile from their action sets. The result of the game is the action profile $a=(a_1,a_2,a_3)$ of actions of three players.

We fix the action profile $a^0 = (a_1{}^0,a_2{}^0,a_3{}^0) \in A$ which corresponds to an obvious (in some sense) action profile and which may reflect such qualities of a player as fairness, discretion etc. (we only consider games that have at least one such an obvious profile).

Definition. Define the best response function of player *i* br_i: $A\to A_i$:

$$br_i(a) = C\left(Arg \max_y u_i(y, a_{-i}) \right),$$

where $a \in A$, a_{-i} is the joint actions of all players except for *i* and *C* is a choice function that uniquely picks one action from a subset of actions.

We say that the player *i* has the zero rank ("rank 0") of reflexive thinking if he chooses the action a_i^0. Some player *i* may have a belief that all his opponents have rank 0. Then he predicts (perhaps, erroneously) opponents' actions, and his subjective optimal action is:

$$a_i^1 = br_i(a^0), \qquad (1)$$

where $a^1{}_i$ is the action of the player *i*, superscript '1' means that we say that such player has rank 1.

If there is two players *i* and *j* of rank 1, then they have wrong beliefs about each other: *i* think that *j* has 0 rank, and *j* think that *i* has 0 rank. In other words we can say that there is the real player *j* and the *phantom* player *j* – existing only in the mind of player *i*. The same true for player *i*.

Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.12.183

Beliefs of the player can be represented as an oriented tree whose nodes are players, real or phantom (i.e. existing only in the mind of other players). An arrow from player A to player B would mean that player B knows all information about player A (hereinafter, we will call a real or phantom player just a "player", unless the opposite is stated).

Such a tree, as an example for $i = 1$, is shown on Fig. 1.

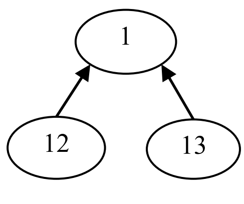

Fig. 1. An example of a belief tree.

On the diagram, node '1' is a real player 1, nodes '12' ("one-two", not twelve) and '13' are players 2 and 3 in beliefs of player 1. Note that nodes '12' and '13' are "hanging", i.e. that phantom players 12 and 13 do not have beliefs about opponents. This fact means that such players have rank 0 in our model.

It is generally assumed (Novikov, Chkhartishvili 2014b) that the rank of a player is one greater than the largest rank that his opponents have, in his opinion. So player 1 on Fig. 1 has rank 1.

If player 1 thinks that his opponents have rank 1, for example, then he has rank 2 and his subjective optimal action is

$$a_1^2 = br_1(a_2^1, a_3^1)\ . \quad (2)$$

Similarly, if player 1 thinks that his opponents 2 and 3 have ranks 0 and 1, then he also has rank 2 and his action is described as

$$a_1^2 = br_1(a_2^0, a_3^1)\ . \quad (3)$$

Graphs for these cases are shown on Fig. 1.

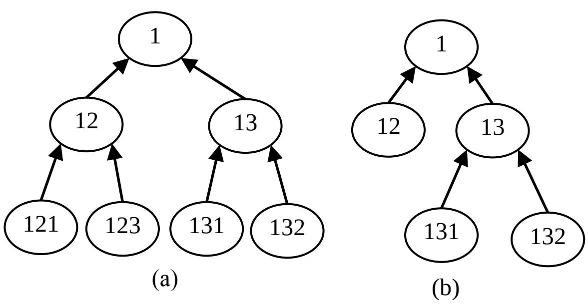

Fig. 2. (a) beliefs of player 1 that his opponents are rank 1 players, (b) beliefs of player 1 that his opponents has rank 0 and 1.

Thus, the rank of a player equals 0, if he is a leaf node of a tree (there are no incoming arrows), otherwise it equals to the number of arrows in maximal length path from terminal nodes to the player.

So players of rank 2 and more must have beliefs about beliefs of their opponents and so on (these are players 121, 123 etc. on Fig. 2).

For a complete description of the situation we need to define beliefs of all players which creates a hierarchic information structure. An example of such a structure is given in Fig. 3.

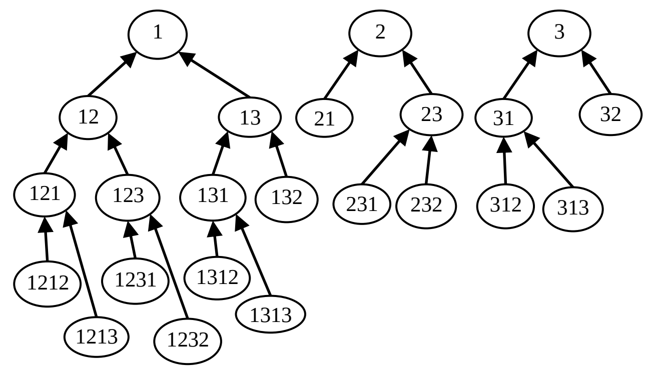

Fig. 3. All players' beliefs example

We are going to merge some players as they have equal beliefs and, consequently, equal actions. For example, 1212 and 132 on Fig. 3 are phantom players 2 with rank 0 who act identically, so we can merge them in one node. In other words, players 121 and 13 have equal beliefs about player 2. Furthermore, if we identify in this way all rank 0 players, we can see that players 12 and 2 have equal beliefs about player 3.

After completing all merging operations, we get a more visual graph of players' beliefs, ***the graph of a reflexive game*** (Novikov Chkhartishvili 2014a), where each node represents a real or a phantom player (Fig. 4).

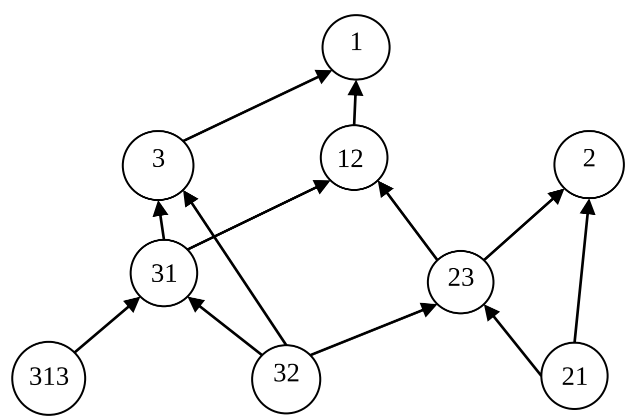

Fig. 4. The graph of a reflexive game.

As we can see, the graph of a reflexive game displays the following information:

- ranks of players;
- adequate or inadeqate beliefs: player 1 is adequately informed about player 3 (i.e. the image of player 3 in mind of player 1 matches with real player 3), but he is not adequately informed about the player 2.

- equivalent beliefs: players 3 and 12 have the same beliefs about the player 1 (though incorrect).

Note that double-sided arrows are not possible in our model unlike in an arbitrary the graph of a reflexive game. This follows from the assumption that each player thinks that his opponents have a lower rank and rank 0 players do not have beliefs about opponents at all.

2.2 Informational reflexion model

To develop our model further, we append the information reflexion to the existing hierarchic beliefs structure. Let $\theta \in \Theta$ be the external system parameter on which players' utility functions are dependent on: u_i: $\Theta \times A \to R$. The second assumption is that players' have common knowledge that the parameter value is not known exactly – each player can estimate it differently. In other words player i have belief about θ: $\theta_i \in \Theta$.

Now, to predict player j best response action a player i must have an opinion about opponent's beliefs about the value of θ: $\theta_{ij} \in \Theta$ and his beliefs about his opponents beliefs and so on. We can put simple information reflexion model to our example case (fig. 4). Let all players believe that $\theta = \theta_a \in \Theta$. Then it can be displayed as the same graph as in fig. 4 – equal beliefs of all are not shown by convention. Here $\theta_1 = \theta_2 = \theta_3 = \theta_{12} = \theta_{23} = \theta_{31} = \theta_{21} = \theta_{32} = \theta_{313} = \theta_a$.

Let all players believe that $\theta = \theta_a$ but player 23 believes that $\theta = \theta_b \neq \theta_a$. Now we cannot merge players 23 and 123 since they have different beliefs (equal strategic but different information reflexion), and so they will be represented by different nodes in the graph of a reflexive game (opposite to the situation in Fig. 4). This is shown on Fig. 5 (here, beliefs of most players $\theta = \theta_a$ are not shown). Now players 2 and 12 have different beliefs about player 3. Despite the fact that player 23 thinks that $\theta = \theta_b$, he believes that his opponents (players 21 and 32) think that $\theta = \theta_a$.

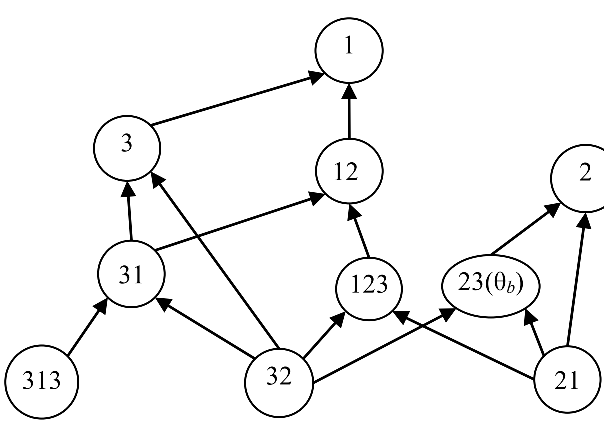

Fig. 5. The graph of a reflexive game with information reflexion.

It is clear that this way we can specify an arbitrary beliefs structure about the parameter θ, i.e. such a structure in which all players have definite beliefs about θ and about the opponents' beliefs about θ and so on, see Novikov, Chkhartishvili (2014a). The resulting model is a combination of the information reflexion model and the strategic reflexion model.

3. EXAMPLES

Consider the classic Cournot oligopoly model for three players, a single-product market with three manufacturers (the "players").

Each player can manufacture the product in the total amount ("volume") $a_i > 0$. Let market price be defined by the expression

$$M - b\sum_{i \in N} a_i \ , \tag{4}$$

where M is the market capacity and b is the price elasticity.

When production costs are not taken into consideration, the player's utility can be expressed as

$$u_i(a_1, a_2, a_3) = a_i(M - b\sum_{i \in N} a_i) . \tag{5}$$

3.1 Example 1. Strategic reflexion

We again start with introducing the strategic reflexion. Since all players are in equal conditions, we can assume that a rank 0 player thinks all players would act the same way and with maximum utility for each. Therefore the action profile of rank 0 players is $a^0 = (a, a, a)$ where a is an action that maximize utility $u_i(a,a,a) = a(M–3ba)$ for all i. So

$$a^0 = \arg\max_{x \geq 0} u_i(x, x, x) = \arg\max_{x \geq 0} x(M - 3bx) = M/(6b) \tag{6}$$

Let the players' beliefs be defined by the graph in Fig. 6.

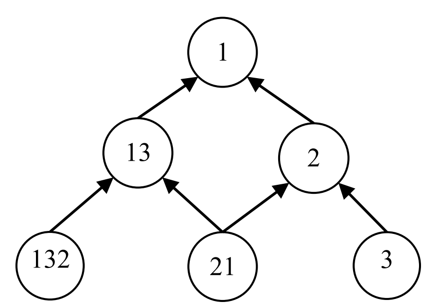

Fig. 6. The graph for the Cournot oligopoly model in Example 1.

We can calculate players' actions moving from the bottom to the top in the given graph of a reflexive game; the results are shown in Table 1. To calculate expected payoff of player 13 we take him action $a_{13} = br_3(a_{21}, a_{132})$ (his best response to his beliefs) and actions of players 21 and 132 (actions that 13 expect from opponents) and put them to utility function $u_3(a_{21}, a_{132}, a_{13})$. Real payoff is calculated only to real players, by put real actions to utility functions: $u_i(a_1, a_2, a_3)$.

Table 1. Example 1

player	action	expected payoff	real payoff
21	$a^0 = M / 6b$	$M^2 / (12b)$	-

132	$a^0 = M / 6b$	$M^2 / (12b)$	-
3	$a^0 = M / 6b$	$M^2 / (12b)$	$M^2 / (18b)$
2	$br_2(a_{21}, a_3) = M / 3b$	$M^2 / (9b)$	$M^2 / (9b)$
13	$br_3(a_{21}, a_{132}) = M / 3b$	$M^2 / (9b)$	-
1	$br_1(a_2, a_{13}) = M / 6b$	$M^2 / (36b)$	$M^2 / (18b)$

It is interesting that actions of players 1 and 21 turned out to be equal. Thus, player 2 appears as if he is adequately informed about player 1, and consequently he wins more than others (player 2 is adequately informed about player 3). Moreover, expectations of player 2 about his payoff and opponents actions (if it is allows by the game setup) will match reality.

3.2 Example 2. Strategic and information reflexion

Now we can introduce the information reflexion. Assume that the value of the parameter M is not known to players and each player has a belief M_i about it. Assume $M_i \in \{M, m\}$, and in reality the market capacity is M. It is possible that the awareness is as given in the table in Fig. 7 (here only wrong beliefs about the market capacity are marked '(m)').

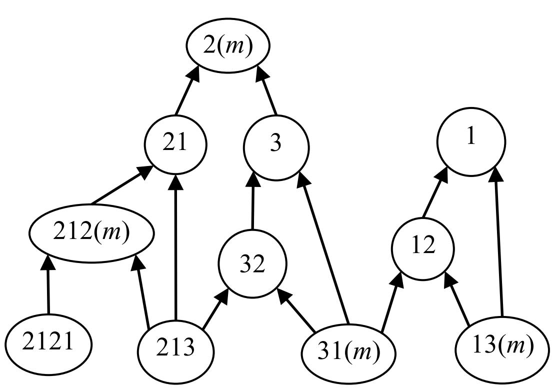

Fig. 7. The graph for the Cournot oligopoly model in example 2.

Table 2. Example 2

player	action	expected profit	real profit
2121	$a^0 = M / 6b$	$M^2 / (12b)$	-
213	$a^0 = M / 6b$	$M^2 / (12b)$	-
31	$a^0(m) = m / 6b$	$m^2 / (12b)$	-
13	$a^0(m) = m / 6b$	$m^2 / (12b)$	-
12	$(3M-m) / 6b$	$(3M-m)^2 / (36b)$	-
32	$(5M-m) / 12b$	$(5M-m)^2 / (12^2b)$	-
212	$(3m-M) / 6b$	$(3m-M)(5M-3m) / (36b)$	-
1	$M / 4b$	$M^2 / (16b)$	$M(41M-29m) / (4*48b)$
21	$(2M-m) / 4b$	$(2M-m)^2 / (16b)$	-
3	$(7M-m) / 24b$	$(7M-m)^2 / (24^2b)$	$(7M-m)(41M-29m) / (24*48b)$
2	$(31m-19M) / 48b$	$(31m-19M)^2 / (48^2b)$	$(31m-19M)(41M-29m) / (48^2b)$

As seen on Table 2, this situation is more complex and in general all players' expectations will not coincide with reality. It is clear that beliefs of each player can be set arbitrarily and that the expected and real payoff can be calculated according to those beliefs.

4. CONCLUSIONS

The general model combining strategic and information reflexion is presented. It turns out that such a combination does not add more complexity; it may only increase the number of nodes of the graph of a reflexive game. We expect that the ideas and results from the models of both reflexive types will extend to this new model. This may include, for example, the common knowledge of two players (which can be represented as a double arrow in the graph of a reflexive game) or unified model of all players of the same rank (such as in level-k, cognitive hierarchies, and reflexive partitions' models). (Novikov Chkhartishvili 2014b; Wright et al. 2012).

The advantage of the model is that it describes a broader class of strategic beliefs that can be modeled in one framework. We hope that this model could be used as a tool for setting up, analyzing and visualizing agents' strategic beliefs in multiagent systems and in game theory.

REFERENCES

Crawford, V., and Iriberri, N. (2007). Level-k auctions: Can boundedly rational strategic thinking explain the winner's curse and overbidding in private-value auctions? *Econometrica*, 75, 1721-1770.

Novikov, D.A., and Chkhartishvili, A.G. (2014a). *Reflexion and control: mathematical models*. CRC Press/Balkema, Leiden.

Novikov, D.A., and Chkhartishvili, A.G. (2014b) Mathematical Models of Informational and Strategic Reflexion: a Survey. *Advances in Systems Science and Applications*, 3, 254-277.

Novikov, D.A., and Korepanov, V.O. (2012). The reflexive partitions method in models of collective behavior and control. *Automation and Remote Control*, 73 (8), 1424-1441.

Östling, R., Wang, J.T., Chou, E.Y., and Camerer, C.F. (2011). Testing Game Theory in the Field: Swedish

LUPI Lottery Games. *American Economic Journal: Microeconomics*, 3(3), 1-33.

Sonu, E., Doshi, P. (2015). Scalable solutions of interactive POMDPs using generalized and bounded policy iteration. *Autonomous Agents and Multi-Agent Systems*, 29(3), 455-494.

Wright J.R., and Leyton-Brown K. (2012). Behavioral game-theoretic models: A Bayesian framework for parameter analysis. *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, 2, 921–928.

Wunder, M., Kaisers, M., Yaros, J. R., & Littman, M. (2011). Using iterated reasoning to predict opponent strategies. In *Proceedings of the 10th International Conference on Autonomous Agents and Multiagent Systems,* 2, 593-600.