RABBIT AND HUNTER :
A Partial Information Game.

P.BERNHARD, O.POURTALLIER
Institut National de Recherche en Informatique et Automatique,
2004 route des Lucioles - Sophia Antipolis
06565 Valbonne Cedex - FRANCE.

1 Introduction

Using “Rabbit and Hunter” as an example of a dynamic partial information game, we will present some important features of strategies.

After a description of the game rules, which are particularly simple, we will explore two ways of solving it. The first one which uses the normal form of the game, provides us with a rather simple method to obtain solutions, but unfortunately, due to the exponential growth of the number of strategies with the final time, we are inefficient to compute them. The second one introduces “behavioral strategies”. We will first compare mixed strategies to behavioral strategies and then we will restrict our interest to this last class of strategies in order to solve the game.

2 The game.

2.1 description

A poor rabbit is moving along an infinite discrete line $\mathbb{Z}$. At each instant he can jump right or left. l is the maximal length of his motion. A bloodthirsty hunter shoots a bullet at every time, aiming at a particular position of the line. The very important fact is that the bullet reaches the line two or more time steps later (two in this paper). As a matter of fact, at each instant of the game one bullet, or more, is flying to a position of the line ; The hunter knows this position, while the rabbit lacks this information. Whatever happens, the game finishes at a fixed final time T. The hunter's wish is to maximize the probability to hit the rabbit while, for some reason, the rabbit wishes to minimize it.

We also studied a second version of the game in which the hunter wants to maximize the number of bullets reaching his opponent, while the rabbit wishes to minimize it. The resolution of this game brings in tangible alteration only in computation, so we will restrict our paper to the first version only.

2.2 Dynamics and notations.

In this paper we consider that the bullet takes two time steps to reache its aim. We denote :
• $y_t \in Z$, the rabbit's position at time t,
• $u_t \in U = \{-l, -l+1, \ldots, +l\}$, the rabbit's control at time t, that is the length of his jump at time t,
• $y^t = y_1.y_2 \ldots .y_t$ the sequence of the rabbit's successive positions until time t ; Y^t the set of such sequences. The “.” denotes the concatenation of elements.
• If the rabbit is at y at time t, it will be on one of the $4l+1$ possible positions : $y-2l, y-2l+1, \ldots, y+2l$ at time $t+2$. For every $y \in Z$, we denote $V(y) = \{y-2l, y-2l+1, \ldots, y+2l\}$; with this definition, if the rabbit is on y at t, $v_t \in V(y)$ is the hunter's control.
• Let y_a and y be two rabbit's successive positions, we define the set :
$V'(y_a, y) = \{y_a - 2l, \ldots, y-l-1, y+l+1, \ldots, y_a+2l\}$.

• At time t we denote z_t, the position the flying bullet is aiming at. $z_1 = \infty$ denotes the fact that there is no flying bullet at the beginning of the game.

At each instant of time the state is defined by the pair $x_t = (y^t, z_t)$, and the following equations describe the dynamics of the game :

$$\begin{cases} y^1 = y_1 \text{ given} \\ y^{t+1} = y^t.(y_t + u_t) = y^t.y_t + u \\ z_1 = \infty \\ z_{t+1} = v_t \ . \end{cases}$$

Let us dwell on the fact that the hunter alone knows perfectly the state at every time, since the rabbit does not know the part of the state z_t.

3 First approach - Normal form of the game.

A first approach to solve this game is to utilize a classical method of game theory, which is the *normal form of the game*. This method uses player's pure strategies that we describe below.

We can define a pure strategy as a sequence of maps, a map for each game instant. Each map associates a player's control to each of his possible observations.

When the rabbit chooses his control the only information he has is the sequence of his past positions. We can describe a rabbit's pure strategy, α, through the sequence of his successive positions from the initial time to the final time T. We call A the set of the rabbit's pure strategies.

At each time, the hunter's available and relevant information is the sequence of the rabbit's past positions, and the position the flying bullet is aiming at. Thus a hunter's pure strategy can be describe as follows : we note B the set of the hunter's pure strategies :

$$\beta \in B, \ \ \beta = (\beta_1, \dots, \beta_{T-2}),$$

$$\begin{array}{rlcl} \beta_t : & Y^t \times \mathbb{Z} & \rightarrow & \mathbb{Z} \\ & y^t, z & \rightarrow & \beta_t(y^t, z) = v \in V(y_t) \end{array}$$

Each pair of pure strategies, $(\alpha, \beta) \in A \times B$, generates a sequence of states $(x_1, \dots, x_T)$, or similarly two sequences $y^T = (y_1, \dots, y_T)$ and $z^T = (z_1, \dots, z_T)$, and to each pair of sequences y^T, z^T we can associate the outcome $\tilde{G}(y^T, z^T)$, such that :

$$\tilde{G}(y^T, z^T) = \begin{cases} 1 \text{ if } \exists t, \ \ y_{t+1} = z_t, \text{ that is if R is killed,} \\ 0 \text{ else.} \end{cases} \tag{1}$$

The payoff associated to the pair of pure strategies (α, β), is then defined by :

$$G(\alpha, \beta) = \tilde{G}(y^T, z^T),$$

where (y^T, z^T) is generated by (α, β).

Obviously there is no pair of optimal pure strategies, that is a pair of pure strategies that realizes a saddle point for G. As a matter of fact, if there was a rabbit's pure optimal strategy, the hunter could compute it and kill the rabbit for sure, contrarily to the optimality notion. That leads us to use *mixed strategies*, which are defined as probability laws respectively on A for the rabbit's mixed strategies and on B for the hunter's one. For a pair of mixed strategies, (p, q), the payoff is then defined by :

$$J(p, q) = \sum_{\alpha \in A, \ \beta \in B} p(\alpha) G(\alpha, \beta) q(\beta).$$

Hence we have to solve a saddle point problem in order to find optimal mixed strategies $p^* \in \pi(A)$, $q^* \in \pi(B)$ such that :

$$
\begin{aligned}
v^* &= \min_{p\in\pi(A)} \max_{q\in\pi(B)} J(p,q)\\
&= \max_{q\in\pi(B)} \min_{p\in\pi(A)} (p,q)\\
&= J(p^*,q^*)
\end{aligned}
$$

The game being finite, we know that the saddle point and consequently mixed optimal strategies exist. This last problem is classically solved using two dual linear programming problems.

Let's notice that this first method does not take in account the dynamical aspect of the game. The dynamics only arises in the strategy description. Each player chooses at random a pure strategy, then during the whole game he plays according to that pure strategy.

Theoretically the problem is well solved, and it would be so, if it were not for some computational difficulties. The number of pure strategies becomes too large as the final time T increases, bringing the size of the linear program beyond practical feasibility.

As seen before a pure rabbit's strategy is a sequence of T successive positions. If l is the maximal length of the rabbit's jump, there are $(2l+1)^{T-1}$ possible pure strategies. Though this number is large it does not prevent us from making computations. Let us compute the number of the hunter's pure strategies. For a given time and a given hunter's observation there are $4l+1$ possible hunter's controls, and since there are $(2l+1)^t$ possible observations at time t, and $T-2$ decision times, we come to the conclusion that the number of the hunter's pure strategies is :

$$
(4l+1)^{1+(2l+1)+(2l+1)^2+\ldots+(2l+1)^{T-3}} = (4l+1)^{((2l+1)^{T-2}-1)/2l},
$$

that is, for example, if $l=1$, 625 for only two shooting decisions and 5^{13} for three shooting decisions. That definitely makes clear that this method is not realistic. Nevertheless the resolution of some "little " games ($T \leq 4$) helped us to understand some important features of optimal solutions.

4 Behavioral strategies.

In view of this impossibility to compute optimal mixed strategies for games with realistic game space and final time, we were obliged to divert to another approach. This second approach is built up on dynamic programming and uses another class of strategies : behavioral strategies.

4.1 Definitions.

We call *rabbit's behavioral strategy* a sequence of maps $\varphi = (\varphi_1, \ldots, \varphi_{T-1})$, where φ_t associates a probability law on possible controls at time t, to each of his possible observations at time t. Consequently $\varphi_t(y^t)(u)$ is the rabbit's probability to jump to $y_t + u$ if the sequence of his past positions is $y^t = (y_1, \ldots, y_t)$. We note Φ the set of these behavioral strategies.

Similarly, we call *hunter's behavioral strategy* a sequence of maps $\psi = (\psi_1, \ldots, \psi_{T-2})$, where ψ_t associates a probability law on $\mathbb{Z}$ to each hunter's possible observation. Then $\psi_t(y^t, z_t)(v)$ is the hunter's probability to aim at v at time t, if the rabbit's past positions are y^t, and if a flying bullet is aiming at z_t. Let us note that $\psi_t(y^t, z_t)(v) = 0$ for each $v \notin V(y_t)$. We note Ψ the set of these strategies.

Any pair, (φ, ψ), of the rabbit and the hunter's behavioral strategies generates some stochastic variables y^T, z^T, and we can define the payoff of the game as the expectation :

$$
J(\varphi, \psi) = E^{\varphi,\psi}(\tilde{G}(y^T, z^T)),
$$

$\tilde{G}$ being defined by (1).

4.2 Notes.

When players opt for a behavioral strategy, they have a dynamical behavior, which is quite natural. Indeed, at each instant they choose their control at random according to a probability law provided by their behavioral strategy, and function of their own information.

On the other hand, let us notice that contrarily to mixed strategies, behavioral strategies do not care with past shots. Indeed only the last shots has influence on the future. All the past shots are included in mixed strategies.

4.3 Behavioral strategies, mixed strategies.

We are tempted to think that the set of behavioral strategies is richer than the set of mixed strategies. A simple example will prove us that this is not so. We will consider the case where $T = 4$, and in which the rabbit has only two available positions, that is $y_t \in \{1, 2\}$. Let us suppose that $y_1 = 1$.

Let us list all pure strategies : For each of the eight pure strategies this table gives us the aim of the shot for each time (1 or 2), and for each possible information.

time	1	2	
information	$y^1 = 1$	$y^2 = (1, 1)$	$y^2 = (1, 2)$
strategy : 1	1	1	1
2	1	1	2
3	1	2	1
4	1	2	2
5	2	1	1
6	2	1	2
7	2	2	1
8	2	2	2

The set of mixed strategies is thus the set of probability laws on a discrete eight element set. Its dimension is then seven.

Let us now look at behavioral strategies. They are defined by the five probabilities $q_1, q_2, \dots, q_5$ of aiming at 1, depending on time and information :

time : 1	$y^1 = (1)$	q_1
time : 2	$y^2 = (1, 1),\ z_2 = 1$	q_2
	$y^2 = (1, 1),\ z_2 = 2$	q_3
	$y^2 = (1, 2),\ z_2 = 1$	q_4
	$y^2 = (1, 2),\ z_2 = 2$	q_5

These five probabilities being free, the behavioral strategy set is then a five dimensional set.

We have shown that two different classes of strategies can be considered. Existence theorem is available only with mixed strategies, *i.e* in the richest class. We thus have to deal with the following problems :

– How to relate mixed and behavioral strategies ?
– Is it enough to consider behavioral strategies ?
– Does there exist an optimal solution, that is a saddle point for J, in this last class of strategies ?

4.4 Definitions.

We define the map *the rabbit's behavior*, γ_R, from the set $\pi(A)$ of mixed strategies to the set Φ of behavioral strategies in the following way :

$$\gamma_R : \pi(A) \longrightarrow \Phi : p \longrightarrow \varphi$$

such that :

$$\varphi_t(y^t)(u) = \frac{\sum_{\alpha\in A_2} p(\alpha)}{\sum_{\alpha\in A_1} p(\alpha)}.$$

where

$$\begin{array}{lcl} A_1 & = & \{\alpha \in A \mid \alpha = (y^t.y'_{t+1}.\dots.y'_T), \quad y'_{t+1},\dots,y'_T \text{ being whatever } \}, \\ \text{and} & & \\ A_2 & = & \{\alpha =\in A \mid \alpha = (y^t.y_t + u.y'_{t+2}.\dots.y'_T), \quad y'_{t+2},\dots,y'_T \text{ being whatever } \}, \end{array}$$

Similarly we define the map *the hunter's behavior*, γ_H, from the set $\pi(B)$ of mixed strategies to the set Ψ of behavioral strategies :

$$\gamma_H(q) : \pi(B) \longrightarrow \Psi : q \longmapsto \psi$$

such that

$$\psi_t(y^t)(v) = \frac{\sum_{\beta\in B_2} q(\beta)}{\sum_{\beta\in B_1} q(\beta)}.$$

where

$$\begin{array}{ll} B_1 = & \{\beta = (\beta_1 \dots, \beta_{T-2}) \in B \mid \beta_1(y_1,\infty) = v_1, \\ & \beta_i(y^i, v_{i-1}) = v_i \ , \ i = 1,\dots,t-2 \text{ et } \beta_{t-1}(y^{t-1}, v_{t-1}) = z \ \} \ , \\ \text{and} & \\ B_2 = & B_1 \bigcap \{\beta = (\beta_1,\dots,\beta_{T-2}) \in B \mid \beta_t(y^t, z) = v\} \ . \end{array}$$

Note that $\gamma_R(p)$ and $\gamma_H(q)$ for $p \in \pi(a)$ and $q \in \pi(B)$ are conditional marginal probabilities. The main property of behaviors is the following :

Theorem 1 :

The payoff $J(p,q)$, $p \in \pi(A)$, $q \in \pi(B)$, only depends on the behaviors $\gamma_R(p)$ and $\gamma_H(q)$.
Another way to say that :
There exists a mapping $\tilde{J} : \Phi \times \Psi \to I\!R$, such that,

$$\forall (p,q) \in \Pi(a) \times \Pi(B), \quad J(p,q) = \tilde{J}(\gamma_R(p), \gamma_H(q)).$$

This allows us to state the result :

Corollary 2 :

The hunter and rabbit game admits a saddle point in behavioral strategies.

Proofs of these two results are developed in a more general set up in [2]. It should be noticed that this theory extends to the case of continuously infinite sets of dicision variables. Our concept of mixed strategies is closer to that of Kuhn than Aumant's, and probably better suited to actually solve (short) games.

The aim of the remainder of this article is to find a pair of behavioral strategies which satisfies the saddle point condition.

5 Optimal behavioral strategies.

5.1 Players' costs.

We will use dynamical programming technics to define players' costs. Let us suppose for a while that the hunter knows the rabbit's strategy, say φ, and we note ψ his own strategy. Using the equation of Kolmogorov, the hunter's costs can be defined recursively from final time T to initial time. Let us note $V_t(y^t, z)$ the hunter's cost at $x_t = (y^t, z)$, that is the probability to touch the rabbit before time T if the actual state is (y^t, z).

$$\begin{aligned} & V_{T-1}(y^{T-1}, z) = \sum_{u \in U} \varphi_{T-1}(y^{T-1})(u) V_T(y^{T-1}.y_{T-1} + u, \infty, z) \\ & \text{and} \\ & V_t(y^t, z) = \sum_{u \in U} \sum_{v \in \mathbb{Z}} \psi_t(y^t, z)(v) \varphi_t(y^t)(u) V_{t+1}(y^t.y_t + u, v, z), \end{aligned} \tag{2}$$

with

$$V_t(y^t, z, z') = \begin{cases} 1 \text{ if } z' = y_t & \text{that is if R is killed,} \\ V_t(y^t, z) \text{ else ,} & \end{cases} \tag{3}$$

Similarly let us suppose the rabbit guessed the hunter's chosen strategy. The rabbit cannot use the entire state (y^t, z) for he does not know the actual value of z. (Here the problem of imperfect information arises). The rabbit builds a function, R^ψ,that provides him with a probability law on possible values of $z \in \mathbb{Z}$:

$$R^\psi_{t+1}(y^t.y_{t+1})(z) = \frac{\sum\limits_{j \neq y_{t+1}} \psi_t(y^t, j)(z) R^\psi_t(y^t)(j)}{\sum\limits_{j \neq y_{t+1}} R^\psi_t(y^t)(j)} \tag{4}$$

Note that $R^\psi_t(y^t)(z) = 0$ for $z \notin V(y_t)$. We took into account a rabbit's extra information : "cogito ergo sum ". The sum is thus computed only in the cases where the rabbit is alive ($j \neq y_{t+1}$) and a normalization term appears.

Now we can compute the rabbit's costs, that is the probabilities to hit the rabbit before time t, knowing only the part y^t of the state. Let us note $\overline{V}_t(y^t)$ these probabilities.

$$\overline{V}_t(y^t) = \sum_{z \in \mathbb{Z}} R^\psi_t(y^t)(z) V_t(y^t, z), \tag{5}$$

that is

$$\begin{aligned} & \overline{V}_{T-1}(y^{T-1}) = \sum_{u \in U} R^\psi_{T-1}(y^{T-1})(y_{T-1} + u) \varphi_{T-1}(y^{T-1})(u), \\ & \text{and} \\ & \overline{V}_t(y^t) = \sum_{z \in \mathbb{Z}} \sum_{u \in U} \sum_{v \in \mathbb{Z}} \psi_t(y^t, z)(v) \varphi_t(y^t)(u) V_{t+1}(y^t.y_t + u, v, z), \end{aligned} \tag{6}$$

Let us note that at the initial time we have :

$$\overline{V}_1(y^1) = V_1(y^1, \infty) = J(\varphi, \psi).$$

5.2 Theorem.

We can now state a theorem that characterizes optimal behavioral strategies. Its proof can be found, in a more general set up, in [2].

Theorem 4 :

Let (φ^*, ψ^*) be a pair of optimal behavioral strategies, that is a pair of strategies satisfying the saddle point condition.

Then

There exists a sequence of optimal cost functions $(\overline{\mathcal{V}}_t)_t$ and $(\mathcal{V}_t)_t$ such that for all (y^t, z) reached with a non zero probability while playing according to $(\varphi^*\ ,\psi^*\ )$ it is true that :

$$\mathcal{V}_{T-1}(y^{T-1}, z) = \sum_{u\in U} \varphi_{T-1}(y^{T-1})(u)\mathcal{V}_T(y^{T-1}.(y_{T-1}+u), \infty, z)$$

$$\begin{aligned}\mathcal{V}_t(y^t, z) &= \max_{v\in\mathbb{Z}} \sum_{u\in U} \varphi_t^*(y^t)(u)\mathcal{V}_{t+1}(y^t.(y_t+u), v, z) \\ &= \max_{q\in\Sigma_m} \sum_{v\in\mathbb{Z}}\sum_{u\in U} \varphi_t^*(y^t)(u)\mathcal{V}_{t+1}(y^t.(y_t+u), v, z)q(v) \\ &= \sum_{v\in\mathbb{Z}}\sum_{u\in U} \varphi_t^*(y^t)(u)\mathcal{V}_{t+1}(y^t.(y_t+u), v, z)\psi_t^*(y^t, z)(v)\end{aligned} \tag{7}$$

and

$$\begin{aligned}\overline{\mathcal{V}}_{T-1}(y^{T-1}) &= \min_{u\in U} R^{\psi^*}_{T-1}(y^{T-1})(y_{T-1}+u) \\ &= \sum_{u\in U} R^{\psi^*}_{T-1}(y^{T-1})(y_{T-1}+u)\psi^*_{T-1}(y^{T-1})(u)\end{aligned}$$

$$\begin{aligned}\overline{\mathcal{V}}_t(y^t) &= \min_{u\in U} \sum_{z\in\mathbb{Z}}\sum_{v\in\mathbb{Z}} R_t^{\psi^*}(y^t)(z)\mathcal{V}_{t+1}(y^t.(y_t+u), v, z)\psi_t^*(y^t, z)(v) \\ &= \min_{p\in\Sigma_n} \sum_{z\in\mathbb{Z}}\sum_{v\in\mathbb{Z}}\sum_{u\in U} R_t^{\psi^*}(y^t)(z)p(u)\mathcal{V}_{t+1}(y^t.(y_t+u), v, z)\psi_t^*(y^t, z)(v) \\ &= \sum_{z\in\mathbb{Z}}\sum_{v\in\mathbb{Z}}\sum_{u\in U} R_t^{\psi^*}(y^t)(z)\varphi_t^*(y^t)(u)\mathcal{V}_{t+1}(y^t.(y_t+u), v, z)\psi_t^*(y^t, z)(v) \\ &= \sum_{z\in\mathbb{Z}} R_t^{\psi^*}(y^t)(z)\mathcal{V}_t(y^t, z)\ ,\end{aligned} \tag{8}$$

with $\mathcal{V}_t(y^t, z, z')$ defined as $V_t(y^t, z, z')$ by (3), and $R_t^{\psi^*}$ by (4).

Furthermore we have :

$$v^* = J(\varphi^*, \psi^*) = \overline{\mathcal{V}}_1(y^1) = \mathcal{V}_1(y^1, \infty).$$

The main difficulty in solving these equations stems from a fixed point problem arisen in them. As a matter of fact the knowledge of R^{ψ^*} is necessary to compute the pair (φ^*, ψ^*) of optimal strategies, and conversely the knowledge of ψ^* is necessary to compute R^{ψ^*}. This precludes any attempt to solve them in a classical way. Some attemps in order to solve them with fixed point algorithms have been done, but up to now they failed.

6 Minimization of the opponent‘s information.

The remainder of this article is grounded on the intuitive idea of the following property on strategies :

“It is quite natural to hope that even if my opponent guesses the strategy I actually play, he cannot derive any advantage from it. In other words, additional information he can obtain knowing my strategy does not help him to choose his own.”

One should notice that, at this level of vagueness, this property is indeed enjoyed by the optimal mixed strategies of a game in normal form.

We have adapted this last idea to the game under study, on both the rabbit and the hunter.

For the rabbit it comes :

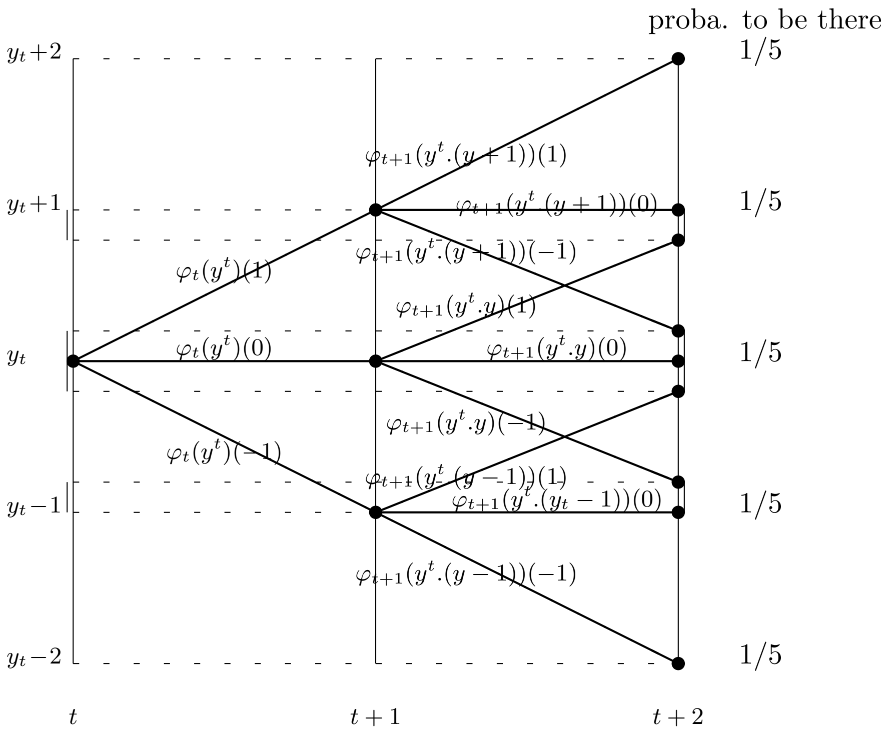
If the hunter guesses the rabbit's strategy, the relevant variables for him, at time t, if the rabbit's past positions are y^t, is the place the rabbit may be two time steps latter, that is at time $t+2$, when the bullet shot at time t reaches its aim. The rabbit will strive to make this further information irrelevant. A way to do that is to choose a strategy $\varphi = (\varphi_1, \ldots, \varphi_{T-1})$ such that for each t in $\{1, \ldots, T-2\}$ and for all $y = y_t$, for all $s \in \{-2l, \ldots, 2l\}$:

$$\text{Proba}^{\varphi_t, \varphi_{t+1}}(y_{t+2} = y + s) = \frac{1}{2l+1},$$

that can be expressed through the system :

$$\begin{cases} \varphi_t(y^t)(-l)\ \varphi_{t+1}(y^t.(y_t - l))(-l) = \frac{1}{2l+1} \\ \varphi_t(y^t)(-l)\ \varphi_{t+1}(y^t.(y_t - l))(-l+1) + \varphi_t(y^t)(-l+1)\ \varphi_{t+1}(y^t.(y_t - l + 1))(-l) = \frac{1}{2l+1} \\ \varphi_t(y^t)(-l)\ \varphi_{t+1}(y^t.(y_t - l))(-l+2) + \varphi_t(y^t)(-l+1)\ \varphi_{t+1}(y^t.(y_t - l + 1))(-l+1) \\ \qquad\qquad + \varphi_t(y^t)(-l+2)\ \varphi_{t+1}(y^t.(y_t - l + 2))(-l) = \frac{1}{2l+1} \\ \vdots \\ \varphi_t(y^t)(l)\ \varphi_{t+1}(y^t.(y_t + l))(l) = \frac{1}{2l+1}, \end{cases}$$

Example : $l = 1$



Hunter's point of view :

If the rabbit guesses the hunter's strategy, say ψ, the only things that can help him to choose his control at time t, in y^t are the probabilities $R^\psi(y^t)(y_t+u)$, $u \in U$. As a matter of fact,

$R^{\psi}(y^t)(y_t+u)$, are the probabilities that a bullet arrives at time $t+1$ on the rabbit's accessible places. The most unfavourable set up for the rabbit is when there exists a value α_t^{ψ} such that for all $u \in U$:

$$R_t^{\psi}(y^t)(y_t+u) = \alpha_t.$$

6.1 A sufficient condition.

We have exhibited a hunter's strategy that satisfies the last property. We then confronted it with a rabbit's strategy that also satisfies a minimized information condition. That allows us to state the following proposition :

Proposition 5.

hypothesis H_1 :

Let $\varphi = (\varphi_1, \ldots, \varphi_{T-1})$ be a rabbit's behavioral strategy that satisfies the following minimized information condition. For each y^t and for each $i \in \{-l, \ldots, l\}$, we define the row vector of length $m = 4l+1$:

$$L_t^i(y^{t-1}) = (\underbrace{0\ldots0}_{i+l}\underbrace{\varphi_t(y^{t-1}.y_{t-1}+i)(-l),\ldots,\varphi_t(y^{t-1}.y_{t-1}+i)(l)}_{2l+1}\underbrace{0\ldots0}_{l-i}),$$

and $N_t(y^{t-1})$ the $n \times m$ matrix made of the row $L_t^i(y^{t-1})$ ($n = 2l+1,\ \ m = 4l+1$) :

$$N_t(y^{t-1}) = \begin{pmatrix} L_t^{-l}(y^{t-1}) \\ L_t^{-l+1}(y^{t-1}) \\ \vdots \\ L_t^{l}(y^{t-1}) \end{pmatrix}$$

For each $t \in \{1, 2, \ldots, T\}$ φ satisfies :

$$\varphi_t(y^t)' N_{t+1}(y^t) = (\underbrace{\frac{1}{m},\ldots,\frac{1}{m}}_{m}).$$

hypothesis H_2 :

Let $\psi = (\psi_1, \ldots, \psi_{T-2})$ be a hunter behavioral strategy such that for each $t \neq 1$:

$$\begin{aligned}
\psi_t(y^t,z)' &= (\underbrace{\nu,\ldots,\nu}_{2l}, 1-4l\nu, \underbrace{\nu,\ldots,\nu}_{2l}) && \text{for each } z \in V'(y_{t-1},y_t)\\
\psi_t(y^t,y_t-l)' &= (\underbrace{0,\ldots,0}_{2l+1}, 1-(2l-1)\mu, \underbrace{\mu,\ldots,\mu}_{2l-1}) && \\
\psi_t(y^t,y_t-l+i)' &= (\underbrace{0,\ldots,0}_{i-1}, \tfrac{1}{2}, \underbrace{0,\ldots,0}_{2l+1}, \tfrac{1}{2}, \underbrace{0,\ldots,0}_{2l-1-i}) && \text{for each } i = 1,\ldots,2l-1\\
\psi_t(y^t,y_t+l)' &= (\underbrace{\mu,\ldots,\mu}_{2l-1}, 1-(2l-1)\mu, \underbrace{0,\ldots,0}_{2l+1})\ , &&
\end{aligned} \tag{9}$$

with

$$\mu = \frac{1}{4l} \ \text{ and } \ \nu = \frac{\beta_t - \frac{1}{m}}{4l\beta_t}, \quad \beta_t = \sum_{z \in V'(y_{t-1},y_t)} R_t^{\psi}(y^t)(z)\ \ ,$$

and for $t = 1$:

$$\psi_1'(y^1,\infty) = (\ \frac{1}{m}\ ,\ \ldots\ ,\ \frac{1}{m}\ ) \in \Sigma_m\ .$$

Conclusion C_1 :

Under H_2 we have :

ψ satisfies the following minized information condition, that is :
Let $(\alpha_t)_{t\in I\!N}$ be the real sequence :

$$\alpha_1 = 0, \quad \alpha_{t+1} = \frac{1}{m(1-\alpha_t)},$$

Whatever the sequence y^t and for each $u \in U$, we have :

$$R_t^\psi(y^t)(y_t + u) = \alpha_t.$$

We notice that these values depend only on time.

Conclusion C_2 :

Under hypothesis H_1 and H_2, the couple of rabbit and hunter's strategies (φ, ψ) when it exists, is optimal.

Conclusion C_3 :

Let the numerical sequences $(\rho_s)_s$ and $(\sigma_s)_s$ be :

$$\begin{cases} \rho_{s+1} &= \frac{1}{m}\sigma_s + \rho_s \\ \sigma_{s+1} &= 1 - \rho_s \\ \rho_1 &= 0 \\ \sigma_1 &= 1\ . \end{cases}$$

The values of rabbit and hunter's criterions are :

$$\mathcal{V}_t(y^t, z) = \begin{cases} \rho_{T-t} + \sigma_{T-t}\varphi_t(y^t)(z - y_t) & \text{if } z - y_t \in U \\ \\ \rho_{T-t} \quad \text{else,} \end{cases}$$

and

$$\overline{\mathcal{V}}_t(y^t) = \rho_{T-t+1} + \sigma_{T-t+1}\alpha_t\ .$$

The payoff is :

$$v^* = J(\varphi, \psi) = \rho_{T-1}\ .$$

6.2 Proof.

Conclusion C_1.

An easy calculation which uses (4), yields the conclusion.

Conclusion C_2 and C_3.

They are obtained from dynamical programming. As a matter of fact, the conclusion C_1 fixes the different values of functions R_t^ψ. The equations of optimality, (7) and (8), then reduce to dynamical equations. The fixed point notion vanishes. The proof then consists of checking that for each $t = T-1, \ldots, 2$, ψ and φ are respectively arguments of the minimum and the maximum of the different costs.

6.3 Restrictions.

It remains to keep a check on conditions which allow existence of strategies satisfying H_1 and H_2.

hypothesis H_1.

A sequence of maps has been exhibited. Nevertheless, the fact that it really is a strategy must be verified. For each sequence y^t, and each z, $\psi_t(y^t, z)$ must belong to the simplex Σ_m, that is to say that μ and ν must respectively stay in the intervals $[0, \frac{1}{2l-1}]$ and $[0, \frac{1}{4l}]$. The only difficulty is to verify that ν is positive, i.e. to prove that $\beta_t \geq \frac{1}{m}$, hence, $\alpha_t \leq \frac{4l}{mn}$, $\quad (n = 2l+1, \quad m = 4l+1)$.

In order to demonstrate this last inequality, we first notice that $(\alpha_t)_{t\in I\!N}$ is an increasing, recurrent sequence, that converges toward $\dfrac{m-\sqrt{m^2-4m}}{2m}$ therefore

$$\alpha_t \leq \frac{m-\sqrt{m^2-4m}}{2m} \; ,$$

and on the other hand :

$$\begin{aligned} & \frac{m-\sqrt{m^2-4m}}{2m} \leq \frac{4l}{nm} \\ \Leftrightarrow \quad & n^2(m^2-4m) \geq (8l-4ln-n)^2 \\ \Leftrightarrow \quad & 16l^3-14l^2-4l-1 \geq 0 \; . \end{aligned}$$

This last equality is true for $l > 1$, (one needs only to consider successive derivatives).In the case $l = 1$, a detailed (and cumbersome) analysis shows that there does not exist a hunter's strategy, satisfying assumption (H_2), for each time $t \leq T-2$, as soon as $T > 6$.
hypothesis H_2.

Attempt to find a general method to study the existence of strategies satisfying H_2 have failed. Nevertheless we have shown that in the case where $l = 1$ there is no such strategy as soon as $T \geq 5$.

6.4 Conclusions.

we can regroup these results in the following table :

Case $l = 1$:

$T < 4$	H_1 and H_2 are feasible.	we obtain optimal strategies.
$T = 4$	Only H_1 is feasible.	We obtain a hunter's strategy that satisfies a lower bound of the payoff.
$T \geq 5$	neither H_1 nor H_2 is feasible.	

Case $l > 1$:

$T \leq 4$	H_1 and H_2 are feasible.	we obtain optimal strategies.
$T \geq 4$	H_1 is feasible. H_2 ?	We obtain a hunter's strategy that satisfies a lower bound of the payoff.

7 A new formulation for the rabbit's costs - Hunter's equalizing strategies.

Although this last proposition does not provide a solution it credits the idea of minimized information and lets us foresee the importance of functions R_t^ψ. In this section we give another formulation for the rabbit's cost. This new formulation emphasises the fundamental role of the functions R_t^ψ. A consequence of this is the characterization of a class of the hunter's equalizing strategies, which is of importance.

7.1 A new formulation of the rabbit's cost.

Proposition 6 :

Let (φ, ψ) be a pair of the rabbit and the hunter's behavioral strategies, and let R_t^ψ be the functions associated to ψ and defined by (4),
then

The payoff $J(\varphi,\psi)=\overline{V}_1(y^1)$ is computed with the following recursive formulas :

$$\begin{cases} \overline{V}_{T-1}(y^{T-1}) = \sum_{u\in U} R^{\psi}_{T-1}(y^{T-1})(y_{T-1}+u)\varphi_{T-1}(y^{T-1})(u) \ , \\ \\ \overline{V}_t(y^t) = \sum_{u\in U} \varphi_t(y^t)(u)\{R^{\psi}_t(y^t)(y_t+u) + (1-R^{\psi}_t(y^t)(y_t+u))\overline{V}_{t+1}(y^t.(y_t{+}u))\} \ , \\ \text{and} \\ \overline{V}_1(y^1) = \sum_{u\in U} \varphi_1(y^1)(u)\overline{V}_2(y^1.(y_1{+}u)) \ . \end{cases} \tag{10}$$

Proof :

- Time $t=T-1$.

  It is the equatity (2) at time $T-1$.

- Time $1<t<T-1$.

  According to (2) we have :

  $$\overline{V}_t(y^t) = \sum_{z\in\mathbb{Z}}\sum_{u\in U}\sum_{v\in\mathbb{Z}} R^{\psi}_t(y^t)(z)\varphi_t(y^t)(u)V_{t+1}(y^t.(y_t{+}u),v,z)\psi_t(y^t,z)(v) \ ,$$

  then we use the definition (3) in order to write :

  $$\begin{aligned}\overline{V}_t(y^t) = \ & \sum_{u\in U}\varphi_t(y^t)(u)\{(\sum_{v\in\mathbb{Z}}\psi_t(y^t,y_t+u)(v)R^{\psi}_t(y^t)(y_t+u)) \\ & +\sum_{v\in\mathbb{Z}}\sum_{z\neq y_t+u} V_{t+1}(y^t.(y_t{+}u),v)\ R^{\psi}_t(y^t)(z)\ \psi_t(y^t,z)(v)\} \ ,\end{aligned}$$

  that is, recognizing the numerator of (4) in the second member of this equality :

  $$\begin{aligned}\overline{V}_t(y^t) = \ & \sum_{u\in U}\varphi_t(y^t)(u)\{R^{\psi}_t(y^t)(y_t+u) \\ & +(1-R^{\psi}_t(y^t)(y_t+u))\sum_{v\in\mathbb{Z}} V_{t+1}(y^t.(y_t{+}u),v)R^{\psi}_{t+1}(y^t.(y_t{+}u))(v)\} \ ,\end{aligned}$$

  and finally the formula claimed :

  $$\overline{V}_t(y^t) = \sum_{u\in U}\varphi_t(y^t)(u)\ \{R^{\psi}_t(y^t)(y_t+u) + (1-R^{\psi}_t(y^t)(y_t+u))\overline{V}_{t+1}(y^t.(y_t{+}u))\} \ .$$

7.2 A consequence : hunter's equalizing strategies.

Proposition 7 :

Let ψ be a hunter's behavioral strategy such that there exists a real sequence $(\alpha^{\psi}_t)_t$ satisfying :

$$\forall y^t,\ \ \forall u\in U, R^{\psi}_t(y^t)(y_t+u) = \alpha^{\psi}_t, \tag{11}$$

α_t only depends on time,
then
ψ is equalizing, that is, for each rabbit's behavioral strategy the payoff $J(\varphi,\psi)$ only depends on ψ :

$$\forall\varphi,\quad J(\varphi,\psi)=J^{\psi}.$$

We note E the set of such hunter's strategies.

Proof :

Let φ be any rabbit's strategy. Let us compute the costs $\overline{V}_t(y^t)$ if the hunter plays a strategy $\overline{\psi}$ that satisfies the equality (11).

• At time $T-1$, according to (10) :

$$\overline{V}_{T-1}(y^{T-1}) = \sum_{u\in U} R^{\overline{\psi}}_{T-1}(y^{T-1})(y_{T-1}+u)\varphi_{T-1}(y^{T-1})(u) ,$$

then using (11) at time $T-1$,

$$\overline{V}_{T-1}(yT-1) = \alpha^{\psi}_{T-1} .$$

this last value is independent of the sequence y^{T-1}, we note :

$$\overline{V}_{T-1}(y^{T-1}) = \overline{V}_{T-1} . \tag{12}$$

• At time t, (10) gives :

$$\overline{V}_t(y^t) = \sum_{u\in U} \varphi_t(y^t)(u)\{R^{\overline{\psi}}_t(y^t)(y_t+u) + (1-R^{\overline{\psi}}_t(y^t)(y_t+u))\overline{V}_{t+1}(y^t.(y_t+u))\} ,$$

then using (11) and (12) we can write :

$$\overline{V}_t(y^t) = \sum_{u\in U} \varphi_t(y^t)(u)\{\alpha^{\psi}_t + (1-\alpha^{\psi}_t)\overline{V}_{t+1}\} ,$$

or

$$\overline{V}_t(y^t) = \alpha^{\psi}_t + (1-\alpha^{\psi}_t)\overline{V}_{t+1} .$$

By induction, $\overline{V}_t(y^t)$ is independent of y^t, we note $\overline{V}_t(y^t) = \overline{V}_t$, and then we have :

$$\overline{V}_t = \alpha^{\psi}_t + (1-\alpha^{\psi}_t)\overline{V}_{t+1} .$$

• Time 1 :

$$\begin{aligned}\overline{V}_1(y^1) &= \sum_{u\in U} \varphi_1(y^1)(u)\overline{V}_2(y^1.(y_1+u)) \\ &= \overline{V}_2 \sum_{u\in U} \varphi_1(y^1)(u) \\ &= \overline{V}_2 .\end{aligned}$$

◇◇

Note that the following property :

$$\exists(\alpha_t(y^t)),\ \ \forall u\in U,\ \ R^{\psi}_t(y^t)(y_t+u) = \alpha_t(y^t),$$

is not enough to be sure that ψ is equalizing, $R^{\psi}_t(y^t)(y_t+u)$ must be independent of both u and y^t.

The main interest of this class of strategies lies in the fact that it provides some strategies that ensure the hunter of a lower bound of the payoff against any rabbit's counter stoke. Note that the strategies which satisfy the last proposition are equalizing. In the cases where the assumption (H_2) is not feasible, we have nevertheless lower bounds for the payoff.

8 A conjecture.

We have just shown the importance of the functions R_t^ψ. In this last section we apply the minimized information notion on the hunter's strategies. We first state a conjecture and give further some justifications. This conjecture allows us to build an algorithm that computes easily the hunter's optimal strategies.

8.1 Statement.

There exists a hunter's optimal behavioral strategy, ψ^* that satisfies the following property :

$$\exists(\alpha_t^{\psi^*})_{t\in N} \quad | \quad \forall y^t,\ \forall u \in U,\ \ R_t^{\psi^*}(y^t)(y_t+u) = \alpha_t^{\psi^*}.$$

8.2 Justifications.

A first justification lies in the fact that the interpretation of the conjecture in terms of "minimized information" is quite natural. Another justification comes from the resolution of short games using the normal form. Indeed, all the examples of solutions we have computed have the right property. Let us note that for a special game, two different optimal strategies provide the same sequence $(\alpha_t^\psi)_t$. A last justification consists of the proposition 5. If the assumptions are feasible then the hunter's optimal strategies satisfy the property (11) and are element of E.

8.3 Consequences : computation of the hunter's optimal strategies.

Corollary 8 :

The hunter's optimal strategies lying in E satisfy the following property :

$$\exists(\alpha_{\max,t})_t \quad \forall y^t,\ \forall u \in U, \quad R_t^{\psi^*}(y^t)(y_t+u) = \alpha_{\max,t}.$$

and

$$\forall t, \quad \alpha_{\max,t} = \max_{\psi\in E} \alpha_t^\psi.$$

Corollary 8 :
Proof :

Let ψ^* be an optimal strategy lying in E, accordingly to the proposition 7 it is equalizing, and the rabbit's cost only depends on time :

$$\begin{aligned}&\overline{\mathcal{V}}_{T-1} = \alpha_{T-1}^{\psi^*}\\&\overline{\mathcal{V}}_t = \alpha_t^{\psi^*} + (1-\alpha_t^{\psi^*})\overline{\mathcal{V}}_{t+1}\\&\overline{\mathcal{V}}_1 = \overline{\mathcal{V}}_2\end{aligned}$$

ψ^* must maximize $J(\varphi^*,\psi) = \overline{\mathcal{V}}_1$ in order to be optimal, and that is to maximize $\alpha_t^{\psi^*}$ at each time step. (The function $x :\rightarrow x + (1-x)a$ is increasing as soon as $a \le 1$).

Finally we have to find a hunter's strategy, ψ^*, lying in E satisfying :

$$(\alpha_t^{\psi^*})_t = (\alpha_{\max,t})_t,$$

where

$$\alpha_{\max,t} = \max_{\psi\in E} \alpha_t^\psi.$$

◇◇

Let us state a last property before presenting the final algorithm.

Proposition 9 :

Let ψ be a hunter's strategy lying in E, and $(\alpha_t^\psi)_t$ be the sequence such that :

$$\alpha_t^\psi = R_t^\psi(y^t)(y_t+u), \forall y^t, \forall u \in U. \tag{13}$$

Then

we can find a strategy, ψ' such that:

$$(\alpha_t^{\psi'})_t = (\alpha_t^\psi)_t,$$

and

$$\begin{array}{l} \forall u \in U : \\ \psi'_t(y^t, y_t+u) = \zeta_t(u) \ , \quad \zeta_t(u) \in \Sigma_m \ , \\ \text{and} \\ \forall z \in V'(y_{t-1}, y_t) \ : \\ \psi'_t(y^t, z) = \eta_t \ , \quad \eta_t \in \Sigma_m \ . \end{array}$$

Proof :

At time 1, for each strategy lying in E we have, accordingly to (4) :

$$R_2(y^2)(y_2+u) = \psi_1(y^1, \infty)(y_2+u) = \alpha_2 \ ,$$

for each $y_2 \in y_1 + U$ and for each $u \in U$ then :

$$\psi_1(y^1, \infty)(v) = \frac{1}{m} \ , \quad \forall v \in V(y_1) \ .$$

and

$$\alpha_2^\psi = \alpha_{\max,t} = \frac{1}{m} \ ,$$

then

$$\psi_1(y^1) = (\frac{1}{m}, \dots, \frac{1}{m}) \ ,$$

and then each strategy of E satisfies the proposition at time 1.

Let us suppose the proposition true up to time $t-1$, Let us prove it at time t. According to (13),

$$\begin{aligned} \alpha_t^\psi &= R_t(y^t, z)(y_t+u) \\ &= \frac{\displaystyle\sum_{j \in \mathbb{Z}), j \neq y_t} R_{t-1}^\psi(y^{t-1})(j)\psi_{t-1}(y^{t-1}, j)(y_t+u)}{\displaystyle\sum_{j \in \mathbb{Z}), j \neq y_t} R_{t-1}^\psi(y^{t-1})(j)} \\ &= \frac{\displaystyle\sum_{j \in \mathbb{Z}), j \neq y_t} R_{t-1}^\psi(y^{t-1})(j)\psi_{t-1}(y^{t-1}, j)(y_t+u)}{1 - R_{t-1}^\psi(y^{t-1})(y_t)} \end{aligned}$$

and ψ lying in E :

$$\alpha_t^\psi = \frac{(\alpha_{t-1}^\psi \displaystyle\sum_{j \in y_{t-1}+U, j \neq y_t} \psi_{t-1}(y^{t-1}, j)(y_t+u)) + (\displaystyle\sum_{j \in V'(y_{t-2}, y_{t-1})} R_{t-1}^\psi(y^{t-1})(j)\psi_{t-1}(y^{t-1}, j)(y_t+u) \ )}{1 - \alpha_{t-1}^\psi} \ .$$

Let us note :

$$a = \sum_{j \in y_{t-1}+U, j \neq y_t} \psi_{t-1}(y^{t-1}, j)(y_t + u)$$

and $\zeta(y_t + u)$ such that :

$$\sum_{j \in V'(y_{t-2}, y_{t-1})} R^{\psi}_{t-1}(y^{t-1})(j)\zeta(y_t + u) = \sum_{j \in V'(y_{t-2}, y_{t-1})} R^{\psi}_{t-1}(y^{t-1})(j)\psi_{t-1}(y^{t-1}, j)(y_t + u) \ ,$$

that is to say, using the fact that :

$$\sum_{j \in V'(y_{t-2}, y_{t-1})} R^{\psi}_{t-1}(y^{t-1})(j) = 1 - n\alpha^{\psi}_{t-1} \ ,$$

$$(1 - n\alpha^{\psi}_{t-1})\zeta(y_t + u) = \sum_{j \in V'(y_{t-2}, y_{t-1})} R^{\psi}_{t-1}(y^{t-1})(j)\psi_{t-1}(y^{t-1}, j)(y_t + u) \ . \tag{14}$$

According to a classical property, it is obvious that $\zeta(y_t + u)$ lies in $[0, 1]$. Thus we have :

$$\alpha^{\psi}_t = \frac{\alpha^{\psi}_{t-1} a + (1 - n\alpha^{\psi}_{t-1})\zeta(y_t + u)}{1 - \alpha^{\psi}_{t-1}} \ .$$

Making the sum of the equality (14) on each $y_t + u$, or similarly on each $z \in V(y_{t-1})$, we obtain :

$$\sum_{z \in \mathbb{Z}} (1 - n\alpha^{\psi}_{t-1})\zeta(z) = \sum_{z \in \mathbb{Z}} \sum_{j \in V'(y_{t-2}, y_{t-1})} R^{\psi}_{t-1}(y^{t-1})(j)\psi_{t-1}(y^{t-1}, j)(z)$$

$$(1 - n\alpha^{\psi}_{t-1}) \sum_{z \in \mathbb{Z}})\zeta(z) = \sum_{j \in V'(y_{t-2}, y_{t-1})} R^{\psi}_{t-1}(y^{t-1})(j) \sum_{z \in \mathbb{Z}} \psi_{t-1}(y^{t-1}, j)(z) \ ,$$

$$(1 - n\alpha^{\psi}_{t-1}) \sum_{z \in \mathbb{Z}} \zeta(z) = 1 - n\alpha^{\psi}_{t-1}$$

and then

$$\sum_{z \in \mathbb{Z})} \zeta(z) = 1 \ ,$$

we note

$$\begin{array}{l} \psi'_t(y^t, y_t + u)(v) = \psi_t(y^t, y_t + u)(v) \ , \\ \text{and} \\ \psi'_t(y^t, z)(y_t + u) = \zeta(y_t + u) \ , \quad \forall z \in V'(y_{t-1}, y_{t-2}) \ , \end{array}$$

and finally we have $\alpha^{\psi'}_{t+1} = \alpha^{\psi}_{t+1}$.

◇◇

Corollary :
There exists a hunter's strategy lying in E that satisfies :

$$\begin{cases} (\alpha^{\psi}_t)_t = (\alpha_{\max,t})_t \ , \\ \psi_t(y^t, y_t + u) = \eta_t(u) \ , \quad \eta_t(u) \in \Sigma_m \\ \psi_t(y^t, z) = \zeta_t \ , \quad \zeta_t \in \Sigma_m \ , \quad \text{for each } z \in V'(y_{t-1}, y_t) \end{cases} \tag{15}$$

This last statement allows us to present an algorithm that computes the hunter's optimal strategies in E, which is overall optimal under our conjecture.

Algorithm :

• At time 1 :

$$\psi_1(y^1, \infty) = (\frac{1}{m}, \dots, \frac{1}{m}) \ .$$

• At time t :

Let us note :

$$\psi_t(y^t, y_t + u)(y_t + \bar{v}) = \zeta_t(u)(\bar{v})$$
et
$$\psi_t(y^t, z)(y_t + \bar{v}) = \eta_t(\bar{v}) \ .$$

We must solve the following linear programming problem in order to find $\eta_t(\bar{v})$ and $\zeta_t(u)(\bar{v})$:

$$\begin{cases} \max\limits_{\zeta,\eta} \alpha^{\psi}_{t+1} \\ \\ \text{under} \\ \dfrac{\alpha^{\psi}_t \sum\limits_{u \in U} \zeta_t(u)(\bar{v}) + (1 - n\alpha^{\psi}_t)\eta_t(\bar{v})}{1 - \alpha^{\psi}_t} = \alpha^{\psi}_{t+1} \ , \\ \\ \zeta_t(u) \in \Sigma_m \ , \quad \eta_t \in \Sigma_m \ . \end{cases}$$

References

[1] T.Başar, "On a class of zero-sum discrete games with delayed information", Lecture notes in control and information sciences 119. pp9–34, Springer, Berlin, 1989.

[2] P.Bernhard, "Information and strategies in dynamic games", SIAM Journal in Control and Optimization, 1991, to appear.

[3] O.Pourtallier, "A two player dynamical game with imperfect information", Lecture notes in control and information sciences 144. pp745–754, Springer-Verlag, 1990.

[4] O.Pourtallier, "Etude d'un jeu dynamique en information incomplète : le jeu du chasseur et du lapin", Thèse, université de Nice Sophia Antipolis, 1991.