Intelligent escalation and the principle of relativity

Pierre Lescanne
University of Lyon, École normale supérieure de Lyon, CNRS (LIP),
46 allée d'Italie, 69364 Lyon, France

May 24, 2015

Abstract

Escalation is the fact that in a game (for instance in an auction), the agents play forever. The $0,1$-game is an extremely simple infinite game with intelligent agents in which escalation arises. At the light of research on cognitive psychology, it shows the difference between intelligence (algorithmic mind) and rationality (algorithmic and reflective mind) in decision processes. It also shows that depending on the point of view (inside or outside) the perception of the rationality of the agent change, which we call the *principle of relativity*.

Keywords: economic game, infinite game, sequential game, extensive game, escalation, speculative bubble, coinduction, auction.

To Bernard Maris
23 September 1946 – 7 January 2015

That "rational agents" should not *engage in such [escalation] behavior seems obvious.*

Wolfgang Leininger [16]

I can calculate the movement of the stars,
but not the madness of men.

Newton[1] in (1720)

Sequential games are the natural framework for decision processes. In this paper we study a decision phenomenon called *escalation*. Infinite sequential games presented here generalize naturally sequential games with perfect information and have been introduced by Lescanne [18] and Lescanne and Perrinel [22] and formalize what is proposed in the literature [27]. Sequential games are games in which each player plays one after the other (or possibly after herself). Here we prove, using coinduction on a simple example, namely *the 0,1 game*, that escalation is not irrational. More precisely, in escalation, agents exhibit only the low part of the rational mind, namely the *algorithmic mind*, assimilated to *intelligence* or *fluid intelligence* to be more specific (see Stanovich [35]). Since intelligent behavior contrasts with the somewhat obvious observation that escalation is irrational, we state a *principle of relativity* which says that *"the view of the insider (the agent) is not the same as the view of the outsider (the observer)"*, a well known fact in computer science when studying distributive system [13]. In addition, the 0,1 game has nice properties which make it an excellent paradigm of escalation, a good domain of application of coalgebras and coinduction and a very nice opportunity to present coinduction, a tool designed to prove properties of complex systems.

[1] Actually probably apocryphal Newton's view on the outcome of the *South Sea Bubble*.

1 The problem of escalation

Escalation in sequential games is a classic of game theory and it is admitted that escalation is irrational. Consider agents able to reason formally that is making choices which are optimal and robust, in other words choosing equilibria. In finite sequential games, a right choice is obtained by a specific equilibrium called *backward induction* (see Appendix). More precisely a consequence of Aumann's theorem [3] says that an agent takes a good decision in a finite sequential game if she makes her choice according to backward induction. In this paper we generalize backward induction into *subgame perfect equilibria* and we explore the kind of reasoning obtained when built on subgame perfect equilibria (SPE in short). Such an appropriate reasoning is said to be based on *an algorithmic mind* (see Section 7).

What is escalation? In a sequential game, escalation is the possibility that agents take adequate decisions forever without stopping. This phenomenon has been evidenced by Shubik [34] in a game called the *dollar auction.* Without being very difficult, the analysis of the dollar auction is relatively involved, because it requires infinitely many infinite strategy profiles indexed by $n \in \mathbb{N}$ [22]. At each step there are two and only two equilibria and therefore two potential intelligent decisions for the agents, namely "stop" or "continue". If both agents choose always "continue", an escalation occurs. In this paper, we propose an example which is simpler theoretically and which offers infinitely many infinite equilibria at each step unlike the dollar auction. Due to the form of the equilibria, the agent has no clue on which strategy is taken by her opponent.

What is coinduction? Wikipedia gives the following definition of coinduction: *"In computer science, coinduction is a technique for defining and proving properties of systems of concurrent interacting objects."* Actually coinduction is more than a technique, since it is a formal approach to correct reasoning on infinite structures and to complex systems which applies far beyond computer science, especially to economics. First traces of coinduction can be found at the beginning of the XXth century when researchers tried to set a mathematical foundation to not well-founded sets [23]. It was revisited by Peter Aczel [2] creating the foundation of coinduction. Davide Sangiorgi [31] gives a historical account of the field. Notice that Lawrence Moss and Ignacio Viglizzo have already apply coalgebras to the theory of normal form game [24].

Escalation and infinite games. Books and articles [9, 11, 26, 16, 25] cover escalation. Following Shubik, all agree that escalation takes place and can only take place in an infinite game, but their argument uses reasoning on finite games. Indeed, if we cut at a finite position the infinite game of the dollar auction in which escalation is supposed to take place, we get a finite game, in which the only right decision is to never start the game, because the only backward induction equilibrium corresponds to not start playing. Then the result is extrapolated by the authors to infinite games by making the size of the game to grow indefinitely. However, it has been known for a long time at least since Weierstraß [40], that the "cut and extrapolate" method is wrong (see Appendix), or said otherwise, there is no continuity at infinite. For Weierstraß this would lead to the conclusion that the infinite sum of differentiable functions would be differentiable whereas he has exhibited a famous counterexample. In the case of infinite structures like infinite games, the right reasoning is coinduction. With coinduction we were able to show that in the dollar auction escalation can be the result of a formal reasoning [22, 20]. Currently, since the tools used generally in economics are pre-coinduction based, they conclude that bubbles and crises are impossible and everybody's experience has witnessed the opposite. Careful analysis done by quantitative economists, like for instance Bouchaud [5, 6], have shown that bursts (aka volatility), which share much similarities with escalation, actually take place at any time scale. Escalation is therefore an intrinsic feature of economics. Consequently, coinduction is probably the tool that economists who call for a rethinking economics [8, 5, 37] are waiting for [41].

Structure of the paper This paper is structured as follows. In Section 2 we present infinite games, infinite strategy profiles and infinite strategies, then we describe the 0,1-game in Section 3. Last, we introduce the concept of equilibrium (Sections 4 and 5) and we discuss escalation (Section 6). In Section 7 we discuss the actual rationality of the agents from a cognitive science point of view. In an appendix, we talk about finite $0,1$-games and finite strategy profiles. Bart Jacobs [13], Jan Rutten [30] and Davide Sangiorgi [32] propose didactic introductions to coalgebras and coinduction. This paper is a deeply modified version of [21].

2 Two choice sequential games

Our aim is not to present a general theory of coalgebras or a theoretical foundation of infinite extensive games. For this the reader is invited to look at [1, 20, 22]. But we want to give a taste of infinite sequential games[2] through a very simple one. This game has two agents and two choices. To support our claim about the intelligence of escalation, we do not need more features.

Assume that the set P of agents is made of two agents called A and B. In this framework, an 'infinite sequential two choice game' has two shapes. First, it can be an ending position in which case it boils down to the distribution of the payoffs to the agents. In other words, an ending game is reduced to a function $f : A \mapsto f_{\mathsf{A}}, B \mapsto f_{\mathsf{B}}$ and we write it $\langle f \rangle$. Second , it can be a generic game with a set Choice made of two potential choices: d or r (d for *down* and r for *right)*. Since the game is potentially infinite, it may continue forever. Thus formally in this most general configuration a game can be seen as a triple:

$$g = \langle p, g_d, g_r \rangle.$$

where p is an agent and g_d and g_r are themselves games. The subgame g_d is for the down choice, i.e., the choice corresponding to go *down* and the subgame g_r is for the *right* choice, i.e., the choice corresponding to go to the right. Therefore, we define a functor (see [30] page 4 and following):

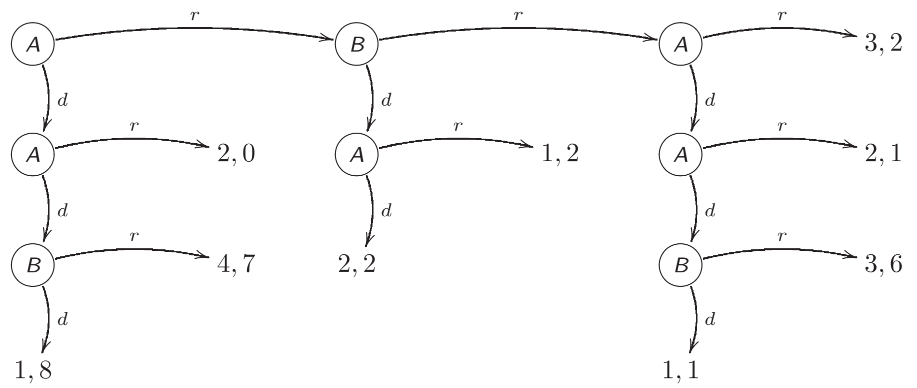
$$\langle\ \rangle : \mathsf{X} \quad \to \quad \mathbb{R}^{\mathsf{P}} \quad + \quad \mathsf{P} \times \mathsf{X} \times \mathsf{X}.$$

introducing a category of coalgebras of which Game is the final coalgebra and where $\mathsf{P} = \{\mathsf{A}, \mathsf{B}\}$. In other words, Game satisfies the isomorphism

$$\mathsf{Game} \quad \simeq \quad \mathbb{R}^{\mathsf{P}} \quad + \quad \mathsf{P} \times \mathsf{Game} \times \mathsf{Game}.$$

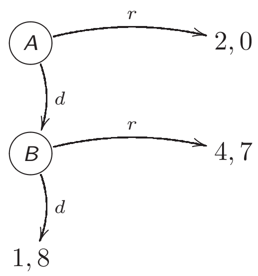
Example 1 *Here is a picture of a typical finite sequential game:*

[2] If the reader feels that this approach is not formal enough, she (he) can look at the "ultra"-formal approach found in the COQ scripts mentioned in Section 5.



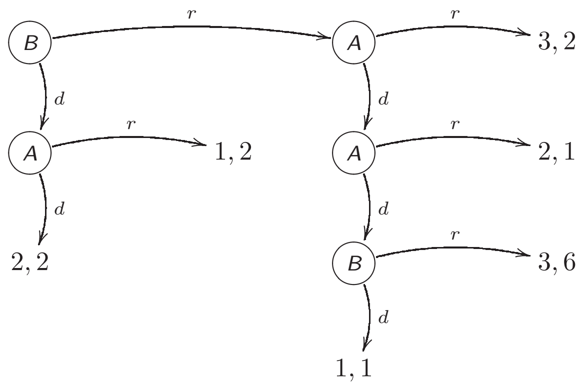
Let us call it gg. In this picture $3, 2$ *represents the game* $\langle \mathsf{A} \mapsto 3, \mathsf{B} \mapsto 2\rangle$ *and* $gg = \langle \mathsf{A}, gg_1, gg_2\rangle$ *where*

$gg_1 \quad =$



and

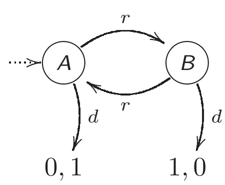
$gg_2 \quad =$



If we write the full decomposition of gg_1 *we get:*

$$gg_1 = \langle A, \langle B, \langle \mathsf{A} \mapsto 1, B \mapsto 8\rangle, \langle \mathsf{A} \mapsto 4, B \mapsto 7\rangle\rangle, \langle A \mapsto 2, B \mapsto 0\rangle\rangle.$$

Example 2

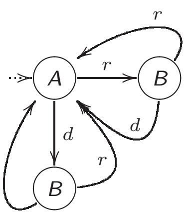


is a picture of an infinite game which will be studied more formally in Section 3. Notice the dotted arrow $\cdots\!\!>$ *which shows the start of the game.*

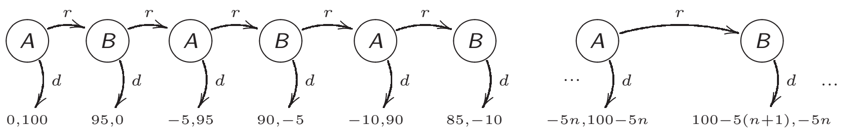
Example 3 *The game which is solution of the equation*

$$g \quad = \quad \langle \mathsf{A}, \langle \mathsf{B}, g, g\rangle, \langle \mathsf{B}, g, g\rangle\rangle$$

is a game infinite in both direction, down *and* right*, alternating agents* A *and* B*. Notice that it has no leaf and that payoffs are not attributed. It can be pictured as:*



Example 4



is the dollar auction *game [34] with bids of* 5¢.

2.1 Strategy profiles

From a game, we can deduce *strategy profiles* (later we will also say sometimes simply *profiles*) obtained by adding a label, at each node, which is a choice made by the agent. In a two choice sequential game, choices belong to the set $\{d, r\}$. Therefore a strategy profile which does not correspond to an ending game is a quadruple:

$$s = \langle\!\langle p, c, s_d, s_r \rangle\!\rangle,$$

where p is an agent (A or B), c is a choice (d or r), and, s_d and s_r are two strategy profiles. The strategy profile which corresponds to an ending position has no choice, namely it is reduced to a function $\langle\!\langle f \rangle\!\rangle = \langle\!\langle A \mapsto f_{\mathsf{A}}, B \mapsto f_{\mathsf{B}} \rangle\!\rangle$. The functor

$$\langle\!\langle\ \rangle\!\rangle : \mathsf{X} \quad \to \quad \mathbb{R}^{\mathsf{P}} \quad + \quad \mathsf{P} \times \mathsf{Choice} \times \mathsf{X} \times \mathsf{X}.$$

where

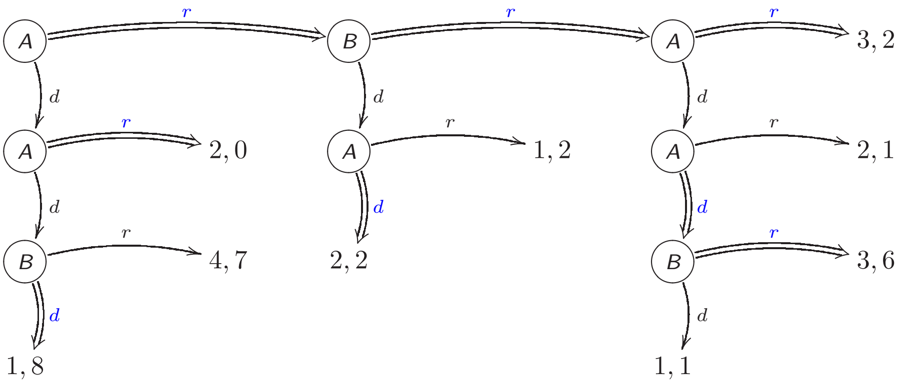
$$\begin{aligned} \mathsf{P} &= \{\mathsf{A}, \mathsf{B}\} \\ \mathsf{Choice} &= \{r, d\} \end{aligned}$$

introduces a category of coalgebras in which the coalgebra StratProf of infinite strategy profiles is the final coalgebra. Hence StratProf satisfies the isomorphism:

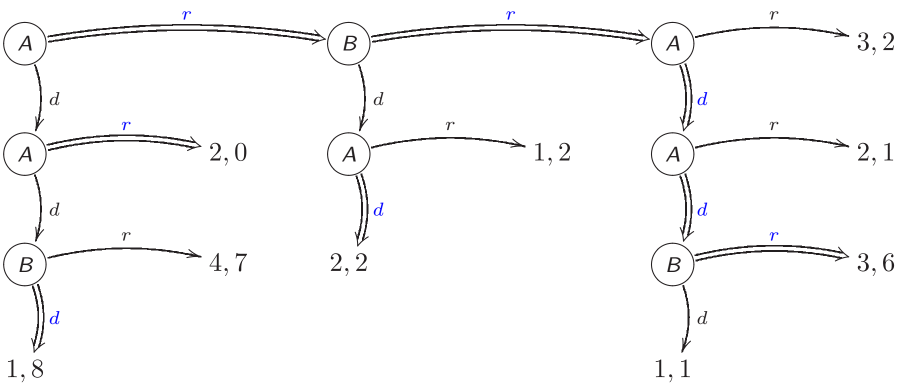
$$\mathsf{StratProf} \quad \simeq \quad \mathbb{R}^{\mathsf{P}} \quad + \quad \mathsf{P} \times \mathsf{Choice} \times \mathsf{StratProf} \times \mathsf{StratProf}.$$

Example 5 *Here are the pictures of three strategy profiles associated with the game* *of Figure 1.*

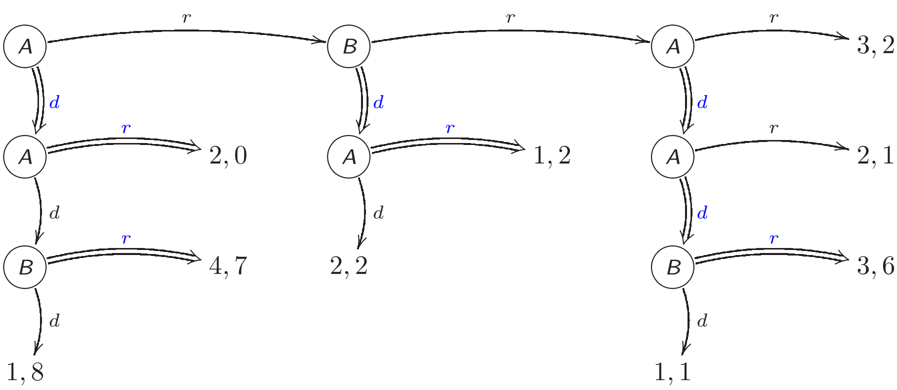
$s_1 \quad =$



$s_2 \quad =$



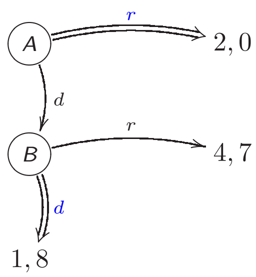
$s_3 \quad =$



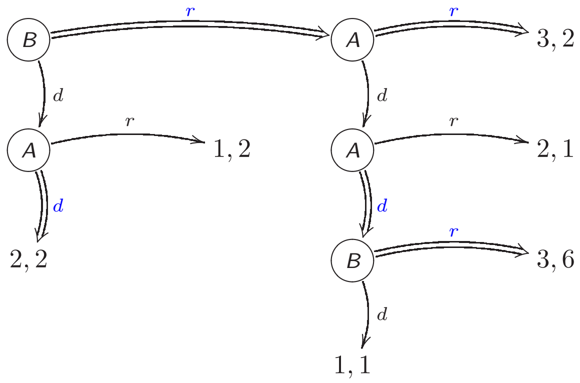
The start of those strategy profiles is taken by player A and she chooses right *in the two first strategy profiles and* down *in the third strategy profile.*

s_1 *is built with the strategy profiles:*

$s_{11} =$



and

$s_{12} =$ 

with

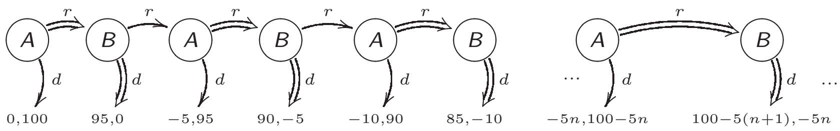
$$s_1 = \langle\!\langle \mathsf{A}, r, s_{11}, s_{12} \rangle\!\rangle.$$

The full decomposition of s_{11} *is*

$$s_{11} = \langle\!\langle \mathsf{A}, r, \langle\!\langle \mathsf{B}, d, \langle\!\langle \mathsf{A} \mapsto 1, \mathsf{B} \mapsto 8 \rangle\!\rangle, \langle\!\langle \mathsf{A} \mapsto 4, \mathsf{B} \mapsto 7 \rangle\!\rangle \rangle\!\rangle, \langle\!\langle \mathsf{A} \mapsto 2, \mathsf{B} \mapsto 0 \rangle\!\rangle \rangle\!\rangle.$$

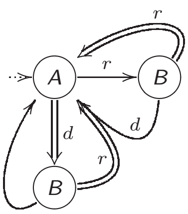
Example 6 *Figure 1 on page 12, Figure 2 on page 12, and Figure 3 on page 14 give strategy profiles of infinite sequential games.*

Example 7



is a strategy profile which is of interest in the dollar auction game [22] for proving that agent reasoning formally can enter escalation.

Example 8



is a strategy profile of the game of Example 3 and is solution of the equation

$$s \quad = \quad \langle\!\langle \mathsf{A}, d, \langle\!\langle \mathsf{B}, r, s, s \rangle\!\rangle, \langle\!\langle \mathsf{B}, r, s, s \rangle\!\rangle \rangle\!\rangle$$

From a strategy profile, we can build a game by removing the choices:

$$\begin{array}{rcl} \mathsf{game} & :: & \mathsf{StratProf} \to \mathsf{Game} \\ \mathsf{game}(\langle\!\langle f \rangle\!\rangle) & = & \langle f \rangle \\ \mathsf{game}(\langle\!\langle p, c, s_d, s_r \rangle\!\rangle) & = & \langle p, \mathsf{game}(s_d), \mathsf{game}(s_r) \rangle \end{array}$$

$\mathsf{game}(s)$ is the game of the strategy profile s. Notice that the function game is not recursive like say the function fib

$$\begin{array}{rcl} \mathsf{fib}(0) & = & 0 \\ \mathsf{fib}(1) & = & 1 \\ \mathsf{fib}(n+2) & = & \mathsf{fib}(n+1) + \mathsf{fib}(n) \end{array}$$

which defines the Fibonacci sequence. In fact, game is corecursive since it works on potentially infinite structures (see the end of this section) whereas fib which works on natural numbers works on finite structure and is recursive. Notice however that if we consider only finite games and finite strategy profiles, then a recursive game function could be defined. In this paper we consider the final coalgebras Game which contains finite and infinite sequential games and the final coalgebra $\mathsf{Stratprof}$ which contains finite and infinite strategy profiles.

Given a strategy profile s, we can associate, by induction, a (partial) payoff function $\widehat{s}$, as follows:

$$\begin{array}{llrcl} \textit{when} & s = \langle\!\langle f \rangle\!\rangle & \widehat{s} & = & f \\ \textit{when} & s = \langle\!\langle p, d, s_d, s_r \rangle\!\rangle & \widehat{s} & = & \widehat{s_d} \\ \textit{when} & s = \langle\!\langle p, r, s_d, s_r \rangle\!\rangle & \widehat{s} & = & \widehat{s_r} \end{array}$$

In the literature on extensive games ([3, 12] for instance), authors have a graph vision and they write $h_p^v(s)$ the payoff obtained by p starting at vertex v. Here we consider only the start vertex (let us call it *start*). For us, the vertices are not primitive objects. The other vertices are start vertices of subgames and of subprofiles and are only considered when dealing with those subgames and those subprofiles. What we write $\widehat{s}(p)$ would be written $h_p^{start}(s)$ in [3, 12].

Example 9 *In Example 5, we have:*

$$\begin{array}{rcl@{\qquad\qquad}rcl@{\qquad\qquad}rcl} \widehat{s_1}(A) & = & 3 & \widehat{s_2}(A) & = & 3 & \widehat{s_3}(A) & = & 2 \\ \widehat{s_1}(B) & = & 2 & \widehat{s_2}(B) & = & 6 & \widehat{s_3}(B) & = & 0 \end{array}$$

$\widehat{s}$ is not defined if its definition runs in an infinite process. For instance, in Example 16, $\widehat{s_{\Box r}}$ is not defined and in Section 6, $\widehat{s_{\mathsf{A},\infty}}$ is not defined. To ensure that we consider only strategy profiles where the payoff function is defined we can impose strategy profiles to be *convergent*[3], written $s \downarrow$ (or prefixed $\downarrow (s)$) and defined as the least predicate satisfying

$$s \downarrow \quad \overset{(i)}{\Longleftrightarrow} \quad s = \langle\!\langle f \rangle\!\rangle \quad \vee \quad (s = \langle\!\langle p, d, s_d, s_r \rangle\!\rangle \wedge s_d \downarrow) \vee (s = \langle\!\langle p, r, s_d, s_r \rangle\!\rangle \wedge s_r \downarrow).$$

Proposition 10 *If $s \downarrow$, then $\widehat{s}$ is defined.*

> Proof: By induction. If $s = \langle\!\langle f \rangle\!\rangle$, then since $\widehat{s} = f$ and f is defined, $\widehat{s}$ is defined.
>
> Assume $s \downarrow$. If $s = \langle\!\langle p, c, s_d, s_r \rangle\!\rangle$, there are two cases: $c = d$ or $c = r$. Let us look at $c = d$. If $c = d$, then $s_d \downarrow$ and $\widehat{s_d}$ is defined by induction and since $\widehat{s} = \widehat{s_d}$, we conclude that $\widehat{s}$ is defined.
>
> The case $c = r$ is similar. $\Box$

As we will consider the payoff function also for subprofiles, we want the payoff function to be defined on subprofiles as well. Therefore we define a property stronger than convergence which we call *strong convergence*[4]. We say that a strategy profile s is *strongly convergent* and we write it $s \Downarrow$ if it is the largest predicate fulfilling the following conditions.

- $\langle\!\langle p, c, s_d, s_r \rangle\!\rangle \Downarrow$ if
  - $\langle\!\langle p, c, s_d, s_r \rangle\!\rangle$ is convergent,
  - s_d is strongly convergent,
  - s_r is strongly convergent.

[3] Called *leads to a leaf* in [22].
[4] Called *always lead to a leaf* in [22].

- $\langle\!\langle f \rangle\!\rangle \Downarrow$ that is that for whatever f, $\langle\!\langle f \rangle\!\rangle$ is strongly convergent

More formally:

$$s \Downarrow \quad \overset{c}{\Longleftrightarrow} \quad s = \langle\!\langle f \rangle\!\rangle \ \vee \ (s = \langle\!\langle p, c, s_d, s_r \rangle\!\rangle \wedge s \downarrow \wedge s_d \Downarrow \wedge s_r \Downarrow).$$

There is however a difference between the definitions of $\downarrow$ and $\Downarrow$. Wherever $s \downarrow$ is defined *by induction*[5], from ending games to the game s, $s \Downarrow$ is defined *by coinduction*[6]. This difference between recursive and corecursive definition is the core of coalgebra theory [13, 30, 32]. However, this is not the aim of the present paper to present coinduction in detail. Both induction and coinduction are based on the fixed-point theorem. The definition of $\Downarrow$ is typical of infinite profiles and means that $\Downarrow$ is invariant along the infinite profile. To make the difference clear and explicit between the definitions, we use the symbol $\overset{i}{\Longleftrightarrow}$ for inductive definitions and the symbol $\overset{c}{\Longleftrightarrow}$ for coinductive definitions. Recall that the definition of the function $\mathsf{game} :: \mathsf{StratProf} \to \mathsf{Game}$ was presented as a coinductive function. We get easily the following proposition.

Proposition 11 $s \Downarrow \quad \Rightarrow \quad s \downarrow$.

Clearly we do not have the opposite implication, as shown by Example 17. Indeed $s \downarrow$ is a local property whereas $s \Downarrow$ is a somewhat global property.

We can define the notion of *subprofile*, written $\precsim$:

$$s' \precsim s \quad \overset{i}{\Longleftrightarrow} \quad s' \sim_s s \ \vee \ s = \langle\!\langle p, c, s_d, s_r \rangle\!\rangle \wedge (s' \precsim s_d \vee s' \precsim s_r),$$

where $\sim_s$ is the bisimilarity[7] among profiles defined as the largest binary predicate $s' \sim_s s$ such that

$$\begin{aligned} s' \sim_s s \quad \overset{c}{\Longleftrightarrow} \quad s' = \langle\!\langle f \rangle\!\rangle = s \quad \vee \quad & (s' = \langle\!\langle p, c, s'_d, s'_r \rangle\!\rangle \wedge \\ & s = \langle\!\langle p, c, s_d, s_r \rangle\!\rangle \wedge \\ & s'_d \sim_s s_d \wedge s'_r \sim_s s_r). \end{aligned}$$

Notice that since we work with infinite objects, we may have $s \not\sim_s s'$ and $s \precsim s' \precsim s$. In other words, an infinite profile can be a strict subprofile of itself. This is the case for $s_{1,0,a}$ and $s_{1,0,b}$ in Section 4. If a profile is strongly convergent, then its subprofiles are strongly convergent as well and the payoffs associated with all its subprofiles are defined.

Proposition 12

1. *If* $s_1 \Downarrow$ *and* $s_2 \precsim s_1$ *then* $s_2 \Downarrow$.
2. *If* $s_1 \Downarrow$ *and if* $s_2 \precsim s_1$, *then* $\widehat{s_2}$ *is defined.*

2.2 The always modality

We notice that $\downarrow$ characterizes a profile by a property of the start vertex, we would say that this property is local. $\Downarrow$ is obtained by distributing the property along the game. In other words we transform the predicate $\downarrow$ and such a predicate transformer is called a *modality*. Here we are interested by the modality *always*, also written $\Box$.

Given a predicate Φ on strategy profiles, the predicate $\Box\,\Phi$ is defined coinductively as follows:

[5] Roughly speaking a definition by induction works from the basic elements, to the constructed elements. For the natural numbers, for 0 and from n to $n+1$. For finite strategy profiles, for $\langle\!\langle f \rangle\!\rangle$ and from s_1 and s_2 to $\langle\!\langle p, c, s_1, s_2 \rangle\!\rangle$.

[6] Roughly speaking a definition by coinduction works on infinite objects, like an invariant.

[7] The reader can consider $\sim_s$ as the equality on $\mathsf{StratProf}$.

$$\Box\,\Phi(s) \quad \overset{c}{\Longleftrightarrow} \quad \Phi(s) \wedge s = \langle\!\langle p, c, s_d, s_r \rangle\!\rangle \Rightarrow (\Box\,\Phi(s_d) \wedge \Box\Phi(s_r)).$$

The predicate "is strongly convergent" is the same as the predicate "is always convergent".

Proposition 13 $s \Downarrow \quad \Leftrightarrow \quad \Box \downarrow (s)$.

2.3 Strategies

The coalgebra Strat of *strategies*[8] is defined by the functor

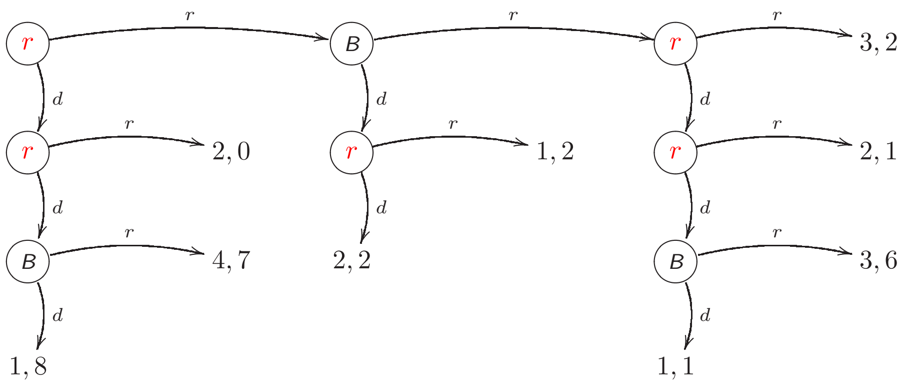
$$[\![\ ]\!] : \mathsf{X} \to \mathbb{R}^{\mathsf{P}} \ + \ (\mathsf{P} + \mathsf{Choice}) \times \mathsf{X} \times \mathsf{X}$$

where $\mathsf{P} = \{\mathsf{A}, \mathsf{B}\}$ and $\mathsf{Choice} = \{d, r\}$. In other words, the coalgebra Strat of strategies is solution of the equation:

$$\mathsf{Strat} \quad \simeq \quad \mathbb{R}^{\mathsf{P}} \ + \ (\mathsf{P} + \mathsf{Choice}) \times \mathsf{Strat} \times \mathsf{Strat}.$$

A strategy of agent p is a game in which some occurrences of p are replaced by choices. A strategy is written $[\![f]\!]$ or $[\![x, s_1, s_2]\!]$.

Example 14 *Consider the strategy for* A *in the game* gg *in which* A *decides to always take the choice* r*.*



By replacing the choice made by agent p by the agent p herself, the function $\mathsf{st2g}$ associates a game with a pair consisting of a strategy and an agent:

$$\begin{array}{lcll}
\mathsf{st2g}([\![f]\!], p) & = & \langle f \rangle & \\
\mathsf{st2g}([\![x, st_1, st_2]\!], p) & = & \textbf{if } x \in \mathsf{P} & \textbf{then } \langle x, \mathsf{st2g}(st_1, p), \mathsf{st2g}(st_2, p) \rangle \\
& & & \textbf{else } \langle p, \mathsf{st2g}(st_1, p), \mathsf{st2g}(st_2, p) \rangle.
\end{array}$$

If a strategy st is really the strategy of agent p it should contain nowhere p and should contain a choice c instead. In this case we say that st is *full for* p and we write it $st \overset{p}{\curlyvee}$.

$$\begin{array}{rcl}
[\![f]\!] \overset{p}{\curlyvee} & & \\
[\![x, st_1, st_2]\!] \overset{p}{\curlyvee} & \overset{c}{\Longleftrightarrow} & (x \notin \mathsf{Choice} \Rightarrow x \neq p) \wedge st_1 \overset{p}{\curlyvee} \wedge st_2 \overset{p}{\curlyvee}.
\end{array}$$

We can sum strategies to make a profile. But for that we have to assume that all strategies are full and have the same game. We say that the strategies are *consistent*. In other words, $(st_p)_{p \in \mathsf{P}}$ is a family of strategies such that:

[8] A strategy is not the same as a strategy profile, which is obtained as the sum of strategies.

- $\forall p \in \mathsf{P}, st_p \overset{p}{\curlyvee}$,
- there exists a game g such that for all p in P, $\mathsf{st2g}$ returns g, more formally $\exists g \in \mathsf{Game}, \forall p \in \mathsf{P}, \mathsf{st2g}(st_p) = g$.

We define the sum $\bigoplus_{p\in\mathsf{P}} st_p$ of consistent strategies as follows:

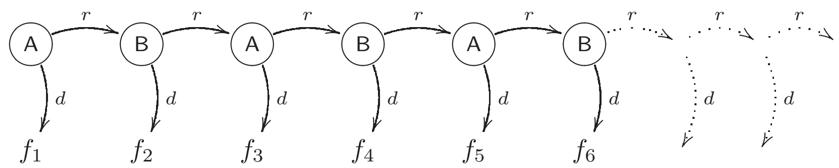
$$\begin{aligned}
\bigoplus_{p\in\mathsf{P}} [\![f]\!] &= \langle\!\langle f \rangle\!\rangle \\
[\![c, st_{p',1}, st_{p',2}]\!] \oplus \bigoplus_{p\in\mathsf{P}\setminus p'} [\![p', st_{p,1}, st_{p,2}]\!] &= \langle\!\langle p', c, \bigoplus_{p\in\mathsf{P}} st_{p,1}, \bigoplus_{p\in\mathsf{P}} st_{p,2} \rangle\!\rangle.
\end{aligned}$$

We can show that the game of all the strategies is the game of the strategy profile which is the sum of the strategies.

Proposition 15 $\mathsf{st2g}(st_{p'}, p') = \mathsf{game}(\bigoplus_{p\in P} st_p)$.

3 Infinipede games and the 0,1-game

We will restrict to simple games which have the shape of combs,



At each step the agents have only two choices, namely to stop or to continue. Let us call such a game, an *infinipede*.

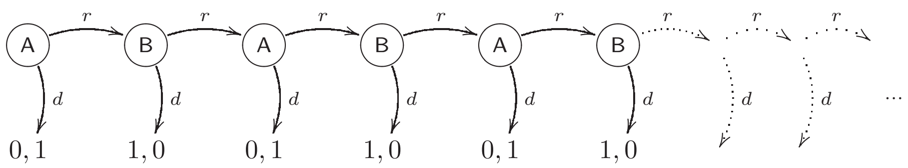
We introduce infinite games by means of equations. Let us see how this applies to define the $0,1$-game. First consider two payoff functions:

$$\begin{aligned}
f_{0,1} &= \mathsf{A} \mapsto 0, \mathsf{B} \mapsto 1 \\
f_{1,0} &= \mathsf{A} \mapsto 1, \mathsf{B} \mapsto 0
\end{aligned}$$

we define two games

$$\begin{aligned}
g_{0,1} &= \langle \mathsf{A}, \langle f_{0,1} \rangle, g_{1,0} \rangle \\
g_{1,0} &= \langle \mathsf{B}, \langle f_{1,0} \rangle, g_{0,1} \rangle
\end{aligned}$$

This means that we define an infinite sequential game $g_{0,1}$ in which agent A is the first player and which has two subgames: the trivial game $\langle f_{1,0} \rangle$ and the game $g_{1,0}$ defined in the other equation. The game $g_{0,1}$ can be pictured as follows:



or more simply in Figure 1.a.

From now on, we assume that we consider only strategy profiles s whose game is the 0,1-game, that is $\mathsf{game}(s) = g_{0,1}$. They are characterized by the following predicates

$$\begin{aligned}
\mathsf{S}_0(s) &\stackrel{c}{\Longleftrightarrow} s = \langle\!\langle \mathsf{A}, c, \langle\!\langle f_{0,1} \rangle\!\rangle, s' \rangle\!\rangle \wedge \mathsf{S}_1(s') \\
\mathsf{S}_1(s) &\stackrel{c}{\Longleftrightarrow} s = \langle\!\langle \mathsf{B}, c, \langle\!\langle f_{1,0} \rangle\!\rangle, s' \rangle\!\rangle \wedge \mathsf{S}_0(s').
\end{aligned}$$

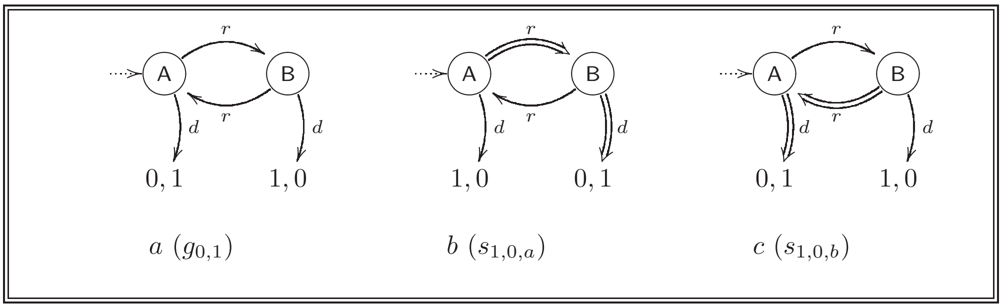


Figure 1: The 0, 1-game and two equilibria seen compactly

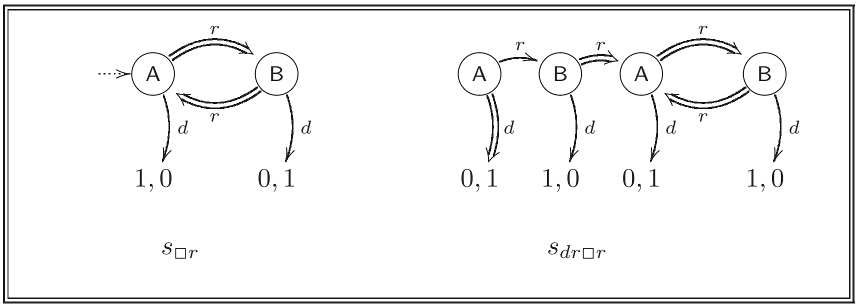


Figure 2: Two examples of strategy profiles for the game $g_{0,1}$

Example 16 *Here is a strategy profile (Figure 2)*

$$s_{\Box r} = \langle\!\langle A, r, \langle\!\langle f_{0,1}\rangle\!\rangle, \langle\!\langle B, r, \langle\!\langle f_{1,0}\rangle\!\rangle, s_{\Box r}\rangle\!\rangle\rangle\!\rangle$$

where both agents continue forever. Notice that $\mathsf{S}_0(s_{\Box r})$, $\neg(s_{\Box r}\downarrow)$ *and a fortiori* $\neg(s_{\Box r}\Downarrow)$*. Said in words,*

1. $s_{\Box r}$ *has game* $g_{0,1}$*,*
2. $s_{\Box r}$ *is not convergent,*
3. $s_{\Box r}$ *is not strongly convergent.*

Example 17 *Consider now the strategy profile (Figure 2)*

$$s_{d\Box r} = \langle\!\langle A, d, \langle\!\langle f_{0,1}\rangle\!\rangle, \langle\!\langle B, r, \langle\!\langle f_{1,0}\rangle\!\rangle, s_{\Box r}\rangle\!\rangle\rangle\!\rangle.$$

This time $\mathsf{S}_0(s_{d\Box r})$, $s_{d\Box r}\downarrow$ *and* $\neg(s_{dd\Box r}\Downarrow)$*. Said in words,*

1. $s_{d\Box r}$ *has game* $g_{0,1}$*,*
2. $s_{d\Box r}$ *is convergent,*
3. $s_{d\Box r}$ *is not strongly convergent.*

We have $\widehat{s_{d\Box r}}(A) = 0$ *and* $\widehat{s_{dd\Box r}}(B) = 1$*. But* $s_{d\Box r}$ *is not strongly convergent since* $\langle\!\langle B, r, \langle\!\langle f_{1,0}\rangle\!\rangle, s_{\Box r}\rangle\!\rangle$ *is not convergent.*

Notice that the 0, 1-game we consider is somewhat a zero-sum game, but we are not interested in this aspect. Moreover, a very specific instance of a 0, 1 game has been considered (by Ummels [38] for instance), but these authors are not interested in the general structure of the game, but in a specific model on a finite graph, which is not general enough for our taste. Therefore for Ummels the 0, 1-game is not a direct generalization of finite sequential games (replacing induction by coinduction) and not a framework to study escalation.

4 Subgame perfect equilibria

Among the strategy profiles, we can select specific ones that are called *subgame perfect equilibria* [33]. Subgame perfect equilibria are specific strategy profiles that fulfill a predicate SPE. This predicate relies on another predicate PE which checks a local property.

$$\begin{array}{lcll} \mathsf{PE}(s) & \Leftrightarrow & s \Downarrow & \wedge\ s = \langle\!\langle p, d, s_d, s_r \rangle\!\rangle \Rightarrow \widehat{s_d}(p) \geq \widehat{s_r}(p) \\ & & & \wedge\ s = \langle\!\langle p, r, s_d, s_r \rangle\!\rangle \Rightarrow \widehat{s_r}(p) \geq \widehat{s_d}(p) \end{array}$$

A strategy profile is a subgame perfect equilibrium if the property PE holds always:

$$\mathsf{SPE} = \Box \mathsf{PE}.$$

Example 18 *In Example 5 we have* $\mathsf{SPE}(s_1)$, $\mathsf{SPE}(s_2)$ *and* $\neg\mathsf{SPE}(s_3)$. *In Example 16,* $\neg\mathsf{SPE}(s_{\Box r})$ *since* $\neg(s_{\Box r}) \Downarrow$.

We may now wonder what the subgame perfect equilibria of the 0,1-game are. We present four of them in Figure 1.b, Figure 1.c and Figure 3. But there are others. To present them, let us define a predicate *"*A *continues and* B *eventually stops"*

$$\begin{array}{lcl} \mathsf{AcBes}(s) \overset{i}{\Longleftrightarrow} s = \langle\!\langle p, c, \langle\!\langle f \rangle\!\rangle, s' \rangle\!\rangle & \Rightarrow & (p = \mathsf{A} \wedge f = f_{0,1} \wedge c = r \wedge \mathsf{AcBes}(s')) \vee \\ & & (p = \mathsf{B} \wedge f = f_{1,0} \wedge (c = d \vee \mathsf{AcBes}(s')) \end{array}$$

Proposition 19 $(\mathsf{S}_1(s) \vee \mathsf{S}_0(s)) \Rightarrow \mathsf{AcBes}(s) \Rightarrow \widehat{s} = f_{1,0}$

> Proof: Assume $\mathsf{S}_1(s) \vee \mathsf{S}_0(s)$. If $s = \langle\!\langle p, c, \langle\!\langle f \rangle\!\rangle, s' \rangle\!\rangle$, then $\mathsf{S}_0(s') \vee \mathsf{S}_1(s')$. Therefore if $\mathsf{AcBes}(s')$, by induction, $\widehat{s'} = f_{1,0}$. By cases:
>
> - If $p = \mathsf{A} \wedge c = r$, then $\mathsf{AcBes}(s')$ and by definition of $\widehat{s}$, we have $\widehat{s} = \widehat{s'} = f_{0,1}$
> - if $p = \mathsf{B} \wedge c = d$, the $\widehat{s} = \widehat{\langle\!\langle f_{1,0} \rangle\!\rangle} = f_{1,0}$.
> - if $p = \mathsf{B} \wedge c = r$, then $\mathsf{AcBes}(s')$ and by definition of $\widehat{s}$, $\widehat{s} = \widehat{s'} = f_{1,0}$.
>
> $\Box$

Like we generalize PE to SPE by applying the modality $\Box$, we generalize AcBes into SAcBes by stating:

$$\mathsf{SAcBes} = \Box \mathsf{AcBes}.$$

There are at least two profiles which satisfies SAcBes namely $s_{1,0,a}$ and $s_{1,0,b}$ which have been studied in [20] and pictured in Figure 1:

$$\begin{array}{lcllcl} s_{1,0,a} & \overset{c}{\Longleftrightarrow} & \langle\!\langle \mathsf{A}, r, \langle\!\langle f_{0,1} \rangle\!\rangle, s_{1,0,b} \rangle\!\rangle & \quad s_{1,0,b} & \overset{c}{\Longleftrightarrow} & \langle\!\langle \mathsf{B}, d, \langle\!\langle f_{1,0} \rangle\!\rangle, s_{1,0,a} \rangle\!\rangle \\ s_{0,1,a} & \overset{c}{\Longleftrightarrow} & \langle\!\langle \mathsf{A}, d, \langle\!\langle f_{0,1} \rangle\!\rangle, s_{0,1,b} \rangle\!\rangle & \quad s_{0,1,b} & \overset{c}{\Longleftrightarrow} & \langle\!\langle \mathsf{B}, r, \langle\!\langle f_{1,0} \rangle\!\rangle, s_{0,1,a} \rangle\!\rangle \end{array}$$

In Figure 3, we give other strategy profiles which fulfill the predicate SAcBes. For the first one we draw only the beginning of the strategy profile, but the reader can imagine that he continues a strategy profile in which A always continues whereas B does not always continue, in other words, B stops infinitely often.

Proposition 20 $\mathsf{SAcBes}(s) \Rightarrow s \Downarrow$.

We may state the following proposition.

Proposition 21 $\forall s, (\mathsf{S}_0(s) \vee \mathsf{S}_1(s)) \Rightarrow (\mathsf{SAcBes}(s) \Rightarrow \mathsf{SPE}(s))$.

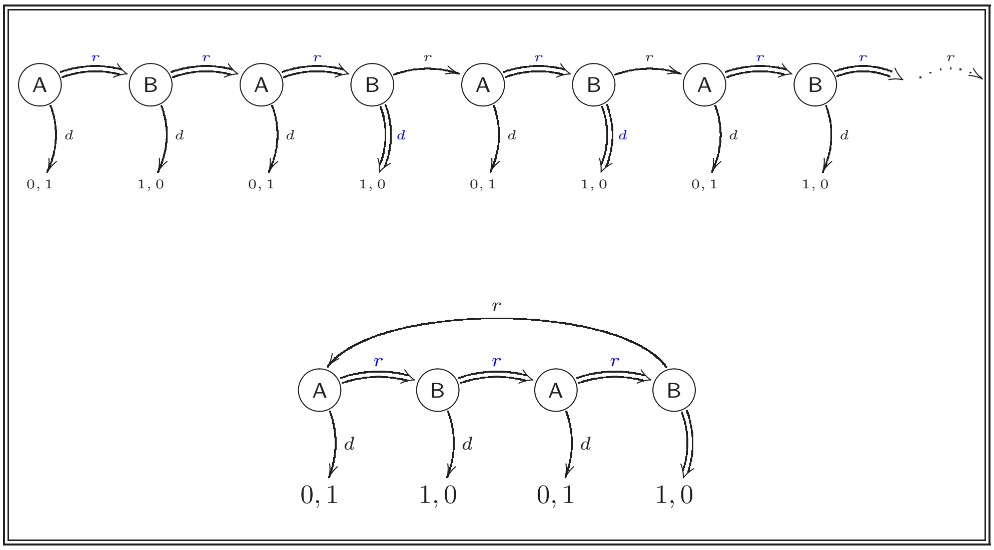


Figure 3: Other equilibria of the 0, 1 game.

Proof: Since SPE is a coinductively defined predicate, the proof is by coinduction.

Given an s, we have to prove $\forall s, \Box \mathsf{AcBes}(s) \wedge (\mathsf{S}_0(s) \vee \mathsf{S}_1(s)) \Rightarrow \Box \mathsf{PE}(s)$.

For that we assume $\Box \mathsf{AcBes}(s) \wedge (\mathsf{S}_0(s) \vee \mathsf{S}_1(s))$ and in addition (coinduction principle) $\Box \mathsf{PE}(s')$ for all strict subprofiles s' of s and we prove $\mathsf{PE}(s)$. In other words, $s \Downarrow \wedge \langle\langle p, d, s_d, s_r \rangle\rangle \Rightarrow \widehat{s_d}(p) \geq \widehat{s_r}(p) \;\; \wedge \langle\langle p, r, s_d, s_r \rangle\rangle \Rightarrow \widehat{s_r}(p) \geq \widehat{s_d}(p)$.

By Proposition 20, we have $s \Downarrow$.

By Proposition 19, we know that for every subprofile s' of a profile s that satisfies $\mathsf{S}_1(s) \vee \mathsf{S}_0(s)$ we have $\widehat{s'} = f_{1,0}$ except when $s' = \langle\langle f_{0,1} \rangle\rangle$. Let us prove $\langle\langle p, d, s_d, s_r \rangle\rangle \Rightarrow \widehat{s_d}(p) \geq \widehat{s_r}(p) \;\wedge\; \langle\langle p, r, s_d, s_r \rangle\rangle \Rightarrow \widehat{s_r}(p) \geq \widehat{s_d}(p)$. Let us proceed by case:

- $s = \langle\langle \mathsf{A}, r, \langle\langle f_{0,1} \rangle\rangle, s' \rangle\rangle$. Then $\mathsf{S}_0(s)$ and $\mathsf{S}_1(s')$. Since $\Box \mathsf{AcBes}(s)$, we have $\mathsf{AcBes}(s')$, therefore $\widehat{s'} = f_{1,0}$ hence $\widehat{s'}(\mathsf{A}) = 1$ and $f_{0,1}(\mathsf{A}) = 0$, henceforth $\widehat{s'}(\mathsf{A}) \geq f_{0,1}(\mathsf{A})$.
- $s = \langle\langle \mathsf{B}, r, \langle\langle f_{1,0} \rangle\rangle, s' \rangle\rangle$. Then $\mathsf{S}_1(s)$ and $\mathsf{S}_0(s')$. Since $\Box \mathsf{AcBes}(s)$, we have $\mathsf{AcBes}(s')$, therefore $\widehat{s'} = f_{1,0}$ hence $\widehat{s'}(\mathsf{B}) = 0$ and $f_{1,0}(\mathsf{B}) = 0$, henceforth $\widehat{s'}(\mathsf{B}) \geq f_{1,0}(\mathsf{B})$.

$\Box$

Symmetrically we can define a predicate BcAes for "B continues and A eventually stops" and a predicate SBcAes which is $\mathsf{SBcAes} = \Box\ \mathsf{BcAes}$ which means that B always continues and A stops infinitely often. With the same argument as for SAcBes we can conclude :

Proposition 22 $\forall s, (\mathsf{S}_0(s) \vee \mathsf{S}_1(s)) \Rightarrow \mathsf{SBcAes}(s) \Rightarrow \mathsf{SPE}(s)$.

Lemma 23 *Assume* $\mathsf{S}_0(s)$ *or* $\mathsf{S}_1(s)$, *then* $\mathsf{SPE}(s) \Rightarrow (\mathsf{SAcBes}(s) \vee \mathsf{SBcAes}(s))$

Proof: By contradiction. Assume $(\neg\mathsf{SAcBes}(s)) \wedge (\neg\mathsf{SBcAes}(s))$. This means that one of the following statements are fulfilled.

- *There exist* s_B *and* s_A *such that* $s_A = \langle\langle A, d, \langle\langle f_{0,1}\rangle\rangle, s'_A\rangle\rangle \precsim s_B = \langle\langle B, d, \langle\langle f_{1,0}\rangle\rangle, s'_B\rangle\rangle \precsim s$ *and there are only "r's" between* B *and* A. Notice that $\widehat{s'_B}(\mathsf{B}) = 1$ while $\widehat{\langle\langle f_{1,0}\rangle\rangle}(\mathsf{B}) = 0$. Since $\mathsf{SPE}(s)$ then $\mathsf{SPE}(s_\mathsf{B})$ therefore $s_\mathsf{B} = \langle\langle \mathsf{B}, d, \langle\langle f_{1,0}\rangle\rangle, s'_\mathsf{B}\rangle\rangle$ implies $\widehat{s'_\mathsf{B}}(\mathsf{B}) \leq \widehat{\langle\langle f_{1,0}\rangle\rangle}(\mathsf{B})$ which is a contradiction.
- *There exist* s_A *and* s_B *such that* $s_B = \langle\langle B, d, \langle\langle f_{1,0}\rangle\rangle, s'_A\rangle\rangle \precsim s_A = \langle\langle A, d, \langle\langle f_{0,1}\rangle\rangle, s'_B\rangle\rangle \precsim s$ *and there are only "r's" between* A *and* B. The contradiction is obtained like above.
- $s_{\Box r} \precsim s$, which means that eventually A and B continue forever and which is in contradiction with $\mathsf{SPE(s)}$ since $\neg\mathsf{SPE}(s_{\Box r})$ (see Example 18).

$\Box$

$\mathsf{SAcBes} \vee \mathsf{SBcAes}$ fully characterizes SPE of 0,1-games, in other words.

Theorem 24 $\forall s, (\mathsf{S}_0(s) \vee \mathsf{S}_1(s)) \Rightarrow (\mathsf{SAcBes}(s) \vee \mathsf{SBcAes}(s) \Leftrightarrow \mathsf{SPE}(s))$.

5 Nash equilibria

Before talking about escalation, let us see the connection between subgame perfect equilibrium and Nash equilibrium in a sequential game. In [26], the definition of a Nash equilibrium is as follows: *A Nash equilibrium is a "pattern[s] of behavior with the property that if every player knows every other player's behavior she has not reason to change her own behavior"* in other words, *"a Nash equilibrium [is] a strategy profile from which no player wishes to deviate, given the other player's strategies."* . The concept of deviation of agent p is expressed by a binary relation we call *convertibility*[9] and we write $\vdash p \dashv$. It is defined inductively as follows:

$$\frac{s \sim_s s'}{s \vdash p \dashv s'}$$

$$\frac{s_1 \vdash p \dashv s'_1 \qquad s_2 \vdash p \dashv s'_2}{\langle\langle p, c, s_1, s_2\rangle\rangle \ \vdash p \dashv \ \langle\langle p, c', s'_1, s'_2\rangle\rangle}$$

$$\frac{s_1 \vdash p \dashv s'_1 \qquad s_2 \vdash p \dashv s'_2}{\langle\langle p', c, s_1, s_2\rangle\rangle \ \vdash p \dashv \ \langle\langle p', c, s'_1, s'_2\rangle\rangle}$$

We define the predicate Nash as follows:

$$\mathsf{Nash}(s) \Leftrightarrow \forall p, \forall s', s \vdash p \dashv s' \Rightarrow \widehat{s}(p) \geq \widehat{s'}(p').$$

The concept of Nash equilibrium is more general than that of subgame perfect equilibrium and we have the following result:

Proposition 25 $\mathsf{SPE}(s) \Rightarrow \mathsf{Nash}(s)$.

The result has been proven in COQ and we refer to the script (see[22]):
`http://perso.ens-lyon.fr/pierre.lescanne/COQ/EscRatAI/`
`http://perso.ens-lyon.fr/pierre.lescanne/COQ/EscRatAI/SCRIPTS/`

Notice that we defined the convertibility inductively, but a coinductive definition is possible. But this would give a more restrictive definition of Nash equilibrium.

[9] This should be called perhaps *feasibility* following [29] and [19]

6 Escalation

Escalation in a game with a set P of agents occurs when there is a tuple of consistent strategies $(st_p)_{p\in\mathsf{P}}$ such that its sum is not convergent, in other words, $\neg\,(\bigoplus_{p\in\mathsf{P}} st_p)\downarrow$. Said differently, it is possible that the agents have all a private strategy which combined with those of the others makes a strategy profile which is not convergent, which means that the strategy profile goes to infinity when following the choices. Notice the two uses of a strategy profile: first, as a subgame perfect equilibrium, second as a combination of the strategies of the agents.

Consider the strategy:

$$\begin{array}{rcl} st_{\mathsf{A},\infty} &=& [\![r, [\![f_{0,1}]\!], st'_{\mathsf{A},\infty}]\!] \\ st'_{\mathsf{A},\infty} &=& [\![\mathsf{B}, [\![f_{1,0}]\!], st_{\mathsf{A},\infty}]\!] \end{array}$$

and its twin

$$\begin{array}{rcl} st_{\mathsf{B},\infty} &=& [\![\mathsf{A}, [\![f_{0,1}]\!], st'_{\mathsf{B},\infty}]\!] \\ st'_{\mathsf{B},\infty} &=& [\![r, [\![f_{1,0}]\!], st_{\mathsf{B},\infty}]\!]. \end{array}$$

Moreover, consider the strategy profile:

$$\begin{array}{rcl} s_{\mathsf{A},\infty} &=& \langle\!\langle \mathsf{A}, r, \langle\!\langle f_{0,1}\rangle\!\rangle, s_{\mathsf{B},\infty}\rangle\!\rangle \\ s_{\mathsf{B},\infty} &=& \langle\!\langle \mathsf{B}, r, \langle\!\langle f_{1,0}\rangle\!\rangle, s_{\mathsf{A},\infty}\rangle\!\rangle. \end{array}$$

Proposition 26

1. $st_{A,\infty} \overset{\mathsf{A}}{\curlyvee\!\!\downarrow}$,
2. $st_{B,\infty} \overset{\mathsf{B}}{\curlyvee\!\!\downarrow}$,
3. $\mathsf{st2g}(st_{A,\infty}, A) = \mathsf{st2g}(st_{B,\infty}, B) = g_{0,1}$,
4. $\mathsf{game}(s_{A,\infty}) = g_{0,1}$,
5. $st_{A,\infty} \oplus st_{B,\infty} = s_{A,\infty}$,
6. $\neg\; s_{A,\infty}\downarrow$.

   Proof: The first statements are proved by coinduction on the definition of $\overset{\mathsf{A}}{\curlyvee\!\!\downarrow}$, $\mathsf{st2g}$, game and $st_{\mathsf{A},\infty}$. The last statement is by induction on the definition of $\downarrow$. $\square$

Proposition 26 can be said in words as follows:

1. $st_{\mathsf{A},\infty}$ is full for A,
2. $st_{\mathsf{B},\infty}$ is full for B,
3. $st_{\mathsf{A},\infty}$ and $st_{\mathsf{B},\infty}$ have game $g_{0,1}$,
4. $s_{\mathsf{A},\infty}$ has game $g_{0,1}$,
5. strategy $st_{\mathsf{B},\infty}$ plus strategy $st_{\mathsf{B},\infty}$ yields profile $s_{\mathsf{A},\infty}$,
6. profile $s_{\mathsf{A},\infty}$ is not convergent.

$st_{\mathsf{A},\infty}$ and $st_{\mathsf{B},\infty}$ are both correct since they are built using choices, namely r, dictated by subgame perfect equilibria[10] which start with r. Another feature of $0,1$-game is that no agent has a clue for what strategy the other agent is using. Indeed after k steps, A does not know if B has used a strategy derived of equilibria in SAcBes or in SBcAes. In other words, A does not know if B will stop eventually or not and vice versa. The agents can draw no conclusion of what they observe. If each agent does not believe in the threat of the other she is naturally led to escalation.

7 Relativity: are agents really rational?

Observers are less cognitively busy and more open to information than actors.

Daniel Kahneman [14]

Rationality as observed by an *outsider* is not the same as rationality seen by an *insider* [14, 15, 7]. Since coalgebras are the mathematical tool for observation [13] a coalgebra approach does not come as a surprise.

In this paper we would like to use the following definition: "An agent is rational[11] if she is motivated by maximizing her own payoff", but since this leads to debate, we prefer to say that an agent is intelligent in this case. In infinite extensive games, this translates in saying that an agent is intelligent if she adopts a strategy consistent with a subgame perfect equilibrium as an extension of backward induction. This explains the behavior of traders especially when they enter escalation. But is this compatible with common sense? More precisely whereas Howie Hubler (who lost 8.67×10^9\$ for Morgan Stanley), Jérôme Kerviel (who lost 6.95×10^9\$ for Société Générale) or Brian Hunter (who lost 6.69×10^9\$ for Amaranth Advisors) were intelligent agents when they acted, they are seen clearly as stupid by an external observer. In other words, agents reason intelligently in an irrational escalation, which means that an agent can be rational in her closed world and seen irrational from outside. Is this consistent? Yes and this allows us drawing three conclusions:

- Insider view differs from outsider observation.
- According to K. E. Stanovich [35, 36], there are *two levels in Kahneman System II* (the effortful and slow part of the mind) [14]: *algorithmic mind* which is the ability of agents to reason perfectly and logically in a deductive system[12] (called *mindware* by Keith Stanovich) and *reflective mind* (or *epistemic mind*) which is the ability of an agent to reconsider her believes. For instance, reflective mind is visible when the agent realizes that the world (the game) is finite or when the agent realizes that maximizing her own profit is no more the main aim and that the survival of her company should be taken into account. In both cases, she changes her belief in an infinite world or in maximizing her own profit and adds or modifies one or more axiom(s) founding her deductive system, as part of her belief.
- Consistently with the previous statement, an agent who is involved in an escalation can be considered as having a *short term vision*, whereas the observer has a *long term vision.* The agent can also be considered as having a local vision, whereas the observer has a *global vision.*

[10] Recall that our concept of intelligent choice is that of a subgame perfect equilibrium, as it generalizes backward induction, which is accepted following Aumann [3] as the criterion of rationality for finite game.

[11] seen from inside

[12] In infinite game theory, the deductive system includes coinduction and SPE's.

8 Mindware revisited

We propose to extend the concept of mindware proposed by Stanovich which itself refines system II (slow thinking) of Kahneman [14]. For Stanovich, the deductive system of the mindware is always the same, namely *classical first order logic*, only its implementation in the mind of the agent may vary and can be incomplete. The reasoning of the mindware relies on believes that can be changed by the reflective mind. For us, the mindware may implement several kinds of deductive systems and may change from one agent to the other. For instance, it can implement classical first order logic, intuitionistic first order logic, classical or intuitionistic first order logic extended with inductive reasoning or extended with coinductive reasoning, higher order logic (intuitionistic or classical) etc. All those logics are equally acceptable, because they are consistent and from Gödel theorem we know that no universal system of deduction exists. Like for Stanovich, believes may also be changed by the reflective mind. In his book *What Intelligence Tests Miss: The Psychology of Rational Thought* [35] Stanovich presents a few examples to set his point. The fact that the mindware may implement several deductive systems leads to reconsider the scope of his examples. This is the case for the example that illustrates the beginning of chapter six:

> *Jack is looking at Anne but Anne is looking at George. Jack is married but George is not. Is a married person looking at a unmarried person?*
> *A) Yes B) No C) Cannot be determined*
> *Answer A, B, or C.*
> *[...]*
> *The vast majority of people answer C.*

In others words, the vast majority of people, me included, show on this example that they have a correct mindware based on a constructive logic, for instance, on intuitionistic logic, because such a logic is easier to use. Therefore they answer C. Stanovich claims that such an answer is incorrect and calls us *cognitive misers*. Actually when reading further [35] I understood that I was supposed to use a more sophisticated logic, especially that I should use the *excluded middle*, in other words, that I should use classical logic, because *"most people can carry fully disjunctive reasoning when they are explicitly* told *that it is necessary"* ([35] p. 71). Therefore I changed my belief by adding *p or not p* and I answered A. But this requires more calculation. Indeed *p or not p* can be specialized into *Anne is married or Anne is not married* from which we draw *Anne is married and George is not married and Anne is looking at George or Jack is married and Anne is not married and Jack is looking a Anne*. We can abstract this into *there exists x and there exists y such that x is married and y is not married and x is looking at y or there exists x and there exists y such that x is married and y is not married and x is looking at y* which simplifies into *there exists x and there exists y such that x is married and y is not married and x is looking at y* and leads to answer A. However if the above question is completed by

> *Choose one of the followings*
> *A) Anne is married*
> *B) Anne is not married*
> *C) I don't know.*

A rational person including a person who answered A at the first question will answer C at this question. More precisely, answering A at the first question does not allow the agent to justify her answer by exhibiting a pair of persons such that one is married and looking at the other who is not married. Unable to justify their answer, the majority of people choose C at the first question as well.

To introduce disrationality, chapter two of [35] starts with the case of John Allen Paulos who was involved in a classical escalation [28]. Actually John Allen Paulos, professor of mathematics at Temple University, is a typical person with a sound algorithmic mind using coinduction knowingly or not[13]. Actually he was disrational because he used improperly his reflective mind and did not change his belief in an "everlasting" Worldcom company (which eventually bankrupted in 2002) and in its eventual restart. Paulos is obviously "foolish", if we say that someone is foolish, if he has a perfect algorithmic mind, but a faulty reflective mind, in other words if he is intelligent, but not rational.

The above comments strengthen Stanovich's distinction between algorithmic mind and reflective mind, but make the delineation of rationality harder and its evaluation difficult, because attributing a "rationality quotient" requires first to determine the deduction system implemented in the mindware and its strength, then to appreciate the ability of the agent of changing her believes.

9 Which deductive system?

I say that it is not illogical to think that the world is infinite.

Jose Luis Borges, *The library of Babel* in [4]

.

We noticed that several deductive systems can be used. We may wonder what features a deductive system should have. Let us tell some of them. First the language should contain a modality to enable agents to express *belief* (B_a) or *knowledge* (K_a). Moreover, the excluded middle ($p \vee \neg p$) is clearly not mandatory. However an agent a should be able to state that she believes in the excluded middle by a statement like

$$B_a((\forall p : \mathsf{Proposition})\ p \vee \neg p).$$

This requires a quantification over propositions, which allows also expressing sentences like: *We know there are known unknowns* (see [17] Section 4):

$$\bigwedge_{a \in \mathsf{Agent}} K_a((\exists p : \mathsf{Proposition})\ K_a(\neg K_a(p))).$$

Besides quantifications over propositions, quantifications over functions and sets are required to express belief in *finiteness* ([20] Section 6):

$$B_a((\forall A : Set)\,(\forall f : A \to A)\,(Surjective(f) \Rightarrow Injectve(f)))$$

where *Surjective* is a predicate which asserts *surjectivity* and *Injective* is a predicate which asserts *injectivity.*

10 Conclusion

In this paper, we have shown how to use coinduction in economics, more precisely in economic game theory where it has not been used yet, or perhaps in a hidden form, which has to be unearthed. We have shown also that rational agents can be seen as irrational by observers since observation changes the point of view, in particular on rationality. When Wolfgang Leininger writes (see citation in front of Section 1)

[13] Actually Paulos uses coinduction unknowingly, rather invoking a kind of invariant

that the fact that rational agents should not engage in an escalation seems obvious, he means, "obvious" for an observer, not for the agents, or perhaps he should have said that rational agents should not engage in an escalation, but that intelligent agents could. If agents are only *intelligent*, the efficiency of the markets should then be revisited at the light of escalation. Therefore coinduction is a possible way for rethinking economics.

References

[1] Samson Abramsky and Viktor Winschel. Coalgebraic analysis of subgame-perfect equilibria in infinite games without discounting. *arXiv*, 2012.

[2] Peter Aczel. *Non-well-founded Sets*, volume 14 of *CSLI Lecture Notes*. CSLI Publications, Stanford, California, 1988.

[3] Robert J. Aumann. Backward induction and common knowledge of rationality. *Games and Economic Behavior*, 8:6–19, 1995.

[4] Jorge Luis Borges. *Ficciones*. Grove Press, 1944.

[5] Jean-Philippe Bouchaud. Economics needs a scientific revolution. *Nature*, 455:1181, oct 2008.

[6] Jean-Philippe Bouchaud and Marc Potters. *Theory of Financial Risk and Derivate Pricing*. Cambridge University Press, 2003.

[7] Jeffrey C. Chandor. Margin call. Movie, May 2012. `http://www.imdb.com/title/tt1615147/`.

[8] David Colander. The future of economics: the appropriately educated in pursuit of the knowable. *Cambridge Journal of Economics*, 29:927–941, 2005.

[9] Andrew M. Colman. *Game theory and its applications in the social and biological sciences*. London New York : Routledge, 1999. Second edition.

[10] Gabrielle Demange. Rational escalation. *Annales de l'INSEE*, pages 227–249, 1992.

[11] Herbert Gintis. *Game Theory Evolving: A Problem-Centered Introduction to Modeling Strategic Interaction*. Princeton University Press, 2000.

[12] Joseph Y. Halpern. Substantive rationality and backward induction. *Games and Economic Behavior*, 37(2):425–435, 2001. available at http://ideas.repec.org/a/eee/gamebe/v37y2001i2p425-435.html.

[13] Bart Jacobs. *Introduction to Coalgebra. Towards Mathematics of States and Observations*. Online book, 2012. Version 2.0.

[14] Daniel Kahneman. *Thinking, Fast and Slow*. Allen Lane, 2011.

[15] Cédric Klapisch. My piece of the pie. Movie, March 2011. `http://www.imdb.com/title/tt1641624/`.

[16] Wolfgang Leininger. Escalation and cooperation in conflict situations. *J. of Conflict Resolution*, 33:231–254, 1989.

[17] Pierre Lescanne. Mechanizing common knowledge logic using COQ. *Ann. Math. Artif. Intell.*, 48(1-2):15–43, 2006.

[18] Pierre Lescanne. Deconstruction of infinite extensive games using coinduction. *CoRR*, abs/0904.3528, 2009.

[19] Pierre Lescanne. Feasibility/desirability games for normal form games, choice models and evolutionary games. *CoRR*, abs/0907.5469, 2009.

[20] Pierre Lescanne. Rationality and escalation in infinite extensive games. *CoRR*, abs/1112.1185, 2011.

[21] Pierre Lescanne. A simple case of rationality of escalation. In Reiko Heckel and Stefan Milius, editors, *CALCO*, volume 8089 of *Lecture Notes in Computer Science*, pages 191–204. Springer, 2013.

[22] Pierre Lescanne and Matthieu Perrinel. "Backward" coinduction, Nash equilibrium and the rationality of escalation. *Acta Inf.*, 49(3):117–137, 2012.

[23] Dimitry Mirimanoff. Les antimonies de Russell et de Burali-Forti et le problème fondamental de la théorie des ensembles. *L'Enseignement Mathématique*, 19:37–52, 1917.

[24] Lawrence S. Moss and Ignacio D. Viglizzo. Harsanyi type spaces and final coalgebras constructed from satisfied theories. *Electr. Notes Theor. Comput. Sci.*, 106:279–295, 2004.

[25] Barry O'Neill. International escalation and the dollar auction. *J. of Conflict Resolution*, 30(33-50), 1986.

[26] Martin J. Osborne. *An Introduction to Game Theory*. Oxford, 2004.

[27] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. The MIT Press, Cambridge, Massachusetts, 1994.

[28] John A. Paulos. *A Mathematician Plays the Stock Market*. Basic Books, 2003.

[29] Ariel Rubinstein. *Lecture Notes in Microecomic Theory*. Princeton University Press, 2006.

[30] Jan J. M. M. Rutten. Universal coalgebra: a theory of systems. *Theor. Comput. Sci.*, 249(1):3–80, 2000.

[31] Davide Sangiorgi. On the origins of bisimulation and coinduction. *ACM Trans. Program. Lang. Syst.*, 31(4), 2009.

[32] Davide Sangiorgi. *An introduction to bisimulation and coinduction*. Cambridge University Press, 2011.

[33] Reinhard Selten. Spieltheoretische Behandlung eines Oligopolmodells mit Nachfrageträgheit. *Zeitschrift für gesamte Staatswissenschaft*, 121, 1965.

[34] Martin Shubik. The dollar auction game: A paradox in noncooperative behavior and escalation. *Journal of Conflict Resolution*, 15(1):109–111, 1971.

[35] Keith E. Stanovich. *What Intelligence Tests Miss: The Psychology of Rational Thought*. Yale University Press, 2010.

[36] Keith E. Stanovich. *Rationality and the reflective mind*. Oxford University Press, 2011.

[37] Adair Turner. *Economics after the Crisis: Objectives and Means*. MIT Press, Cambridge, MA, USA, 2012.

[38] Michael Ummels. The complexity of Nash equilibria in infinite multiplayer games. In Roberto M. Amadio, editor, *FoSSaCS*, volume 4962 of *Lecture Notes in Computer Science*, pages 20–34. Springer, 2008.

[39] René Vestergaard. A constructive approach to sequential Nash equilibria. *Inf. Process. Lett.*, 97:46–51, 2006.

[40] Karl Weierstrass. Über continuirliche Funktionen eines reellen Arguments, die für keinen Werth des letzteren einen bestimmten Differentialquotienten besitzen. *in Karl Weiertrass Mathematische Werke, Abhandlungen II Gelesen in der Königl. Akademie der Wisseschaften, am 18 Juli 1872*, pages 71–74, 1872.

[41] Viktor Winschel. Private communication, 2012.

A Finite 0,1 games and the "cut and extrapolate" method

We spoke about the "cut and extrapolate" method, applied in particular to the dollar auction. Let us see how it would work on the 0,1-game. Finite games, finite strategy profiles and payoff functions of finite strategy profiles are the inductive equivalent of infinite games, infinite strategy profiles and infinite payoff functions which we presented. Notice that payoff functions of finite strategy profiles are always defined. Despite we do not speak of the same types [14] of objects, we use the same notations, but this does not lead to confusion. Consider two infinite families of finite games, that could be seen as approximations of the 0,1-game:

$$F_{0,1} \quad = \quad \langle \mathsf{A}, \langle f_{0,1}\rangle, \langle \mathsf{B}, \langle f_{1,0}\rangle, F_{0,1}\rangle\rangle \cup \{\langle f_{0,1}\rangle\}$$

$$\begin{array}{rcl} K_{0,1} & = & \langle \mathsf{A}, \langle f_{0,1}\rangle, K'_{0,1}\rangle \\ K'_{0,1} & = & \langle \mathsf{B}, \langle f_{1,0}\rangle, K_{0,1}\rangle \cup \{\langle f_{1,0}\rangle\} \end{array}$$

In $F_{0,1}$ we cut after B and replace the tail by $\langle f_{0,1}\rangle$. In $K_{0,1}$ we cut after A and replace the tail by $\langle f_{1,0}\rangle$. Recall [39] the predicate *backward induction* shortened in BI, which is the finite and inductive version of PE.

$$\begin{array}{rcl} & & \mathsf{BI}(\langle f\rangle) \\ \mathsf{BI}(\langle p, c, s_d, s_r\rangle) & = & \mathsf{BI}(s_l) \wedge \mathsf{BI}(s_r) \wedge \\ & & \langle\!\langle p, d, s_d, s_r\rangle\!\rangle \Rightarrow \widehat{s_d}(p) \geq \widehat{s_r}(p) \ \wedge \\ & & \langle\!\langle p, r, s_d, s_r\rangle\!\rangle \Rightarrow \widehat{s_r}(p) \geq \widehat{s_d}(p) \end{array}$$

Example 27 *In Example 5 we have* $\mathsf{BI}(s_1)$, $\mathsf{BI}(s_2)$ *and* $\neg\mathsf{BI}(s_3)$.

We consider the two families of strategy profiles:

$$\begin{array}{rcl} \mathsf{SF}_{0,1}(s) & \overset{i}{\Longleftrightarrow} & (s = \langle\!\langle \mathsf{A}, d, \langle\!\langle f_{0,1}\rangle\!\rangle, \langle\!\langle \mathsf{B}, r, \langle\!\langle f_{1,0}\rangle\!\rangle, s'\rangle\!\rangle\rangle\!\rangle \wedge \mathsf{SF}_{0,1}(s')) \quad \vee \\ & & (s = \langle\!\langle \mathsf{A}, r, \langle\!\langle f_{0,1}\rangle\!\rangle, \langle\!\langle \mathsf{B}, r, \langle\!\langle f_{1,0}\rangle\!\rangle, s'\rangle\!\rangle\rangle\!\rangle \wedge \mathsf{SF}_{0,1}(s')) \quad \vee \\ & & s = \langle\!\langle f_{0,1}\rangle\!\rangle \end{array}$$

$$\begin{array}{rcl} \mathsf{SK}_{0,1}(s) & \overset{i}{\Longleftrightarrow} & s = \langle\!\langle \mathsf{A}, r, \langle\!\langle f_{0,1}\rangle\!\rangle, s'\rangle\!\rangle \wedge \mathsf{SK}'_{0,1}(s') \\ \mathsf{SK}'_{0,1}(s) & \overset{i}{\Longleftrightarrow} & (s = \langle\!\langle \mathsf{B}, d, \langle\!\langle f_{1,0}\rangle\!\rangle, s'\rangle\!\rangle \vee s = \langle\!\langle \mathsf{B}, r, \langle\!\langle f_{1,0}\rangle\!\rangle, s'\rangle\!\rangle) \wedge \mathsf{SK}_{0,1}(s') \quad \vee \\ & & s = \langle\!\langle f_{1,0}\rangle\!\rangle \end{array}$$

[14] In the sense of type theory

In $\mathsf{SF}_{0,1}$, B continues and A does whatever she likes and in $\mathsf{SK}_{0,1}$, A continues and B does whatever she likes. The following proposition characterizes the backward induction equilibria for games in $F_{0,1}$ and $K_{0,1}$ respectively and is easily proved by induction:

Proposition 28

- $\mathsf{game}(s) \in F_{0,1} \wedge \mathsf{SF}_{0,1}(s) \Leftrightarrow \mathsf{BI}(s)$,
- $\mathsf{game}(s) \in K_{0,1} \wedge \mathsf{SK}_{0,1}(s) \Leftrightarrow \mathsf{BI}(s)$.

This shows that cutting at an even or an odd position does not give the same strategy profile by extrapolation. Consequently the "cut and extrapolate" method does not anticipate all the subgame perfect equilibria. Let us add that when cutting we decide which leaf to insert, namely $\langle f_{0,1} \rangle$ or $\langle f_{1,0} \rangle$, but we could do another way obtaining different results.

0,1 game and limited payroll. To avoid escalation in the dollar auction, people require a *limited payroll*, i.e., a bound on the amount of money handled by the agents, but this is inconsistent with the fact that the game is infinite. Said otherwise, to avoid escalation, they forbid escalation. We can notice that, in the 0,1-game, a limited payroll would not prevent escalation, since the payoffs are anyway limited by 1. In the same vein, Demange [10] adds, to justify escalation, a new feature called *joker* which is not necessary as we have shown.