Swansea University
Prifysgol Abertawe

Cronfa
Setting Research Free

Cronfa - Swansea University Open Access Repository

_____

This is an author produced version of a paper published in:
Electronic Proceedings in Theoretical Computer Science

Cronfa URL for this paper:
http://cronfa.swan.ac.uk/Record/cronfa36013

_____

Paper:

Le Roux, S. & Pauly, A. (2016). A Semi-Potential for Finite and Infinite Sequential Games (Extended Abstract).
Electronic Proceedings in Theoretical Computer Science, 226, 242-256.
http://dx.doi.org/10.4204/EPTCS.226.17

_____

A Semi-Potential for Finite and Infinite Sequential Games
(Extended Abstract)

Stéphane Le Roux
Département d'informatique
Université libre de Bruxelles, Belgique
Stephane.Le.Roux@ulb.ac.be

Arno Pauly
Département d'informatique
Université libre de Bruxelles, Belgique
Arno.Pauly@cl.cam.ac.uk

We consider a dynamical approach to sequential games. By restricting the convertibility relation over strategy profiles, we obtain a semi-potential (in the sense of Kukushkin), and we show that in finite games the corresponding restriction of better-response dynamics will converge to a Nash equilibrium in quadratic time. Convergence happens on a per-player basis, and even in the presence of players with cyclic preferences, the players with acyclic preferences will stabilize. Thus, we obtain a candidate notion for rationality in the presence of irrational agents. Moreover, the restriction of convertibility can be justified by a conservative updating of beliefs about the other players strategies.

For infinite sequential games we can retain convergence to a Nash equilibrium (in some sense), if the preferences are given by continuous payoff functions; or obtain a transfinite convergence if the outcome sets of the game are Δ_2^0-sets.

1 Introduction

Nash equilibria are the terminal strategy profiles of the better (or best) response dynamics. In general, though, these dynamics do not terminate (and do not even converge). A particular exception is found in the potential games [20]: A potential is an acyclic joint order extension of the individual players improvement relations. In a potential game, better-response dynamics thus will always improve the potential, and hence terminates (at a Nash equilibrium) if the game is finite.

The notion of semi-potential was introduced by KUKUSHKIN ([10], also [11]) in order to salvage some of the nice properties of potential games for a larger class of games. Here, the players freedom to change their strategies is restricted – however, only in such a way that if they can change the current outcome to a particular one, they can also do so in a way that is consistent with the restriction. In a *generic* normal form game this is equivalent to a potential, as there different strategies will induce different outcomes. Nevertheless, several classes of non-generic games have no potential but have a semi-potential: [10, Theorem 3] proved that it is the case for sequential games.

We study the restriction of the convertibility relation (we call it *lazy convertibility*) as well as the resulting better-response dynamics (*lazy improvement*) in some more detail. We give two alternative proofs of the termination at a Nash equilibrium in finite games, one of which yields a tight quadratic bound on the number of steps required. Moreover, our two proofs of termination works on a per player basis: Thus, any player with acyclic preferences will converge, even in the presence of players with cyclic preferences. Then we explore several possible extensions to infinite sequential games, and one to directed acyclic graphs. A very specific infinite setting was explored in [3], and lazy improvement in infinite sequential games with continuous payoff functions was investigated by the authors in [16].

Some relevant properties of lazy improvement are:

- The dynamics are uncoupled: Each player bases her decisions only on her own preference and the strategies of the other players, but does not need to know the other players' preferences.

- The dynamics are history-independent: Unlike e.g. fictitious play or typical regret-minimization approaches (e.g. [9]), the next step in the dynamics depends only on the current strategies of the players. In particular, players do not need additional memory for *learning*.

- We consider pure strategies, not stochastic ones. Thus, our approach has a very different flavour from the usual evolutionary game theory one (e.g. [6, 5, 24]).

- No restrictions akin to *generic payoffs* are required, we merely need acyclic preferences to guarantee termination at a Nash equilibrium in finite games (and anyway this requirement cannot be avoided for existence of Nash equilibrium [12, 13]).

- In a finite game, the dynamics stabilizes at a Nash equilibrium after a quadratic number of steps.

- The stabilization result for the rational players, *i.e.* with acyclic preferences, remains unaffected, if unpredictable players, *i.e.* with cyclic preferences, are added.

- Under some conditions, even in an infinite sequential game we can ensure stabilization at a Nash equilibrium after a transfinite number of steps.

The rest of the paper is organized as follows: Section 3 introduces the core concept of *lazy improvement*. Section 4 proves that in a finite game, lazy improvement terminates at a Nash equilibrium. Section 5 gives an alternative proof also showing that termination occurs after a quadratic number of improvement steps. Section 6 gives a basic epistemic justification for lazy convertibility. In Section 7 we discuss extensions to infinite games. Finally, Section 8 provides a number of (counter)examples showing that, to some extent, our definitions have to be the way they are. An extended preprint is available as [15].

2   Background and Notation

We start by introducing notation for both games in normal form and sequential games.

Definition 1 (Games in normal form). *A game in normal form is a tuple $\langle A, (S_a)_{a \in A}, O, v, (\prec_a)_{a \in A} \rangle$ satisfying the following:*

- *A is a non-empty set (of players, or agents),*

- $\prod_{a \in A} S_a$ *is a non-empty Cartesian product (whose elements are the* strategy profiles *and where S_a represents the strategies available to player a),*

- *O is a non-empty set (of possible outcomes),*

- $v : \prod_{a \in A} S_a \to O$ *(the outcome function that values the strategy profiles),*

- *Each $\prec_a$ is a binary relation over O (modelling the preference of player a).*

Definition 2 (Nash equilibrium). *Let $\langle A, (S_a)_{a \in A}, O, v, (\prec_a)_{a \in A} \rangle$ be a game in normal form. A strategy profile (profile for short) s in $S := \prod_{a \in A} S_a$ is a Nash equilibrium if it makes every player a stable, i.e. $v(s) \not\prec_a v(s')$ for all $s' \in S$ that differ from s at most at the a-component.*

$$NE(s) \quad := \quad \forall a \in A, \forall s' \in S, \quad \neg(v(s) \prec_a v(s') \wedge \forall b \in A - \{a\}, s_b = s'_b)$$

Implicit in the concept of Nash equilibrium is the notion of *convertibility*: An agent can convert one strategy profile to another, if they differ only in her actions. As lazy improvement will be introduced in Section 3 by restricting the convertibility relation, we provide a formal definition:

Definition 3 (Convertibility, induced preference over profiles, and improvement)**.**

- *Let* $\langle A, (S_a)_{a \in A}, O, v, (\prec_a)_{a \in A} \rangle$ *be a game in normal form. For* $s, s' \in \prod_{a \in A} S_a$, *let* $s \xrightarrow{c}_a s'$ *denote the ability of player* a *to convert* s *to* s' *by changing her own strategy, formally* $s \xrightarrow{c}_a s' := \forall b \in A - \{a\}, s_b = s'_b$.

- *Given a game* $\langle A, (S_a)_{a \in A}, O, v, (\prec_a)_{a \in A} \rangle$, *let* $s \prec_a s'$ *denote* $v(s) \prec_a v(s')$. *So in this article* $\prec_a$ *may also refer to the induced preference over the profiles.*

- *Let* $\twoheadrightarrow_a := \prec_a \cap \xrightarrow{c}_a$ *be the individual improvement relations of the players and let* $\twoheadrightarrow := \cup_{a \in A} \twoheadrightarrow_a$ *be the collective[1] improvement relation.*

Observation 4 below is a direct consequence of Definitions 2 and 3.

Observation 4. *The Nash equilibria of a game are exactly the sinks, i.e., the terminal profiles of the collective improvement* $\twoheadrightarrow$.

A (generalized) *potential* is an acyclic relation containing $\twoheadrightarrow$. Clearly a game has a potential iff $\twoheadrightarrow$ itself is acyclic. If $\prod_{a \in A} S_a$ is finite, this is equivalent to the better-response dynamics' always converging. A less restrictive notion is a semi-potential (introduced in [10]). A semi-potential is an acyclic relation $\hookrightarrow$ contained in $\twoheadrightarrow$, such that whenever $s \twoheadrightarrow s'$ then there is some s'' with $s \hookrightarrow s''$ and $v(s) = v(s'')$. In words, if a strategy profile can be reached by an improvement step, then there is an equivalent strategy profile (w.r.t outcome) reachable via a step in the semi-potential. It follows that the sinks of a semi-potential are exactly the sinks of the collective improvement. Thus, in a finite setting, the existence of a semi-potential in particular implies the existence of sinks, i.e. Nash equilibria.

Our setting will be sequential games, rather than games in normal form. The idea here is that the players collectively choose a path through a tree, with each player deciding the direction at the vertices that she is controlling. The preferences refer only to the path created, choices off the chosen path are irrelevant. Thus, the evaluation map v is highly non-injective[2], which in turn gives room for the notion of a semi-potential to be interesting. Formally, we define sequential games as follows:

Definition 5 (Sequential games)**.** *A sequential game is a tuple* $(A, T, O, d, v, (\prec_a)_{a \in A})$ *where*

- *A is the non-empty set of players,*

- *T is a rooted tree (finite or infinite),*

- *O is the non-empty set of outcomes,*

- *d associates a player with each vertex in the tree,*

- *v associates an outcome with each maximal path from the root through the tree,*

- *and for each player* $a \in A$, *$\prec_a$ is a relation on O (the preference relation of a).*

The corresponding game in normal form is obtained as follows: Let a strategy of player a associate an outgoing edge with each vertex controlled by a. If a strategy per player is given, the collective choices identify some maximal path p through the tree, called the *induced play*. Applying v to that path yields the outcome of the game; i.e. the valuation of the game in normal form is the composition of the map that identifies the induced play and the valuation of the sequential game.

In our concrete examples, the outcomes will be tuples of natural numbers, and the n-th player will prefer a tuple $(x_1, \ldots, x_{|A|})$ to $(y_1, \ldots, y_{|A|})$ iff $x_n > y_n$.

---

[1]The word *collective* should not be misread to indicate any form of coordination between the players.

[2]

3 Defining *lazy improvement*

The idea underlying lazy improvement is that we do not let a player change their irrelevant choices, i.e. those choices not along the play induced after the improvement. Equivalently, we require a player to change as few choices as possible when changing the induced play.

Definition 6 (Lazy convertibility and improvement).

- *For two strategy profiles s, s' in a sequential game let $s \overset{c}{\dashrightarrow}_a s'$ (read: a can lazily convert s into s'), if for any vertex $t \in T$, if $s(t) \neq s'(t)$, then $d(t) = a$ and t lies along the play induced by s'.*

- *Let $\dashrightarrow_a := \prec_a \cap \overset{c}{\dashrightarrow}_a$ be the lazy improvement of player a and let $\dashrightarrow := \cup_{a \in A} \dashrightarrow_a$ be the (collective) lazy improvement.*

Let us exemplify the notion of lazy convertibilty, which has nothing to do with the preferences or the outcomes: player a can lazily convert the leftmost strategy profile below into each of the profiles below, but not into any other profile. Player a is written bold face at nodes where changes occur, and double lines represent strategy choices.
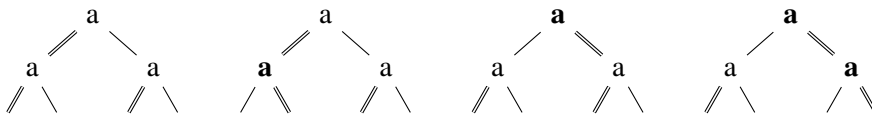
Contrary to the convertibility relations $\overset{c}{\rightarrow}_a$ which are equivalence relations, the lazy convertibility relations $\overset{c}{\dashrightarrow}_a$ are certainly reflexive but in general neither symmetric nor transitive. For instance, player a cannot lazily convert the rightmost profile above back into the leftmost one. In the additional example below, player a can convert the leftmost profile to the middle profile but not to the rightmost profile.
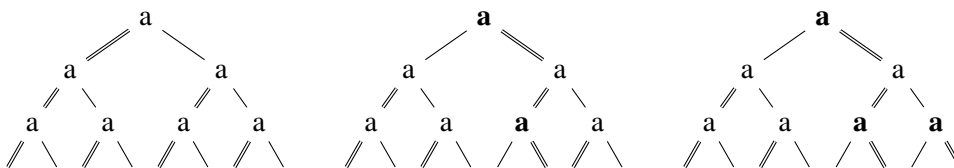
The lazy convertibility enjoys a useful property nonetheless, that the usual convertibility does not: if a player avoids a play during a sequence of lazy conversion, only the very same player is later able to make the last step to induce the same play again, possibly induced by a different profile. This phenomenon is more formally stated by Lemma 7.

Lemma 7. *If $s \overset{c}{\dashrightarrow}_a s_0 \overset{c}{\dashrightarrow} \cdots \overset{c}{\dashrightarrow} s_n \overset{c}{\dashrightarrow}_b s'$ where s and s' induce the same play, and if this play is different from the plays that are induced by the s_i, then $a = b$.*

Proof. Let us prove the claim by induction on the underlying game. Since the play induced by s_0 is different from the play induced by s, these profiles are not just leaves, but proper trees instead. During the assumed $\overset{c}{\dashrightarrow}$ reduction of s, its subprofile that is chosen by the root owner in s undergoes a $\overset{c}{\dashrightarrow}$ reduction too, say $t \overset{c}{\dashrightarrow}_a t_0 \overset{c}{\dashrightarrow} \cdots \overset{c}{\dashrightarrow} t_n \overset{c}{\dashrightarrow}_b t'$, where t and t' induce the same play (and the root owner chooses t' in s'). If all these subprofiles are equal, player a must be the root owner (of s), since s and s_0 induce different plays by assumption, and b is also the root owner since s_n and s' induce different plays, so $a = b$. Now let t_j be the first subprofile different from t, so t_j induces a play different from t and t'. For all k such that $j \leq k < n$, if t_k and t' induce different plays but t_{k+1} and t' induce the same play, then s_{k+1} and s' induce the same play by definition of $\overset{c}{\dashrightarrow}$, contradiction with the assumptions of the lemma, so all $t_j, \ldots t_n$ induce plays different from that of t'. If $t_1 \neq t$, then $a = b$ by the induction hypothesis, else a must be the root

owner and does not choose t_1 in s_1. The first time that a chooses some t_i again must be in s_j: indeed if it were before, s and s_i would induce the same play, and if it were after, a could not change t_{j-1} into t_j. Therefore $t_{j-1} \xrightarrow{c}_a t_j \xrightarrow{c} \cdots \xrightarrow{c} t_n \xrightarrow{c}_b t'$ and $a = b$ by the induction hypothesis. $\square$

Despite the restrictive property from Lemma 7, the lazy convertibility is as effective as the usual convertibility, in the same sense as used in the definition of a semi-potential. (Thus, it will only remain to prove that lazy improvement is acyclic in order to establish lazy improvement as a semi-potential).

Observation 8. *If $s \twoheadrightarrow s'$, then there is some strategy profile s'' such that $s \rightharpoonup s''$ and $v(s') = v(s'')$.*

Proof. By definition, lazy convertibility does not restrict the choice of the new induced play, merely the ability to alter the strategy off the new induced play. $\square$

Corollary 9. *The Nash equilibria of a game are exactly the terminal profiles of the lazy improvement $\rightharpoonup$.*

## 4   Termination in finite games, first proof

Theorem 10. *Consider a sequential game played on a finite tree, and some sequence $(s_n)_{n \in \mathbb{N}}$ such that $s_n \rightharpoonup s_{n+1}$ for all $n \in \mathbb{N}$. Assume that for a player a there are infinitely many n with $s_n \rightharpoonup_a s_{n+1}$. Then a has a cyclic preference.*

Proof. Towards a contradiction let us assume that a's preference is acyclic. Among the profiles s such that $s = s_n \rightharpoonup_a s_{n+1}$ for infinitely many n, let s be minimal for a's preference, and let M be large enough such that every profile s_n with $M < n$ occurs infinitely often in the sequence. Let $s = s_n$ for some $n > M$, and let $k > n$ be the least k such that s_n and s_k induce the same play. Lemma 7 implies that $s_{k-1} \rightharpoonup_a s_k$, so player a prefers the outcome of s_n over that of s_{k-1}, contradiction. $\square$

Together with Corollary 9 the following corollary shows the equivalence between all preferences being acyclic and universal existence of NE.

Corollary 11. *Consider outcomes O, players A, and their preferences $(\prec_a)_{a \in A}$: All $\prec_a$ are acyclic iff for all finite sequential games built from O, A and $(\prec_a)_{a \in A}$ the collective lazy improvement terminates.*

Proof. The difficult implication of the equivalence is a corollary of Theorem 10. For the other implication, note that if $x_0 \prec_a x_1 \prec_a \cdots \prec_a x_n \prec_a x_0$, then $\rightharpoonup_a$ does not terminate on the profile below.

$\square$
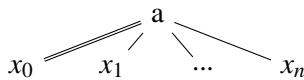
Corollary 12. *In a finite sequential game where every player has acyclic preferences, lazy improvement is a semi-potential.*

Proof. Combine Corollary 9 and Corollary 11. $\square$

KUKUSHKIN ([10, Theorem 3]) proved Corollary 12 in the case where the preferences are derived from payoffs. In this specific (yet usual) setting, it is not possible to consider players with cyclic preferences, so Theorem 10 or Corollary 11 cannot even be stated.

Based on Corollary 11 we obtain a reasonable candidate for rational behaviour in sequential games played with an unpredictable nature or erratic players: Perform lazy improvement until the players with acyclic preferences no longer change their strategies. It is always consistent with the observations to

assume that the changes in another player's strategy are based on lazy convertibility. This argument is explored in more detail in Section 6. Nature can then be modelled as a player with the full relation as preferences, such that any convertible step for nature becomes an improvement step.

5   Termination in finite games, second proof

The proof of Theorem 10, by contradiction, gives a quick argument but no deep insight on how and how fast the relation terminates. A stronger statement can proven by using the multiset of outcomes avoided by a player a (i.e. the outcomes obtained in a subgame, where the decision not to play into that subgame was made by a, see Definition 13) to construct a measure that will decrease on any lazy improvement step by a (Lemma 16), and remain unchanged by any lazy convertibility step by a different player (Lemma 15). Thus, we are in a situation very similar to potential games [20] – however, in a potential game a player *can* increase the potential (which is common to all the players) but does not *want* to, whereas here the players *cannot* impact the measure of another player as long as they are restricted to lazy convertibility.

Definition 13 (Avoided outcomes of a game and of a profile). *The avoided outcomes of a game g is a function $\Delta(g)$ of type $A \to \mathbb{N}$, and it is defined inductively below.*

- $\Delta(g, a) := 0$ *if g is a leaf game.*

- *If player a owns the root of a game g whose children are $g_0, \ldots, g_n$ then*
  - $\Delta(g, b) := \sum_{j=0}^{n} \Delta(g_j, b)$ *for all $b \neq a$.*
  - $\Delta(g, a) := \left(\sum_{j=0}^{n} \Delta(g_j, a) \right) + n$

The avoided outcomes of a profile s is a function $\delta(s)$ of type $A \to O \to \mathbb{N}$, or equivalently in this case, of type $A \times O \to \mathbb{N}$, and it is defined inductively below.

- $\delta(s, a, o) := 0$ *if s is a leaf profile.*

- *If player a owns the root of a profile s and chooses the subprofile s_i among $s_0, \ldots, s_n$ then*
  - $\delta(s, b, o) := \sum_{j=0}^{n} \delta(s_j, b, o)$ *for all $b \neq a$.*
  - $\delta(s, a, o) := \left(\sum_{j=0}^{n} \delta(s_j, a, o) \right) + |\{ j \in \{0, \ldots, n\} - \{i\} \mid v(s_j) = o \}|$

The smaller array below describes the function $\Delta(g)$, where g is the underlying game of the left-hand profile s below, and the right-hand array describes the function $\delta(s)$. For instance $\delta(s, b, y) = 2$ because player b avoids the outcome y twice: once at the left-most internal node, after two leftward moves, when choosing outcome x rather than y, and also once after one rightward move, also when choosing x rather than y. Note that the only leaf that is not accounted for by the function of the avoided outcome of a profile/game is the leaf that is induced by the profile.

$\Delta(g, \cdot)$
$a \mapsto 5$
$b \mapsto 3$

$\delta(s, \cdot, \cdot)$	x	y	z	t
a	1	0	1	3
b	0	2	1	0
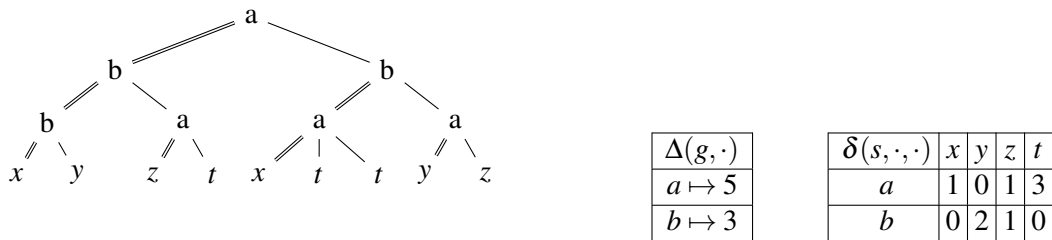
Observation 14 below relates the two functions from Definition 13. It refers to $s2g$, a function that returns the underlying game of a given profile, see [12] or [13] for a proper definition.

Observation 14.   *1. Let s be a profile and a be a player, then $\Delta(s2g(s), a) = \sum_{o \in O} \delta(s, a, o)$.*

  2. Let g be a game, then $1 + \sum_{a \in A} \Delta(g, a)$ equals the number of leaves of g.

Proof.   1. By induction on s. If s is a leaf profile, the claim holds since $\Delta(s2g(s), a) = 0 = \delta(s, a, o)$ by definition, so now let s be a profile where the root owner a chooses s_i among subprofiles $s_0, \ldots, s_n$. For $b \neq a$ Definition 13 and the induction hypothesis yield $\Delta(s2g(s), b) = \sum_{j=0}^n \Delta(s2g(s_j), b) \overset{I.H.}{=} \sum_{j=0}^n \sum_{o \in O} \delta(s_j, b, o) = \sum_{o \in O} \sum_{j=0}^n \delta(s_j, b, o) = \sum_{o \in O} \delta(s, b, o)$. Similarly we have $\Delta(s2g(s), a) = \sum_{j=0}^n \Delta(s2g(s_j), a) + n \overset{I.H.}{=} \sum_{j=0}^n \sum_{o \in O} \delta(s_j, a, o) + |\{j \in \{0, \ldots, n\} - \{i\} \mid v(s_j) \in O\}| = \sum_{o \in O} \left(\sum_{j=0}^n \delta(s_j, a, o) + |\{j \in \{0, \ldots, n\} - \{i\} \mid v(s_j) = o\}| \right) = \sum_{o \in O} \delta(s, a, o)$.

  2. By induction on g. This holds for every leaf game g since $\Delta(g, a) = 0$ by definition. Let g be a game whose root is owned by player a and whose subgames are $g_0, \ldots, g_n$. The number of leaves in g is the sum of the numbers of leaves in the g_j, that is, $\sum_{j=0}^n \left(1 + \sum_{b \in A} \Delta(g_j, b) \right)$ by induction hypothesis. This, equals $1 + \sum_{j=0}^n \sum_{b \in A - \{a\}} \Delta(g_j, b) + n + \sum_{j=0}^n \Delta(g_j, a)$, which, in turn, equals $1 + \sum_{b \in A - \{a\}} \Delta(g, b) + \Delta(g, a)$ by definition.

$\square$

Lemma 15 below states conservation of the outcomes that are avoided by a player in a profile during a lazy conversion of another player. Intuitively, it is because a lazy conversion of a player cannot modify the subtrees that are avoidd by the other players, even though she owns node therein.

Lemma 15. $s \overset{c}{\to}_a s' \wedge b \neq a \quad \Rightarrow \quad \delta(s, b) = \delta(s', b)$

Proof. By induction on the profile. It holds for leaves, so let $s \overset{c}{\to}_a s'$ with subprofiles $s_0, \ldots, s_n$ and $s'_0, \ldots, s'_n$, respectively. By definition of $\overset{c}{\to}_a$ we have $s_j \overset{c}{\to}_a s'_j$ for all j, and therefore $\delta(s_j, b, o) = \delta(s'_j, b, o)$ by induction hypothesis. If the root owner is different from b, then $\delta(s, b, o) = \sum_{j=0}^n \delta(s_j, b, o) = \sum_{j=0}^n \delta(s'_j, b, o) = \delta(s', b, o)$ by definition of δ. If b is the root owner, she chooses the i-th subprofile in both s and s' since $b \neq a$, and moreover $s'_j = s_j$ for all j distinct from i. So $\delta(s, b, o) = \sum_{j=0}^n \delta(s_j, b, o) + |\{j \in \{0, \ldots, n\} - \{i\} \mid v(s_j) = o\}| = \sum_{j=0}^n \delta(s'_j, b, o) + |\{j \in \{0, \ldots, n\} - \{i\} \mid v(s'_j) = o\}| = \delta(s', b, o)$.

$\square$

However, the conservation does not fully hold for the player who converts the profile, unless the induced outcomes are the same for both profiles. The difference is little though, only depending on both induced outcomes. In Lemma 16 below, *eq* is just a boolean representation of equality: $eq(x, x) := 1$ and $eq(x, y) := 0$ for $x \neq y$.

Lemma 16. $s \overset{c}{\to}_a s' \quad \Rightarrow \quad \delta(s, a) + eq(v(s)) = \delta(s', a) + eq(v(s'))$

Proof. By induction on the profile s. It holds for leaves, so let $s \overset{c}{\to}_a s'$ with subprofiles $s_0, \ldots, s_n$ and $s'_0, \ldots, s'_n$, respectively. If the root owner is distinct from a, she chooses the same i-th subprofile in both s and s', therefore $\delta(s, a, o) + eq(v(s), o) = \sum_{0 \leq j \leq n \wedge j \neq i} \delta(s_j, a, o) + \delta(s_i, a, o) + eq(v(s_i), o) = \sum_{0 \leq j \leq n \wedge j \neq i} \delta(s'_j, a, o) + \delta(s'_i, a, o) + eq(v(s'_i), o) = \delta(s', a, o) + eq(v(s'), o)$ by definition of δ, since $s_j = s'_j$ for $j \neq i$, and by induction hypothesis.

  If a is the root owner, let a choose the i-th and k-th subprofiles in s and s', respectively. Let $N := \delta(s, a, o) + eq(v(s), o)$, so $N = \sum_{0 \leq j \leq n \wedge j \neq k} \delta(s'_j, a, o) + |\{j \in \{0, \ldots, n\} - \{i\} \mid v(s_j) = o\}| + \delta(s_k, a, o) + eq(v(s_i), o)$ by unfolding Definition 13, since $s'_j = s_j$ for all $j \neq k$, and since $v(s) = v(s_i)$ by the choice at the root. Rewriting N twice with the easy-to-check equality $|\{j \in \{0, \ldots, n\} - \{x\} \mid v(s_j) = o\}| + eq(v(s_x), o) = |\{j \in \{0, \ldots, n\} \mid v(s_j) = o\}|$, first with $x := i$ and then with $x := k$ yields the equality $N =$

$\sum_{0 \leq j \leq n \wedge j \neq k}^{n} \delta(s'_j, a, o) + |\{j \in \{0, \ldots, n\} - \{k\} \mid v(s_j) = o\}| + \delta(s_k, a, o) + eq(v(s_k), o)$. Since $s_k \xrightarrow{c}_a s'_k$ by definition of lazy convertibility, and by the induction hypothesis, let us further rewrite $\delta(s_k, a) + eq(v(s_k))$ with $\delta(s'_k, a) + eq(v(s'_k))$ in N. Folding Definition 13 yields $N = \delta(s', a, o) + eq(v(s'), o)$. $\square$

The two lemmas above suggest that whenever a player lazily converts a profile to obtain a better outcome, some measure decreases a bit with respect to her preference, but does not change for the other players. The lazy improvement should therefore terminate, and even quite quickly, as proved below. Recall that a finite preference relation $\prec$ has height at most h if there is no chain $s_1 \prec s_2 \prec \ldots \prec s_{h+1}$.

Theorem 17. *Consider a game g where player a has an acyclic preference of height h. Let $\Delta(g, a)$ be the total number of choices available to player a, minus the number of vertices where a is choosing. Then in any sequence (possibly infinite) of lazy improvement, the number of lazy improvement steps performed by player a is bounded by $(h-1) \cdot \Delta(g, a)$.*

Proof. For every outcome o let $h(a, o)$ be the maximal cardinality of the $\prec_a$-chains whose $\prec_a$-maximum is o, and note that $o \prec_a o'$ implies $h(a, o) < h(a, o')$. For every profile s let $M(s, a) := \sum_{o \in O} (h(a, o) - 1) \cdot \delta(s, a, o)$ and note that $0 \leq M(s, a) \leq (h-1) \cdot \Delta(g, a)$ by Observation 14.1. Let $s \rightarrow_a s'$ be a lazy improvement step, so $s \xrightarrow{c}_a s'$ and $v(s) \prec_a v(s')$ by definition, then $M(s, a) - M(s', a) = \sum_{o \in O} (h(a, o) - 1) \cdot (\delta(s, a, o) - \delta(s', a, o)) = h(a, v(s')) - h(a, v(s)) > 0$ by Lemma 16. Let $s \rightsquigarrow_a s'$ be a lazy equilibrium perturbation step. By Lemma 16 and the definition of $s \sim s'$, we have $M(s, a) = M(s', a)$. Let $s \xrightarrow{c}_b s'$ be a lazy conversion step where $b \neq a$, then $M(s, a) = M(s', a)$ by Lemma 15. This shows that the $\rightarrow_a$ steps are at most $(h-1) \cdot \Delta(g, a)$ in any sequence of $\rightarrow \cup \rightsquigarrow$. $\square$

Corollary 18. *The lazy improvement terminates for all games iff all preferences are acyclic, in which case the number of sequential lazy improvement steps is at most $(h-1) \cdot (l-1)$ where h bounds the cardinality of the preference chains and l is the number of leaves.*

Observation 19.   *1. The maximal length of a lazy improvement sequence is bounded in a quadratic manner in the size of the game in general and linearly when h from Corollary 18 is fixed.*

   2. The quadratic and linear bounds are tight.

Proof. (of 19.2.) For the linear bound, let us consider the figure below and set $x := x_0 = \cdots = x_n$ and $y \prec_a x$ and $x \prec_b y$. There is clearly a lazy improvement sequence starting from the figure and visiting each leaf exactly once.
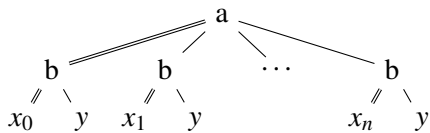

It is similar for the quadratic bound, but we need to be a bit more careful. For $n \in \mathbb{N}$, consider the game in the above figure, where $y \prec_a x_0 \prec_a x_1 \cdots \prec_a x_n$ and $x_i \prec_b y$ for all i. Let us prove by induction on n the existence of a sequence of $\frac{(n+2)(n+3)}{2} - 2$ lazy improvement steps when starting from the strategy profile above. For the base case $n = 0$, there are $1 = \frac{(0+2)(0+3)}{2} - 2$ lazy improvement steps. For the inductive case, let player a make n lazy improvements in a row, by choosing x_1, then x_2, and so on until x_n. At that point, let player b improve from x_n to y and then let player a come back to x_0. So far, $n + 2$ lazy improvement steps have been performed. Now let us ignore the substrategy profile involving x_n (and y). By induction hypothesis, $\frac{(n+1)(n+2)}{2} - 2$ additional lazy improvement steps can be performed in a row. Since $(n+2) + \frac{(n+1)(n+2)}{2} - 2 = \frac{(n+2)(n+3)}{2} - 2$, we are done. $\square$

6 Lazy convertibility as belief updating

Let us discuss whether we should expect players to conform to lazy convertibility when playing a sequential game repeatedly. Observation 8 tells us that in the short term, a player has no incentive to deviate from lazy convertibility: If she desires some outcome she can reach by some deviation from her current strategy, she can obtain this outcome by converting a strategy in a lazy way. There is a caveat, though, in that restricting convertibility to lazy convertibility changes the overall reachability structure, as the following example shows.

Example 20. *The last profile of the three-step improvement relation below is a Nash equilibrium that cannot be reached from the first profile under lazy improvement.*
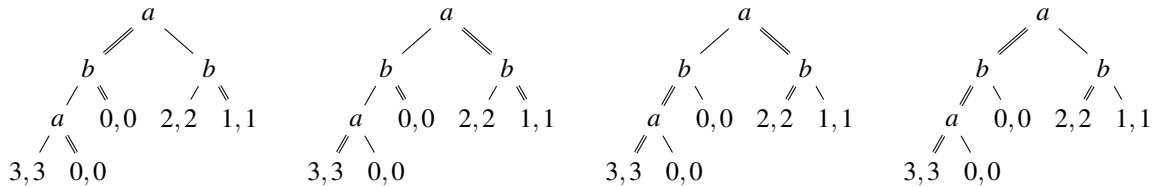
From the perspective of any given player, it however makes a lot of sense to assume that all other players are updating their own strategies only in a lazy way – assuming that only relevant choices of the other players can be observed. The latter seems to be crucial in order to make the game truly sequential: If all players announced their entire strategy simultaneously, it would be a game in normal form after all.

To formalize this idea, let us fix a player a and consider the game from her perspective. She may consider the game as a two-player game played by her against all other players aggregated into a single player b. She starts with some initial strategy $s_a^{(0)}$, and some prior assumption $s_b^{(0)}$ on the strategy of her opponent(s). She then updates her own strategy via lazy improvement to $s_a^{(1)}$. Then the game is actually played, and player a observes the actual moves (but not the strategy) of her opponents. As she only observes the moves along the path actually taken, it is consistent with her observations to assume that the aggregated opponent player lazily converted $s_b^{(0)}$ into some $s_b^{(1)}$. Then the player a again performs a lazy improvement step to $s_a^{(2)}$, plays the game, etc. Provided that the player a has acyclic preferences, Theorem 10 implies that her own strategy stabilizes to some strategy s_a eventually.

Note that this procedure requires no assumptions on knowledge of rationality of other players or their payoff functions, not to speak of common knowledge. There is in general no reason to assume that the aggregated player b acts according to acyclic preferences (given that the different players making up b may have partially antagonistic preferences). However, if each player has a cyclic preference and performs the same procedure as a above, then each players actual strategy will stabilize. As any change in what a player assumes her aggregated opponents are playing has to be caused by either a change in her own, or someone else's strategy, this implies that also the believed strategy of the aggregated players s_b will stabilize. Furthermore, all the strategy profiles constructed in this way induce the same play, and combining them as follows yields a Nash equilibrium:

Proposition 21. *Let a set of players play a finite sequential game by converting their own strategies lazily based on beliefs about the other players strategies in order to maximize an acyclic preference relation. Then a Nash equilibrium can be obtained from the stable strategies they will settle to as follows: Along the common path chosen by their stable strategies, everyone follows their own strategy. In any subgame that is not reached, each player plays according to the beliefs held by the player controlling access to the subgame about their strategies.*

Proof. At any vertex reached during the final play, the choice facing the current player is the same one she was anticipating due to her beliefs on her opponents strategies. As her choice is consistent with the stable choice made during the dynamical updating, she has no incentive to change. □

In comparison, the investigation of the epistemic foundations of Nash equilibria by AUMANN and BRANDENBURGER [2] identified mutual knowledge of rationality, knowledge of the game and (in case of more than two players) a common prior as the prerequisite for playing a Nash equilibrium. A subgame perfect equilibrium requires even stronger assumptions, namely well-aware players [1].

7   Convergence in infinite games

Infinite sequential games with win/lose preferences are generalizations of Gale-Stewart games [7], and are of great relevance for logic. That any two-player game with antagonistic preferences and a Borel winning set actually has a Nash equilibrium is a highly non-trivial result by MARTIN [19]. This is generalized to multi-player games with non-necessarily antagonistic preferences in [14]. Moreover, subgame-perfect equilibria are not always guaranteed to exist (cf. [16, 23]).

The definition of lazy improvement applies to infinite sequential games as well, and we can adapt the results on finite games to see that it still constitutes a semi-potential:

Proposition 22. *Consider an infinite sequential game where each player (there might be countably many) has acyclic preferences. Then lazy improvement is a semi-potential.*

Proof. As argued in Section 3, we only need to show that lazy improvement is acyclic. Assume the contrary, then there is some finite cycle $s_1 \rightharpoonup s_2 \rightharpoonup \ldots \rightharpoonup s_n \rightharpoonup s_1$. Let $p_1, \ldots, p_n$ be the paths induced by these strategy profiles, and choose $k \in \mathbb{N}$ such that $p_i|_{\leq k} = p_j|_{\leq k} \Leftrightarrow p_i = p_j$.

By choice of k, the path chosen inside any subgame rooted at depth k remains unchanged throughout the improvement cycle. Thus, replacing any such subgame with a leaf carrying the outcome induced by this path has no impact on the improvement cycle. We have obtained a finite sequential game with the same preferences and a cycle built from lazy improvement step, contradicting Theorem 10. □

Of course, in an infinite game acyclicity does not suffice to ensure termination or even convergence. In fact, [16, Example 26] (reproduced below as Example 33) shows that lazy improvement in infinite games will not always converge, and that even accumulation points do not have to be Nash equilibria. There are however several potential ways to extend the results on lazy improvement to infinite sequential games:

1. We can consider games where the preferences are expressed via continuous payoff functions. For some fixed $\varepsilon > 0$, we can then consider ε-lazy improvement (where only lazy convertibility is allowed, and improvement steps are only taken if the player can improve by more than ε). Then Theorems 10 and 17 carry over, and as a counterpart to Corollary 9 we find that the terminal profiles of ε-lazy improvement are precisely the ε-Nash equilibria. See [15] for details.

2. We can employ lazy improvement being done in a finitary way with increasing precision, and find that any accumulation point of particular subsequence is guaranteed to be a Nash equilibrium. This was discussed in [16, Section VI].

3. We can consider games where the players have win/lose objectives (i.e. their preference relations have height 2), and the winning sets are Δ_2^0-sets. Then transfinite iteration of lazy improvement will reach a Nash equilibrium.

4. We can generalize the measure employed in the proof of Theorem 17 to infinite games with Lipschitz payoff functions.

The rest of the section provides some details on the third and fourth approaches. We start by formalizing what it means to do a transfinite number of improvement steps. The following definition generalises the notion of finite sequence or ω-sequence induced by a binary relation to α-sequence for some ordinal α: at limit ordinals, following a valid sequence amounts to picking an "accumulation point".

Definition 23 (ordinal sequence of a relation)**.** *Let $\to$ be a binary relation on some topological space S, and let α be an ordinal number. An α-sequence of $\to$ is a family $(s_\beta)_{\beta < \alpha}$ of elements in S such that for all $\beta < \alpha$, if $\beta + 1 < \alpha$ then $s_\beta \to s_{\beta+1}$, and if β is a limit ordinal, then for every $\beta' < \beta$ and every neighbourhood U of s_β there exists $\gamma \in]\beta', \beta[$ such that $s_\gamma \in U$.*

Observation 24 below says that, given a binary relation over a sequentially compact set, the only reason why an ordinal sequence cannot be further extended is when a sink has been reached.

Observation 24. *Let $(s_\beta)_{\beta < \alpha}$ be a countable ordinal sequence of $\to$ over a sequentially compact set S. If $(s_\beta)_{\beta \leq \alpha}$ is not a sequence of $\to$ for any $s_\alpha \in S$, then $\alpha = \alpha' + 1$ for some α' and $s_{\alpha'}$ is a sink of $\to$.*

In this paper by countable we mean at most countable. We do find that even in very simple games, we can have improvement sequences of any countable length:

Proposition 25. *For every countable ordinal α there exists a win-lose two-player game on a binary tree with open winning set for one player, and an α-sequence of lazy improvement in the game.*

Lemma 26 below is the base case of the proof of Theorem 27, which is proved by transfinite induction in the difference hierarchy.

Lemma 26. *Let g be a game with finitely many players who have win-lose objectives. If every winning set is open or closed, every sequence of lazy improvement in g is countable.*

Recall that a subset S of a metric space is called a Δ_2^0-set, if it is expressible both as $S = \bigcap_{i \in \mathbb{N}} U_i$ with open U_i, and as $S = \bigcup_{i \in \mathbb{N}} A_i$ with closed A_i.

Theorem 27. *Let g be a game with finitely many players who have win-lose objectives. If every winning set is Δ_2^0, every sequence of lazy improvement in g is countable.*

Corollary 28. *Let g be a game with finite branching and finitely many players who have win-lose objectives. If every winning set is Δ_2^0, every sequence of lazy improvement in g is countable and ends at a Nash equilibrium.*

Proof. By Theorem 27 and Observation 24, since finite branching implies compactness.   □

Regarding a potential extension of Corollary 28 to winning sets beyond Δ_2^0 we shall make a tangential remark: The computational task of finding a Nash equilibrium in a two-player sequential game with Δ_2^0 winning sets is just as hard as iterating the task of finding an accumulation point of a sequence over some countable ordinal. This follows from results in [18, 17, 21, 4]. Finding a Nash equilibrium of a game with Σ_2^0 winning sets is strictly more complicated. Thus, Δ_2^0 seems to be a natural boundary for results of the form of Corollary 28.

We shall now proceed to discuss the fourth approach, based on Lipschitz payoff functions.

Definition 29 (Fair improvement[3])**.** *An infinite sequence of improvements is fair if the following holds: for all positive real numbers r, if improvements by more than r are possible infinitely often during the sequence, they also occur infinitely often.*

---

[3]This is *fair* as in *fair scheduler*, not as in *fair division of cake*.

Proposition 30. *If the game tree is binary and if for each player there exists $\eta > 2$ such that her payoff function is Lipschitz-continuous for the distance d defined by $d(h0\rho, h1\rho') = \frac{1}{\eta^{|h|}}$, then all the accumulation points of a fair lazy improvement sequence are Nash equilibria.*

Proof. To all strategy profiles s and all players a let us associate a real number:

$$M_a(s) := \sum_{h \in d^{-1}(a)} v_a\big(h \cdot (1 - s(h)) \cdot \rho(h \cdot (1 - s(h)), s)\big) - \min_{\rho \in \{0,1\}^\omega} (v_a(h\rho))$$

where $d(h)$ is the player that plays at history h, and $v_a(\rho)$ is the payoff for player a and run ρ, and $s(h)$ is the choice in $\{0,1\}$ that is prescribed by s at h, and $\rho(h,s)$ is the run induced by strategy profile s from h on. Similarly to the finite case $v_a(h \cdot (1 - s(h)) \cdot \rho(h \cdot (1 - s(h)), s))$ is the payoff that is avoided by a at history h. Note that the summands of $M_a(s)$ are all non-negative by definition of the minimum, and that the sum converges absolutely: indeed, by assumption $|v_a(h0\rho) - v_a(h1\rho')| \leq \frac{L_a}{\eta^{|h|}}$ for some $L_a > 0$ and for all h, ρ, and ρ', so $M_a(s) \leq \sum_{h \in \{0,1\}^*} \frac{L_a}{\eta^{|h|}} = L_a \sum_{l=0}^{+\infty} (\frac{2}{\eta})^l = \frac{L_a}{1 - \frac{2}{\eta}}$. Also, each M_a is continuous.

Similarly to the finite case, M_a decreases by x when player a performs a lazy improvement by x, and M_a is left unchanged when another player performs a lazy convertibility step. Let $(s_n)_{n \in \mathbb{N}}$ be a fair lazy improvements sequence, and let s be some accumulation point that is not an Nash equilibrium. So $s \rightharpoonup_a t$ for some profile t and player a. By continuity of the payoffs we can assume wlog that s and t only disagree on a finite set H_0 of histories shorter than some k. Let $r := v_a(t) - v_a(s) > 0$ and $\delta > 0$ be such that $|v_a(\sigma) - v_a(\sigma')| < \frac{r}{4}$ whenever $d'(\sigma, \sigma') < \delta$ with $d'(\sigma, \sigma') := \frac{1}{2^i}$ where i is maximal such that σ and σ' coincide for histories not longer than i. By the Bolzano-Weierstrass theorem there is a subsequence $(s_{\varphi(n)})$ converging towards s, and $d'(s, s_{\varphi(n)}) < \min(\delta, \frac{1}{2^k})$ for all but finitely many n. When the inequality holds, let t_n coincide with $s_{\varphi(n)}$ outside of H_0 and with t on H_0. So $s_{\varphi(n)} \xrightarrow{c}_a t_n$ and $d'(t, t_n) = d'(s, s_{\varphi(n)})$, so $|v_a(t_n) - v_a(t)| < \frac{r}{4}$, and $\frac{r}{2} < v_a(t_n) - v_a(s_{\varphi(n)})$. Thus improvement by at least $\frac{r}{2}$ are infinitely often possible for player a in the sequence. By fairness infinitely many of them must occur, so M_a diverges to $-\infty$, contradiction. $\qquad\square$

8   Some counterexamples

In order to obtain the termination result in the finite case (Theorem 10), some restriction on how players can improve is indeed necessary. We shall show below that the collective improvement $\twoheadrightarrow$ may fail to terminate even for very simple games in extensive form:

Example 31. *An improvement cycle:*
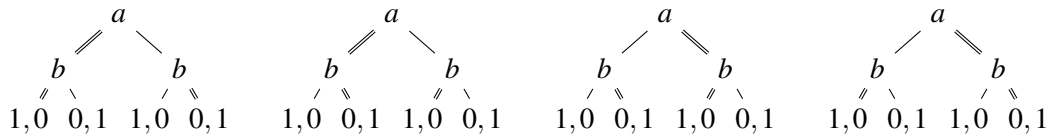
The technical notion of strategy that is used in this article to represent the intuitive concept of a strategy (in games in extensive form) is not the only possible notion. An alternative notion does not require choices from a player at every node that she owns, but only at nodes that are not ruled out by the strategy of the same player. The three objects in Example 32 are such minimalist, alternative strategy profiles, where double lines still represent choices. Up to symmetry, they constitute from left to right a cycle of improvements that could be intuitively described as lazy, so an actual cycle of length eight can easily be inferred from the short pseudo cycle. This may happen because, although the improvements

may look lazy, player *a* forgets about her choices in a subgame (of the root) when leaving it, and may settle for different choices when coming back to the subgame. This suggests that even counter-factual choices are sometimes relevant. In particular, this means that lazy improvement is not a *natural* dynamics in the sense of HART [8]; or a *simple* model in the sense of ROTH and EREV [22].

**Example 32.** *Let W be winning for player a and L be losing; and vice versa for player b.*
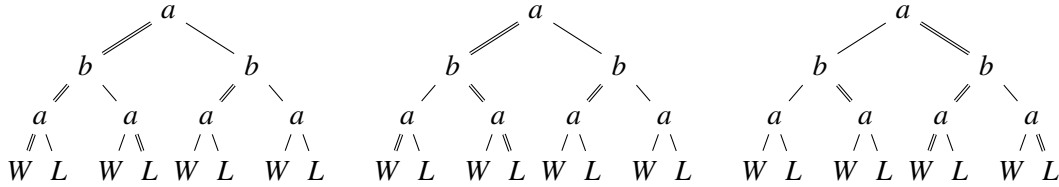
The example below shows that for infinite games, a sequence of lazy improvement steps may have multiple accumulation points even for continuous payoff functions; and moreover, that not all accumulation points have to be Nash equilibria.

Example 33 ([16, Example 26])**.**

Let us consider games with four players *a*, *b*, *c*, and *d*. Given four real-valued sequences $\mathscr{A} = (\alpha_n)_{n\in\mathbb{N}}$, $\mathscr{B} = (\beta_n)_{n\in\mathbb{N}}$, $\mathscr{C} = (\gamma_n)_{n\in\mathbb{N}}$, and $\mathscr{D} = (\delta_n)_{n\in\mathbb{N}}$ converging towards α, β, γ, and δ, let $T(\mathscr{A},\mathscr{B},\mathscr{C},\mathscr{D})$ be the following game and strategy profile. Note that apart from the payoffs, the underlying game effectively involves players *c* and *d* only. If $\mathscr{C}$ and $\mathscr{D}$ are increasing, the lazy improvement dynamics sees players *c* and *d* alternating in switching their top left-move to a right-move.
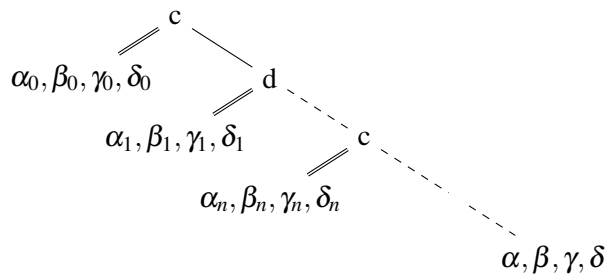
Let $\mathscr{A} := \mathscr{B} := (1 + \frac{1}{n+1})_{n\in\mathbb{N}}$ and let $\mathscr{C} := \mathscr{D} := (1 - \frac{1}{n+1})_{n\in\mathbb{N}}$. Starting from the profile below, players *c* and *d* will continue to unravel the subgame currently chosen jointly by *a* and *b*. Player *b* will keep alternating her choices to pick the least-unraveled subgame available to her. Player *a* will prefer to chose a subgame where player *b* currently chooses right, and also prefers less-unraveled subgames.
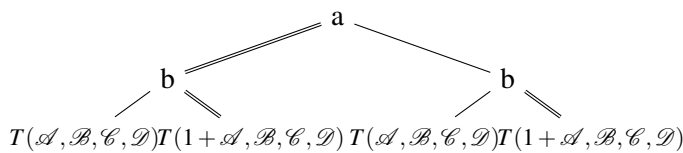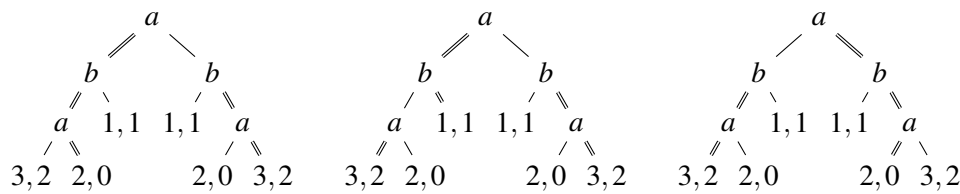
First of all, already the subgame where *b* moves first demonstrates that the lazy improvement dynamics will not always converge, hence we have to consider accumulation points rather than limit points. For the next feature, note that there is an infinite sequence of lazy improvement where players *a* and *b* (at both nodes that she owns) switch infinitely often, and where player *a* switches only when player *b* chooses the right subgame (on the induced play). Then the following strategy profile is an accumulation point, but it is clearly not a Nash equilibrium.
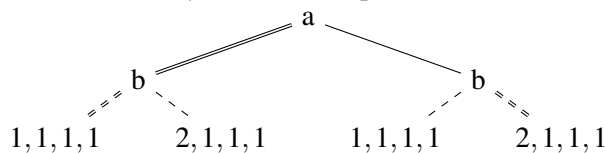
In our current model the players perform lazy improvement updates in a sequential manner. If simultaneity were allowed (yet not compulsory), cycles could occur, as shown in the example below.

Example 34. *It is a cycle up to symmetry only, a proper cycle of length* 4 *may be easily derived from it.*

```
        a                        a                        a
      ⫽   ╲                    ⫽   ╲                    ╱   ⫽
    b       b                b       b                b       b
  ⫽ ╲     ╱ ⫽              ╱ ⫽     ╱ ⫽              ⫽ ╲     ╱ ╲
 a  1,1  1,1  a           a  1,1  1,1  a           a  1,1  1,1  a
╱ ⫽        ╱ ⫽           ⫽ ╲        ╱ ⫽           ⫽ ╲        ╱ ╲
3,2 2,0   2,0 3,2       3,2 2,0   2,0 3,2       3,2 2,0   2,0 3,2
```

This behaviour can be avoided by considering lazy best-response dynamics, rather than merely lazy better-response. In the sequential case, clearly the termination of the latter implies termination of the former. In the simultaneous case, we have:

Proposition 35. *The synchronous lazy best-response sequences in a game with n internal nodes have length at most 2^n, provided that the players have acyclic preferences.*

Proof. It suffices to prove the claim for preferences that are linear orders, which we prove by induction on the number of internal nodes of the game g. (It holds for zero.) Let v be an internal node in g whose children are all leaves, let a be the owner of v, and let us consider a sequence where a always chooses the same outcome x at v. Let g' be the game derived from g by replacing v with a leaf enclosing the outcome x. The synchronous lazy best-response sequence in g corresponds, by restriction of the profiles, to a sequence in g', so it has length at most 2^{n-1} by I.H. Now let us consider an arbitrary sequence, and note that a can change choices only once at v, from some non-preferred outcome to her preferred one (among the outcomes occurring below v). So the length of a sequence in g is at most $2^{n-1} + 2^{n-1} = 2^n$. □

Acknowledgements

References

[1] Robert J. Aumann (1995): *Backward induction and common knowledge of rationality*. Games and Economic Behavior 8(1), pp. 6 – 19, doi:10.1016/S0899-8256(05)80015-6.

[2] Robert J. Aumann & Adam Brandenburger (1995): *Epistemic Conditions for Nash Equilibrium*. Econometrica 63(5), pp. 1161–1180, doi:10.2307/2171725.

[3] Endre Boros, Khaled Elbassioni, Vladimir Gurvich & Kazuhisa Makino (2012): *On Nash equilibria and improvement cycles in pure positional strategies for Chess-like and Backgammon-like n-person games*. Discrete Mathematics 312(4), pp. 772 – 788, doi:10.1016/j.disc.2011.11.011.

[4] Vasco Brattka, Guido Gherardi & Alberto Marcone (2012): *The Bolzano-Weierstrass Theorem is the Jump of Weak König's Lemma*. Annals of Pure and Applied Logic 163(6), pp. 623–625, doi:10.1016/j.apal.2011.10.006. Also arXiv:1101.0792.

[5]   R. Cressman & K.H. Schlag (1998): *The Dynamic (In)Stability of Backwards Induction*. Journal of Economic Theory 83(2), pp. 260 – 285, doi:10.1006/jeth.1996.2465.

[6]   Ross Cressman (2003): *Evolutionary Dynamics and Extensive Form Games*. MIT Press.

[7]   D. Gale & F.M. Stewart (1953): *Infinite games with perfect information*. In: *Contributions to the theory of games, Annals of Mathematical Studies* 28, Princeton University Press, pp. 245–266, doi:10.1515/9781400881970-014.

[8]   Sergiu Hart (2008): *Dynamics and Equilibria*. GAMES 2008.

[9]   Sergiu Hart & Andreu Mas-Colell (2000): *A simple adaptive procedure leading to correlated equilibrium*. Econometrica 68, pp. 1127–1150, doi:10.1111/1468-0262.00153.

[10] Nikolai S. Kukushkin (2002): *Perfect Information and Potential Games*. Games and Economic Behavior 38(2), pp. 306 – 317, doi:10.1006/game.2001.0859.

[11] Nikolai S. Kukushkin (2011): *Acyclicity of improvements in finite game forms*. International Journal of Game Theory 40(1), pp. 147–177, doi:10.1007/s00182-010-0231-0.

[12] Stéphane Le Roux (2008): *Generalisation and formalisation in game theory*. Ph.D. thesis, Ecole Normale Supérieure de Lyon.

[13] Stéphane Le Roux (2009): *Acyclic preferences and existence of sequential Nash equilibria: a formal and constructive equivalence*. In: *TPHOLs, International Conference on Theorem Proving in Higher Order Logics*, Lecture Notes in Computer Science, Springer, pp. 293–309, doi:10.1007/978-3-642-03359-9_21.

[14] Stéphane Le Roux (2013): *Infinite Sequential Nash Equilibria*. Logical Methods in Computer Science 9(2), doi:10.2168/LMCS-9(2:3)2013.

[15] Stéphane Le Roux & Arno Pauly (2013): *A semi-potential for finite and infinite sequential games*. arXiv:1309.2798. Available at `http://arxiv.org/abs/1309.2798`.

[16] Stéphane Le Roux & Arno Pauly (2014): *Infinite Sequential Games with Real-valued Payoffs*. In: *CSL-LICS '14*, ACM, pp. 62:1–62:10, doi:10.1145/2603088.2603120.

[17] Stéphane Le Roux & Arno Pauly (2014): *Weihrauch degrees of finding equilibria in sequential games*. arXiv:1407.5587. Available at `http://arxiv.org/abs/1407.5587`.

[18] Stéphane Le Roux & Arno Pauly (2015): *Weihrauch Degrees of Finding Equilibria in Sequential Games*. In Arnold Beckmann, Victor Mitrana & Mariya Soskova, editors: *Evolving Computability, Lecture Notes in Computer Science* 9136, Springer, pp. 246–257, doi:10.1007/978-3-319-20028-6_25.

[19] Donald A. Martin (1975): *Borel Determinacy*. Annals of Mathematics 102(2), pp. pp. 363–371, doi:10.2307/1971035.

[20] D. Monderer & L.S. Shapley (1996): *Potential games*. Games and Economic Behavior 14(124), pp. 124–143, doi:10.1006/game.1996.0044.

[21] Arno Pauly (2015): *Computability on the countable ordinals and the Hausdorff-Kuratowski theorem*. arXiv 1501.00386. Available at `http://arxiv.org/abs/1501.00386`.

[22] Alvin E. Roth & Ido Erev (1995): *Learning in extensive form games: Experimental data and simple dynamic models in the intermediate term*. Games and Economic Behaviour 8, pp. 164–212, doi:10.1016/S0899-8256(05)80020-X.

[23] Eilon Solan & Nicolas Vieille (2003): *Deterministic multi-player Dynkin games*. Journal of Mathematical Economics 39(8), pp. 911 – 929, doi:10.1016/S0304-4068(03)00021-1.

[24] Zibo Xu (2013): *Convergence of best-response dynamics in extensive-form games*. SSE/EFI Working Paper Series in Economics and Finance 745, Stockholm School of Economics.