



**HAL**  
open science

# Mixed Nash Equilibria in Concurrent Terminal-Reward Games

Patricia Bouyer, Nicolas Markey, Daniel Stan

► **To cite this version:**

Patricia Bouyer, Nicolas Markey, Daniel Stan. Mixed Nash Equilibria in Concurrent Terminal-Reward Games. FSTTCS 2014, Dec 2014, Delhi, India. 10.4230/LIPIcs.FSTTCS.2014.351 . hal-01105081

**HAL Id: hal-01105081**

**<https://hal.science/hal-01105081>**

Submitted on 19 Jan 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Mixed Nash Equilibria in Concurrent Terminal-Reward Games

Patricia Bouyer, Nicolas Markey, and Daniel Stan

LSV – CNRS & ENS Cachan – France

---

## Abstract

We study mixed-strategy Nash equilibria in multiplayer deterministic concurrent games played on graphs, with terminal-reward payoffs (that is, absorbing states with a value for each player). We show undecidability of the existence of a constrained Nash equilibrium (the constraint requiring that one player should have maximal payoff), with only three players and 0/1-rewards (*i.e.*, reachability objectives). This has to be compared with the undecidability result by Ummels and Wojtczak for turn-based games which requires 14 players and general rewards. Our proof has various interesting consequences: (i) the undecidability of the existence of a Nash equilibrium with a constraint on the social welfare; (ii) the undecidability of the existence of an (unconstrained) Nash equilibrium in concurrent games with terminal-reward payoffs.

**1998 ACM Subject Classification** F.3.1 (Specifying and Verifying and Reasoning about Programs); D.2.4 (Software/Program Verification); G.3 (Probability and statistics)

**Keywords and phrases** concurrent games, randomized strategy, Nash equilibria, undecidability

**Digital Object Identifier** 10.4230/LIPIcs.xxx.yyy.p

## 1 Introduction

Games (especially games played on graphs) have been intensively used in computer science as a powerful way of modelling interactions between several computerised systems [11, 8]. Until recently, more focus had been put on the study of purely antagonistic games (a.k.a. zero-sum games, where the aim of one player is to prevent the other player from achieving her objective), which conveniently model systems evolving in a (hostile) environment.

Over the last ten years, games with non-zero-sum objectives have come into the picture: they allow for conveniently modelling complex infrastructures where each individual system tries to fulfill its own objectives, while still being subject to actions of the surrounding systems. As an example, consider (a simplified version of) the team formation problem [4], an example of which is presented in Fig. 1: several agents are trying to achieve tasks; each task requires some resources, which are shared by the players. Achieving a task thus requires the formation of a team that have all required resources for that task: each player selects the task she wants to achieve (and so proposes her resources for achieving that task), and if a task receives enough resources, the associated team receives the corresponding payoff (to be divided among the players in the team). In such a game, there is a need of cooperation (to gather enough resources), and an incentive to selfishness (to maximise the payoff).

In that setting, focusing only on optimal strategies for one single agent is not relevant. In game theory, several solution concepts have been defined, which more accurately represents rational behaviours of these multi-player systems; Nash equilibrium [9] is the most prominent such concept: a Nash equilibrium is a strategy profile (that is, one strategy to each player) where no player can improve her own payoff by unilaterally changing her strategy. In other terms, in a Nash equilibrium, each individual player has a satisfactory strategy with regards to the other players' strategies. Notice that Nash equilibria need not exist or be



© Patricia Bouyer, Nicolas Markey, Daniel Stan;  
licensed under Creative Commons License CC-BY

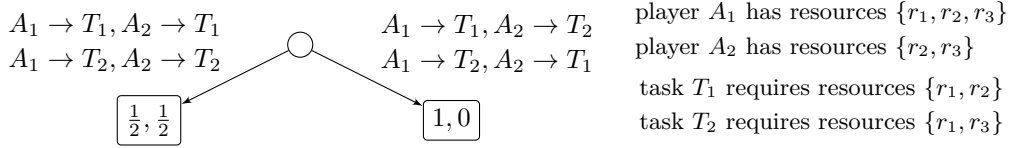
Conference title on which this volume is based on.

Editors: Billy Editor and Bill Editors; pp. 1–13



Leibniz International Proceedings in Informatics

Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany



■ **Figure 1** An instance of the team-formation problem. For any deterministic choice of actions, one of the players has an incentive to change her choice: there is no pure Nash equilibrium. However there is one mixed Nash equilibrium, where each player plays  $T_1$  and  $T_2$  uniformly at random.

unique, and are not necessarily optimal: Nash equilibria where all players lose may coexist with more interesting Nash equilibria. Therefore, looking for constrained Nash equilibria (*e.g.* equilibria in which some players are required to win, or equilibria with maximal social welfare) is an interesting and important problem to study, which has been suggested both in the game-theory community [5] and in the computer-science community [12].

In this paper, we study (deterministic) concurrent games played on graphs. Such games are indeed a general and relevant model for interactive systems, where the agents take their decision simultaneously (which is the case for instance in distributed systems). Concurrent games subsume turn-based games, where in each state, only one player has the decision for the next move, and which have attracted more focus until now in the computer science community. Notice also that in game theory, models are almost exclusively based on concurrent actions (*e.g.* games in normal form given as matrices indicating the payoff of each player for each concurrent choice of actions, and extensions thereof, such as repeated games).

In this paper we are interested in randomized (a.k.a. mixed) strategies for the players. A mixed strategy consists in choosing, at each step of the game, a probability distribution over the set of available actions; the game then proceeds following the product distribution of the strategies of all players. Strategies may depend on the history of the game, *i.e.*, the sequence of visited states, but do not require to see the actions played by the other agents. In previous works, the first two authors have focused on pure strategies, where at each step, each player proposes exactly one action, and developed algorithms for deciding the existence of constrained Nash equilibria in various settings [1]. In the present paper, we focus on terminal-reward payoffs (where some designated states are absorbing, and each player has a value—or reward—attached to each of these states): the payoff of a player is then her expected reward. We will also consider the subclass of games with terminal-reachability objectives, where the reward in each absorbing state is either 0 or 1 (hence the expected reward for a player is the probability to reach her winning states). The game in Fig. 1 has terminal-reward payoffs: they are given by the values labelling the two absorbing states (1 for player  $A_1$  and 0 for player  $A_2$  in the right-most state). This game can be shown to have no pure-strategy Nash equilibria, but it has a mixed-strategy one.

**Our results.** Our main result is the undecidability of the existence of a 0-optimal Nash equilibrium in concurrent games with terminal-reachability payoff functions, with only three players and strategies insensitive to actions. A 0-optimal Nash equilibrium is a Nash equilibrium in which one designated player is required to have maximal payoff (that is, 1 in the case of terminal-reachability payoffs). A corollary of our result is the undecidability of the existence of unconstrained Nash equilibria in concurrent games with terminal payoffs. We believe that these results are important, as they solve natural questions for basic objectives. Moreover, our constructions give new insight in the understanding of concurrent games and their algorithmics, and contain several intermediary tools that can be interesting on their own in different contexts.

Several results already exist in related settings:

- our result should first be compared with the undecidability of the existence of a 0-optimal Nash equilibrium in turn-based games with terminal-reward payoffs [13], which requires 14 players and general rewards. It should be noticed that this result requires more than 0/1 rewards (contrary to our result), since the existence of a 0-optimal equilibrium can be decided in polynomial time in turn-based games with terminal-reachability payoffs (by combining the reduction to pure 0-optimal Nash equilibria of [12] and the algorithm in [13] for computing such equilibria);
- our result should also be compared with polynomial-time algorithm for deciding the existence of a 0-optimal *pure* Nash equilibrium in concurrent terminal-reward games [15];
- our result has several corollaries, that we develop at the end of the paper:
  - the existence of a (unconstrained) Nash equilibrium in terminal-reward games with three players; on the opposite, stationary  $\epsilon$ -Nash equilibria do always exist in concurrent games for terminal-reachability (and terminal-reward) games [3];
  - the existence of a Nash equilibrium that maximizes the social welfare in games with terminal-reachability payoffs is undecidable with three players. This should be compared with the NP-completeness of the existence of such equilibria for two-player normal-form games [6];
  - the existence of a constrained finite-memory Nash equilibrium in terminal-reachability games is undecidable with three players;
  - the existence of a constrained Nash equilibrium in safety games is undecidable with three players. This can be compared to the result of [10], which states that there always exists a Nash equilibrium (with little memory) in a safety game.

By lack of space, only some sketches of proofs could be included in this paper. We refer the interested reader to [2] for full details.

## 2 Definitions

► **Definition 1.** A concurrent arena  $\mathcal{A}$  is a tuple  $\mathcal{A} = \langle \text{States}, \text{Agt}, \text{Act}, \text{Tab}, (\text{Allow}_i)_{i \in \text{Agt}} \rangle$  where **States** is a finite set of states; **Agt** is a finite set of players; **Act** is a finite set of actions; For all  $i \in \text{Agt}$ ,  $\text{Allow}_i: \text{States} \rightarrow 2^{\text{Act}} \setminus \{\emptyset\}$  is a function describing authorized actions in a given state for Player  $i$ ;  $\text{Tab}: \text{States} \times \text{Act}^{\text{Agt}} \rightarrow \text{States}$  is the transition function.

A state  $s \in \text{States}$  is said *terminal* (or *final*) if  $\text{Tab}(s, \cdot) \equiv s$ . We write  $F_{\mathcal{A}}$  (or simply  $F$  when the underlying arena is clear from the context) for the set of terminal states of  $\mathcal{A}$ .

A *history* of such an arena  $\mathcal{A}$  is a finite, non-empty word  $h \in \text{States}^+$ . We denote by  $\text{first}(h)$  and  $\text{last}(h)$  respectively the first and last states of the word  $h$ . During a play, players in **Agt** choose their next moves concurrently and independently from each others, according to the current history  $h$  and what they are allowed to do in the current state  $\text{last}(h)$ .

► **Definition 2.** A *strategy* for Player  $i$  is a function  $\sigma_i: \text{States}^+ \rightarrow \text{Dist}(\text{Act})$  with the requirement that  $\sigma_i(h)^{-1}(\mathbb{R}_{>0}) \subseteq \text{Allow}_i(\text{last}(h))$  for all history  $h$ .

Let  $\alpha \in \text{Act}$ . We write  $\sigma_i(\alpha \mid h)$  for the probability mass  $\sigma_i(h)(\alpha)$  of action  $\alpha$  in the distribution  $\sigma_i(h)$ . In the sequel, we sometimes write  $\sigma_i(h) = \alpha$  when  $\sigma_i(\alpha \mid h) = 1$ . When  $\sigma_i(h) \in \text{Act}$  for all  $h$ , the strategy  $\sigma_i$  for player  $i$  is said to be *pure*. Otherwise it is said to be *mixed*. We denote by  $S_i$  (resp.  $\mathbb{S}_i$ ) the set of pure (resp. mixed) strategies of Player  $i$ . A *strategy profile*  $\sigma$  is a mapping assigning one strategy to each player. We write  $\mathbb{S}$  for the set of all strategy profiles, and for  $\sigma \in \mathbb{S}$ , we will write  $\sigma_i$  in place of  $\sigma(i)$  for the strategy of Player  $i$ .

► **Remark.** While strategies are aware of the sequence of actions played in a turn-based game, we can notice this is generally not the case in the concurrent setting depicted here, since strategies only depend on the sequence of visited states. This is realistic when considering multi-agent systems, where only the global effect of the actions of the players is assumed to be observable. However this partial-information hypothesis makes the detection of strategy deviations (and therefore the computation of Nash equilibria) harder.

Consider a strategy profile  $\sigma \in \mathbb{S}$  and an initial state  $s_0$ . For any history  $h \in \mathbf{States}^+$  and any player  $i \in \mathbf{Agt}$ , we construct the random variable  $\alpha_i(h) \in \mathbf{Act}$  with distribution  $\sigma_i(h)$  such that  $(\alpha_i(h))_{i \in \mathbf{Agt}, h \in \mathbf{States}^+}$  is a family of independent random variables.

We define the stochastic process  $(X_n)_{n \in \mathbb{N}}$  inductively by  $X_0 = s_0$  and for every  $n$ ,  $X_{n+1} = X_n \cdot \text{Tab}(\text{last}(X_n), (\alpha_i(X_n))_i)$ . For each  $n$ , the random variable  $X_n$  takes value in  $\mathbf{States}^{n+1}$ :  $(X_n)_n$  is an increasing sequence of prefixes whose limit is an infinite random run  $X_\infty \in \mathbf{States}^\omega$ .

We now consider the standard Borel  $\sigma$ -algebra over  $\mathbf{States}^\omega$  from  $s_0$ , and define the probability measure  $\mathbb{P}^\sigma$  as the probability distribution induced by  $X_\infty$ , that is, if  $B$  is a Borel subset of  $\mathbf{States}^\omega$ ,  $\mathbb{P}^\sigma(B) = \mathbb{P}(X_\infty \in B)$ . It coincides with the standard construction based on cylinders. In the following, to make explicit the initial state, we may write  $\mathbb{P}^\sigma(B \mid s_0)$  instead of simply  $\mathbb{P}^\sigma(B)$ . In the sequel, we sometimes also abusively write  $h$  for the cylinder  $h \cdot \mathbf{States}^\omega$ : then, when we write  $\mathbb{P}^\sigma(h \mid s_0)$ , we mean  $\mathbb{P}^\sigma(X_{|h|} = h)$ . If  $\mathbb{P}^\sigma(h \mid s_0) > 0$ , we say that  $\sigma$  enables  $h$  from  $s_0$ : in that case we can define the conditional probability  $\mathbb{P}^\sigma(B \mid h) = \mathbb{P}^\sigma(B \mid X_{|h|} = h)$ .

Finally we say that a node  $n$  is activated by a strategy profile whenever it is visited with positive probability under that profile.

► **Definition 3.** A *terminal-reward game*  $\mathcal{G} = \langle \mathcal{A}, s, (\phi_i)_{i \in \mathbf{Agt}} \rangle$  is given by an arena  $\mathcal{A}$ , an initial state  $s$ , and for every player  $i \in \mathbf{Agt}$ , a real-valued function  $\phi_i$  ranging over terminal states of  $\mathcal{A}$ . In the following, we extend  $\phi_i$  to every  $r \in \mathbf{States}^\omega$ , by  $\phi_i(r) = \phi_i(s)$  if  $r$  is an infinite path ending in a state  $s \in \mathbf{F}$ , and  $\phi_i(r) = 0$  otherwise.

The game  $\mathcal{G}$  will be said a *terminal-reachability game* whenever each function  $\phi_i$  only takes values 0 or 1.

► **Remark.** In the sequel, we represent terminal-reward games as graphs with circle states representing non-terminal states, and rectangle states representing terminal states, decorated with the associated rewards for all players. The self-loop on terminal states will be omitted. The transition table of the underlying arena is encoded by decorating the transitions with the move vectors that trigger it. Move vectors are written as words over  $\mathbf{Act}$ , by identifying  $\mathbf{Agt}$  with the subset  $\llbracket 0, |\mathbf{Agt}| - 1 \rrbracket$ . We will use  $\cdot$  as a special symbol representing any action. Also, for a set  $S$  of words in  $(\mathbf{Act} \cup \{\cdot\})^k$ , with  $k < |\mathbf{Agt}|$ , and for a letter  $a \in \mathbf{Act} \cup \{\cdot\}$ , we write  $aS$  for the words  $\{aw \mid w \in S\}$ . See Fig. 2 (and the subsequent figures) for an example.

Consider a terminal-reward game  $\mathcal{G}$ , a strategy profile  $\sigma$ , and an enabled history  $h$ . One can easily check that  $\phi_i$  is a measurable function under  $\mathbb{P}^\sigma$ . The expected payoff of Player  $i$  under  $\sigma$  after  $h$  is defined as

$$\mathbb{E}^\sigma(\phi_i \mid h) = \sum_{x \in \text{Im}(\phi_i)} x \cdot \mathbb{P}^\sigma(\phi_i^{-1}(\{x\}) \mid h).$$

In case  $\mathcal{G}$  is a terminal-reachability game, the expected payoff of Player  $i$  is the probability of reaching terminal states with value 1 under  $\phi_i$ .

Let  $\mathcal{G}$  be a terminal-reward game. Let  $\sigma \in \mathbb{S}$  be a (mixed) strategy profile in  $\mathcal{G}$ , and  $h$  be a history. A *single-player deviation* (simply called *deviation* hereafter, as we only consider

deviations of a single player at a time) of  $\sigma$  for Player  $i$  after history  $h$  is another strategy profile  $\sigma'$  for which there exists  $\sigma''_i \in \mathbb{S}_i$  satisfying

$$\forall h' \in \text{States}^+. \begin{cases} h \sqsubseteq h' \Rightarrow \sigma'_i(h') = \sigma''_i(h') \\ \forall j \in \text{Agt}. (h \not\sqsubseteq h' \vee j \neq i) \Rightarrow \sigma'_j(h') = \sigma_j(h') \end{cases}$$

where  $\sqsubseteq$  is the prefix relation. We then write  $\sigma' = \sigma[i/\sigma''_i]^h$ . The deviation  $\sigma[i/\sigma''_i]^h$  is said *deterministic* if  $\sigma''_i$  is.

► **Definition 4.** Let  $\mathcal{G}$  be a terminal-reward game. A strategy profile  $\sigma$  forms a *Nash equilibrium* after a history  $h$  when the following conditions are met:

- $h \in \text{States}^+$  is enabled by  $\sigma$  from  $\text{first}(h)$ ;
- No player has a profitable deviation; in other terms, for all  $i \in \text{Agt}$  and for all  $\sigma'_i \in \mathbb{S}_i$ , it holds  $\mathbb{E}^{\sigma[i/\sigma'_i]^h}(\phi_i | h) \leq \mathbb{E}^\sigma(\phi_i | h)$ .

We then write that  $\langle \sigma, h \rangle$  is a Nash equilibrium.

A Nash equilibrium  $\langle \sigma, h \rangle$  is said *0-optimal* whenever the expected payoff of Player 0 is optimal, that is,  $\mathbb{E}^\sigma(\phi_0 | h) = \max(\text{Img}(\phi_0))$ . In case of a terminal-reachability game, it amounts to saying that the payoff of Player 0 is 1.

The following result will be useful all along the paper:

► **Lemma 5.** Let  $\mathcal{G}$  be a terminal-reward game, and  $\langle \sigma, h \rangle$  be a Nash equilibrium. If  $\langle \sigma, h \rangle$  enables  $h'$ , then  $\langle \sigma, h' \rangle$  is a Nash equilibrium.

In general, several Nash equilibria may coexist. It is therefore very relevant to look for *constrained Nash equilibria*, that is, Nash equilibria that satisfy a constraint on the expected payoff. In this paper, we only consider 0-optimality as the constraint, and we prove that the existence of a 0-optimal Nash equilibrium in a three-player terminal-reachability game is undecidable. To prove this result, we will first show undecidability in the case of terminal-reward games, and then extend the result to terminal-reachability games. Those results will have interesting corollaries, like the undecidability of the existence of a Nash equilibrium (with no constraint) in terminal-reward games, when the rewards are in  $\{-1, 0, 1\}$ , or the existence of a Nash equilibrium with optimal social welfare.

### 3 Tools

In this section, we develop several intermediary results that will be useful for our reduction. We first show that we can equivalently define Nash equilibria by considering only deterministic deviations (for non-negative terminal-reward games). We then study a few simple games and constructions which will be used in the encoding.

#### 3.1 Deterministic deviations

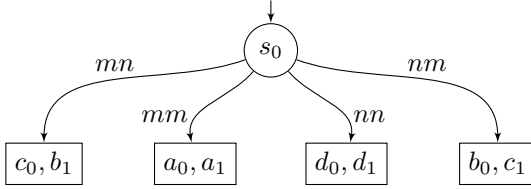
We explain in this section that it is enough to consider deterministic deviations in the characterization of a Nash equilibrium.

► **Proposition 6.** Let  $\mathcal{G}$  be a terminal-reward game with non-negative rewards. Pick a history  $h \in \text{States}^+$ , and a strategy profile  $\sigma$ . Then  $\langle \sigma, h \rangle$  is a Nash equilibrium if, and only if, for all  $i \in \text{Agt}$  and all *deterministic* deviation  $\sigma''_i \in \mathbb{S}_i$ , it holds  $\mathbb{E}^{\sigma[i/\sigma''_i]^h}(\phi_A | h) \leq \mathbb{E}^\sigma(\phi_i | h)$ .

► **Remark.** A similar result was proven in [15, Proposition 3.1] for turn-based games with qualitative Borel objectives (the payoff is 1 if the run belongs to the designed objective, and it is 0 otherwise).

### 3.2 One-stage games

We analyse two-player two-action one-stage games (that is, games that end up in a terminal state in one step), and obtain useful properties of their Nash equilibria. Such games can be represented by a graph as shown in Fig. 2a. Alternatively, these games, also known as one-shot games, can be represented as a matrix as in Table 2b (this is the standard representation in the game-theory community).



(a) A generic two-player two-action one-stage game

	0	m	n
1		m	n
m		$a_0, a_1$	$b_0, c_1$
n		$c_0, b_1$	$d_0, d_1$

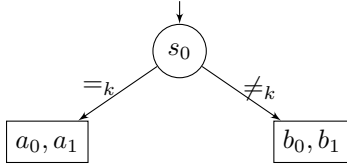
(b) Associated matrix representation

■ **Figure 2** Representations of a one-shot game

► **Lemma 7.** Consider the two-player two-action one-shot concurrent game  $\mathcal{G}$  of Fig. 2, and pick some strategy profile  $\sigma$ . If  $\langle \sigma, s_0 \rangle$  is a Nash equilibrium, then for every player  $i \in \{0, 1\}$ , it holds

$$\begin{cases} \sigma_i(m | s_0) < 1 & \Rightarrow [(d_i - c_i) + (a_i - b_i)] \cdot \sigma_{1-i}(m | s_0) \leq d_i - c_i \\ \sigma_i(m | s_0) > 0 & \Rightarrow [(d_i - c_i) + (a_i - b_i)] \cdot \sigma_{1-i}(m | s_0) \geq d_i - c_i \end{cases}$$

### 3.3 $k$ -action matching-pennies games



■ **Figure 3** Matching-pennies game

The classical matching-pennies games are a special case of one-stage games, where  $a_i = d_i$  and  $b_i = c_i$ : basically, there are two outcomes, depending on whether the players propose the same action or not. This game can be generalized to  $k$  ( $\geq 2$ ) actions, as depicted on Fig. 3. In this figure (and in the sequel),  $=_k$  (resp.  $\neq_k$ ) is a shorthand for pairs of identical (resp. different) actions taken from a set of  $k$  actions  $\Sigma_k = \{c_1, \dots, c_k\}$ . In other terms,  $=_k$  represents the set of words  $\{c_i c_i \mid 1 \leq i \leq k\}$ , and  $\neq_k$  is the complement in  $\Sigma_k^2$ .

► **Lemma 8.** In the  $k$ -action matching-pennies game, playing uniformly at random for both players defines a Nash equilibrium. Moreover, this is the unique Nash equilibrium of the game if, and only if, either  $a_0 < b_0$  and  $a_1 > b_1$ , or  $a_0 > b_0$  and  $a_1 < b_1$ . The payoff of this Nash equilibrium is  $(\frac{1}{k} \cdot a_0 + (1 - \frac{1}{k}) \cdot b_0, \frac{1}{k} \cdot a_1 + (1 - \frac{1}{k}) \cdot b_1)$ .

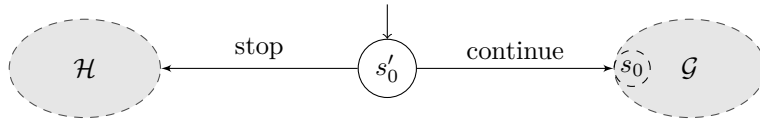
### 3.4 Games without equilibrium

In this section, we show that there are games that admit no Nash equilibria. We then explain how these games can be used to impose constraints on payoffs.

Consider the game *hide-or-run*, depicted in Fig. 4a. Player 0 can either hide (h) or run home (r), while Player 1 can either shoot him (s), or wait (w). If Player 1 shoots while Player 0 is hiding, she loses her bullet and loses the game. If Player 1 shoots when Player 0



■ **Figure 4** Hide-or-run games



■ **Figure 5** A game that has a Nash equilibrium if, and only if,  $\mathcal{G}$  has a 0-optimal Nash equilibrium

is running, she wins. This game has been shown to have no optimal almost-sure strategy [7], and we adapt the proof to show that it has no Nash equilibria.

► **Lemma 9.** *The game  $\mathcal{H}$  has no Nash equilibria.*

The payoff function of  $\mathcal{H}$  takes negative value. In order to only have nonnegative payoffs, we could shift the values by 1, which yields the game  $\mathcal{H}'$  depicted on Fig. 4b. But then one easily sees that the strategies  $\sigma_0(h \mid s_0^n) = 1$  and  $\sigma_1(s \mid s_0^n) = 1$  form a Nash equilibrium, contrary to a claim in [3, 13]. The difference is that when shifting the payoffs, we did not modify the payoff of the run that never reaches a terminal state: while this run was a positive deviation for Player 1 in  $\mathcal{H}$ , this is not the case in  $\mathcal{H}'$  anymore.

We now explain how we use the game  $\mathcal{H}$  to impose a 0-optimality constraint on the payoff. In the sequel, we restrict<sup>1</sup> to games where  $\max_{s \in F} \phi_0(s) = 1$ . Then:

► **Lemma 10.** *Let  $\mathcal{G}$  be a terminal-reward game. Then we can build a terminal-reward game  $\mathcal{G}'$  (see Fig. 5) such that  $\mathcal{G}$  has a 0-optimal Nash equilibrium if, and only if,  $\mathcal{G}'$  has a Nash equilibrium.*

This lemma will be useful for extending the undecidability result from the constrained existence to the existence problem (Corollary 14).

► **Remark.** Note that in the above construction, game  $\mathcal{H}$  can be replaced by any game with no Nash equilibria, such that Player 0 can secure a payoff  $1 - \varepsilon$  for every  $\varepsilon > 0$ . For instance, one could use a game with limit-average payoff and nonnegative rewards only [14].

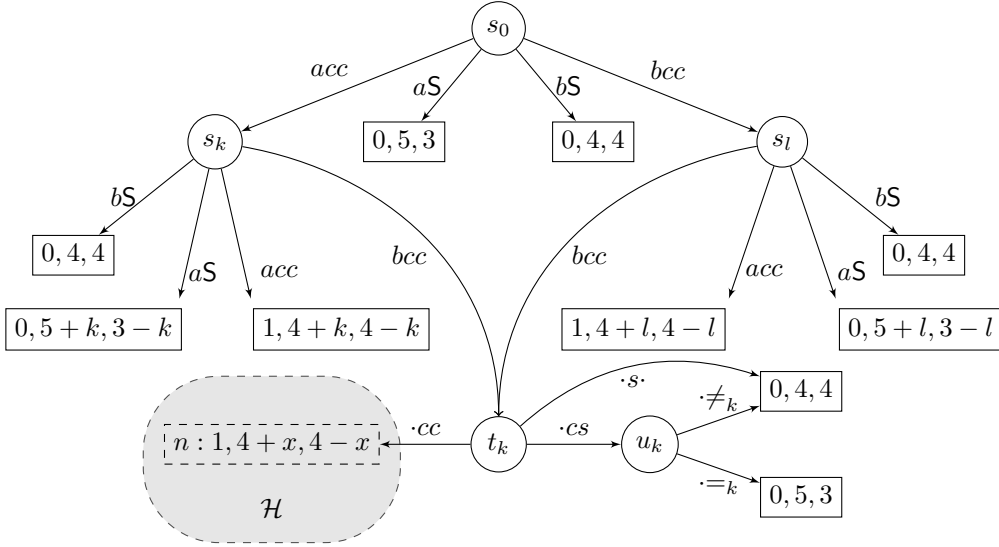
## 4 Updating values

Our undecidability proof will be based on an encoding of a two-counter machine. In this section and in the next one, we present games that will be building blocks for our proof.

Consider the game  $\mathcal{G}_k^r$  depicted on Fig. 6: in this game, Player 0 has two available actions  $a$  and  $b$  from  $s_0, s_k$  and  $s_l$ , while the other two players can either continue (action  $c$ ),

<sup>1</sup> This is no loss of generality, since we can linearly rescale payoffs of each player without changing profitable deviations.





■ **Figure 6** The rescale game  $\mathcal{G}_k^r$  for  $k \in \{1, 2, 3\}$  and  $l = k - 1$

or unilaterally decide to stop the game (action  $s$ ) and go to a terminal state (where Player 0 will have payoff 0). In node  $t_k$ , only players 1 and 2 have a choice: they can either continue to the game  $\mathcal{H}$  (when both of them play  $c$ ), or decide to stop and go to a  $k$ -action matching-pennies game (when one of them plays  $s$ ). In Fig. 6, we write  $S$  as a shorthand representing any combination of moves of players 1 and 2 where at least one of them decides to stop (action  $s$ ). Node  $n$  is the initial node of a game  $\mathcal{H}$  (which is unknown for the moment).

The interesting property of game  $\mathcal{G}_k^r$  is that we can relate 0-optimal Nash equilibria from  $s_0$  and those from  $n$ : (roughly) there is a Nash equilibrium from  $n$  of expected payoff  $(1, 4 + x, 4 - x)$  if, and only if, there is a Nash equilibrium from  $s_0$  of expected payoff  $(1, 4 + k \cdot x, 4 - k \cdot x)$ .

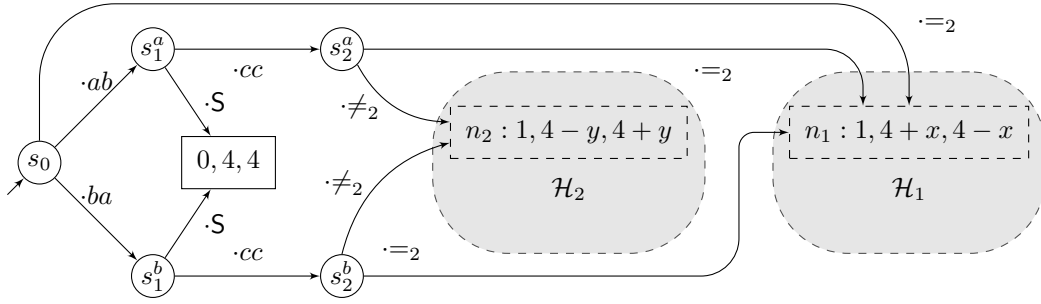
This is because, from  $s_0$  and  $s_k$ , there is a threat for Player 0 that one of the players 1 and 2 stops the game immediately, leading to a state with payoff 0 for her. Hence, Player 0 is forced to “collaborate” with players 1 and 2 and help them be satisfied with their payoffs, either by joining one of the interesting terminal states of  $\mathcal{G}_k^r$ , or in the next game  $\mathcal{H}$  after  $n$ . Some technical calculations show that Player 0 has to play  $a$  with probability  $k \cdot x$  at  $s_0$ , and with probability  $x/(x + 1)$  at  $s_k$  and  $s_l$ . The gadget to the right of  $t_k$  is just for ensuring that  $0 \leq k \cdot x \leq 1$  (this condition is required for having the above-mentioned equivalence between Nash equilibria from  $s_0$  and Nash equilibria from  $n$ ).

## 5 Comparing values

### 5.1 Testing game

We present in this section the construction of a game for comparing the expected payoffs in different nodes. This will be useful in our reduction to encode the zero-tests of our two-counter machine.

Consider the game  $\mathcal{G}_t$  depicted on Fig. 7. This game has the very interesting property that if we assume there are 0-optimal Nash equilibria from  $n_1$  and  $n_2$  of respective payoffs  $(1, 4 + x, 4 - x)$  and  $(1, 4 - y, 4 + y)$ , then there is a 0-optimal Nash equilibrium from  $s_0$  if, and only if,  $x = y$ , and the payoff is then  $(1, 4 + x/2, 4 - x/2)$ . Indeed, unless  $x = 0$  or  $y = 0$ ,



■ **Figure 7** The testing game  $\mathcal{G}_t$

it should be the case that a 0-optimal Nash equilibrium activates all states  $s_j^\alpha$  in the game, and then, as players 1 and 2 have zero-sum objectives, the best way is to play uniformly at random in all states where this makes sense (when actions  $a$  and  $b$  are available), and to play deterministically action  $c$  in all states where  $c$  is available.

This gadget allows, by plugging in  $n_2$  a game with known payoffs (the games on the next subsection), to check that the payoff at  $s_0$  has some particular value (which depends on that after  $n_1$ ).

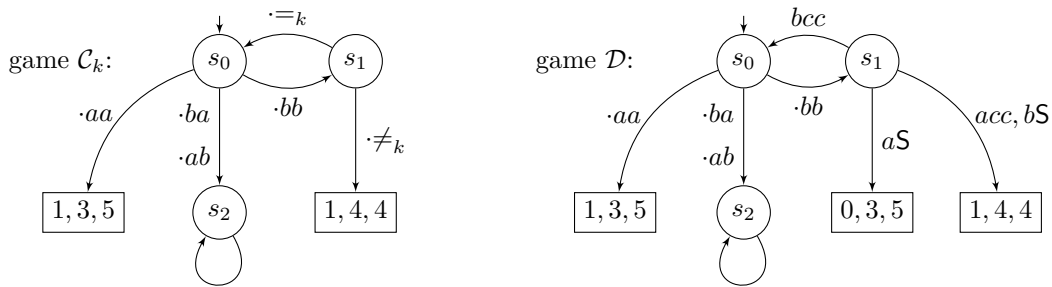
## 5.2 Counting modules

We now present games that generate a family of Nash equilibria with a particular expected payoffs. These modules will later be plugged at node  $n_2$  of game  $\mathcal{G}_t$ , and will ensure that the payoff of an Nash equilibrium in  $\mathcal{G}_t$  will have a predefined form.

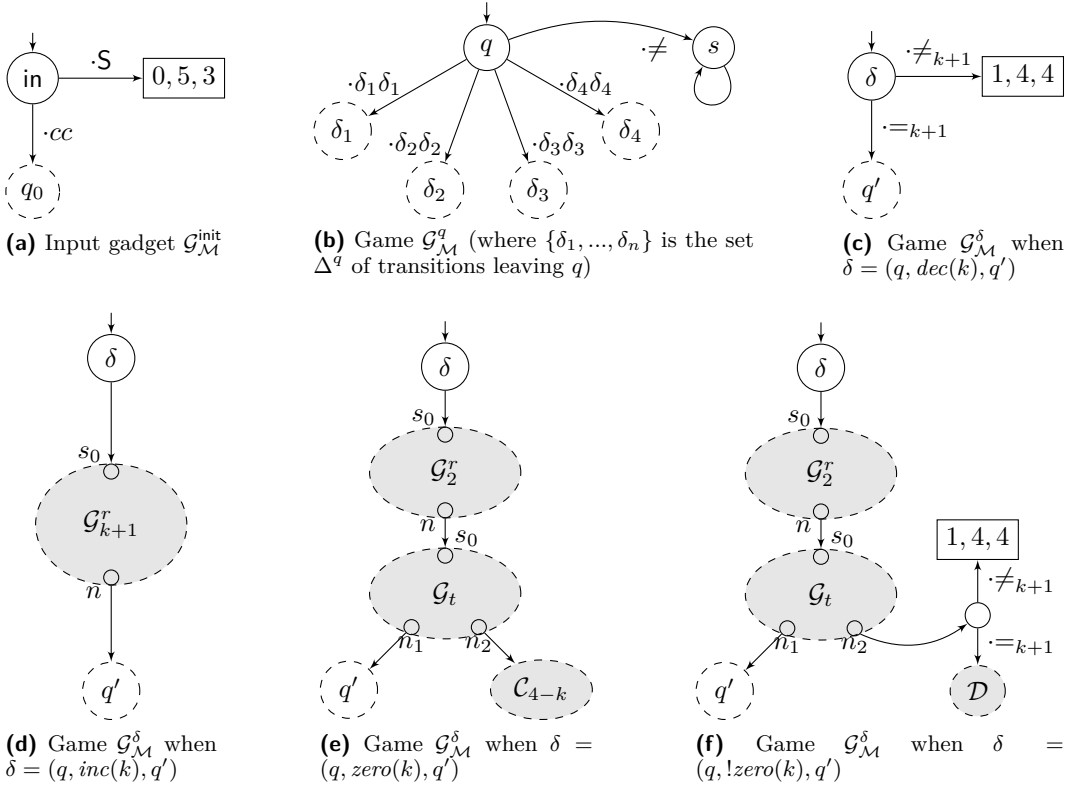
► **Lemma 11.** *Consider the games of Fig. 8. For  $n \in \mathbb{N} \cup \{+\infty\}$ , define*

$$r_k(n) = \left(1, 4 - \frac{1}{k^n}, 4 + \frac{1}{k^n}\right) \quad s(n) = \left(1, 4 - \frac{1}{n+1}, 4 + \frac{1}{n+1}\right)$$

*Then the set of 0-optimal Nash equilibrium values is  $\{s(n) \mid n \in \mathbb{N} \cup \{\infty\}\}$  in  $\mathcal{D}$ , and  $\{r_k(n) \mid n \in \mathbb{N} \cup \{\infty\}\}$  in  $\mathcal{C}_k$  for all  $k \geq 2$ .*



■ **Figure 8** The modules  $\mathcal{C}_k$  (for  $k \geq 2$ ) and  $\mathcal{D}$ . Notice that state  $s_2$  should be considered terminal, as it only carries a self-loop. We could replace it by a two-state loop. We could also see it as a terminal state with reward  $(0, 0, 0)$ , but for the proof of Corollary 16, we want the terminal rewards of players 1 and 2 to always sum to 8, which we could not achieve easily in this case.



■ **Figure 9** Description of the subgames  $\mathcal{G}_{\mathcal{M}}^q$  and  $\mathcal{G}_{\mathcal{M}}^\delta$

## 6 Undecidability proof

We now turn to the global undecidability proof of the constrained-existence problem in three-player games. The proof is a reduction from the recurring problem of a two-counter machine. We encode the behaviour of a two-counter machine  $\mathcal{M}$  as a concurrent game  $\mathcal{G}_{\mathcal{M}}$ , which connects the various subgames depicted on Figures 9 (one initial gadget, one per state  $q$ , one per transition  $\delta$ ). Roughly, this game will encode a configuration  $(q, c_1, c_2)$  of  $\mathcal{M}$  using a Nash equilibrium  $\sigma \in \mathbb{S}$  from  $q$  such that  $\mathbb{E}^\sigma(\phi \mid s) = (1, 4 + \frac{1}{2c_1 3c_2}, 4 - \frac{1}{2c_1 3c_2})$  (property  $P(q, c_1, c_2)$ ). Using the various constructions we have made previously, we can show that if  $P(q, c_1, c_2)$  is satisfied, then there is a transition  $(q, c_1, c_2) \rightarrow_\delta (q', c'_1, c'_2)$  in  $\mathcal{M}$  such that  $P(q', c'_1, c'_2)$  is satisfied as well, which allows to progress ‘along’ a Nash equilibrium while building a computation in  $\mathcal{M}$ .

The correspondence between  $\mathcal{M}$  and  $\mathcal{G}_{\mathcal{M}}$  is made precise as follows:

► **Proposition 12.** The two-counter machine  $\mathcal{M}$  has an infinite valid computation if, and only if, there is a 0-optimal Nash equilibrium from state  $\text{in}$  in game  $\mathcal{G}_{\mathcal{M}}$ .

This immediately entails:

► **Theorem 13.** We cannot decide whether there exists a 0-optimal Nash equilibrium in three-player games with non-negative terminal-reward payoffs.

We now consider several extensions of this result. We first state two straightforward corollaries. First, applying Lemma 10, we can enforce the 0-optimality constraint in the game by inserting an initial module. It follows:

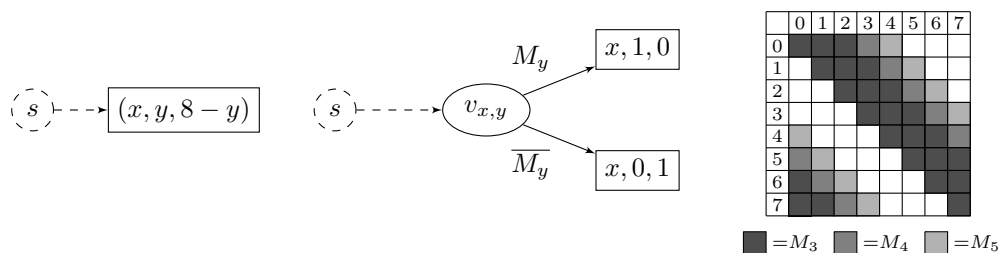


Figure 10 Transformation of a terminal node  $(x, y, 8 - y)$  with an intermediate node  $v_{x,y}$ . The table on the right gives the value of  $M_y$  for some values of  $y$  (notice that  $M_y \subseteq M_{y'}$  when  $y \leq y'$ ).

► **Corollary 14.** *We cannot decide whether there exists a Nash equilibrium in three-player games with (possibly negative) terminal-reward payoffs.*

Now we realize that in this reduction, there is a 0-optimal Nash equilibrium from in if, and only if, there is a Nash equilibrium with social welfare larger than or equal to 9, where the social welfare is defined as the sum of the expected payoffs of all players. As an immediate corollary, we get:

► **Corollary 15.** *We cannot decide whether there exists a Nash equilibrium with some lower bound on the social welfare (or with optimal social welfare) in three-player terminal-reward games with non-negative payoffs.*

We now explain briefly how the main theorem can be extended to terminal-reachability payoffs. We indeed realize that the payoffs of players 1 and 2 always sum up to 8 in the reduction (the game between those two players is zero-sum). The idea is then to replace each terminal state with a simple gadget in which the payoffs of players 1 and 2 are  $(8, 0)$  or  $(0, 8)$ , and to use an adequate set of actions which decomposes runs into two sets with proportions mimicking the normal rewards of the terminal state. For instance, for a reward  $(x, y, 8 - y)$ , the set of actions  $M_y = \{i \cdot j \mid \exists 0 \leq r < y. i - j = r \pmod{8}\}$  will lead to  $(x, 8, 0)$  and its complement to  $(x, 0, 8)$ , as illustrated on Fig. 10. In the game from  $v_{x,y}$ , there is a unique Nash equilibrium which consists in playing uniformly at random for both players, yielding a payoff of  $(x, y, 8 - y)$ . It remains to normalize and replace each  $(x, 8, 0)$  (resp.  $(x, 0, 8)$ ) by  $(x, 1, 0)$  (resp.  $(x, 0, 1)$ ).

► **Corollary 16.** *We cannot decide whether there exists a 0-optimal Nash equilibrium in three-player games with terminal-reachability payoffs.*

Finally, (roughly) by dualizing reachability and safety conditions, we can prove that the constrained existence of Nash equilibrium in safety games cannot be decided. This is to be compared with the fact that there always exists a Nash equilibrium in safety games [10].

► **Corollary 17.** *We cannot decide whether there exists a Nash equilibrium in three-player safety games with payoff 0 assigned to Player 0.*

## 7 Conclusion and future work

In this paper we have shown the undecidability of the existence of a constrained Nash equilibrium in a three-player concurrent game with terminal-reachability objectives. We believe this result is surprising, since it applies to very simple payoff functions, and with very few players. This result has to be compared with the undecidability result of [13], which

on one hand, applies to turn-based games, but requires 14 players and the full power of terminal-reward payoffs. Furthermore, in turn-based games with terminal-reachability payoffs, constrained Nash equilibria can be computed (in polynomial time) through a reduction to pure Nash equilibria [12] and algorithms for computing pure Nash equilibria [13]. We have also mentioned a couple of interesting corollaries that we do not repeat here.

This work lets open the decidability status of the constrained-existence problem in two-player games with terminal-reward and terminal-reachability payoffs. In fact, even the existence of Nash equilibria in such games is an open problem: it was believed until recently that there are two-player games with nonnegative terminal rewards having no Nash equilibrium [3, 13], but the proposed example was actually wrong (as we explained in Section 3.4). If one can find such a game with no Nash equilibrium, then our Corollary 14 extends to nonnegative terminal-reward games, and possibly to terminal-reachability games. Notice that two-player games have been studied quite a lot in the literature, and we know for instance that (uniform)  $\epsilon$ -Nash equilibria always exist in terminal-reward games [16, 17].

---

## References

- 1 P. Bouyer, R. Brenguier, N. Markey, and M. Ummels. Pure Nash equilibria in concurrent games. *Logical Methods in Computer Science*, 2014. Submitted, under revision.
- 2 P. Bouyer, N. Markey, and D. Stan. Mixed Nash equilibria in concurrent terminal-reward games. Research Report LSV-14-11, Lab. Spécification & Vérification, ENS Cachan, France, Oct. 2014.
- 3 K. Chatterjee, M. Jurdziński, and R. Majumdar. On Nash equilibria in stochastic games. In *CSL'04*, LNCS 3210, p. 26–40. Springer, 2004.
- 4 T. Chen, M. Kwiatkowska, D. Parker, and A. Simaitis. Verifying team formation protocols with probabilistic model checking. In *CLIMA XII*, LNAI 6814, p. 190–207. Springer, 2011.
- 5 V. Conitzer and T. Sandholm. Complexity results about Nash equilibria. In *IJCAI'03*, p. 765–771, 2003.
- 6 V. Conitzer and T. Sandholm. New complexity results about Nash equilibria. *Games and Economic Behavior*, 63(2):621–641, 2008.
- 7 L. de Alfaro, T. Henzinger, and O. Kupferman. Concurrent reachability games. *Theoretical Computer Science*, 386(3):188–217, 2007.
- 8 T. A. Henzinger. Games in system design and verification. In *TARK'05*, p. 1–4, 2005.
- 9 J. F. Nash. Equilibrium points in  $n$ -person games. *Proceedings of the National Academy of Sciences of the United States of America*, 36(1):48–49, 1950.
- 10 P. Secchi and W. D. Sudderth. Stay-in-a-set games. *International Journal of Game Theory*, 30:479–490, 2001.
- 11 W. Thomas. Infinite games and verification. In *CAV'02*, LNCS 2404, p. 58–64. Springer, 2002. Invited Tutorial.
- 12 M. Ummels. The complexity of Nash equilibria in infinite multiplayer games. In *FoSSaCS'08*, LNCS 4962, p. 20–34. Springer, 2008.
- 13 M. Ummels and D. Wojtczak. The complexity of Nash equilibria in limit-average games. In *CONCUR'11*, LNCS 6901, p. 482–496. Springer, 2011.
- 14 M. Ummels and D. Wojtczak. The complexity of Nash equilibria in limit-average games. Technical Report abs/1109.6220, CoRR, 2011.
- 15 M. Ummels and D. Wojtczak. The complexity of Nash equilibria in stochastic multiplayer games. *Logical Methods in Computer Science*, 7(3), 2011.
- 16 N. Vieille. Equilibrium in two-person stochastic games I: A reduction. *Israel Journal of Mathematics*, 119:55–91, 2000.
- 17 N. Vieille. Equilibrium in two-person stochastic games II: The case of recursive games. *Israel Journal of Mathematics*, 119:93–126, 2000.