Journal of Dynamics and Games doi:10.3934/jdg.2021028
©American Institute of Mathematical Sciences
Volume **9**, Number **1**, January **2022** pp. **117–122**

A NOTE ON THE NASH EQUILIBRIA OF SOME MULTI-PLAYER REACHABILITY/SAFETY GAMES

Athanasios Kehagias

Department of Electrical and Computer Engineering
Aristotle University of Thessaloniki
Thessaloniki, Greece

(Communicated by Panayotis Mertikopoulos)

Abstract. In this short note we study a class of multi-player, turn-based games with deterministic state transitions and reachability / safety objectives (this class contains as special cases "classic" two-player reachability and safety games as well as multi-player and ""stay–in-a-set" and "reach-a-set" games). Quantitative and qualitative versions of the objectives are presented and for both cases we prove the existence of a deterministic and memoryless Nash equilibrium; the proof is short and simple, using only Fink's classic result about the existence of Nash equilibria for *multi-player discounted stochastic games*

1. **Introduction.** The simplest ω-regular games are, arguably, two-player turn-based safety and reachability games [7]. *Multiplayer* variants of these are the "*Stay–in-a-set*" (SIAS) games [6, 8] and "*Reach-a-set*" (RAS) games [1, 2]. The existence of *Nash equilibria* (NE) has been proved: for SIAS games in [8] and for RAS games in [1, 2]; in particular in the special case of turn-based games with deterministic state transitions and Borel objectives (these include SIAS and RAS objectives) the existence of a pure *Nash equilibrium* (NE) is proved in [2, Corollary 1]. In both cases the state space is assumed finite and the NE are not, in general, memoryless.[1] A stronger result is proved in [9], namely: every turn-based multi-player game with deterministic state transitions and Borel objectives possesses a pure *sub-game perfect* (and hence memoryless) equilibrium. These results are quite general but their proofs are rather involved.

In the current note our main goal is to provide a *short and simple* proof of a special case: every turn-based SIAS and RAS game with deterministic state transitions possesses a deterministic and memoryless NE. This is proved using only Fink's classic result on the existence of NE for *multi-player discounted stochastic games* [5].

Our result is actually a little more general, in that it applies to the class of multi-player, turn-based games with deterministic state transitions, reachability objectives for some players and safety objectives for others. For brevity, we will henceforth refer to these as *multi-player reachability / safety games* (MPRS games);

2020 *Mathematics Subject Classification.* Primary: 91A06, 91A15; Secondary: 68R10.
Key words and phrases. Stochastic Games, N-player Games, Reachability, Nash Equilibrium, Graph Games.

[1] In [2, Theorem 1] is also proved the existence of memoryless ε-NE for a broader class, which contains SIAS and RAS games.

they contain as special cases classic reachability and safety games as well as SIAS and RAS games.

Informally, the MPRS game can best be visualized as a *graphical game*, in which N players move a token along the arcs of a digraph $G=(V,E)$. The vertices of G are partitioned into N sets: $V=V_1\cup V_2\cup ...\cup V_N$; if at the t-th turn the token is located on a vertex $v_t\in V_n$, then it is moved by the n-th player (henceforth denoted by P_n) into some vertex v_{t+1} such that (v_t,v_{t+1}) is an arc of G. In general we have two type of players: *reachers* and *avoiders*. To each P_n is associated a *nonempty* set $R_n\subseteq V$, related to his objective. If P_n is a reacher, he wins iff the token enters some vertex $v\in R_n$; if he is an avoider, he wins iff the token never enters a vertex $v\in R_n$.

In Section 2 we define the *quantitative* MPRS game and prove that every such game has a NE in deterministic memoryless strategies. In Section 3 we do the same things for the *qualitative* MPRS game.

2. **The quantitative MPRS game.** We now formulate MPRS as a *discounted stochastic game.*[2] In what follows the quantities N, V, E, V_1, ..., V_N, R_1, ..., R_N are the ones presented in the previous section.

1. The *player set* is $\{P_1,P_2,...,P_N\}$ or, for simplicity, $\{1,2,...,N\}$.
2. The *state set* is $S:=V\cup\{\overline{s}\}$, where V is the vertex set of the previously mentioned $G=(V,E)$ and $\overline{s}$ is the *terminal state.*
3. We define $\{S_1,...,S_N\}$, a partition of S, as follows: $S_1:=V_1\cup\{\overline{s}\}$, $S_2:=V_2$, ..., $S_N:=V_N$.
4. For $n\in\{1,2,...,N\}$, P_n's *target set* is R_n; the *total target set* is $R:=\cup_{m=1}^{N}R_m$.
5. $A_n(s)$ denotes P_n's *action set* when the game is at state s and is defined by (λ is the "trivial" move):
$$\begin{aligned}&\text{when } s\in S_n\backslash R: A_n(s):=\{s':(s,s')\in E\};\\&\text{when } s\in S_m\backslash R, m\neq n: A_n(s):=\{\lambda\};\\&\text{when } s\in R\cup\{\overline{s}\}: A_n(s):=\{\lambda\}.\end{aligned}$$
P_n's "total" action set is $A_n:=\cup_{s\in S}A_n(s)$.
6. The *law of motion* is deterministic and has the following form:
   (a) when $s\in S_n\backslash R$ and $a=(\lambda,...,a^n,...,\lambda)$:
$$\Pr(s_{t+1}=s'|s_t=s,a_t=a):=\begin{cases}1 & \text{when } s'=a^n,\\ 0 & \text{else;}\end{cases}\tag{1}$$
   (b) when $s\in R\cup\{\overline{s}\}$ and $a=(\lambda,...,\lambda,...,\lambda)$
$$\Pr(s_{t+1}=s'|s_t=s,a_t=a):=\begin{cases}1 & \text{when } s'=\overline{s},\\ 0 & \text{else.}\end{cases}\tag{2}$$
   All admissible state/action combinations are covered by (1)-(2), from which we see the following.
   (a) If the current state s "belongs" to P_n (i.e., $s\in S_n$) and is not a target state, then he is the only player who can perform a non-trivial action $a^n\in V$; the next state is, with certainty, a^n.

[2] We follow the formulation of [4], expanded to the multi-player case.

(b) If the current state s is either target or terminal, then the only admissible action vector is $a = (\lambda, ..., \lambda, ..., \lambda)$; the next and all subsequent states are the terminal $\overline{s}$.

It is convenient to describe the *deterministic* state transitions in terms of a state transition function $\mathbf{T} : S \times A \to S$, defined by

$$\mathbf{T}(s, a^n) := \begin{cases} a^n & \text{when } s \in S_n \backslash R \text{ and } a^n \in A^n(s) \backslash \lambda, \\ \overline{s} & \text{when } s \in R \cup \{\overline{s}\} \text{ and } a^n = \lambda. \end{cases} \tag{3}$$

All admissible state/action combinations are covered by (3).

7. P_n's *turn payoff function* depends only on the current game state s (but not on the current action vector) and can be either of the following:

$$q^n(s) := \begin{cases} 1 & \text{when } s \in R_n \\ 0 & \text{when } s \notin R_n \end{cases} \quad (P_n \text{ is a reacher});$$

$$q^n(s) := \begin{cases} -1 & \text{when } s \in R_n \\ 0 & \text{when } s \notin R_n \end{cases} \quad (P_n \text{ is an avoider}).$$

8. P_n's *total payoff* function is (with *discount factor* $\gamma \in (0, 1)$): $Q^n(s_0, s_1, ...) = \sum_{t=0}^{\infty} \gamma^t q^n(s_t)$.

The game starts at an initial state $s_0 = s \in S \backslash \overline{s}$ and, at the t-th turn ($t \in \{0, 1, 2, ...\}$) all players perform "trivial" moves, except for the player who "owns" s_t. Two possibilities exist.

1. If a target state is entered at some time t' ($s_{t'} = s' \in R = \cup_{m=1}^{N} R_m$) the next and all subsequent states are the terminal ($\forall t > t' : s_t = \overline{s}$).[3] For each $n \in \{1, ..., N\}$, P_n receives total payoff:

$$Q^n(s_0, s_1, ...) = \begin{cases} \gamma^{t'} & \text{if } s' \in R_n \text{ and he is a reacher}; \\ -\gamma^{t'} & \text{if } s' \in R_n \text{ and he is an avoider}; \\ 0 & \text{if } s' \notin R_n. \end{cases}$$

2. If a target state is never entered ($\forall t : s_t \notin R$), the game continues ad infinitum and all players receive zero payoff.

A reacher (resp. avoider) P_n wants the game to enter R_n in the shortest (resp. longest) possible time. Hence the above defined discounted stochastic game will be called "*quantitative MPRS game*".

A *finite-length history* is a finite sequence of states (we omit player actions, since they will not be needed in our proof[4]):

$$h = s_0 s_1 ... s_k \in \underbrace{S \times S \times ... \times S}_{k \text{ times}} \qquad \text{for some } k \in \{1, 2, ...\};$$

the set of all finite-length histories is denoted by H^*. A *deterministic strategy* for the n-th player is a function σ^n which assigns an action to each finite-length history: $\sigma^n : H^* \to A_n$. A strategy σ^n is called *memoryless* if it only depends on the current state, in which case we write (with a slight notation abuse) $\sigma^n(s_0 s_1 ... s_k) = \sigma^n(s_k)$. A *strategy profile* is a tuple $\sigma = (\sigma^1, \sigma^2, ..., \sigma^N)$ which specifies one strategy for each player. As usual, $\sigma^{-n} = (\sigma^j)_{j \in \{1,2,...,N\} \backslash \{n\}}$, so we can write $\sigma = (\sigma^n, \sigma^{-n})$. Since an initial state s_0 and a deterministic strategy profile σ determine fully the

---

[3]Hence, while the game lasts an infinite number of turns, it *effectively* ends at t'.

[4]Besides they are directly inferred from the states, due to the deterministic law of motion.

history $s_0 s_1 s_2 \ldots$, the payoff function $Q^n(s_0, s_1, \ldots)$ will also be written as $Q^n(s_0, \sigma)$, $Q^n(s_0, \sigma^1, \ldots, \sigma^N)$ or $Q^n(s_0, \sigma^n, \sigma^{-n})$.

Theorem 2.1. *Every quantitative MPRS game has a deterministic memoryless NE. In other words, there exists a profile of deterministic memoryless strategies* $\widehat{\sigma} = (\widehat{\sigma}^1, \widehat{\sigma}^2, \ldots, \widehat{\sigma}^N)$ *such that*

$$\forall n \in \{1, 2, \ldots, N\}, \forall s_0 \in S, \forall \sigma^n : Q^n\left(s_0, \widehat{\sigma}^n, \widehat{\sigma}^{-n}\right) \geq Q^n\left(s_0, \sigma^n, \widehat{\sigma}^{-n}\right). \tag{4}$$

For every s *and* n, *let* $u^n(s) := Q^n(s, \widehat{\sigma})$. *Then the following equations are satisfied*

$$\forall n, \forall s \in S_n : \widehat{\sigma}^n(s) = \arg \max_{a^n \in A^n(s)} \left[q^n(s) + \gamma u^n(\mathbf{T}(s, a^n))\right], \tag{5}$$

$$\forall n, m, \forall s \in S_n : u^m(s) = q^m(s) + \gamma u^m(\mathbf{T}(s, \widehat{\sigma}^n(s))). \tag{6}$$

Proof. Fink has proved in [5] that every N-player discounted stochastic game has a memoryless NE in *probabilistic* strategies; this result holds for the general game (i.e., with *concurrent* moves and probabilistic strategies and state transitions). According to [5], at equilibrium the following equations must be satisfied for all m and s:

$$\mathfrak{u}^m(s) = \left(\max_{\mathbf{p}^m(s)} \sum_{a^1 \in A^1(s)} \ldots \sum_{a^N \in A^N(s)} p^1\left(a^1|s\right) \ldots p^N\left(a^N|s\right)\right) \cdot A \tag{7}$$

with

$$A = \left[q^m(s) + \gamma \sum_{s'} \Pi\left(s'|s, a^1, \ldots, a^N\right) \mathfrak{u}^m(s')\right].$$

In the above we have modified Fink's original notation to fit our own; in particular:

1. $\mathfrak{u}^m(s)$ is the expected value of $u^m(s)$;
2. $p^m(a^m|s)$ is the probability that, given the current game state is s, the m-th player plays action a^m;
3. $\mathbf{p}^m(s) = (p^m(a^m|s))_{a^m \in A^m(s)}$ is the vector of all such probabilities (one probability per available action);
4. $\Pi\left(s'|s, a^1, a^2, \ldots, a^N\right)$ is the probability that, given the current state is s and the player actions are $a^1, a^2, \ldots, a^N$, the next state is s' .

Now choose any n and any $s \in S_n$. For all $m \neq n$, the m-th player has a single move: $A^m(s) = \{\lambda\}$, and so $p^m(a^m|s) = 1$. Also, since transitions are deterministic,

$$\sum_{s'} \Pi\left(s'|s, a^1, a^2, \ldots, a^N\right) \mathfrak{u}^n(s') = \mathfrak{u}^n(\mathbf{T}(s, a^n)).$$

Hence, for $m = n$, (7) becomes

$$\mathfrak{u}^n(s) = \max_{\mathbf{p}^n(s)} \sum_{a^n \in A^n(s)} p^n(a^n|s) \left[q^n(s) + \gamma \mathfrak{u}^n(\mathbf{T}(s, a^n))\right]. \tag{8}$$

Furthermore let us define $\widehat{\sigma}^n(s)$ (for the specific s and n) by

$$\widehat{\sigma}^n(s) = \arg \max_{a^n \in A^n(s)} \left[q^n(s) + \gamma \mathfrak{u}^n(\mathbf{T}(s, a^n))\right]. \tag{9}$$

If (8) is satisfied by more than one a^n, we set $\widehat{\sigma}^n(s)$ to one of these arbitrarily. Then, to maximize the sum in (8) the n-th player must set $p^n(\widehat{\sigma}^n(s)|s) = 1$ and $p^n(a^n|s) = 0$ for all $a^n \neq \widehat{\sigma}^n(s)$. Since this is true for all states and all players (i.e., every player can, without loss, use deterministic strategies) we also have $\mathfrak{u}^n(s) = u^n(s)$. Hence (8) becomes

$$u^n(s) = \max_{a^n \in A^n(s)} \left[q^n(s) + \gamma u^n(\mathbf{T}(s, a^n))\right] = q^n(s) + \gamma u^n\left(\mathbf{T}\left(s, \widehat{\sigma}^n(s)\right)\right). \quad (10)$$

For $m \neq n$, the m-th player has no choice of action and (8) becomes

$$u^m(s) = q^m(s) + \gamma u^m\left(\mathbf{T}\left(s, \widehat{\sigma}^n(s)\right)\right). \quad (11)$$

We recognize that (9)-(11) are (5)-(6); replacing $\mathfrak{u}^n(\mathbf{T}(s, a^n))$ with $u^n(\mathbf{T}(s, a^n))$ in (9) defines $\widehat{\sigma}^n(s)$ for every n and s and so yields the required deterministic memoryless strategies $\widehat{\sigma} = \left(\widehat{\sigma}^1, \widehat{\sigma}^2, ..., \widehat{\sigma}^3\right)$. □

3. **The qualitative MPRS game.** The qualitative MPRS game elements are identical to those of the quantitative one, except for the payoff functions. The qualitative game: (i) does *not* have a turn payoff function; (ii) has total payoff function

$$\widetilde{Q}^n(s_0, \sigma) = \begin{cases} 1 & \text{if} \quad Q^n(s_0, \sigma) > 0, \\ -1 & \text{if} \quad Q^n(s_0, \sigma) < 0, \\ 0 & \text{if} \quad Q^n(s_0, \sigma) = 0; \end{cases}$$

It is easily checked that: $\widetilde{Q}^n(s_0, \sigma) = 1$ (resp. $\widetilde{Q}^n(s_0, \sigma) = 0$) iff P_n is a reacher (resp. an avoider) and his target set is entered (resp. not entered). Accordingly, in the qualitative MPRS game P_n *wins* (resp. *loses*) iff he achieves the maximum (resp. minimum) possible value of $\widetilde{Q}^n$. More specifically, we have the following.

1. When P_n is a reacher, he wins (resp. loses) iff $\widetilde{Q}^n(s_0, \sigma) = 1$ (resp. $\widetilde{Q}^n(s_0, \sigma) = 0$).
2. When P_n is an avoider, he wins (resp. loses) iff $\widetilde{Q}^n(s_0, \sigma) = 0$ (resp. $\widetilde{Q}^n(s_0, \sigma) = -1$).

In short, the quantitative Q^n's defines the qualitative $\widetilde{Q}^n$'s which are used to formalize win/lose criteria analogous to these of reachability, safety, RAS and SIAS games; these are special cases of qualitative MPRS:

1. two-player reachability games ($N = 2$, P_1 a reacher with $R_1 \neq \emptyset$ and P_2 an avoider with $R_2 = R_1$);
2. safety games (same as the reachability game, with player roles interhanged);
3. SIAS games ($\forall n : P_n$ is an avoider);
4. RAS games ($\forall n : P_n$ is a reacher). [5]

The most general MPRS game involves N_1 reachers and N_2 avoiders; we can have more than one winners (e.g., if P_m and P_n are reachers, both win if the token enters some $v \in R_m \cap R_n \neq \emptyset$) and the same is true for losers.

It is easily checked that every $\widehat{\sigma} = \left(\widehat{\sigma}^1, ..., \widehat{\sigma}^N\right)$ which is a NE of the Q^n's is also a NE of the $\widetilde{Q}^n$'s. Hence, by Theorem 2.1, we have the following.

Corollary 1. *Every qualitative MPRS game has a deterministic memoryless NE.*

[5] Note that, while the RAS game can be seen as a *variant* of the classic reachability game, it is not a generalization thereof, because it does not involve a player with safety objectives [1, 2]. Similarly, the classic safety game is not a SIAS game.

REFERENCES

[1] K. Chatterjee, R. Majumdar and M. Jurdziński, *On Nash Equilibria in Stochastic Games*, Report No.UCB/CSD-3-1281, 2003, Computer Science Division (EECS), Univ. of California at Berkeley.

[2] K. Chatterjee, R. Majumdar and M. Jurdziński, *On Nash Equilibria in Stochastic Games,* International Workshop on Computer Science Logic, Springer, Berlin, Heidelberg, 2004.

[3] K. Chatterjee and T. A. Henzinger, A survey of stochastic ω-regular games, *Journal of Computer and System Sciences*, **78** (2012), 394–413.

[4] J. Filar and K. Vrieze, *Competitive Markov Decision Processes: Heory, Algorithms, and Applications*, Springer-Verlag, New York, 1997.

[5] A. M. Fink, Equilibrium in a stochastic n-person game, *Journal of science of the Hiroshima University, Series Ai (Mathematics)*, **28** (1964), 89–93.

[6] A. Maitra and W. D. Sudderth, Borel stay-in-a-set games, *International Journal of Game Theory*, **32** (2003), 97–108.

[7] R. Mazala, Infinite games, in *Automata Logics, and Infinite Games*, Springer, **2500** (2002), 23–38.

[8] P. Secchi and W. D. Sudderth, Stay-in-a-set games, *International Journal of Game Theory*, **30** (2002), 479–490.

[9] M. Ummels, *Stochastic Multiplayer Games: Theory and Algorithms*, Amsterdam University Press, 2010.

Received for publication November 2020; early access November 2021.

E-mail address: kehagiat@ece.auth.gr