Playing Games with Counter Automata

Antonín Kučera*

Faculty of Informatics, Masaryk University
Botanická 68a, 60200 Brno
Czech Republic
kucera@fi.muni.cz

Abstract. We survey recent results about subclasses of multi-counter games that are either equipped with more than one counter or allow for stochastic control states.

1 Introduction

Markov decision processes (MDPs) and stochastic games (SGs) are standard models for systems that exhibit both stochastic and non-deterministic behaviour. The algorithmic theory of finite-state MDPs and SGs is well-developed (see, e.g., [27, 23]), and in recent years, the scope of this study has been extended to certain classes of finitely representable but *infinite-state* MDPs and SGs. Not surprisingly, these classes are usually obtained as extensions of well-known classes of abstract computational devices, such as pushdown automata [19–21, 18, 9, 8] or lossy channel systems [2, 1]. In this paper, we survey recent results about MDPs and games over *counter automata*, which seem to represent a particularly convenient trade-off between modelling power and computational tractability.

A *multi-counter game* with n counters is a directed finite-state graph whose states are partitioned into three subsets of *stochastic*, *Player* $\Box$, and *Player* $\Diamond$ states, and each transition is labeled by an *update vector* $\boldsymbol{u} \in \mathbb{Z}^n$, where $\boldsymbol{u}_i$ represents the change in the i-th counter caused by the transition. For every stochastic state, there is a fixed probability distribution over its outgoing transitions. A *play* of a multi-counter game starts in some state for some initial values of the counters. In a current configuration $p\boldsymbol{v}$ (where p is a state and $\boldsymbol{v}$ a vector of counter values), the next transition is chosen either randomly or by Player $\Box/\Diamond$, depending on whether p is stochastic or belongs to the respective player. The aim of Player $\Box$ is to *maximize* the expected value of a certain *payoff function* which assigns a real payoff to every run of a play, while Player $\Diamond$ aims at minimizing this expectation. Intuitively, the counters represent various resources that are produced or consumed along a play, and the payoff function specifies how well are these resources treated by a given run. For example, we may be interested whether Player $\Box$ can play *safely*, i.e., so that the resources are never exhausted and remain positive in all of the visited configurations. In this case, the associated payoff functions assigns to every run either 1 or 0, depending on whether or not the above

* Supported by the Czech Science Foundation, grant No. P202/12/G061.

A. Finkel, J. Leroux, and I. Potapov (Eds.): RP 2012, LNCS 7550, pp. 29–41, 2012.

safety condition holds. The expected payoff of a run then corresponds to the *probability* of all runs satisfying the safety condition.

The modelling power of counter automata is further extended by allowing some components in update vectors to take ω-values, which intuitively corresponds to "resource reloading". The ω's may appear only in the outgoing transitions of non-stochastic states. If $\boldsymbol{u}_i = \omega$, then the responsible player chooses an arbitrarily large non-negative integer which is added to the i-th counter when performing the transition. In reality, the capacity of storage devices (such as tanks or batteries) is bounded, and hence we may ask, for example, what is the least capacity sufficient for Player $\Box$ to play safely.

The general model of multi-counter games is hard to analyze, particularly due to the interplay between stochastic control states and counters (note that one can easily simulate multi-dimensional partially controlled random walks using multi-counter MDPs). Hence, the existing works concern only subclasses of multi-counter games. In this paper, we survey recent results of two types:

– Results about subclasses of multi-counter games *with more than one counter but without stochastic control states*.
– Results about subclasses of multi-counter games *with stochastic states and only one counter*.

In particular, we do not attempt to survey results about (non-stochastic) energy games and closely related mean payoff games (see, e.g., [14]), because these areas are quite rich on their own and deserve a bit more space and attention than can be provided in the framework of this paper.

2 Preliminaries

We use $\mathbb{Z}$ and $\mathbb{N}$ to denote the sets of integers and non-negative integers, respectively. Further, $\mathbb{Z}_\omega$ denotes the set $\mathbb{Z} \cup \{\omega\}$, and we use $\boldsymbol{u}, \boldsymbol{v}, \ldots$ to range over $\mathbb{Z}_\omega^n$, where $n \geq 1$.

We assume familiarity with basic notions of probability theory (such as probability space, random variable, the expected value of a random variable, etc.) In particular, we call a probability distribution f over a discrete set A *positive* if $f(a) > 0$ for all $a \in A$, and *rational* if $f(a)$ is rational for every $a \in A$.

Definition 1 (Stochastic games). *A* stochastic game *is a tuple* $\mathcal{G} = (S, (S_\bigcirc, S_\Box, S_\Diamond), \rightsquigarrow, Prob)$*, consisting of a finite or countably infinite set S of* states*, partitioned into the set $S_\bigcirc$ of* stochastic *states, and the sets $S_\Box$, $S_\Diamond$ of states owned by Player $\Box$ (Max) and Player $\Diamond$ (Min), respectively. The* edge relation *$\rightsquigarrow \subseteq S \times S$ is total, i.e., for every $r \in S$ there is $s \in S$ such that $r \rightsquigarrow s$. Finally, Prob assigns to every $s \in S_\bigcirc$ a positive probability distribution over its outgoing edges.*

A *finite path* in $\mathcal{G}$ is a sequence $w = s_0 s_1 \cdots s_k$ of states such that $s_i \rightsquigarrow s_{i+1}$ for all i where $0 \leq i < k$. A *run* is an infinite sequence of states every finite prefix of which is a path. For a finite path w, we denote by $Run(w)$ the set of runs having w as a prefix. These generate the standard σ-algebra on the set of runs.

Let $\mathcal{G}$ be a stochastic game. A *strategy* for Player $\Box$ in $\mathcal{G}$ is a function σ assigning to each finite path w ending in a state $s \in S_\Box$ a distribution on edges leaving the state s. A strategy σ is *pure* if it always assigns 1 to some edge and 0 to the others, and *memoryless* if $\sigma(w) = \sigma(s)$ where s is the last state of w. A strategy π for Player $\Diamond$ is defined analogously. Fixing a pair (σ, π) of strategies for Player $\Box$ and $\Diamond$, respectively, and an initial state s, we obtain in a standard way a probability measure $\mathbb{P}_s^{\sigma,\pi}(\cdot)$ on the subset of runs starting in s.

For a stochastic game $\mathcal{G}$, a *payoff* is a Borel measurable real-valued function f over the runs in $\mathcal{G}$. Player $\Box$ is trying to *maximize* the expected value of f, while Player $\Diamond$ is trying to *minimize* it. For given pair of strategies (σ, π) and a state $s \in S$, we use $\mathbb{E}_s^{\sigma,\pi}[f]$ to denote the expected value of f over $Run(s)$, where $\mathbb{P}_s^{\sigma,\pi}(\cdot)$ is the underlying probability measure. If f is bounded, then

$$\sup_\sigma \inf_\pi \mathbb{E}_s^{\sigma,\pi}[f] \quad = \quad \inf_\pi \sup_\sigma \mathbb{E}_s^{\sigma,\pi}[f]$$

for every $s \in S$, and we say that $(\mathcal{G}, f)$ is *determined* [26, 25]. If $(\mathcal{G}, f)$ is determined, then the above equality defines the *value* of s, denoted by $Val(f, s)$.

For a given $\varepsilon \geq 0$, a strategy σ^* of Player $\Box$ is *ε-optimal* in s, if

$$\mathbb{E}_s^{\sigma^*,\pi}[f] \quad \geq \quad Val(f, s) - \varepsilon$$

for every strategy π of Player $\Diamond$. An ε-optimal strategy for Player $\Diamond$ is defined analogously. 0-optimal strategies are called *optimal*. Note that $(\mathcal{G}, f)$ is determined iff both players have ε-optimal strategies for every $\varepsilon > 0$.

An important subclass of payoff functions are *ω-regular payoffs* which assign to each run either 1 or 0 depending on whether or not the run satisfies a certain ω-regular condition R. Note that for every pair of strategies (σ, π), the expectation $\mathbb{E}_s^{\sigma,\pi}[f]$ is then equal to the probability of all runs satisfying R. Simple examples of ω-regular conditions are *reachability* and *safety*, satisfied by all runs that visit and avoid visiting a given subset of target states, respectively.

Definition 2 (Multi-counter games). *A* multi-counter game with $n \geq 1$ counters *is a tuple* $\mathcal{A} = (Q, (Q_\bigcirc, Q_\Box, Q_\Diamond), \delta, P)$ *consisting of*

- *a finite non-empty set Q of* control states, *partitioned into stochastic and players' states $Q_\bigcirc$, $Q_\Box$, and $Q_\Diamond$ as in the case of stochastic games;*
- *a set of* transition rules $\delta \subseteq Q \times \mathbb{Z}_\omega^n \times Q$. *For a given $(q, \boldsymbol{u}, r) \in \delta$, we call q the* source, *r the* target, *and $\boldsymbol{u}$ the* update vector *of the rule. We require that for every $q \in Q$ there is at least one outgoing transition rule of the form $(q, \boldsymbol{u}, r)$ in δ, and for every transition $(q, \boldsymbol{u}, r) \in \delta$ where q is stochastic we have that $\boldsymbol{u} \in \mathbb{Z}^n$ (i.e., there are no ω's in $\boldsymbol{u}$ when q is stochastic);*
- *the* probability assignment *P that assigns to every $q \in Q_\bigcirc$ a positive rational probability distribution over its outgoing transition rules.*

By $||\mathcal{A}|| := |Q| + ||\delta|| + ||P||$ we denote the encoding size of $\mathcal{A}$, where all rational constants are encoded as fractions of binary numbers.

Every multi-counter game $\mathcal{A}$ determines an infinite-state stochastic game where the states are *configurations* of $\mathcal{A}$, i.e., pairs of the form $q\boldsymbol{v}$ where $q \in Q$ and $\boldsymbol{v}$ a vector of counter values, and the edges are determined by applying the rules of $\mathcal{A}$. Here, every ω component in $\boldsymbol{v}$ is interpreted as an unbounded increase in the respective counter value, i.e., the responsible player selects a non-negative integer which is added to the current counter value. However, the precise definition of this infinite-state stochastic game depends on whether or not the counters are allowed to take negative values, which also influences the treatment of transitions that decrease (some) counters below zero. This difference can be important as well as irrelevant, depending on the choice of payoff function and the studied subclass of multi-counter games. Here we present both approaches.

$\mathbb{N}$-Semantics of Multi-counter Games. Let $\mathcal{A} = (Q, (Q_\bigcirc, Q_\Box, Q_\Diamond), \delta, P)$ be a multi-counter game with $n \geq 1$ counters. We define a stochastic game $\mathcal{G}_{\mathcal{A},\mathbb{N}}$ where

- the states are elements of $Q \times \mathbb{N}^n$;
- the edges are defined as follows: for a given $q\boldsymbol{v}$, we put $q\boldsymbol{v} \rightsquigarrow r\boldsymbol{t}$ for all $r\boldsymbol{t} \in Q \times \mathbb{N}^n$ such that there is $(q, \boldsymbol{u}, r) \in \delta$ satisfying the following:
  - $\boldsymbol{t}_i = \boldsymbol{v}_i + \boldsymbol{u}_i \geq 0$ for every $1 \leq i \leq n$ such that $\boldsymbol{u}_i \neq \omega$,
  - $\boldsymbol{t}_i \geq \boldsymbol{v}_i$ for every i such that $\boldsymbol{u}_i = \omega$.

  If there is no such $r\boldsymbol{t}$, then the configuration $q\boldsymbol{v}$ has only one outgoing edge $q\boldsymbol{v} \rightsquigarrow q\boldsymbol{v}$;
- the probability assignment *Prob* is derived from P as follows: for a given $q\boldsymbol{v} \in Q_\bigcirc \times \mathbb{N}^n$, let $\beta \subseteq \delta$ be the set of all $(q, \boldsymbol{u}, r) \in \delta$ that are enabled in $q\boldsymbol{v}$, i.e., $\boldsymbol{v}_i + \boldsymbol{u}_i \geq 0$ for every $1 \leq i \leq n$. If $\beta = \emptyset$, then $q\boldsymbol{v} \rightsquigarrow q\boldsymbol{v}$ with probability one. Otherwise, let y be the sum of the probabilities of all rules in β. For every $(q, \boldsymbol{u}, r) \in \beta$, we put $q\boldsymbol{v} \rightsquigarrow r(\boldsymbol{v}+\boldsymbol{u})$ with probability x/y, where $x = P((q, \boldsymbol{u}, r))$.

Hence, if a transition rule requires decreasing some counter below zero in $q\boldsymbol{v}$, then it is *disabled*. This is similar to the standard semantics of Petri nets, where the places can become empty but cannot hold negative values. Note that the probability distribution over the outgoing edges of stochastic configurations is derived from P by conditioning on enabled transition rules.

One can also extend multi-counter games with *zero test*, i.e., add special transitions enabled only when a given counter holds zero. Although this model is obviously Turing powerful for $n \geq 2$, it can still be considered for one-counter games.

$\mathbb{Z}$-Semantics of Multi-counter Games. Let $\mathcal{A} = (Q, (Q_\bigcirc, Q_\Box, Q_\Diamond), \delta, P)$ be a multi-counter game with $n \geq 1$ counters. We define a stochastic game $\mathcal{G}_{\mathcal{A},\mathbb{Z}}$ where

- the states are elements of $Q \times \mathbb{Z}^n$;
- the edges are defined by $q\boldsymbol{v} \rightsquigarrow r\boldsymbol{t}$ iff there is a transition $(q, \boldsymbol{u}, r) \in \delta$ such that
  - $\boldsymbol{t}_i = \boldsymbol{v}_i + \boldsymbol{u}_i$ for every $1 \leq i \leq n$ such that $\boldsymbol{u}_i \neq \omega$,
  - $\boldsymbol{t}_i \geq \boldsymbol{v}_i$ for every i such that $\boldsymbol{u}_i = \omega$;
- the probability assignment *Prob* is derived naturally from P.

Note that for *every* configuration $q\boldsymbol{v} \in Q \times \mathbb{Z}^n$, there is always at least one transition enabled in $q\boldsymbol{v}$.

A special type of pure memoryless strategies that are applicable both in $\mathcal{G}_{\mathcal{A},\mathbb{N}}$ and $\mathcal{G}_{\mathcal{A},\mathbb{Z}}$ are *counterless* strategies which depend only of the control state of the currently visited configuration.

3 Existing Results about Multi-counter Games

In this section we give an overview of the existing results about subclasses of non-stochastic multi-counter games with more than one counter (Sections 3.1, 3.2, and 3.3) and stochastic one-counter games (Sections 3.4, 3.5, and 3.6).

3.1 eVASS Games

Games over extended vector addition systems with states (*eVASS games*) have been introduced and studied in [11]. eVASS games are multi-counter games with $\mathbb{N}$-semantics such that

- there are no stochastic control states (i.e., $Q_\bigcirc = \emptyset$),
- the update vectors are elements of $\{-1, 0, 1, \omega\}^n$.

The main results of [11] concern eVASS games with *zero reachability* payoff functions, which are considered in two variants:

- *selective zero reachability*, denoted by Z_T, where $T \subseteq Q$. The function Z_T assigns to every run either 1 or 0, depending on whether or not the run visits a configuration $q\boldsymbol{v}$ such that $q \in T$ and $\boldsymbol{v}_i = 0$ for some $1 \leq i \leq n$.
- *non-selective zero reachability*, denoted by Z, which is defined in the same way as Z_Q. That is, Z assigns 1 to those runs that decrease some counter to zero.

One can easily show that in eVASS games with Z_T payoff function, both players have pure memoryless strategies that are optimal in every configuration. Hence, the value of every configuration is either 1 or 0. For *selective* zero reachability, the following is observed:

Theorem 3. *Let $\mathcal{A}$ be an eVASS game. The problem whether $Val(Z_T, q\boldsymbol{v}) = 1$ is undecidable, even if $\mathcal{A}$ has just two counters, no ω-components in update vectors, and $\boldsymbol{v} = (0, 0)$. Further, the problem is highly undecidable (beyond the arithmetical hierarchy) even if $\mathcal{A}$ has just three counters and $\boldsymbol{v} = (0, 0, 0)$.*

Theorem 3 is obtained by straightforward reductions from the halting problem and the recurrence problem for two-counter Minsky machines.

The properties of eVASS games with Z payoff function are different. The set Val_0 of all configurations with value 0 is obviously upwards closed in the sense that if $p\boldsymbol{v} \in Val_0$, then also $p(\boldsymbol{v}+\boldsymbol{t}) \in Val_0$ for all $\boldsymbol{t} \in \mathbb{N}^n$. Hence, the set Val_0 is fully described by a finite set of its minimal elements, and the set Val_1 of all configurations with value 1 is just a complement of Val_0. In [11], the following theorem is proven:

Theorem 4. *Let $\mathcal{A}$ be an eVASS game with $n \geq 1$ counters.*

- *The set of minimal elements of Val_0 is computable in $(n-1)$-exponential time[1]. In particular, the problem whether the value of a given configuration is 0 (or 1) is solvable in $(n-1)$-exponential time.*

[1] Here, 0-exponential time means polynomial time.

– *The problem whether the value of a given configuration is* 0 *is* ***EXPSPACE****-hard, even if* $Q_\Box = \emptyset$.
– *Optimal strategies for both players are finitely and effectively representable.*

An optimal strategy for Player $\Diamond$ can be specified just by the moves in all of the finitely many minimal configurations of Val_0 (observe that in a non-minimal configuration $p(\boldsymbol{v}+\boldsymbol{u}) \in Val_0$ such that $p\boldsymbol{v} \in Val_0$ is minimal, Player $\Diamond$ can safely make a move $p(\boldsymbol{v}+\boldsymbol{u}) \rightarrow q(\boldsymbol{v}'+\boldsymbol{u})$ where $p\boldsymbol{v} \rightarrow q\boldsymbol{v}'$ is the move associated to $p\boldsymbol{v}$. This also implies that there is a finite and effectively computable constant c such that Player $\Diamond$ can always replace every ω with c when performing a transition whose update vector contains some ω components (obviously, Player $\Box$ can always choose zero for every ω). A finite description of an optimal winning strategy for Player $\Box$ is more complicated. We refer to [11] for details.

It is worth noting that in the special case of two-counter eVASS games where update vectors do not contain any ω components, the complexity of the problem whether $p\boldsymbol{v} \in Val_0$ (or whether $p\boldsymbol{v} \in Val_1$) can be improved from **EXPTIME** to **P** [13].

3.2 Consumption Games

Consumption games, introduced in [10], are multi-counter games with $\mathbb{Z}$-semantics such that

– there are no stochastic control states (i.e., $Q_\bigcirc = \emptyset$),
– the update vectors are elements of $(\mathbb{Z}_\omega^{\leq 0})^n$, where $\mathbb{Z}_\omega^{\leq 0}$ is the set of all non-positive integers together with ω.

Hence, in consumption games, the counters can be increased only by performing transitions with ω components in update vectors. Intuitively, the counters in consumption games model resources of various types that can be only consumed or "reloaded" to some finite amount.

The payoff functions studied for consumption games in [10] are *zero safety* and *zero safety with upper bound* $\boldsymbol{u}$, where $\boldsymbol{u} \in \mathbb{N}^n$ (here n is the number of counters). Formally, let

– S be a function which to every run assigns either 1 or 0 depending on whether or not the run avoids visiting configurations of the form $q\boldsymbol{v}$ where $\boldsymbol{v}_i \leq 0$ for some $1 \leq i \leq n$;
– $S^{\boldsymbol{u}}$ be a function which to every run assigns either 1 or 0 depending on whether or not the run avoids visiting configurations of the form $q\boldsymbol{v}$ where $\boldsymbol{v}_i \leq 0$ or $\boldsymbol{v}_i > \boldsymbol{u}_i$ for some $1 \leq i \leq n$.

Hence, zero safety is dual to zero reachability discussed in Section 3.1. Again, the value of every configuration is either 1 or 0. For every control state p, let

– $safe(p)$ be the set of all $\boldsymbol{v} \in \mathbb{N}^n$ such that $Val(S, p\boldsymbol{v}) = 1$;
– $cover(p)$ be the set of all $\boldsymbol{v} \in \mathbb{N}^n$ such that $Val(S^{\boldsymbol{v}}, p\boldsymbol{v}) = 1$.

Intuitively, *safe*(p) contains all $\boldsymbol{v} \in \mathbb{N}^n$ such that Player $\Box$ can play "safely" in $p\boldsymbol{v}$, i.e., without ever running out of any resource. The set *cover*(p) contains all $\boldsymbol{v} \in \mathbb{N}^n$ such that Player $\Box$ can play safely in $p\boldsymbol{v}$ without ever reloading any resource above the capacity specified by $\boldsymbol{v}$.

Obviously, both *safe*(p) and *cover*(p) are upwards-closed with respect to component-wise ordering, and hence these sets are fully described by the corresponding finite sets of minimal elements. In [10], the following is proven:

Theorem 5. *Let $\mathcal{A}$ be a consumption game with n counters. Further, let ℓ be the maximal $|\boldsymbol{v}_i| \neq \omega$, where $\boldsymbol{v}$ is an update vector used in $\mathcal{A}$. Then*

- *the emptiness problems for safe(s) and cover(s) are **co-NP**-complete and solvable in $O(n! \cdot |Q|^{n+1})$ time;*
- *the membership problem for safe(p) is **PSPACE**-hard, and the set of all minimal elements of safe(p) is computable in time $(n \cdot \ell \cdot |Q|)^{O(n)}$;*
- *the membership problem for cover(p) is **PSPACE**-hard, and the set of all minimal elements of cover(p) is computable in time $(n \cdot \ell \cdot |Q|)^{O(n \cdot n!)}$.*

Note that all of the problems considered in Theorem 5 are solvable in polynomial time when n and ℓ are fixed.

For the special cases of *one-player* and *decreasing* consumption games, it is possible to design even more efficient algorithms (a consumption game is one-player if $Q_\Diamond = \emptyset$, and decreasing if every counter is either reloaded or decreased along every cycle in the graph of $\mathcal{A}$). We refer to [10] for details.

3.3 Multiweighted Energy Games

Multiweighted energy games [22], also known as *generalized energy games* [16], are multi-counter games with $\mathbb{Z}$-semantics such that

- there are no stochastic control states (i.e., $Q_\bigcirc = \emptyset$),
- the update vectors are elements of $\mathbb{Z}^n$.

Further, there is a special variant of this model called *multiweighted energy games with weak upper bound* [22], where the counters are constrained by a given vector $\boldsymbol{b} \in \mathbb{N}^n$. Whenever a counter i should exceed $\boldsymbol{b}_i$, it is immediately truncated to $\boldsymbol{b}_i$.

The payoff functions studied in [22] are closely related to *zero safety* and *zero safety with upper bound* $\boldsymbol{u}$ that have been defined in Section 3.2. The only difference is that Player $\Box$ should avoid decreasing a counter *strictly below* zero. Formally, let

- S_0 be a function which to every run assigns either 1 or 0 depending on whether or not the run avoids visiting configurations of the form $q\boldsymbol{v}$ where $\boldsymbol{v}_i < 0$ for some $1 \leq i \leq n$;
- $S_0^{\boldsymbol{u}}$ be a function which to every run assigns either 1 or 0 depending on whether or not the run avoids visiting configurations of the form $q\boldsymbol{v}$ where $\boldsymbol{v}_i < 0$ or $\boldsymbol{v}_i > \boldsymbol{u}_i$ for some $1 \leq i \leq n$.

For the subclass of multiweighted energy games with only *one counter*, the following results can be derived from [4]:

Theorem 6. *Let $\mathcal{A}$ be a multiweighted energy game with one counter. Then*

- *the problem whether $Val(S_0, p(0)) = 1$ for a given $p \in Q$ is in* ***UP*** $\cap$ ***co-UP****; if $Q_\diamond = \emptyset$, then the problem is in* ***P****;*
- *the problem whether $Val(S_0^{\boldsymbol{u}}, p(0)) = 1$ for a given $p \in Q$ is* ***EXPTIME****-complete; if $Q_\diamond = \emptyset$, then the problem is* ***NP****-hard and in* ***PSPACE****.*

Further, for multiweighted energy games with one counter and weak upper bound, the problem whether $Val(S_0, p(0)) = 1$ for a given $p \in Q$ is in **NP** $\cap$ **co-NP**, and if $Q_\diamond = \emptyset$, then it is solvable in polynomial time [4].

By applying the results of [11] (see also Theorem 4), one can deduce the following:

Theorem 7. *Let $\mathcal{A}$ be a multiweighted energy game with n counters. The problem whether $Val(S_0, p\mathbf{0}) = 1$ for a given $p \in Q$ is* ***EXPSPACE****-hard and in n-***EXPTIME****.*

Recall that update vectors in multiweighted energy games may contain arbitrarily large integers encoded in binary, and hence the $(n-1)$-**EXPTIME** upper bound of Theorem 4 increases to n-**EXPTIME** in Theorem 7.

The main results about multiweighted energy games proven in [22] concern the $S_0^{\boldsymbol{u}}$ payoff function and can be summarized as follows:

Theorem 8. *Let $\mathcal{A}$ be a multiweighted energy game with n counters. The problem whether $Val(S_0^{\boldsymbol{u}}, p\mathbf{0}) = 1$ for given $p \in Q$ and $\boldsymbol{u} \in \mathbb{N}^n$ is* ***EXPTIME****-complete. If $Q_\diamond = \emptyset$, then the problem is* ***PSPACE****-complete.*

In [22], it is also shown that Theorem 8 remains valid for multiweighted energy games with weak upper bound and S_0 payoff function.

The complexity of *initial credit problem* for multiweighted energy games with S_0 payoff function is studied in greater detail in [16]. An instance of the initial credit problem is a control state p of a multiweighted energy game $\mathcal{A}$ with n counters, and the question is whether there is some $\boldsymbol{v} \in \mathbb{Z}^n$ such that $Val(S_0, p\boldsymbol{v}) = 1$. It follows from the results of [11] that the initial credit problem is solvable in **PSPACE** for eVASS games, and hence in **EXPSPACE** for multiweighted energy games (cf. the comments after Theorem 7). In [16], the following is shown:

Theorem 9. *The initial credit problem for multiweighted energy games is* ***co-NP****-complete.*

3.4 One-Counter Games and MDPs

A *one-counter game* is a multi-counter game $\mathcal{A}$ with $\mathbb{Z}$-semantics where $\mathcal{A}$ has only one counter and the counter updates range over $\{-1, 0, 1\}$. If we also have that $Q_\diamond = 0$ (or $Q_\Box = \emptyset$), then $\mathcal{A}$ is a *maximizing* (or *minimizing*) *one-counter MDP*.

One-counter games and MDPs have so far been studied with the following payoff functions:

- *cover negatives*, denoted by *CN*, which to every run assigns either 1 or 0 depending on whether or not *lim inf* of all counter values visited along the run is equal to $-\infty$;

- *zero rechability*, denoted by Z, which to every run assigns either 1 or 0 depending on whether or not the run visits a configuration with zero counter;
- *selective zero rechability*, denoted by Z_T, where $T \subseteq Q$. The function Z_T assigns to every run either 1 or 0 depending on whether or not the run visits a configuration $q(0)$ where $q \in T$ and the counter value remains non-negative in all configurations preceding this visit;
- *termination time*, denoted by T, which to every run assigns the number of transitions performed before visiting a configuration with zero counter for the first time. If a run does not visit a configuration with zero counter at all, then T returns ∞.

Maximizing MDPs have been first studied in [7], where the following results are proven:

Theorem 10. *Let $\mathcal{A}$ be a maximizing one-counter MDP.*

- *For all $p \in Q$ and $i \in \mathbb{Z}$, the value $Val(CN, p(i))$ is rational, independent of i, and computable in polynomial time. Further, there is a counterless strategy σ constructible in polynomial time which is optimal in every configuration of $\mathcal{A}$.*
- *The problem whether $Val(Z, p(i)) = 1$ for a given configuration $p(i)$ is in* ***P***. *Further, there is counterless strategy σ constructible in polynomial time which is optimal in every configuration $q(j)$ such that $Val(Z, q(j)) = 1$.*
- *The problem whether Player $\Box$ has a strategy σ for a given configuration $p(i)$ such that $\mathbb{E}^{\sigma}_{p(i)}[Z_T] = 1$ is* ***PSPACE****-hard and solvable in exponential time. Moreover, a finite description of σ (if it exists) is computable in exponential time.*

The last item of Theorem 10 requires some comment. First, $Val(Z_T, p(i)) = 1$ does *not* necessarily imply the existence of an optimal strategy in $p(i)$ (as opposed to *zero reachability* payoff considered in the second item of Theorem 10). In fact, the decidability of the problem whether $Val(Z_T, p(i)) = 1$ for a given $p(i)$ is still open. The strategy σ considered in the last item of Theorem 10 (if it exists) can be constructed so that it is *ultimately periodic* in the sense that for a sufficiently large counter value k, the behaviour of σ in a configuration $q(k)$ depends only on $k \bmod c$, where c is a constant depending only on $\mathcal{A}$ whose value is at most exponential in $||\mathcal{A}||$.

The results about maximizing one-counter MDPs with *zero reachability* payoff have been extended to one-counter games in [5] as follows:

Theorem 11. *Let $\mathcal{A}$ be a one-counter game. The problem whether $Val(Z, p(i)) = 1$ for a given configuration $p(i)$ is in* ***NP*** $\cap$ ***co-NP***. *For one-counter MDPs (both maximizing and minimizing), the same problem is in* ***P***.

Improving the **NP** $\cap$ **co-NP** upper bound of Theorem 11 would require a breakthrough, because the problem whether $Val(Z, p(i)) = 1$ (in one-counter games) is at least as hard as Condon's [17] quantitative reachability problem for finite-state simple stochastic games. Furthermore, it is also shown in [5] that if $Val(Z, p(i)) = 1$, then Player $\Box$ has a counterless optimal strategy in $p(i)$. Similarly, if $Val(Z, p(i)) < 1$, then Player $\Diamond$ has a simple strategy π^* (using finite memory, linearly bounded in the number of control states) that ensures $\mathbb{E}^{\sigma,\pi^*}_{p(i)}[Z] < 1 - \delta$ for some $\delta > 0$, regardless of σ. Such strategies for both players are shown computable in non-deterministic polynomial time for one-counter games, and in deterministic polynomial time for (both maximizing and minimizing) one-counter MDPs.

In general, $Val(Z, p(i))$ may be irrational, even if $Q_\Box = Q_\Diamond = \emptyset$. Hence, the value cannot be computed precisely in general, but it can be effectively approximated up to an arbitrarily small additive error $\varepsilon > 0$, as the following result of [6] shows:

Theorem 12. *There is an algorithm which inputs a one-counter game $\mathcal{A}$, a configuration $p(i)$ of $\mathcal{A}$, and a rational $\varepsilon > 0$, and outputs a rational number v such that $|Val(Z, p(i)) - v| \leq \varepsilon$, and (a finite description of) ε-optimal strategies for both players. The algorithm runs in non-deterministic exponential time; if $\mathcal{A}$ is a maximizing one-counter MDP, then it runs in deterministic exponential time.*

A similar result for maximizing one-counter MDPs with *termination time* payoff was achieved in [12], together with a lower bound showing that approximating the value in one-counter MDPs with termination time payoff is computationally difficult. More precisely, the following holds:

Theorem 13. *There is a deterministic exponential-time algorithm which inputs a maximizing one-counter MDP $\mathcal{A}$, a configuration $p(i)$ of $\mathcal{A}$, and a rational $\varepsilon > 0$, and outputs a rational number v such that $|Val(T, p(i)) - v| \leq \varepsilon$, and (a finite description of) an ε-optimal strategy for Player $\Box$.*

Further, $Val(T, p(1))$ cannot be approximated up to the additive error $1/3$ in polynomial time unless $\mathbf{P} = \mathbf{NP}$.

The lower bound of Theorem 13 is proven in two phases, which are relatively independent. First, it is shown that given a propositional formula φ, one can efficiently compute a one-counter MDP $\mathcal{A}$, a configuration $p(K)$ of $\mathcal{A}$, and a number N such that $Val(T, p(K))$ is either $N - 1$ or N depending on whether φ is satisfiable or not, respectively. Interestingly, an optimal strategy for Player $\Box$ in the configurations of $\mathcal{A}$ is not counterless but *ultimately periodic* (cf. the comments after Theorem 10). The numbers K and N are exponential in $||\varphi||$, which means that their encoding size is polynomial. Here, the technique of encoding propositional assignments into counter values presented in [24] is used, but some specific gadgets need to be invented to deal with termination time payoff. The first part already implies that approximating $Val(p(i))$ is computationally hard. In the second phase, it is shown that the same holds also for configurations where the counter is initiated to 1. This is achieved by employing another gadget which just increases the counter to an exponentially high value with a sufficiently large probability.

The question whether the results of Theorem 13 can be extended to one-counter games is open. It is also not clear whether Player $\Box$ *always* has an ultimately periodic optimal strategy in one-counter MDPs (and games) with termination time payoff.

3.5 Energy Markov Decision Processes

An *energy MDP* is a multi-counter game $\mathcal{A}$ with $\mathbb{Z}$-semantics where $\mathcal{A}$ has only one counter, $Q_\Diamond = \emptyset$, and the update vectors are elements of $\mathbb{Z}^n$. Energy MDPs have been studied in [15] with a payoff function which combines *zero safety* and *parity* requirements. Assume that every control state is assigned a positive integer *priority*. We define a payoff function *ZP* which assigns to every run either 1 or 0, depending on whether or not the following conditions are satisfied:

- the run avoids visiting a configuration with negative counter value;
- the minimum priority of a control state visited infinitely often along the run is even.

The following theorem is proven in [15].

Theorem 14. *Let $\mathcal{A}$ be an energy MDP and $p \in Q$. The problem whether there exists $k \in \mathbb{N}$ and a strategy σ such that $\mathbb{E}^{\sigma}_{p(k)}[ZP] = 1$ is in **NP** $\cap$ **co-NP**.*

It is open whether this theorem can be extended to energy games (with stochastic states).

3.6 Solvency Games

A *solvency game*, introduced in [3], is a multi-counter game $\mathcal{A}$ with $\mathbb{Z}$-semantics where $\mathcal{A}$ has only one counter, $Q_{\Box} = \{p\}$, $Q_{\bigcirc} = \{s_1, \ldots, s_m\}$, the outgoing transitions of p are precisely $(p, 0, s_j)$ for all $1 \leq j \leq m$, and all outgoing transitions of every s_j are of the form (s_j, d, p) where $d \in \mathbb{Z}$. Hence, a solvency game can be seen as a simple maximizing one-counter MDP except that counter updates are arbitrary integers.

Solvency games have been studied with *survival* payoff function, which is similar to *zero safety* of Section 3.2. Formally, we define a function S which to every run assigns either 1 or 0 depending on whether or not the run avoids visiting a configuration with a non-positive (i.e., ≤ 0) counter value.

Due to the simplicity of solvency games, one may be tempted to conclude that for a sufficiently large i, an optimal move in $p(i)$ is to select $s_j(i)$ with a maximal *expected counter change* given by

$$\sum_{(s_j,d,p)\in\delta} d \cdot P((s_j, d, p)) .$$

In [3], it is shown that this hypothesis is *incorrect* for general solvency games, and holds only under a suitable technical condition. It is also shown how to compute an optimal strategy for Player $\Box$ if this technical condition holds.

The results about one-counter MDPs are of course applicable to solvency games. Since counter updates in solvency games are arbitrary integers encoded in binary, a straightforward translation of solvency games into one-counter MDPs is exponential (cf. the comments after Theorem 7). In some cases, this translation can be avoided and the results for one-counter MDPs carry over to solvency games immediately. In particular, in [7] it is explicitly mentioned that the questions whether $Val(S, p(i))$ is >0, $=1$, $=0$, or <1, where $p(i)$ is a configuration of a solvency game, are all solvable in polynomial time.

In the light of the previously presented results about one-counter MDPs and games, a natural conjecture is that Player $\Box$ has an ultimately periodic optimal strategy in solvency games with survival payoff. This conjecture is open.

4 Conclusions

The results summarized in the previous section show that interesting subclasses of multi-counter games admit efficient algorithmic analysis, at least if some of the crucial parameters (such as the number of counters) are fixed. The existing works are scattered over various variants of multi-counter games which makes them difficult to compare because of subtle differences in definitions. On the other hand, the area is rather dynamic and offers a number of challenging open problems.

References

1. Abdulla, P.A., Ben Henda, N., de Alfaro, L., Mayr, R., Sandberg, S.: Stochastic Games with Lossy Channels. In: Amadio, R.M. (ed.) FOSSACS 2008. LNCS, vol. 4962, pp. 35–49. Springer, Heidelberg (2008)
2. Baier, C., Bertrand, N., Schnoebelen, P.: On Computing Fixpoints in Well-Structured Regular Model Checking, with Applications to Lossy Channel Systems. In: Hermann, M., Voronkov, A. (eds.) LPAR 2006. LNCS (LNAI), vol. 4246, pp. 347–361. Springer, Heidelberg (2006)
3. Berger, N., Kapur, N., Schulman, L., Vazirani, V.: Solvency games. In: Proceedings of FST&TCS 2008. Leibniz International Proceedings in Informatics, vol. 2, pp. 61–72. Schloss Dagstuhl–Leibniz-Zentrum für Informatik (2008)
4. Bouyer, P., Fahrenberg, U., Larsen, K., Markey, N., Srba, J.: Infinite Runs in Weighted Timed Automata with Energy Constraints. In: Cassez, F., Jard, C. (eds.) FORMATS 2008. LNCS, vol. 5215, pp. 33–47. Springer, Heidelberg (2008)
5. Brázdil, T., Brožek, V., Etessami, K.: One-counter stochastic games. In: Proceedings of FST&TCS 2010. Leibniz International Proceedings in Informatics, vol. 8, pp. 108–119. Schloss Dagstuhl–Leibniz-Zentrum für Informatik (2010)
6. Brázdil, T., Brožek, V., Etessami, K., Kučera, A.: Approximating the Termination Value of One-Counter MDPs and Stochastic Games. In: Aceto, L., Henzinger, M., Sgall, J. (eds.) ICALP 2011, Part II. LNCS, vol. 6756, pp. 332–343. Springer, Heidelberg (2011)
7. Brázdil, T., Brožek, V., Etessami, K., Kučera, A., Wojtczak, D.: One-counter Markov decision processes. In: Proceedings of SODA 2010, pp. 863–874. SIAM (2010)
8. Brázdil, T., Brožek, V., Forejt, V., Kučera, A.: Reachability in recursive Markov decision processes. Information and Computation 206(5), 520–537 (2008)
9. Brázdil, T., Brožek, V., Kučera, A., Obdržálek, J.: Qualitative reachability in stochastic BPA games. Information and Computation 208(7), 772–796 (2010)
10. Brázdil, T., Chatterjee, K., Kučera, A., Novotný, P.: Efficient Controller Synthesis for Consumption Games with Multiple Resource Types. In: Madhusudan, P., Seshia, S.A. (eds.) CAV 2012. LNCS, vol. 7358, pp. 23–38. Springer, Heidelberg (2012)
11. Brázdil, T., Jančar, P., Kučera, A.: Reachability Games on Extended Vector Addition Systems with States. In: Abramsky, S., Gavoille, C., Kirchner, C., Meyer auf der Heide, F., Spirakis, P.G. (eds.) ICALP 2010, Part II. LNCS, vol. 6199, pp. 478–489. Springer, Heidelberg (2010)
12. Brázdil, T., Kučera, A., Novotný, P., Wojtczak, D.: Minimizing Expected Termination Time in One-Counter Markov Decision Processes. In: Czumaj, A., Mehlhorn, K., Pitts, A., Wattenhofer, R. (eds.) ICALP 2012, Part II. LNCS, vol. 7392, pp. 141–152. Springer, Heidelberg (2012)
13. Chaloupka, J.: Z-Reachability Problem for Games on 2-Dimensional Vector Addition Systems with States Is in P. In: Kučera, A., Potapov, I. (eds.) RP 2010. LNCS, vol. 6227, pp. 104–119. Springer, Heidelberg (2010)
14. Chatterjee, K., Doyen, L.: Energy Parity Games. In: Abramsky, S., Gavoille, C., Kirchner, C., Meyer auf der Heide, F., Spirakis, P.G. (eds.) ICALP 2010, Part II. LNCS, vol. 6199, pp. 599–610. Springer, Heidelberg (2010)
15. Chatterjee, K., Doyen, L.: Energy and Mean-Payoff Parity Markov Decision Processes. In: Murlak, F., Sankowski, P. (eds.) MFCS 2011. LNCS, vol. 6907, pp. 206–218. Springer, Heidelberg (2011)

16. Chatterjee, K., Doyen, L., Henzinger, T., Raskin, J.F.: Generalized mean-payoff and energy games. In: Proceedings of FST&TCS 2010. Leibniz International Proceedings in Informatics, vol. 8, pp. 505–516. Schloss Dagstuhl–Leibniz-Zentrum für Informatik (2010)
17. Condon, A.: The complexity of stochastic games. Information and Computation 96(2), 203–224 (1992)
18. Etessami, K., Wojtczak, D., Yannakakis, M.: Recursive Stochastic Games with Positive Rewards. In: Aceto, L., Damgård, I., Goldberg, L.A., Halldórsson, M.M., Ingólfsdóttir, A., Walukiewicz, I. (eds.) ICALP 2008, Part I. LNCS, vol. 5125, pp. 711–723. Springer, Heidelberg (2008)
19. Etessami, K., Yannakakis, M.: Recursive Markov Decision Processes and Recursive Stochastic Games. In: Caires, L., Italiano, G.F., Monteiro, L., Palamidessi, C., Yung, M. (eds.) ICALP 2005. LNCS, vol. 3580, pp. 891–903. Springer, Heidelberg (2005)
20. Etessami, K., Yannakakis, M.: Efficient Qualitative Analysis of Classes of Recursive Markov Decision Processes and Simple Stochastic Games. In: Durand, B., Thomas, W. (eds.) STACS 2006. LNCS, vol. 3884, pp. 634–645. Springer, Heidelberg (2006)
21. Etessami, K., Yannakakis, M.: Recursive Concurrent Stochastic Games. In: Bugliesi, M., Preneel, B., Sassone, V., Wegener, I. (eds.) ICALP 2006, Part II. LNCS, vol. 4052, pp. 324–335. Springer, Heidelberg (2006)
22. Fahrenberg, U., Juhl, L., Larsen, K., Srba, J.: Energy Games in Multiweighted Automata. In: Cerone, A., Pihlajasaari, P. (eds.) ICTAC 2011. LNCS, vol. 6916, pp. 95–115. Springer, Heidelberg (2011)
23. Filar, J., Vrieze, K.: Competitive Markov Decision Processes. Springer (1996)
24. Kučera, A.: The complexity of bisimilarity-checking for one-counter processes. Theoretical Computer Science 304(1-3), 157–183 (2003)
25. Maitra, A., Sudderth, W.: Finitely additive stochastic games with Borel measurable payoffs. International Journal of Game Theory 27, 257–267 (1998)
26. Martin, D.: The determinacy of Blackwell games. Journal of Symbolic Logic 63(4), 1565–1581 (1998)
27. Puterman, M.: Markov Decision Processes. Wiley (1994)