4.1. Observations and labeled game graphs. In partial-observation games, a coloring of the state space defines classes of indistinguishable states called observations. Player 1 does not see the current state of the game, but only its color. Edges of the game graph carry a label which is used by player 1 to select edges. Player 2 resolves the non-determinism.

A partial-observation game $G = (Q, \Sigma, \Delta)$ with weight function $w : \Delta \to \mathbb{Z}$ and observations $\text{Obs} \subseteq 2^Q$ consists of

- $Q$ a finite set of states,
- $\Sigma$ a finite alphabet of actions,
- $\Delta \subseteq Q \times \Sigma \times Q$ a set of labeled transitions such that for all $q \in Q$ and $\sigma \in \Sigma$, there exists (at least one) $q' \in Q$ such that $(q, \sigma, q') \in \Delta$,
- $\text{Obs}$ a partition of $Q$, and for each $q \in Q$, let $\text{obs}(q)$ the unique observation $o \in \text{Obs}$ such that $q \in o$.

For $s \subseteq L$ and $\sigma \in \Sigma$, we denote by $\text{post}^G(s) = \{q' \in Q \mid \exists q \in s : (q, \sigma, q') \in \Delta\}$ the set of $\sigma$-successors of $s$. A game with perfect observation is such that $\text{Obs} = \{\{q\} \mid q \in Q\}$. A partial-observation game is blind if $\text{Obs} = \{Q\}$.

The game is played in rounds. In each round, if the current state is $q$, player 1 does not see the state $q$ but gets the observation $\text{obs}(q)$. Player 1 selects an action $\sigma \in \Sigma$, and then player 2 chooses a state $q'$ such that $(q, \sigma, q') \in \Delta$. The game proceeds to the next round in state $q'$.

4.2. Example. In the following (unweighted) partial-observation game, the observations are $o_1 = \{q_1\}$, $o_2 = \{q_2, q_2'\}$, $o_3 = \{q_3, q_3'\}$, and $o_4 = \{q_4\}$. From the initial state $q_1$, there is no winning strategy for player 1 to reach $T = \{q_4\}$. This is because no matter the observation-based strategy $\alpha$ for player 1, there exists a play $\rho$ compatible with $\alpha$ that never visits $q_3$. The play $\rho$ is of the form $(q_1, \Sigma, q_x, q_3, \Sigma)^{n}$ where $q_x = q_2$ if $\sigma_x = a$, and $q_x = q_2'$ if $\sigma_x = b$. Note that this definition has no circularity because the value of $\sigma_x$ (chosen by $\alpha$) is independent of $q_x \in \{q_2, q_2'\}$ since $\text{obs}(q_2) = \text{obs}(q_2')$.

4.3. Winning strategy. A strategy for player 1 is a function $\alpha : (Q \cdot \Sigma)^{n}Q \to \Sigma$ such that for all $\rho = q_0 \sigma_0 q_1 \sigma_1 q_2 \ldots q_n$ and $\rho' = q_0' \sigma_0 q_1' \sigma_1 q_2' \ldots q_n'$, if $\text{obs}(q_i) = \text{obs}(q_i')$ for all $0 \leq i \leq n$, then $\alpha(\rho) = \alpha(\rho')$. We say that strategies are observation-based.
An infinite play \( \rho = q_0 \sigma_0 q_1 \sigma_1 q_2 \ldots \) is compatible with a strategy \( \alpha \) if \( \sigma_i = \alpha(q_0 \sigma_0 q_1 \ldots q_i) \) and \((q_i, \sigma_i, q_{i+1}) \in \Delta\) for all \( i \geq 0 \).

Given an initial credit \( c_0 \in \mathbb{N} \), the energy level of a play \( \rho = q_0 \sigma_0 q_1 \sigma_1 q_2 \ldots \) at position \( k \geq 0 \) is \( \text{EL}(\rho, k) = \Sigma_{i=0}^{k-1} w(q_i, \sigma_i, q_{i+1}) \).

A strategy \( \alpha \) for player 1 is winning from state \( q \) with initial credit \( c_0 \) for the energy objective if for all plays \( \rho \) from \( q \) compatible with \( \alpha \), we have \( c_0 + \text{EL}(\rho, k) \geq 0 \) for all \( k \geq 0 \).

The unknown initial credit problem asks to decide, given a partial-observation energy game, an initial state \( q \) and initial credit \( c_0 \), whether there exists a winning strategy for player 1 for the energy objective.

### 4.4. Fixed initial credit.

For an initial state \( q \in Q \) and a fixed initial credit \( c_0 \in \mathbb{N} \), we solve energy games by a reduction to safety games of perfect observation.

Let \( F \) be the set of functions \( f : Q \to \mathbb{Z} \cup \{ \bot \} \). The support of \( f \) is \( \text{supp}(f) = \{ q \in Q \mid f(q) \neq \bot \} \). A function \( f \in F \) stores the possible current states of the game \( G \) together with their worst-case energy level.

We say that a function \( f \) is nonnegative if \( f(q) \geq 0 \) for all \( q \in \text{supp}(f) \). Initially, we set \( f_{c_0}(q_0) = c_0 \) and \( f_{c_0}(q) = \bot \) for all \( q \neq q_0 \). The set \( F \) is ordered by the relation \( \preceq \) such that \( f_1 \preceq f_2 \) if \( \text{supp}(f_1) = \text{supp}(f_2) \) and \( f_1(q) \leq f_2(q) \) for all \( q \in \text{supp}(f_1) \).

For \( \sigma \in \Sigma \), we say that \( f_2 \in F \) is a \( \sigma \)-successor of \( f_1 \in F \) if there exists an observation \( o \in \text{Obs} \) such that \( \text{supp}(f_2) = \text{post}_G^o(\text{supp}(f_1)) \cap o \) and \( f_2(q) = \min \{ f_1(q') + w(q', \sigma, q) \mid q' \in \text{supp}(f_1) \land (q', \sigma, q) \in \Delta \} \) for all \( q \in \text{supp}(f_2) \). Given a sequence \( x = f_0 \sigma_0 f_1 \sigma_1 \ldots f_n \), let \( f_x = f_n \) be the last function in \( x \). Define the safety game \( H = (Q^H, \Sigma, \Delta^H) \) with initial state \( f_{c_0} \) where \( Q^H \) is the smallest subset of \((F \times \Sigma)^* \cdot F\) such that

1. \( f_{c_0} \in Q^H \), and

2. for each sequence \( x \in Q^H \), if (i) \( f_x \) is nonnegative, and (ii) there is no strict prefix \( y \) of \( x \) such that \( f_y \preceq f_x \), then \( x \cdot \sigma \cdot f_2 \in Q^H \) for all \( \sigma \)-successors \( f_2 \) of \( f_x \).

The transition relation \( \Delta^H \) contains the corresponding triples \( (x, \sigma, x \cdot \sigma \cdot f_2) \), and the game is made total by adding self-loops \( (x, \sigma, x) \) to sequences \( x \) without outgoing transitions. We call such sequences the leaves of \( H \). Note that the game \( H \) is acyclic, except for the self-loops on the leaves.

Since the relation \( \preceq \) on nonnegative functions is a well quasi order, the state space \( Q^H \) is finite by König’s Lemma.

Define the safety objective \( \text{Safe}(T) \) in \( H \) where \( T = \{ x \in Q^H \mid f_x \) is nonnegative \} \). Intuitively, a winning strategy in the safety game \( H \) can be extended to an observation-based winning strategy in the energy game \( G \) because whenever a leaf of \( H \) is reached, there exists a \( \preceq \)-smaller ancestor that Player 1 can use to go on in \( G \) using the strategy played from the ancestor in \( H \). The correctness argument is based on the fact that if Player 1 is winning from state \( f \) in \( H \), then he is also winning from all \( f' \preceq f \).

**Theorem 4A.** Let \( G \) be an energy game with partial observation, and let \( c_0 \in \mathbb{N} \) be an initial credit. There exists a winning observation-based strategy in \( G \) for the energy objective with initial credit \( c_0 \) if and only if there exists a winning strategy in \( H \) for the objective \( \text{Safe}(T) \). Hence the fixed initial credit problem is decidable.

### 4.5. Unknown initial credit.

We show that the unknown initial credit problem is undecidable using a reduction from the halting problem for deterministic 2-counter Minsky machines.

**Theorem 4B.** The unknown initial credit problem for energy games with partial observation is undecidable, even for blind games.

Given a (deterministic) 2-counter machine \( M \), we construct a blind energy game \( G_M \) such that \( M \) has an accepting run if and only if there exists an initial credit \( c_0 \in \mathbb{N} \) such that Player 1 has a winning strategy
in $G_M$ for the energy objective. In particular, a strategy that plays a sequence $\bar{\pi}_0 \bar{\pi}_1 \ldots$ (where $\bar{\pi}_i$’s are run traces of $M$) is winning in $G_M$ if and only if all but finitely many $\bar{\pi}_i$’s are accepting run traces of $M$.

The alphabet of $G_M$ is $\Sigma = \delta_M \cup \{\#\}$. The game $G_M$ consists of an initial nondeterministic choice between several gadgets described below. Each gadget checks one property of the sequence of actions played in order to ensure that a trace of an accepting run in $M$ is eventually played. Since the game is blind, it is not possible for player 1 to see which gadget is executed, and therefore the strategy has to fulfill all properties simultaneously.

The gadget in Figure 1 with $\sigma_1 = \#$ checks that the first symbol is a #. If the first symbol is not #, then the energy level drops below 0 no matter the initial credit. The gadget in Figure 2 checks that a certain symbol $\sigma_1$ is always followed by a symbol $\sigma_2$, and it is used to ensure that # is followed by an instruction $(q_I, \cdot, \cdot, \cdot)$, and that every instruction $(q, \cdot, \cdot, q')$ is followed by an instruction $(q', \cdot, \cdot, q'')$, or by # if $q' = q_F$.

The gadget in Figure 3 ensures that # is played infinitely often (and a bit more). If # is played only finitely many times, then the gadget can guess the last # and jump to the middle state where no initial credit would allow to survive.

Finally, we use the gadget in Figure 4 to check that the tests on counter $c$ are correctly executed. It can accumulate in the energy level the increments and decrements of a counter $c$ between the start of a run (i.e., when # occurs) and a zero test on $c$. A **positive cheat** occurs when $(\cdot, 0?, c, \cdot)$ is played while the counter $c$ has positive value. Likewise, a **negative cheat** occurs when $(\cdot, \text{dec}, c, \cdot)$ is played while the counter $c$ has value 0. On reading the symbol #, the gadget can guess that there will be a positive or negative cheat by moving to the states $q_1$ and $q_2$, respectively. In $q_1$, the energy level simulates the operations on the counter $c$ but with opposite effect, thus accumulating the opposite of the counter value. When a positive cheat occurs, the gadget returns to the initial state, thus decrementing the energy level. The state $q_2$ of the gadget is symmetric. A negative cheat costs one unit of energy. Note that the gadget has to go back to its initial state before the next #, as otherwise Player 1 wins. This ensures that the gadget does not monitor a zero-test across two different runs.

The game $G_M$ has such gadgets for each counter. Thus, a strategy in $G_M$ which cheats infinitely often on a counter would not survive no matter the value of the initial credit.

The correctness of this construction is established as follows. First, assume that $M$ has an accepting run.
Figure 4: Gadget to check the zero tests on counter $c$ (assuming $\sigma$ ranges over $\Sigma \setminus \{\#\}$).

$\pi$ with trace $\bar{\pi}$. Then, the strategy playing $(\#\bar{\pi})^\omega$ is winning for the energy objective with initial credit $|\bar{\pi}|$ because an initial credit $|\bar{\pi}|$ is sufficient to survive in the “$\infty$-many $\#$” gadget of Figure 3, as well as in the zero-test gadget of Figure 4 because all zero tests are correct in $\pi$ and the counter values are bounded by $|\bar{\pi}|$.

Second, if there exists a winning strategy in $G_M$ with some finite initial credit, then the sequence played by this strategy can be decomposed into run traces separated by $\#$, and since the strategy survived in the gadget of Figure 4, there must be a point where all run traces played correspond to faithful simulation of $M$ with respect to counter values, thus $M$ has an accepting run. \hfill \Box