Stackelberg Mean-Payoff Games With a Rationally Bounded Adversarial Follower

Shibashis Guha[1]

Jointly with: Mrudula Balachander[2] Jean-François Raskin[3]

[1]Tata Institute of Fundamental Research, India

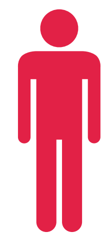
[2]Chennai Mathematical Institute, India

[3]Université libre de Bruxelles, Belgium

September 17th 2020

Stackelberg Games

Leader

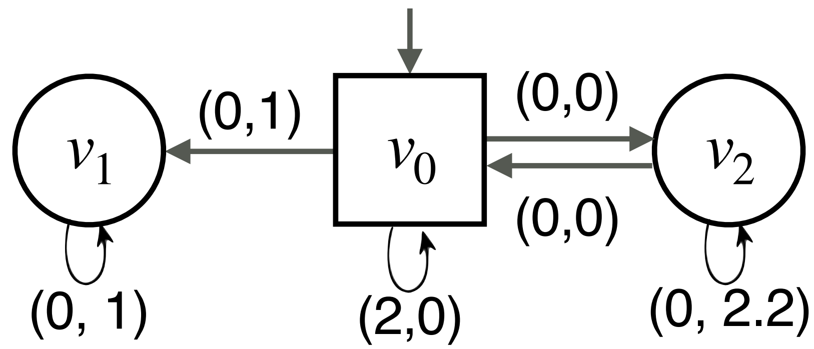


Follower

Two (types of) Players:

Game:

Mean-Payoff Games



Bi-Matrix Games

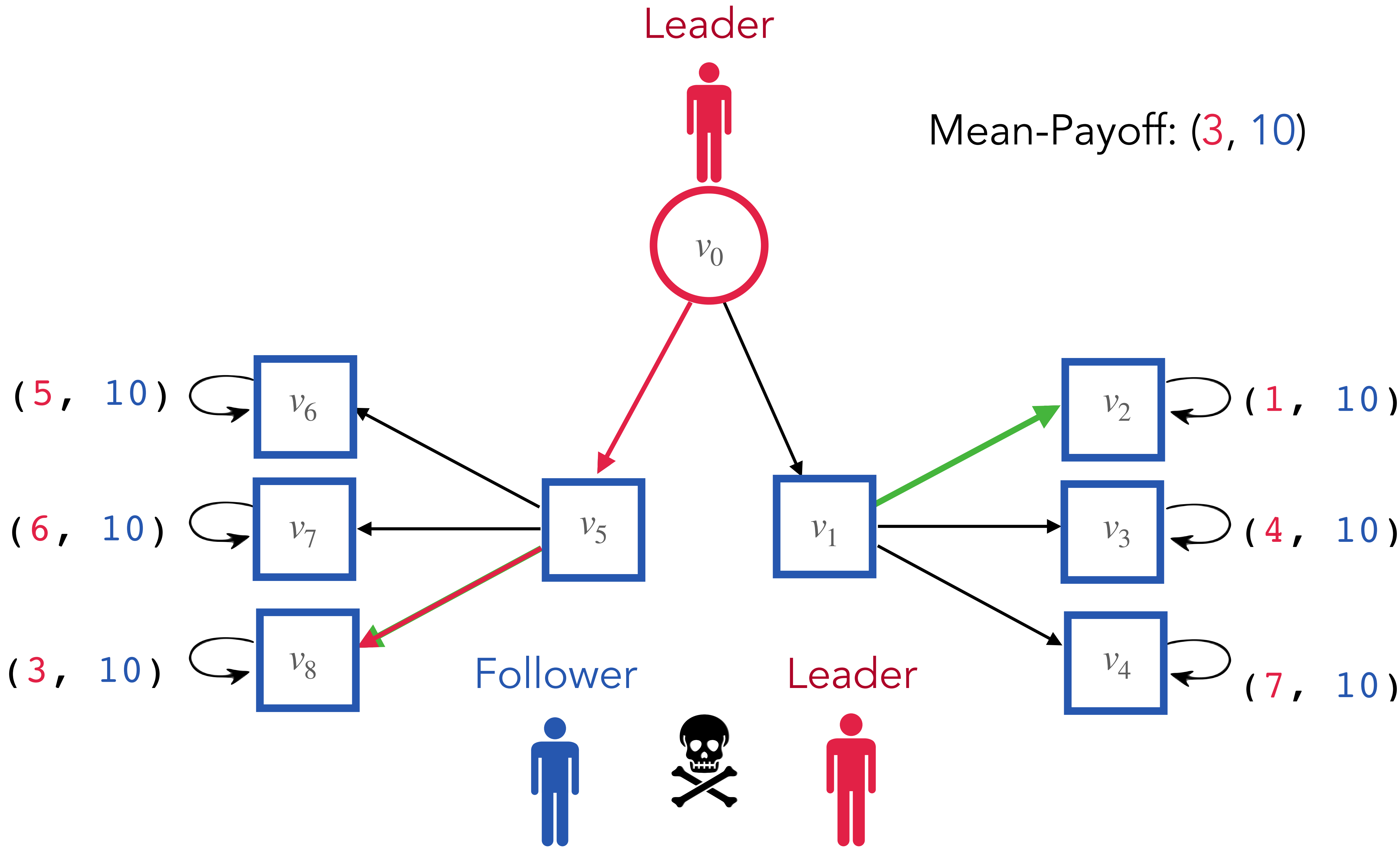
	I	II
I	(1,4)	(4,2)
II	(1,3)	(3,5)

Sequential Move:

1. Leader announces her strategy

2. Follower announces his response to leader's strategy

Cooperative vs Adversarial

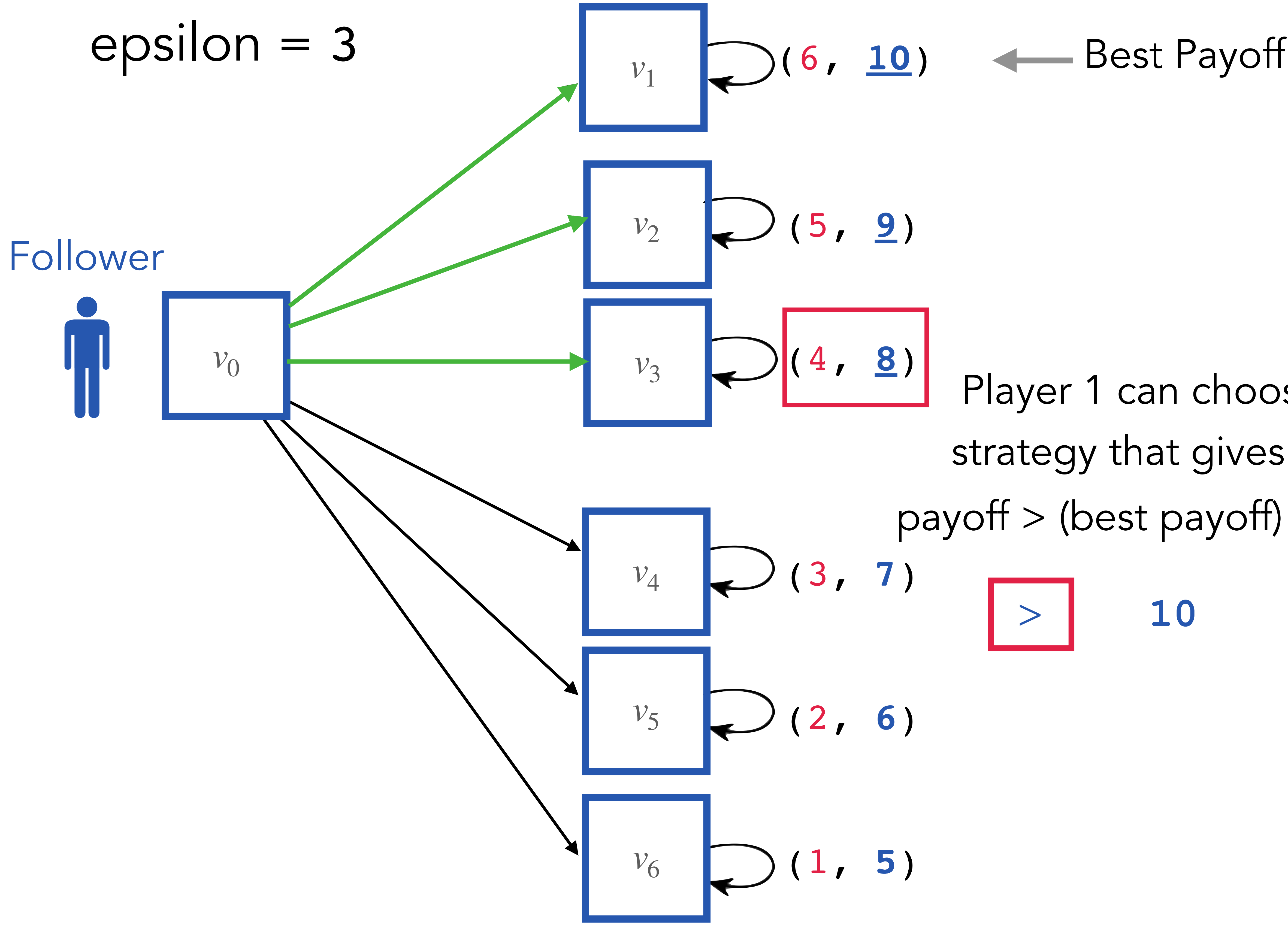


In the adversarial setting, Follower chooses Best-Response which `minimises` payoff of Leader

Epsilon-Optimal Best Response

(Best response may not exist: Filiot, Gentilini and Raskin - ICALP 2020)

epsilon = 3

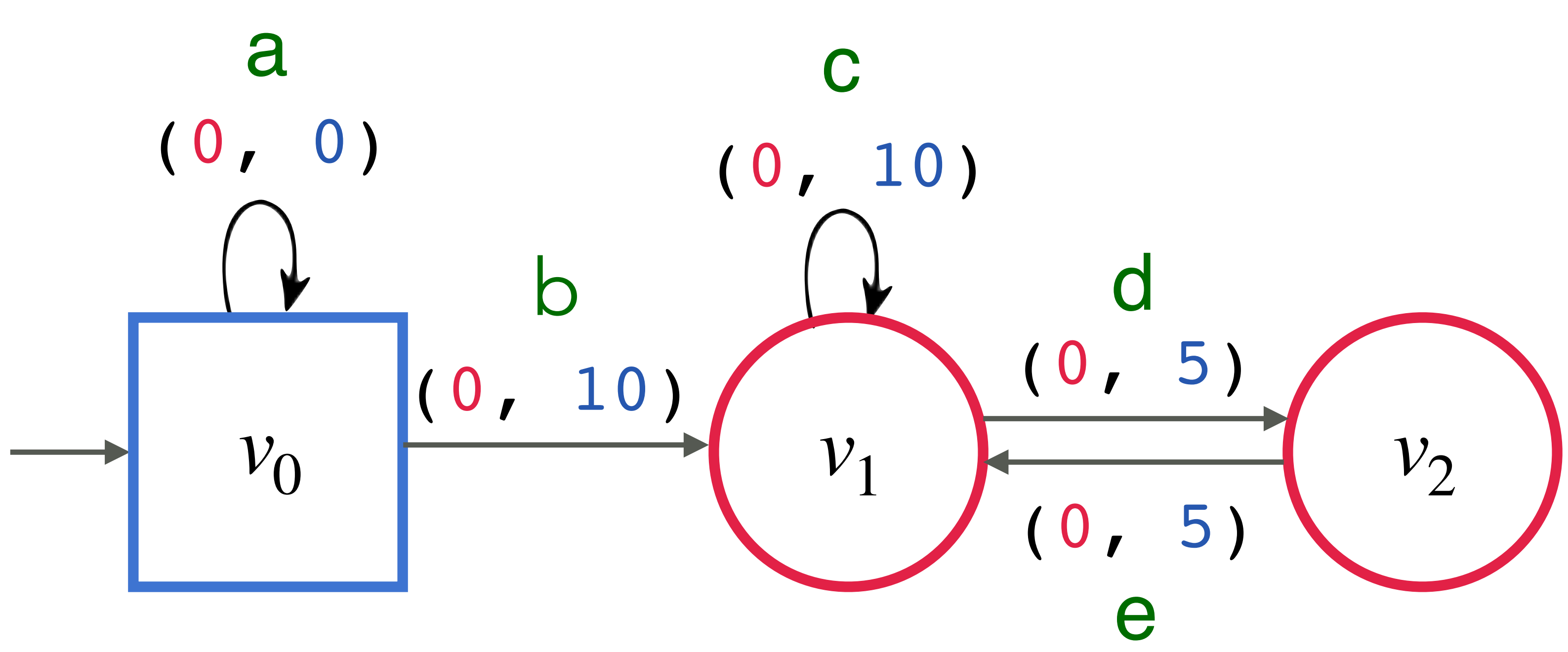


Player 1 can choose any strategy that gives him a payoff > (best payoff) - epsilon

> 10 - 3

Epsilon-Best Responses Always Exist

(Filiot, Gentilini and Raskin - ICALP 2020)



Leader strategy: If $a^k b$, then $(c^k de)^\omega$

Follower strategy: For $\epsilon = 0.1$, play $a^{1000}b$

For $\epsilon = 0.001$, play $a^{100000}b$

Follower is adversarial, bounded rational, i.e. chooses the epsilon-optimal best response

ϵ is fixed

Epsilon-Optimal Adversarial Stackelberg Value ($\mathbf{ASV}^{\epsilon}$)

$\mathbf{ASV}^{\epsilon}$ is the largest mean-payoff value the Leader can obtain when the Follower plays an **adversarial epsilon-best** response.

$$\mathbf{ASV}^{\epsilon}(\sigma_0)(v) = \inf_{\sigma_1 \in \mathbf{BR}^{\epsilon}(\sigma_0)} \text{Mean-Payoff}_0 \, [\text{Outcome}(\sigma_0, \sigma_1)]$$

$$\mathbf{ASV}^{\epsilon}(v) = \sup_{\sigma_0} \mathbf{ASV}^{\epsilon}(\sigma_0)(v)$$

σ_0 : Leader Strategy

σ_1 : Follower Strategy

$\mathbf{BR}^{\epsilon}(\sigma_0)$: Epsilon-Best Response of Follower to Leader's strategy σ_0

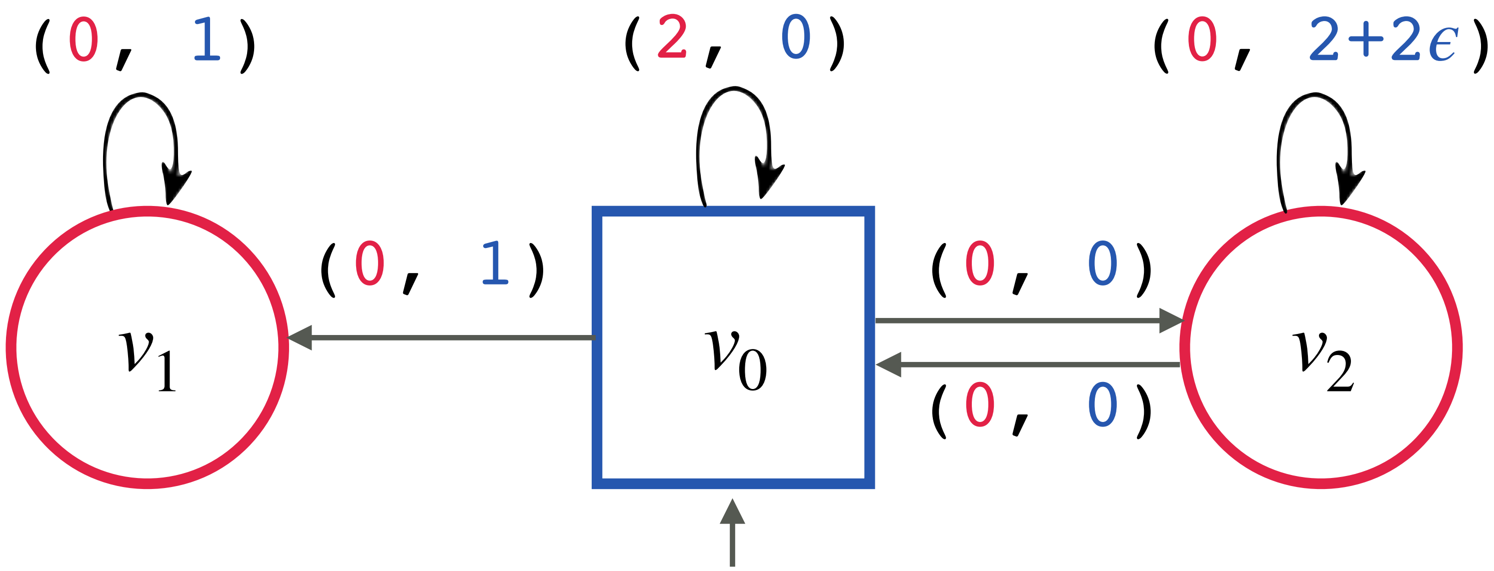
RESULT 1:

$\mathbf{ASV}^{\epsilon}$ is always achievable

There exists a Leader Strategy σ_0 such that

$$\mathbf{ASV}^{\epsilon}(v) = \mathbf{ASV}^{\epsilon}(\sigma_0)(v)$$

Infinite Memory Required for Leader



Leader strategy: $\left((v_0 \to v_0)^k (v_0 \to v_2) \cdot (v_2 \to v_2)^k (v_2 \to v_0)\right)_{k \in \mathbb{N}}$

(Finite Memory) If any deviation, then play $v_2 \to v_0$

The effects of edges (0, 0) become non-negligible and

decrease Leader's mean-payoff

$$\mathbf{ASV}^{\epsilon}(\text{Leader Strategy})(v_0) = 1$$

RESULT 2:

Infinite memory might be required for Leader strategies to achieve the $\mathbf{ASV}^{\epsilon}$

RESULT 3:

Infinite memory might be required for the Follower to play an epsilon-optimal best-response

Threshold Problem:

Is $\mathbf{ASV}^{\epsilon}(v) > c$?

Witnesses and Bad Vertices

$$\Lambda^{\epsilon}(v) = \left\{ (c, d) \in \mathbb{R}^2 \mid \begin{array}{l} \text{From vertex } v\text{, the Follower can ensure that} \\ \text{Leader's payoff} \leqslant c \text{ and Follower's payoff} > d - \epsilon \end{array} \right\}$$

A vertex v is $(c, d)^{\epsilon}$-bad if $(c, d) \in \Lambda^{\epsilon}(v)$

A path π starting from v is a witness for $\mathbf{ASV}^{\epsilon}(v) > c$ if Mean-Payoff of π is (c', d), where $c' > c$ and π does not cross a $(c, d)^{\epsilon}$-bad vertex.

RESULT 4:

$\mathbf{ASV}^{\epsilon}$(v)> c if and only if

there exists a witness

RESULT 6:

We can guess

a regular-witness

in NP-Time

Results

- Results in our work
- Results by Filiot, Gentilini and Raskin, ICALP'20

	Threshold Problem	Computing ASV	Achievability
General Case	NP-Time Finite Memory Strategy	Theory of Reals	No
Fixed Epsilon	NP-Time Finite Memory Strategy	Theory of Reals/ Solving LP in EXPTime	Yes (Requires Infinite Memory)

Stackelberg Mean-payoff Games with a Rationally Bounded Adversarial Follower: https://arxiv.org/abs/2007.07209