Stochastic Best-Effort Strategies for Borel Goals

Florian Zuleger, TU Wien

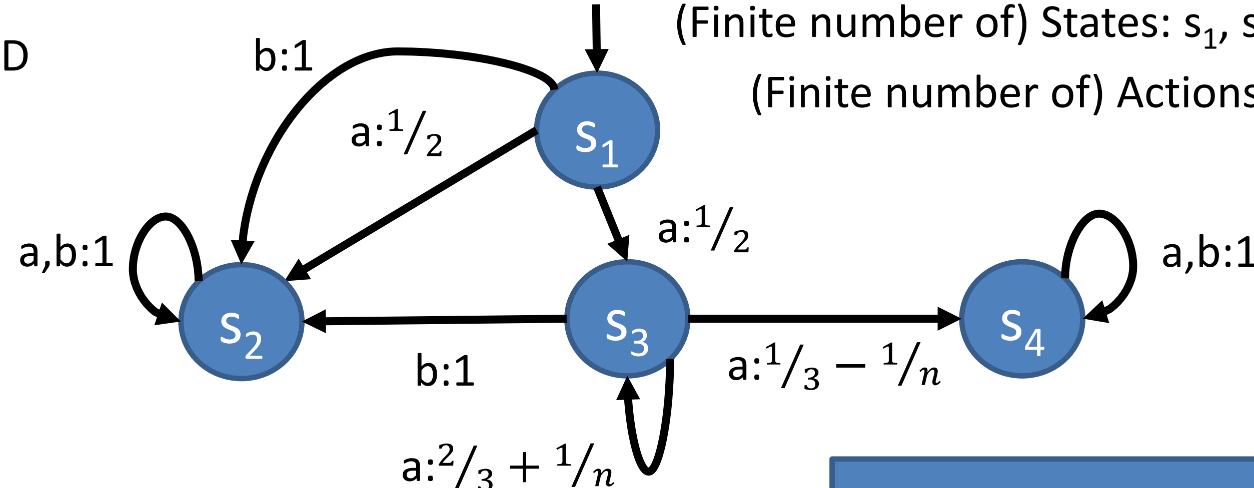
LICS, Boston, 27.6.2023

Joint work with Benjamin Aminof, University of Rome ”La Sapienza”, Giuseppe De Giacomo, University of Rome ”La Sapienza”, University of Oxford,

Sasha Rubin, University of Sydney

Stochastic Domain

Domain D



(Finite number of) States: s_1, s_2, s_3, s_4

(Finite number of) Actions: a, b

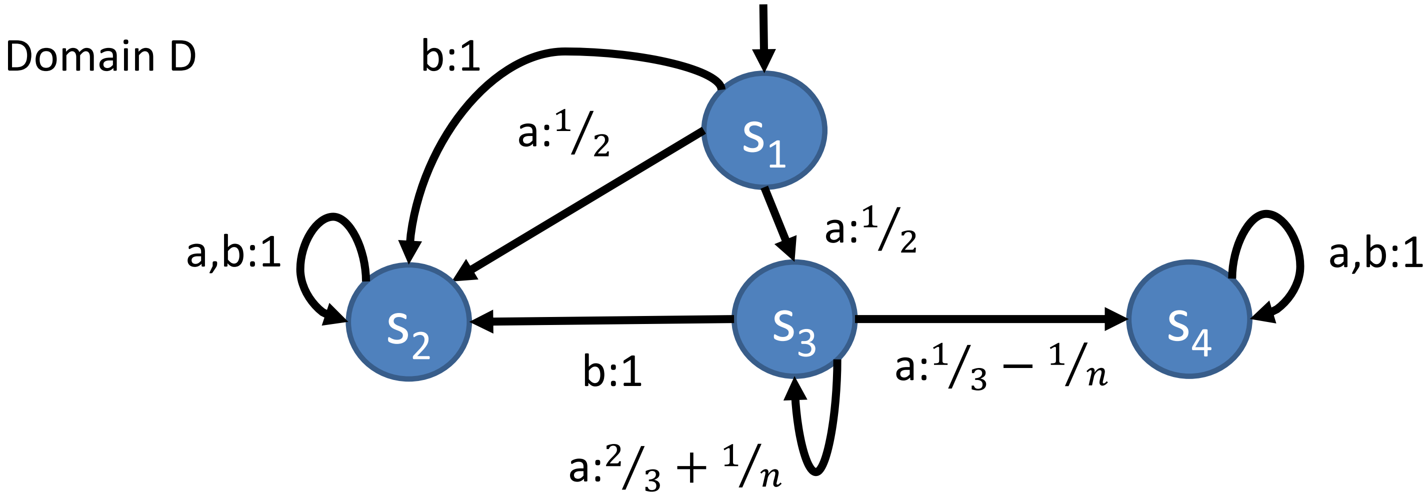
History = ($state_1$,act_1,…,$state_k$,act_k,$state_{k+1}$)
Play = ($state_1$,act_1,…)
Prob = History x Actions → Dist(State)
Strategy = History → Actions
Goal = measureable set of plays
Strategy σ enforces goal G, if the σ-consistent plays satisfy G with probability 1.

e.g., always play action a (this is a memory-less policy)

e.g., set of plays that
- reach s_4
-satisfy the LTL formula FGs_4

strategy σ = "always play action a" does not enforce goal G = "reach s_4",
i.e., $\mu_D(\sigma,G) < 1$

Similar Domains

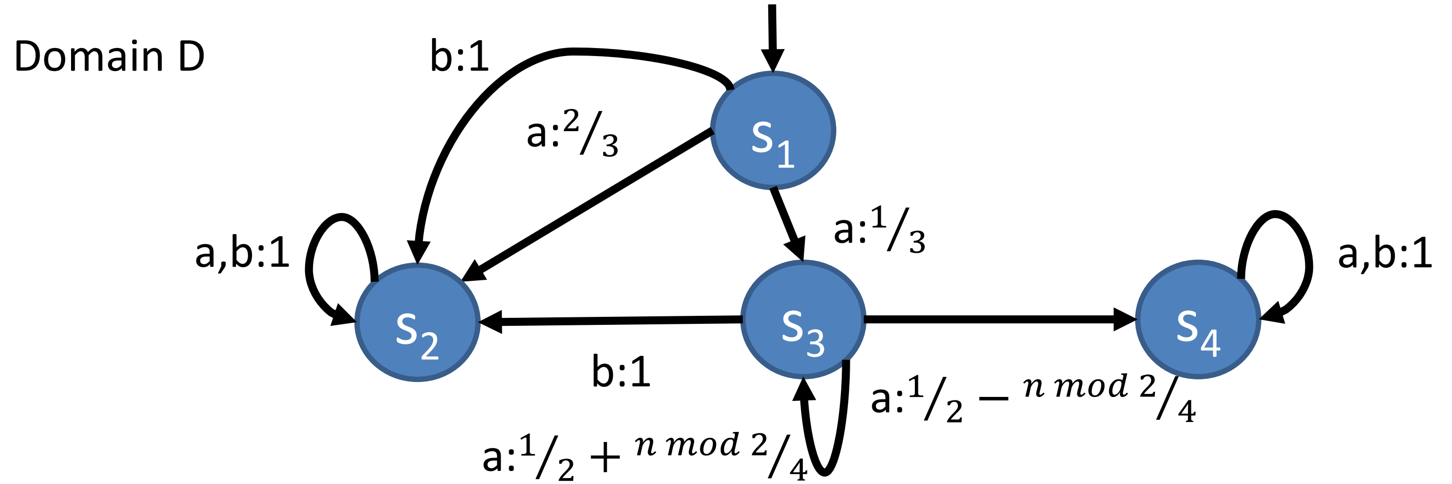


$\text{Supp}_D(\text{hist,act}) = \{ \text{ state } \mid \text{Prob(hist,act)(state)} > 0 \}$

Similar Domains = all domains D and D’ such that

$\text{Supp}_D(\text{hist,act}) = \text{Supp}_{D'}(\text{hist,act})$ for all hist and act

Similar Domains



$Supp_D$(hist,act) = { state | Prob(hist,act)(state) > 0 }

Similar Domains = all domains D and D' such that

$Supp_D$(hist,act) = $Supp_{D'}$(hist,act) for all hist and act

Stochastic Best-effort Strategies

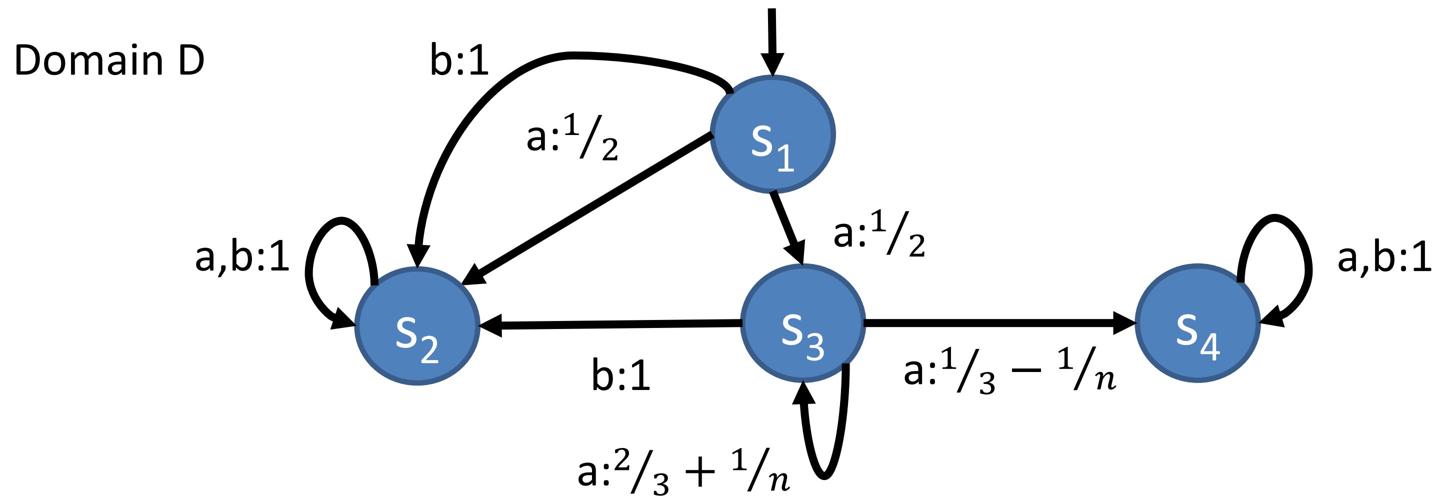
Given some set of similar domains Δ and goal G.

Strategy σ_1 **dominates** σ_2, written $\sigma_1 \geq \sigma_2$, iff $\mu_D(\sigma_1,G) \geq \mu_D(\sigma_2,G)$ for every $D \in \Delta$.

Our proposal:

A strategy σ is **stochastic best-effort** for D and G iff it is maximal wrt $\geq$.

Abstracted Values



stochastic domain D, goal G, strategy σ, history h

$$val_{D,G}(\sigma,h) = \begin{cases} win, & \text{if } \sigma \text{ enforces G assuming history h} \\ lose, & \text{if } \sigma \text{ enforces not G assuming history h} \\ pend, & \text{otherwise} \end{cases}$$

We set $lose < pend < win$.

E.g.,
$val_{D,G}(\sigma,s_1) = pend$
$val_{D,G}(\sigma,s_1 a s_2) = lose$
$val_{D,G}(\sigma,s_1 a s_3) = win$
for the policy σ = "always play a"

A Characterization Result

Lemma
Let G be a goal and Δ be the set of **all domains** similar to some domain.
Then, $\sigma_1 \geq \sigma_2$ iff for all split points h and domains $D \in \Delta$

1. $val_{D,G}(\sigma_1,h) \geq val_{D,G}(\sigma_2,h)$, and
2. not both values are $pend$.

split point =

a history $h=(state_1,act_1,\ldots,state_k,act_k,state_{k+1})$ that is consistent with σ_1 and σ_2, and a state s such that $\sigma_1(h) \neq \sigma_2(h)$

A Characterization Result

Lemma
Let G be a goal and Δ be the set of **all domains** similar to some domain.
Then, $\sigma_1 \geq \sigma_2$ iff for all split points h and domains $D \in \Delta$
1. $val_{D,G}(\sigma_1,h) \geq val_{D,G}(\sigma_2,h)$, and
2. not both values are $pend$.

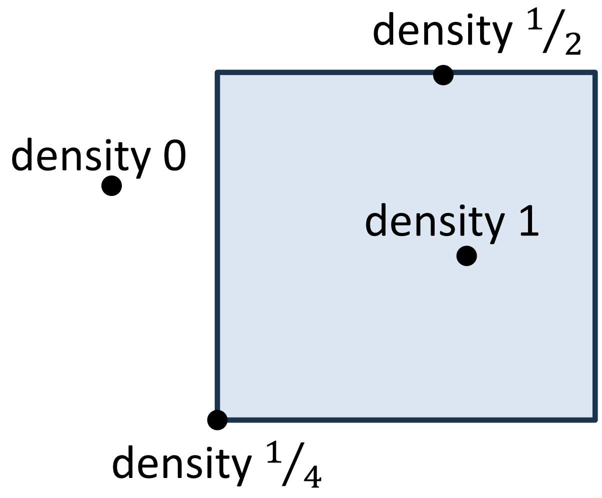
split point =

a history $h=(state_1,act_1,\ldots,state_k,act_k,state_{k+1})$ that is consistent with σ_1 and σ_2, and a state s such that $\sigma_1(h) \neq \sigma_2(h)$

Uses that Δ is closed under **amplifications**:

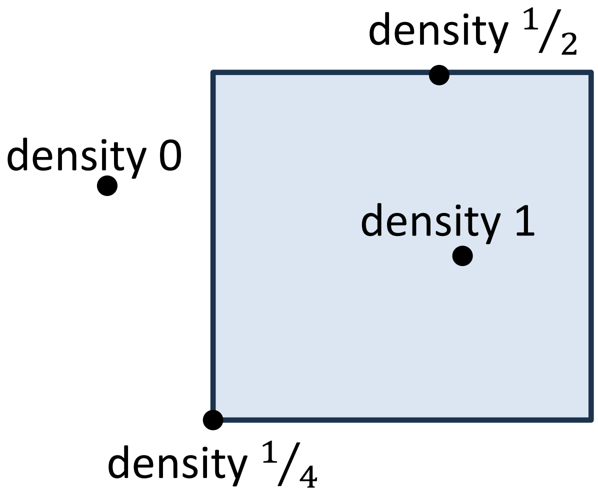
Lemma
Let G be a goal, $D \in \Delta$ a domain, and h a history. Let $0 < \varepsilon < 1$ be some value.
Then, there is some $E \in \Delta$ with
- $\mu_E(\sigma,C_h) \geq \varepsilon$ for any strategy σ consistent with h,
- $\mu_E(\sigma,G|C_{h'}) = \mu_D(\sigma,G|C_{h'})$ for any strategy σ and any history h' such that h'σ (h') is not a prefix of h

The Lebesgue Density Theorem



Theorem
The set of points with density strictly between 0 and 1 has measure zero.

The Lebesgue Density Theorem



Corollary
Let D ∈ Δ be a domain, G a goal and h a history. If $\mu_D(\sigma, G | C_h) > 0$, then there exists a play extending h with density 1.

Theorem
The set of points with density strictly between 0 and 1 has measure zero.

A Characterization Result

Lemma
Let Δ be the set of all domains similar to some domain.
Then, $\sigma_1 \geq \sigma_2$ iff for all split points h and domains $D \in \Delta$

1. $val_{D,G}(\sigma_1,h) \geq val_{D,G}(\sigma_2,h)$, and
2. not both values are $pend$.

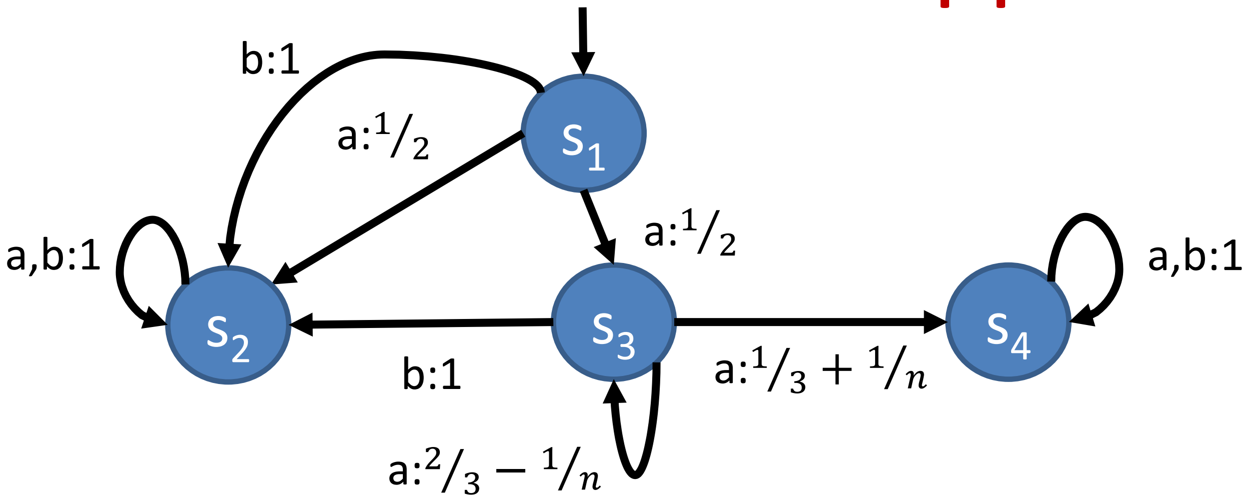
Uses that Δ is closed under **shifts (not defined here)**:

Theorem
Let Δ be the set of all domains similar to some domain.
Then, $\sigma_1 \geq \sigma_2$ iff for all split points h either

1. $\min_{D \in \Delta} val_{D,G}(\sigma_1,h) = win$, or
2. $\max_{D \in \Delta} val_{D,G}(\sigma_2,h) = lose$.

Moreover, a stochastic best-effort strategy for Δ always exists.

Bounded Domains with Markovian Support



Supp(hist,act) = { state | Prob(hist,act)(state) > 0 }

Markovian support = Supp(hist,act) only depends on the last state in hist for all hist, act

Bounded = there is some ε > 0 such that Prob(hist,act)(state) ≥ ε for all state ∈ Supp(hist,act)

A Characterization Result

As bounded domains are also closed under **amplifications and shifts** we obtain:

Theorem

Let Δ be the set of all bounded domains similar to some domain with Markovian support.

Then, $\sigma_1 \geq \sigma_2$ iff for all split points h either

1. $\min_{D \in \Delta} val_{D,G}(\sigma_1, h) = win$, or
2. $\max_{D \in \Delta} val_{D,G}(\sigma_2, h) = lose$.

Moreover, a stochastic best-effort strategy for Δ always exists.

Remark:

The theorem does not holds for sets of domains not closed under amplifications and shifts. For example, for a singleton set = {D}, stochastic best-effort strategies are exactly the optimal strategies, but optimal strategies do not always exist!

Synthesis of Stochastic Best-effort Stragegies for LTL Goals

Theorem
Let Δ be the set of all bounded domains similar to some domain with Markovian support, and let G be some LTL goal. Then, the synthesis of a best-effort strategy for G wrt to Δ is 2EXPTIME-complete.

The synthesis algorithm (IJCAI2022):

1. Fix a Markov decision process M ∈ Δ (e.g., by equipped with the uniform probability distribution).
2. Compute the optimal policy for M and goal G.

Correctness: Relies on the fact that plays in bounded domains are fair with probability one.

Lower bound: Reduces from almost-sure satisfaction of an LTL formula in an MDP and the complement of this problem!

Contributions

- Definition of **stochastic best-effort** strategies for **general (but similar) stochastic domains**.
- Characterization of stochastic best-effort strategies for similar sets of domains that satisfy some **closure properties.**
- Synthesis of stochastic best-effort strategies is **2EXPTIME complete** for **LTL goals** in bounded domains with Markovian support.