# Approximating Fixpoints of Approximated Functions

Florian Wittbold
Universität Duisburg-Essen

Joint work with Paolo Baldan, Sebastian Gurke, Barbara König, Tommaso Padoan

# Fixpoints in Computer Science

**General Task**

Given a function $f\colon L \to L$ over a partially ordered set $(L, \sqsubseteq)$, compute its least fixpoint $x$, i.e., the least element $x \in L$ fulfilling

$$f(x) = x.$$

# Fixpoints in Computer Science

## General Task

Given a function $f : L \to L$ over a partially ordered set $(L, \sqsubseteq)$, compute its least fixpoint $x$, i.e., the least element $x \in L$ fulfilling

$$f(x) = x.$$

Applications in

- concurrency theory (behavioural equivalences and metrics)

# Fixpoints in Computer Science

## General Task

Given a function $f : L \to L$ over a partially ordered set $(L, \sqsubseteq)$, compute its least fixpoint $x$, i.e., the least element $x \in L$ fulfilling

$$f(x) = x.$$

Applications in

- concurrency theory (behavioural equivalences and metrics)
- model checking ($\mu$-calculus)

# Fixpoints in Computer Science

**General Task**

Given a function $f : L \to L$ over a partially ordered set $(L, \sqsubseteq)$, compute its least fixpoint $x$, i.e., the least element $x \in L$ fulfilling

$$f(x) = x.$$

Applications in

- concurrency theory (behavioural equivalences and metrics)
- model checking ($\mu$-calculus)
- program analysis (dataflow analysis)
- Markov decision processes and games (computation of value functions)
- ...

# Fixpoints in Computer Science

**General Task**

Given a function $f : L \to L$ over a partially ordered set $(L, \sqsubseteq)$, compute its least fixpoint $x$, i.e., the least element $x \in L$ fulfilling

$$f(x) = x.$$

- There exist numerous results concerning fixpoints: Banach, Knaster-Tarski, Kleene,. . .

# Fixpoints in Computer Science

## General Task

Given a function $f\colon L \to L$ over a partially ordered set $(L, \sqsubseteq)$, compute its least fixpoint $x$, i.e., the least element $x \in L$ fulfilling

$$f(x) = x.$$

- There exist numerous results concerning fixpoints: Banach, Knaster-Tarski, Kleene,...
- But: What if the target function $f$ can only be approximated?

# Fixpoints in Computer Science

**General Task**

Given a function $f \colon L \to L$ over a partially ordered set $(L, \sqsubseteq)$, compute its least fixpoint $x$, i.e., the least element $x \in L$ fulfilling

$$f(x) = x.$$

- There exist numerous results concerning fixpoints: Banach, Knaster-Tarski, Kleene,. . .
- But: What if the target function $f$ can only be approximated?

  $\implies$ reinforcement learning

  Will fixpoint iterations still work? For which types of functions? Do we need new techniques?

# Solution Techniques

## Knaster-Tarski and Kleene theorems

- Knaster-Tarski theorem: Existence of least fixpoints ($\mu f$) and greatest fixpoints ($\nu f$) for *monotone* functions $f$ over a *complete lattice*.

# Solution Techniques

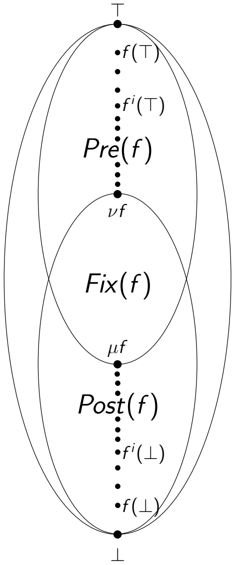
## Knaster-Tarski and Kleene theorems

- Knaster-Tarski theorem: Existence of least fixpoints ($\mu f$) and greatest fixpoints ($\nu f$) for *monotone* functions $f$ over a *complete lattice*.
- Kleene's theorem: constructive algorithm to find $\mu f$ and $\nu f$ whenever $f$ is (co-)continuous
  - Least fixpoint: $\mu f = \bigsqcup_{n \in \mathbb{N}} f^n(\bot)$
  - Greatest fixpoint: $\nu f = \bigsqcap_{n \in \mathbb{N}} f^n(\top)$

# Solution Techniques

## Knaster-Tarski and Kleene theorems

- Knaster-Tarski theorem: Existence of least fixpoints ($\mu f$) and greatest fixpoints ($\nu f$) for *monotone* functions $f$ over a *complete lattice*.
- Kleene's theorem: constructive algorithm to find $\mu f$ and $\nu f$ whenever $f$ is (co-)continuous
  - Least fixpoint: $\mu f = \bigsqcup_{n \in \mathbb{N}} f^n(\bot)$
  - Greatest fixpoint: $\nu f = \bigsqcap_{n \in \mathbb{N}} f^n(\top)$



# Solution Techniques

## Contractive and non-expansive functions

A function $f$ over a metric space $(X, d)$ is called

# Solution Techniques

## Contractive and non-expansive functions

A function $f$ over a metric space $(X, d)$ is called

- a contraction if there exists $0 \leq q < 1$ such that

$$d(f(x), f(y)) \leq q \cdot d(x, y) \quad \text{for all } x, y \in X,$$

# Solution Techniques

## Contractive and non-expansive functions

A function $f$ over a metric space $(X, d)$ is called

- a contraction if there exists $0 \leq q < 1$ such that

$$d(f(x), f(y)) \leq q \cdot d(x, y) \quad \text{for all } x, y \in X,$$

- non-expansive if this holds for $q = 1$,

# Solution Techniques

## Contractive and non-expansive functions

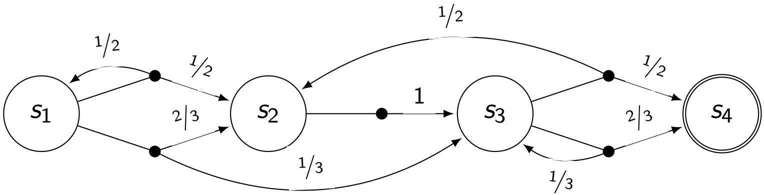
A function $f$ over a metric space $(X, d)$ is called

- a contraction if there exists $0 \leq q < 1$ such that

$$d(f(x), f(y)) \leq q \cdot d(x, y) \quad \text{for all } x, y \in X,$$

- non-expansive if this holds for $q = 1$,
- a power contraction if there exists $n \in \mathbb{N}$ such that $f^n$ is a contraction.

# Case Study: Markov decision processes



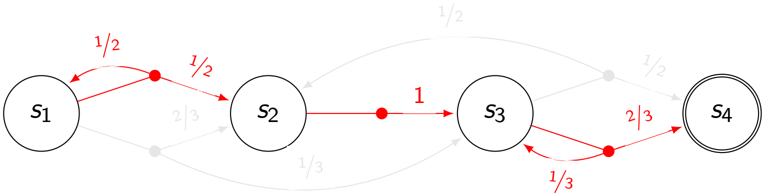

## Markov decision process

A Markov decision process (MDP) is a tuple $(S, T)$ where

- $S$ is a finite set of *states* and
- $T : S \to \mathcal{P}_f(\mathcal{D}(S))$ is a *transition* function.

We let $T(s)$ be indexed over (pairwise disjoint) sets $A(s)$ called *actions*, writing $F = \{s \in S \mid A(s) = \emptyset\}, A = \bigcup_{s \in S} A(s)$, and $T(s' \mid s, a) = T(s)_a(s')$.

# Case Study: Markov decision processes

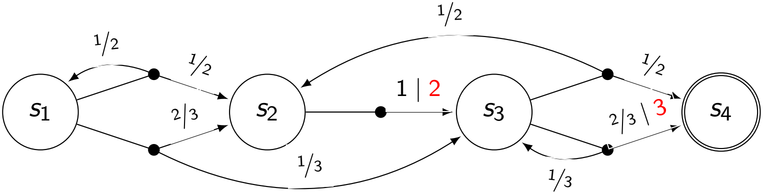



## Definition (Policy)

Given an MDP $M = (S, T)$, a policy is a function $\pi : S \setminus F \to A$ with $\pi(s) \in A(s)$ for all $s \in S \setminus F$. The set of policies will be denoted by $\Pi(M)$.

A policy defines the Markov chain $M^\pi = (S, T^\pi)$ where $T^\pi(s) = T(s)_{\pi(s)}$.

# Case Study: Markov decision processes




The objective for the agent is determined by a (step-wise) reward

$$R : S \times A \times S \to \mathbb{R}_{\geq 0}.$$

Given a discount factor $\gamma \in [0, 1]$, we try to solve

$$\max_{\pi \in \Pi(M)} \quad \mathbb{E}^{\pi} \left[ \sum_{n=0}^{\infty} \gamma^n R(\mathbf{s}_n, \pi(\mathbf{s}_n), \mathbf{s}_{n+1}) \right]$$

# Case Study: Markov decision processes

Given an MDP $M = (S, T)$ and reward function $R$, the optimal value function is given as least fixpoint of the function $f : \mathbb{R}^S_{\geq 0} \to \mathbb{R}^S_{\geq 0}$ (Bellman optimality operator) given by

$$f(v)(s) = \max_{a \in A(s)} \sum_{s' \in S} T(s' \mid s, a) \cdot \big(R(s, a, s') + \gamma v(s')\big)$$

Note that, for $\gamma < 1$, $f$ is contractive.

# Case Study: Approximating MDPs

Question: How to determine $\mu f$ if the probabilities (given by $T$) and possibly the rewards (given by $R$) are not known precisely, but can only be approximated by sampling through interaction with the MDP?

# Case Study: Approximating MDPs

Question: How to determine $\mu f$ if the probabilities (given by $T$) and possibly the rewards (given by $R$) are not known precisely, but can only be approximated by sampling through interaction with the MDP?

$\Longrightarrow$ (model-based) reinforcement learning

# Case Study: Approximating MDPs

Question: How to determine $\mu f$ if the probabilities (given by $T$) and possibly the rewards (given by $R$) are not known precisely, but can only be approximated by sampling through interaction with the MDP?

$\implies$ (model-based) reinforcement learning

But: Reinforcement learning typically concentrates on the contractive case ($\gamma < 1$), where fixpoints are unique and errors made in early stages are removed by the contraction.

# Case Study: Approximating MDPs

Question: How to determine $\mu f$ if the probabilities (given by $T$) and possibly the rewards (given by $R$) are not known precisely, but can only be approximated by sampling through interaction with the MDP?

$\Longrightarrow$ (model-based) reinforcement learning

But: Reinforcement learning typically concentrates on the contractive case ($\gamma < 1$), where fixpoints are unique and errors made in early stages are removed by the contraction.

Here: We address the non-contractive case ($\gamma = 1$) of which the contractive case can be constructed as a special case.

# Introducing Approximations

Sources of approximation:

- Unknown / partially known system
- Changes to the system
- Approximated partial results
- ...

# Introducing Approximations

Sources of approximation:

- Unknown / partially known system
- Changes to the system
- Approximated partial results
- ...

## Nested Fixpoint Equations

For $f : L^d \to L^d$, find $x \in L^d$ such that

$$\begin{aligned} x_1 &=_\mu f_1(x_1, \ldots, x_d) \\ x_2 &=_\nu f_2(x_1, \ldots, x_d) \\ &\vdots \\ x_d &=_\mu f_d(x_1, \ldots, x_d) \end{aligned}$$

# Approximations in the Non-expansive Case

**Task**

Given a sequence of monotone and non-expansive (w.r.t. the supremum distance) functions $f_0, f_1, \dots : \mathbb{R}^d_{\geq 0} \to \mathbb{R}^d_{\geq 0}$ that (uniformly) converges to $f : \mathbb{R}^d_{\geq 0} \to \mathbb{R}^d_{\geq 0}$.
Compute a sequence $x_0, x_1, x_2, \dots$ that converges to $\mu f$ (if it exists).

# Approximations in the Non-expansive Case

**Task**

Given a sequence of monotone and non-expansive (w.r.t. the supremum distance) functions $f_0, f_1, \dots : \mathbb{R}^d_{\geq 0} \to \mathbb{R}^d_{\geq 0}$ that (uniformly) converges to $f : \mathbb{R}^d_{\geq 0} \to \mathbb{R}^d_{\geq 0}$.
Compute a sequence $x_0, x_1, x_2, \dots$ that converges to $\mu f$ (if it exists).

Remarks:

- Non-expansiveness w.r.t. the supremum distance covers many interesting cases: termination probabilities and non-discounted rewards in Markov chains, MDPs, stochastic games, behavioural metrics,...

# Approximations in the Non-expansive Case

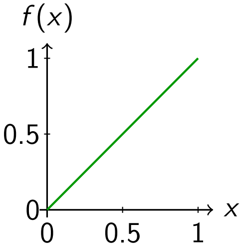
## Task

Given a sequence of monotone and non-expansive (w.r.t. the supremum distance) functions $f_0, f_1, \dots : \mathbb{R}^d_{\geq 0} \to \mathbb{R}^d_{\geq 0}$ that (uniformly) converges to $f : \mathbb{R}^d_{\geq 0} \to \mathbb{R}^d_{\geq 0}$.
Compute a sequence $x_0, x_1, x_2, \dots$ that converges to $\mu f$ (if it exists).

Remarks:

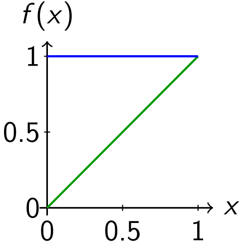
- Non-expansiveness w.r.t. the supremum distance covers many interesting cases: termination probabilities and non-discounted rewards in Markov chains, MDPs, stochastic games, behavioural metrics,...
- In a compact space, uniform convergence for non-expansive functions follows from pointwise convergence.
  $\implies$ uniform convergence of $(f_n)_n$ can be replaced by boundedness of $(x_n)_n$

# Introducing Approximation



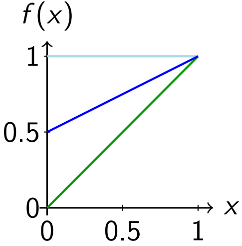
# Introducing Approximation

$f, f_n \colon [0,1] \to [0,1]$ with $f_n(x) = 1/n + (1 - 1/n) \cdot x$, $f(x) = x$



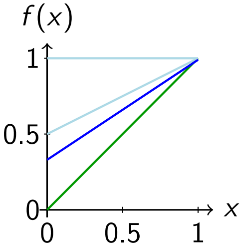
# Introducing Approximation

$f, f_n \colon [0,1] \to [0,1]$ with $f_n(x) = 1/n + (1 - 1/n) \cdot x$, $f(x) = x$



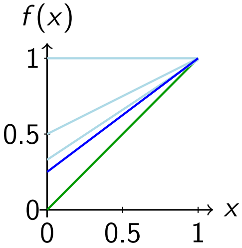
# Introducing Approximation

$f, f_n \colon [0,1] \to [0,1]$ with $f_n(x) = 1/n + (1 - 1/n) \cdot x,\ f(x) = x$



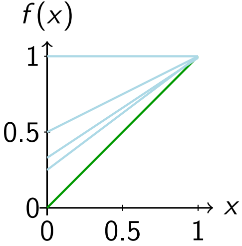
# Introducing Approximation

$f, f_n\colon [0,1] \to [0,1]$ with $f_n(x) = 1/n + (1 - 1/n) \cdot x$, $f(x) = x$



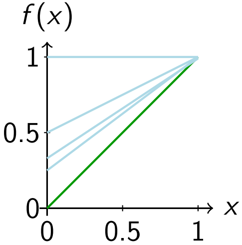
# Introducing Approximation

$f, f_n\colon [0,1] \to [0,1]$ with $f_n(x) = {}^1\!/\!_n + (1 - {}^1\!/\!_n) \cdot x,\ f(x) = x$



# Introducing Approximation

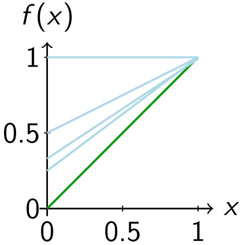
$f, f_n \colon [0,1] \to [0,1]$ with $f_n(x) = 1/n + (1 - 1/n) \cdot x$, $f(x) = x$




$\mu f_n = 1$ (for all $n$), while $\mu f = 0$.
Hence $\lim_{n\to\infty} \mu f_n = 1 \neq 0 = \mu f$.

# Introducing Approximation

$f, f_n \colon [0,1] \to [0,1]$ with $f_n(x) = {}^1/_n + (1 - {}^1/_n) \cdot x$, $f(x) = x$




$\mu f_n = 1$ (for all $n$), while $\mu f = 0$.
Hence $\lim_{n\to\infty} \mu f_n = 1 \neq 0 = \mu f$.

The fixpoints of the approximations need not converge to the fixpoint of $f$.

# Dampened Mann Iteration

Inspired by a paper from Kim/Xu, we now consider the following form of iteration:

$$x_{n+1} = (1 - \beta_n) \cdot ((1 - \alpha_n) \cdot x_n + \alpha_n \cdot f_n(x_n))$$

which is

- a Mann iteration $(1 - \alpha_n) \cdot x_n + \alpha_n \cdot f_n(x_n)$
- together with a dampening factor $1 - \beta_n$.

# Dampened Mann Iteration

$$x_{n+1} = (1-\beta_n) \cdot ((1-\alpha_n) \cdot x_n + \alpha_n \cdot f_n(x_n))$$

Choice of parameters:

- dampening factors $\beta_n$ fulfill

$$\lim_{n\to\infty} \beta_n = 0 \quad \text{and} \quad \sum_{n=0}^{\infty} \beta_n = \infty.$$

  Intuition: a form of 'vanishing discount', but there is always enough 'power' left to decrease a current over-estimation to the true least fixpoint.

- learning rates $\alpha_n$ fulfill

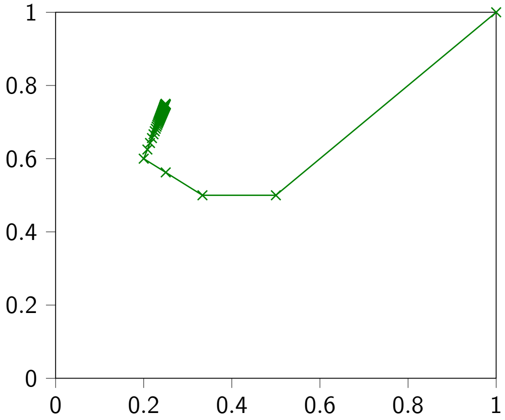
$$\lim_{n\to\infty} \alpha_n = 1$$

  These provide extra flexibility and give the option to generalize the results.

# Dampened Mann Iteration: Example

Dampened Mann iteration (with $\alpha_n = 1, \beta_n = 1/n$) for

$$f_n \colon [0,1]^2 \to [0,1]^2, f_n(x) = \max(x, \begin{pmatrix} 3/4 \\ 1/4 \end{pmatrix} \cdot x^n + \begin{pmatrix} 1/4 \\ 3/4 \end{pmatrix})$$



# Dampened Mann Iteration: Results

## Main Theorem

Given $f_n \to f$ uniformly, the sequence $(x_n)_n$ converges to the correct solution $\mu f$ from any starting point in the following cases:

# Dampened Mann Iteration: Results

## Main Theorem

Given $f_n \to f$ uniformly, the sequence $(x_n)_n$ converges to the correct solution $\mu f$ from any starting point in the following cases:

1. When iterating with the exact function, i.e., when $f_n = f$ for all $n$,

# Dampened Mann Iteration: Results

## Main Theorem

Given $f_n \to f$ uniformly, the sequence $(x_n)_n$ converges to the correct solution $\mu f$ from any starting point in the following cases:

1. When iterating with the exact function, i.e., when $f_n = f$ for all $n$,
2. when $f$ is a power contraction,

# Dampened Mann Iteration: Results

## Main Theorem

Given $f_n \to f$ uniformly, the sequence $(x_n)_n$ converges to the correct solution $\mu f$ from any starting point in the following cases:

1. When iterating with the exact function, i.e., when $f_n = f$ for all $n$,
2. when $f$ is a power contraction,
3. when $\mu f_n \to \mu f$ and $f_n \to f$ monotonically,

# Dampened Mann Iteration: Results

**Main Theorem**

Given $f_n \to f$ uniformly, the sequence $(x_n)_n$ converges to the correct solution $\mu f$ from any starting point in the following cases:

1. When iterating with the exact function, i.e., when $f_n = f$ for all $n$,
2. when $f$ is a power contraction,
3. when $\mu f_n \to \mu f$ and $f_n \to f$ monotonically,
4. when $f_n \to f$ normally, i.e.

$$\sum \|f_n - f\|_\infty < \infty$$

# Dampened Mann Iteration: Results

**Main Theorem**

Given $f_n \to f$ uniformly, the sequence $(x_n)_n$ converges to the correct solution $\mu f$ from any starting point in the following cases:

1. When iterating with the exact function, i.e., when $f_n = f$ for all $n$,
2. when $f$ is a power contraction,
3. when $\mu f_n \to \mu f$ and $f_n \to f$ monotonically,
4. when $f_n \to f$ normally, i.e.

$$\sum \|f_n - f\|_\infty < \infty$$

There also exist counterexamples for the case $\mu f_n \to \mu f$.

# Dampened Mann Iteration for MDPs: End Components

## End component of an MDP

Let $M = (S, T)$ be an MDP. Writing for $E \subseteq S$

$$A_E(s) = \{a \in A(s) \mid \{s' \in S \mid T(s' \mid s, a) > 0\} \subseteq E\},$$

for $s \in E$, $E$ is called an end component of $M$ if for all $s, s' \in E$, there exists a strategy in $A_E$ with positive probability of reaching $s'$ from $s$.

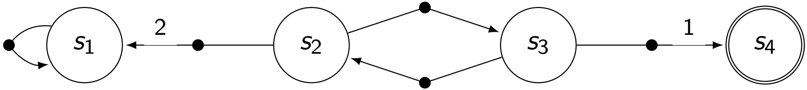
# Dampened Mann Iteration for MDPs: End Components

## End component of an MDP

Let $M = (S, T)$ be an MDP. Writing for $E \subseteq S$

$$A_E(s) = \{a \in A(s) \mid \{s' \in S \mid T(s' \mid s, a) > 0\} \subseteq E\},$$

for $s \in E$, $E$ is called an end component of $M$ if for all $s, s' \in E$, there exists a strategy in $A_E$ with positive probability of reaching $s'$ from $s$.



# Dampened Mann Iteration for Terminating MDPs

A Markov chain is terminating if for any starting state its probability of eventually reaching a terminal state is 1. An MDP $M$ is terminating if for all policies $\pi$, the induced Markov chain $M^\pi$ is terminating.

# Dampened Mann Iteration for Terminating MDPs

A Markov chain is terminating if for any starting state its probability of eventually reaching a terminal state is 1. An MDP $M$ is terminating if for all policies $\pi$, the induced Markov chain $M^\pi$ is terminating.

For an MDP $M$, the following are equivalent:

- $M$ is terminating.
- $M$ has no end components.
- The corresponding function $f_M$ is a power contraction.

# Dampened Mann Iteration for Terminating MDPs

A Markov chain is terminating if for any starting state its probability of eventually reaching a terminal state is 1. An MDP $M$ is terminating if for all policies $\pi$, the induced Markov chain $M^\pi$ is terminating.

For an MDP $M$, the following are equivalent:

- $M$ is terminating.
- $M$ has no end components.
- The corresponding function $f_M$ is a power contraction.

**Main Theorem**

2. The sequence $(x_n)_n$ converges to the correct solution $\mu f$ from any starting point when $f$ is a power contraction and $f_n \to f$.

# Dampened Mann Iteration for MDPs

In a general MDP $M$ (with non-negative rewards), the value of every state is bounded (i.e., $\mu f_M \in \mathbb{R}^S_{\geq 0}$) if and only if no reward is given when staying inside an end component.

# Dampened Mann Iteration for MDPs

In a general MDP $M$ (with non-negative rewards), the value of every state is bounded (i.e., $\mu f_M \in \mathbb{R}^S_{\geq 0}$) if and only if no reward is given when staying inside an end component.

The approximating sequence of MDPs being the result of sampling ensures that

- the sequence of approximating MDPs will almost surely converge to the correct MDP.

# Dampened Mann Iteration for MDPs

In a general MDP $M$ (with non-negative rewards), the value of every state is bounded (i.e., $\mu f_M \in \mathbb{R}^S_{\geq 0}$) if and only if no reward is given when staying inside an end component.

The approximating sequence of MDPs being the result of sampling ensures that

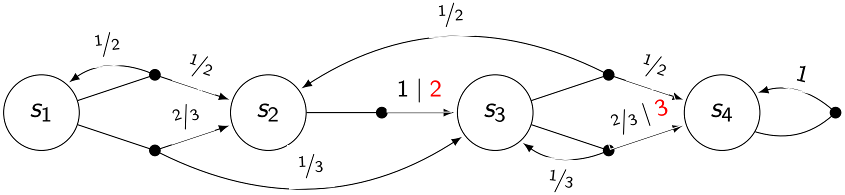
- the sequence of approximating MDPs will almost surely converge to the correct MDP.
- eventually the underlying graph will be the one of the correct MDP (once all transitions have been discovered).

In particular, the end components of the approximated MDPs will eventually be exactly the ones of the correct MDP.

# Dampened Mann Iteration for MDPs

## Lower Bound ($\mu f_M \leq \liminf x_n$)

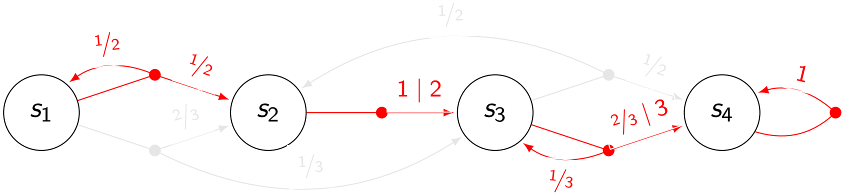
There exists an optimal policy $\pi$ in an MDP $M$, i.e., one for which $\mu f_M$ is equal to $\mu f_{M^\pi}$.



# Dampened Mann Iteration for MDPs

**Lower Bound ($\mu f_M \leq \liminf x_n$)**

There exists an optimal policy $\pi$ in an MDP $M$, i.e., one for which $\mu f_M$ is equal to $\mu f_{M^\pi}$.

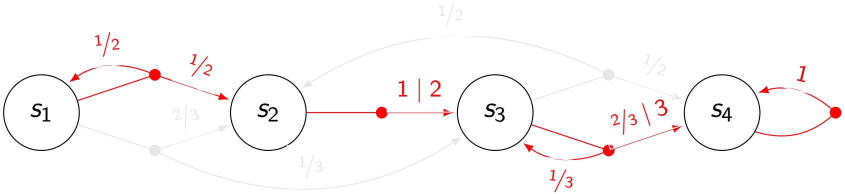


# Dampened Mann Iteration for MDPs

## Lower Bound ($\mu f_M \leq \liminf x_n$)

There exists an optimal policy $\pi$ in an MDP $M$, i.e., one for which $\mu f_M$ is equal to $\mu f_{M^\pi}$.

In a Markov chain, all states in an end component have value 0. Fixing the value to 0 in these states gives a power contraction with the same fixpoint.



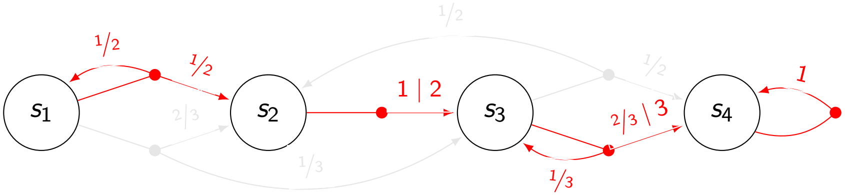
# Dampened Mann Iteration for MDPs

## Lower Bound ($\mu f_M \leq \liminf x_n$)

There exists an optimal policy $\pi$ in an MDP $M$, i.e., one for which $\mu f_M$ is equal to $\mu f_{M^\pi}$.

In a Markov chain, all states in an end component have value 0. Fixing the value to 0 in these states gives a power contraction with the same fixpoint.

From this, one can deduce that the true state values ($\mu f_M$) are always a lower bound for the result of the dampened Mann iteration of the MDP.



# Dampened Mann Iteration for MDPs

## Upper Bound ($\limsup x_n \leq \mu f_M$)

Idea: Derive a monotonically decreasing sequence of 'sampled' Bellman operators that bound the error made in each iteration from above.

# Dampened Mann Iteration for MDPs

## Upper Bound ($\limsup x_n \leq \mu f_M$)

Idea: Derive a monotonically decreasing sequence of 'sampled' Bellman operators that bound the error made in each iteration from above.

### Main Theorem

3. The sequence $(x_n)_n$ converges to the correct solution $\mu f$ from any starting point when $\mu f_n \to \mu f$ and $f_n \to f$ monotonically.

# Dampened Mann Iteration for Other Cases

What can we do in other cases where none of the above conditions (directly) apply?

# Dampened Mann Iteration for Other Cases

What can we do in other cases where none of the above conditions (directly) apply?

**Main Theorem**

4. The sequence $(x_n)_n$ converges to the correct solution $\mu f$ from any starting point when $f_n \to f$ normally, i.e.

$$\sum \|f_n - f\|_\infty < \infty.$$

# Dampened Mann Iteration for Other Cases

What can we do in other cases where none of the above conditions (directly) apply?

**Main Theorem**

4. The sequence $(x_n)_n$ converges to the correct solution $\mu f$ from any starting point when $f_n \to f$ normally, i.e.

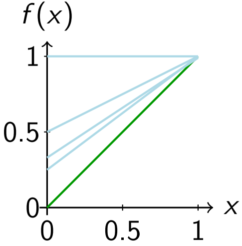
$$\sum \|f_n - f\|_\infty < \infty.$$

Idea: Fix a subsequence of functions $f_{n_1}, f_{n_2}, \ldots$ that converges normally.

Intuition: Approximate fast / far enough (e.g., by performing enough sampling steps) before performing the next iteration.

# Dampened Mann Iteration for Other Cases

Example: $f, f_n \colon [0,1] \to [0,1]$ with $f_n(x) = 1/n + (1 - 1/n) \cdot x$, $f(x) = x$




Perform dampened Mann iteration with the sequence $(f_{n^2})_n$.

This works, although the sequence of least fixpoints of these functions does not converge to the least fixpoint of $f$!

# Dampened Mann Iteration for Other Cases

In many cases, deterministic error bounds are not available, but the approximation errors can be estimated.

Choose $n_i$ such that

$$\mathbb{P}\big[\|f_{n_i} - f\| > \gamma_i\big] \leq \delta_i,$$

where $\sum_i \gamma_i < \infty$ and $\sum_i \delta_i < \infty$ (for instance $\gamma_i = \delta_i = 1/i^2$).

By the Borel-Cantelli Lemma we get that

$$\mathbb{P}\big[\|f_{n_i} - f\| > \gamma_i \text{ for infinitely many } i\big] = 0$$

Hence, almost surely we have $\|f_{n_i} - f\| \leq \gamma_i$ eventually and by our fixpoint results (normal convergence) the sequence produced by the algorithm converges to the solution vector of its input in that case.

# Future Work

## Future and ongoing work

# Future Work

## Future and ongoing work

- Apply to other use cases
  - Stochastic games
  - Partially observable systems (extend to infinite dimensions)
  - Nested fixpoint equations
  - ...

# Future Work

## Future and ongoing work

- Apply to other use cases
  - Stochastic games
  - Partially observable systems (extend to infinite dimensions)
  - Nested fixpoint equations
  - ...
- Chaotic iteration

# Future Work

## Future and ongoing work

- Apply to other use cases
  - Stochastic games
  - Partially observable systems (extend to infinite dimensions)
  - Nested fixpoint equations
  - ...
- Chaotic iteration
- Negative values

# Future Work

## Future and ongoing work

- Apply to other use cases
  - Stochastic games
  - Partially observable systems (extend to infinite dimensions)
  - Nested fixpoint equations
  - ...
- Chaotic iteration
- Negative values
- What if the learning rates $\alpha_n$ converge to 0? (Mann iteration converging to the identity)

# Future Work

## Future and ongoing work

- Apply to other use cases
  - Stochastic games
  - Partially observable systems (extend to infinite dimensions)
  - Nested fixpoint equations
  - ...
- Chaotic iteration
- Negative values
- What if the learning rates $\alpha_n$ converge to 0? (Mann iteration converging to the identity)
- Model-free learning
  Idea: Sequence $f_1, f_2, \ldots$ of functions approximates $f$ only in the average: $^1/_n \sum_{k=1}^{n} f_k \to f$
  Aim: Obtain model-free reinforcement learning algorithms as special cases

# Solution Techniques

**Banach fixpoint theorem**

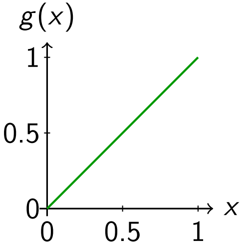
Given a complete metric space $X$ and a (power) contraction $f : X \to X$, $f$ admits a unique fixed-point $x^* = \mu f = \nu f$. Any sequence $(x_n)_n$ given by

$$\begin{cases} x_0 \in X \\ x_{n+1} = f(x_n) \end{cases}$$

converges to $x^*$.

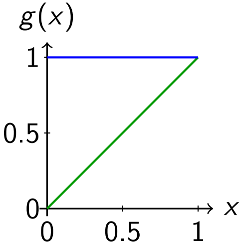
# Introducing Approximation

$g, g_n \colon [0,1] \to [0,1]$ with $g_n(x) = 1/n$ if $x \leq 1/n$, $g_n(x) = x$ otherwise, $g(x) = x$



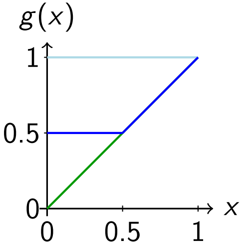
# Introducing Approximation

$g, g_n \colon [0,1] \to [0,1]$ with $g_n(x) = 1/n$ if $x \leq 1/n$, $g_n(x) = x$ otherwise, $g(x) = x$



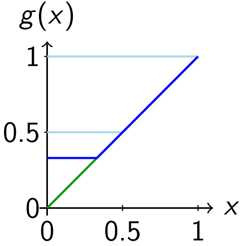
# Introducing Approximation

$g, g_n \colon [0,1] \to [0,1]$ with $g_n(x) = 1/n$ if $x \leq 1/n$, $g_n(x) = x$ otherwise, $g(x) = x$



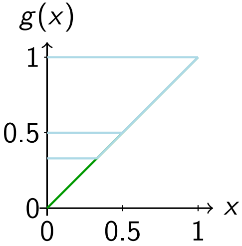
# Introducing Approximation

$g, g_n \colon [0,1] \to [0,1]$ with $g_n(x) = {}^1\!/_n$ if $x \leq {}^1\!/_n$, $g_n(x) = x$ otherwise, $g(x) = x$



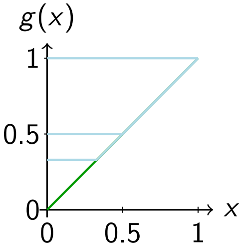
# Introducing Approximation

$g, g_n \colon [0,1] \to [0,1]$ with $g_n(x) = 1/n$ if $x \leq 1/n$, $g_n(x) = x$ otherwise, $g(x) = x$



# Introducing Approximation

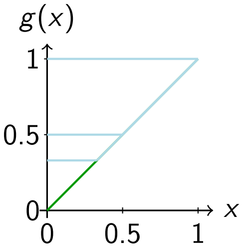
$g, g_n \colon [0,1] \to [0,1]$ with $g_n(x) = 1/n$ if $x \leq 1/n$, $g_n(x) = x$ otherwise, $g(x) = x$




In this case $\lim_{n\to\infty} \mu f_n = 0 = \mu f$.

# Introducing Approximation

$g, g_n \colon [0,1] \to [0,1]$ with $g_n(x) = 1/n$ if $x \leq 1/n$, $g_n(x) = x$ otherwise, $g(x) = x$




In this case $\lim_{n\to\infty} \mu f_n = 0 = \mu f$.

But: Starting a Kleene iteration with some $f_n$ will over-estimate the least fixpoint and further iterations with functions $f_m$ $(m > 0)$ will never decrease it.

# Approximations in the Contractive Case

## Generalized Banach fixpoint theorem

Let a complete metric space $X$, a (power) contraction $f : X \to X$, and a sequence $(f_n)_n$ of functions uniformly converging to $f$ be given. Any sequence $(x_n)_n$ given by

$$\begin{cases} x_0 \in X \\ x_{n+1} = f(x_n) \end{cases}$$

converges to the unique fixpoint $x^*$ of $f$.

# Dampened Mann Iteration for MDPs: Exact Case

## Main Theorem

1. The sequence $(x_n)_n$ converges to the correct solution $\mu f$ from any starting point when iterating with the exact function, i.e., when $f_n = f$ for all $n$.

# Dampened Mann Iteration for MDPs: Exact Case

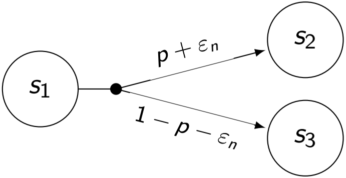
**Main Theorem**

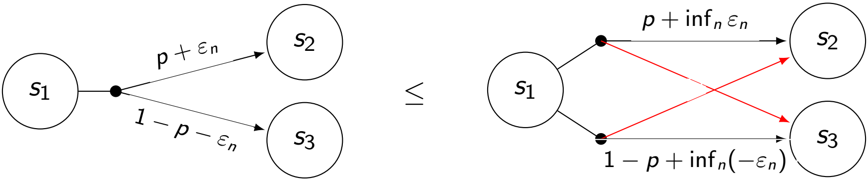
1. The sequence $(x_n)_n$ converges to the correct solution $\mu f$ from any starting point when iterating with the exact function, i.e., when $f_n = f$ for all $n$.

**Dampened Mann Iteration for MDPs (Exact Case)**

Assuming knowledge of the exact function $f_M$ of an MDP (i.e., its transition probabilities and rewards), the dampened Mann iteration converges to the optimal value function.

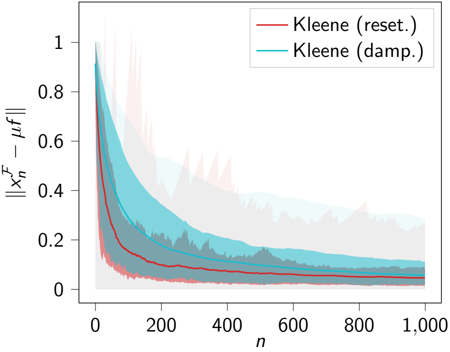
But: None of the other conditions are satisfied for general *approximated* MDPs.


$s_1$
$s_2$
$s_3$
$p + \varepsilon_n$
$1 - p - \varepsilon_n$



$s_1$
$p + \varepsilon_n$
$1 - p - \varepsilon_n$
$s_2$
$s_3$
$\leq$
$s_1$
$p + \inf_n \varepsilon_n$
$1 - p + \inf_n(-\varepsilon_n)$
$s_2$
$s_3$

# Implementation



# Implementation