Stochastic dominance and the bijective ratio of online algorithms

Spyros Angelopoulos

CNRS & LIP6--Univ. Pierre and Marie Curie

Joint work with Marc Renault and Pascal Schweitzer

PGMOdays, November 2017

Decision-making under limited information

We often decide:

Decision-making under limited information

©Ron Leishman * illustrationsOf.com/1047168

We often decide:

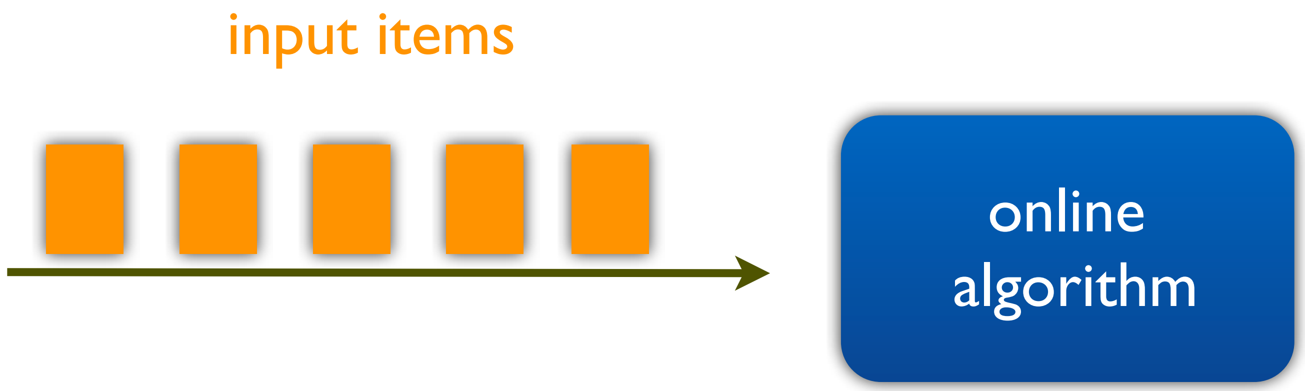
- irrevocably
- with no knowledge of the future

Online computation

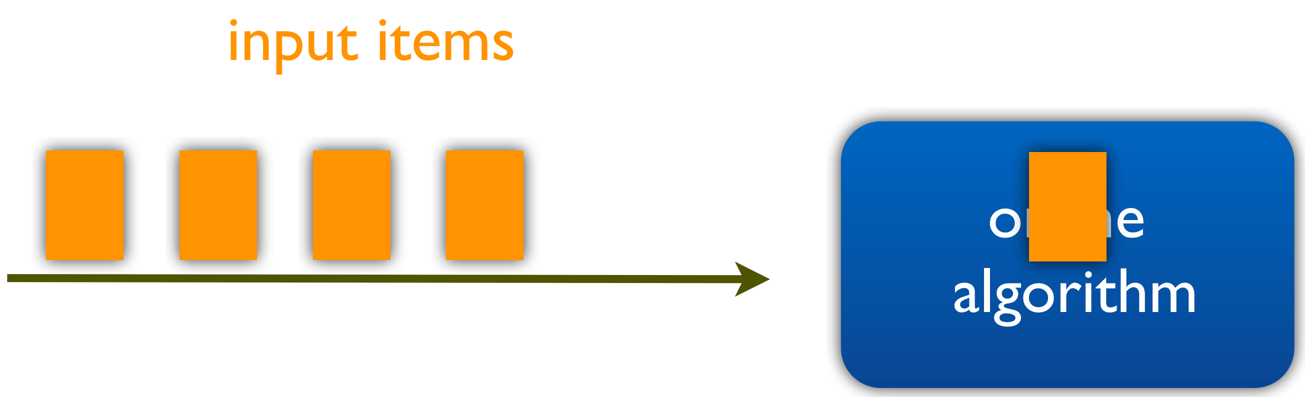
Online computation

online algorithm

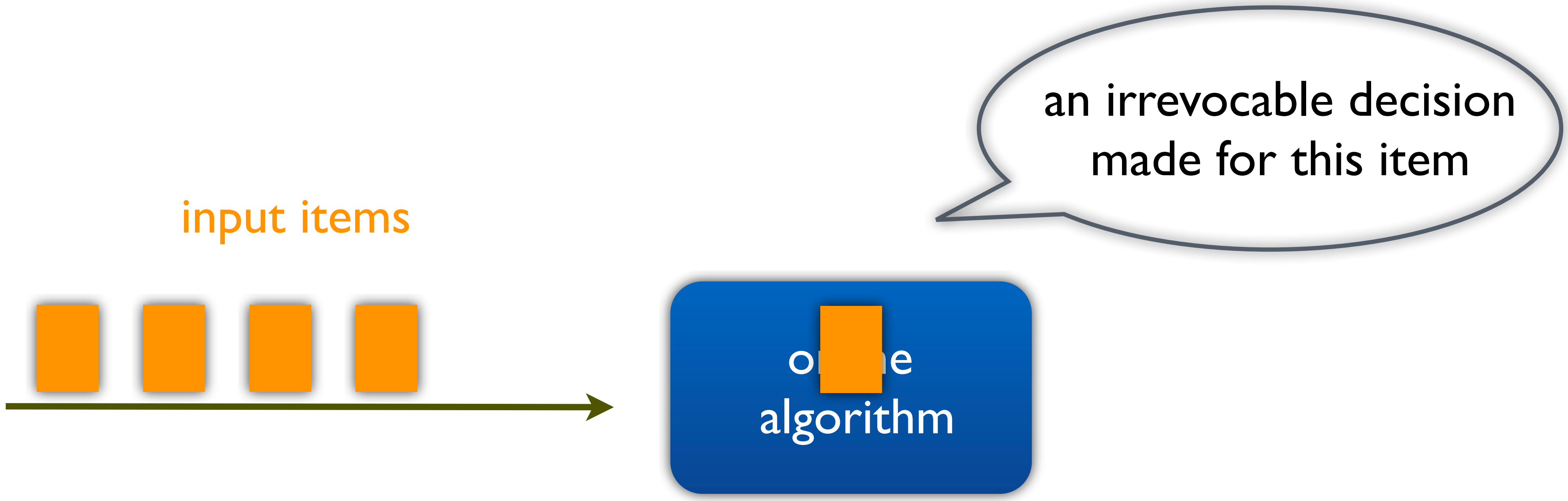
Online computation



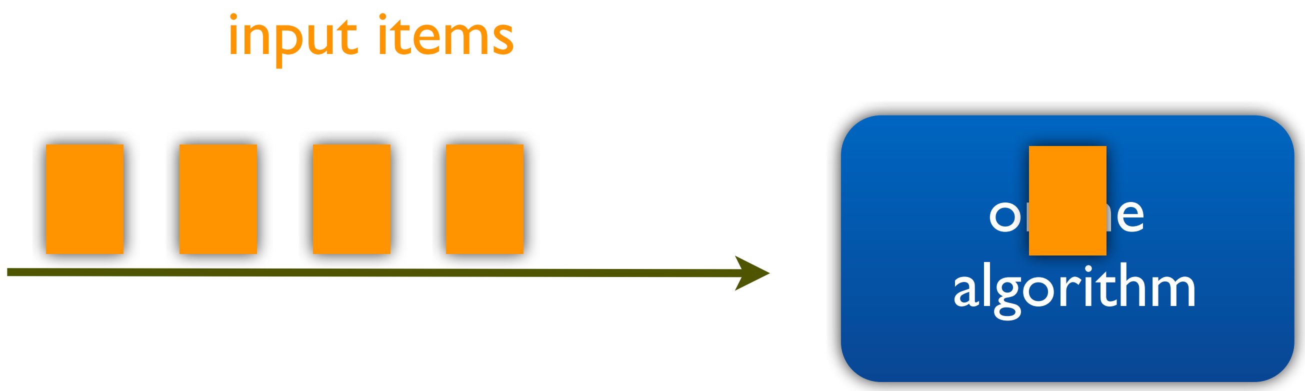
Online computation



Online computation



Online computation



Online computation

Competitive analysis of online algorithms

Competitive analysis of online algorithms

- Given a sequence of requests σ and an online algorithm ALG

$ALG(\sigma)$	cost of ALG on sequence σ
$OPT(\sigma)$	optimal cost for sequence σ

Competitive analysis of online algorithms

- Given a sequence of requests σ and an online algorithm ALG

$ALG(\sigma)$ cost of ALG on sequence σ

$OPT(\sigma)$ optimal cost for sequence σ

- ALG has competitive ratio at most ρ if for every request sequence σ,

$$ALG(\sigma) \leq \rho \cdot OPT(\sigma)$$

[Sleator and Tarjan 85]

Competitive analysis of online algorithms

- Given a sequence of requests σ and an online algorithm ALG

$ALG(\sigma)$ cost of ALG on sequence σ

$OPT(\sigma)$ optimal cost for sequence σ

- ALG has competitive ratio at most ρ if for every request sequence σ,

$$ALG(\sigma) \leq \rho \cdot OPT(\sigma)$$

[Sleator and Tarjan 85]

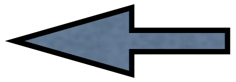
- An online algorithm with best-possible competitive ratio is called optimal (and this determines the competitive ratio of the problem)

Why competitive analysis?

Why competitive analysis?

- Worst-case guarantees
- No need for probabilistic assumptions
- Insights into the development of new algorithms
- Amenability of problems to analysis

Why competitive analysis?

- Worst-case guarantees
- No need for probabilistic assumptions
- Insights into the development of new algorithms
- Amenability of problems to analysis 

Why competitive analysis?

- Worst-case guarantees
- No need for probabilistic assumptions
- Insights into the development of new algorithms
- Amenability of problems to analysis

However: Some significant deficiencies

Some of the many alternatives

- Max/Max ratio [Ben-David, Borodin 94]
- Loose Competitiveness [Young 94]
- Diffuse Adversary [Koutsoupias Papadimitriou 00]
- Random Order Ratio [Kenyon 96]
- Relative Worst Order Ratio [Boyar Favrholdt 03]
- Paging with locality of reference
    - Access Graph Model [BIRS 95]
    - Concave Analysis [Albers, Farvhold, Giel 05]
    - Adequate Analysis [Panagiotou, Souza 06]

[Dorrigiv Lopez-Ortiz 05]: A survey of Performance Measures for On-line Algorithms. ACM SIGACT News

Bijective analysis [A., Dorrigiv, Lopez-Ortiz 07]

Bijective analysis [A., Dorrigiv, Lopez-Ortiz 07]

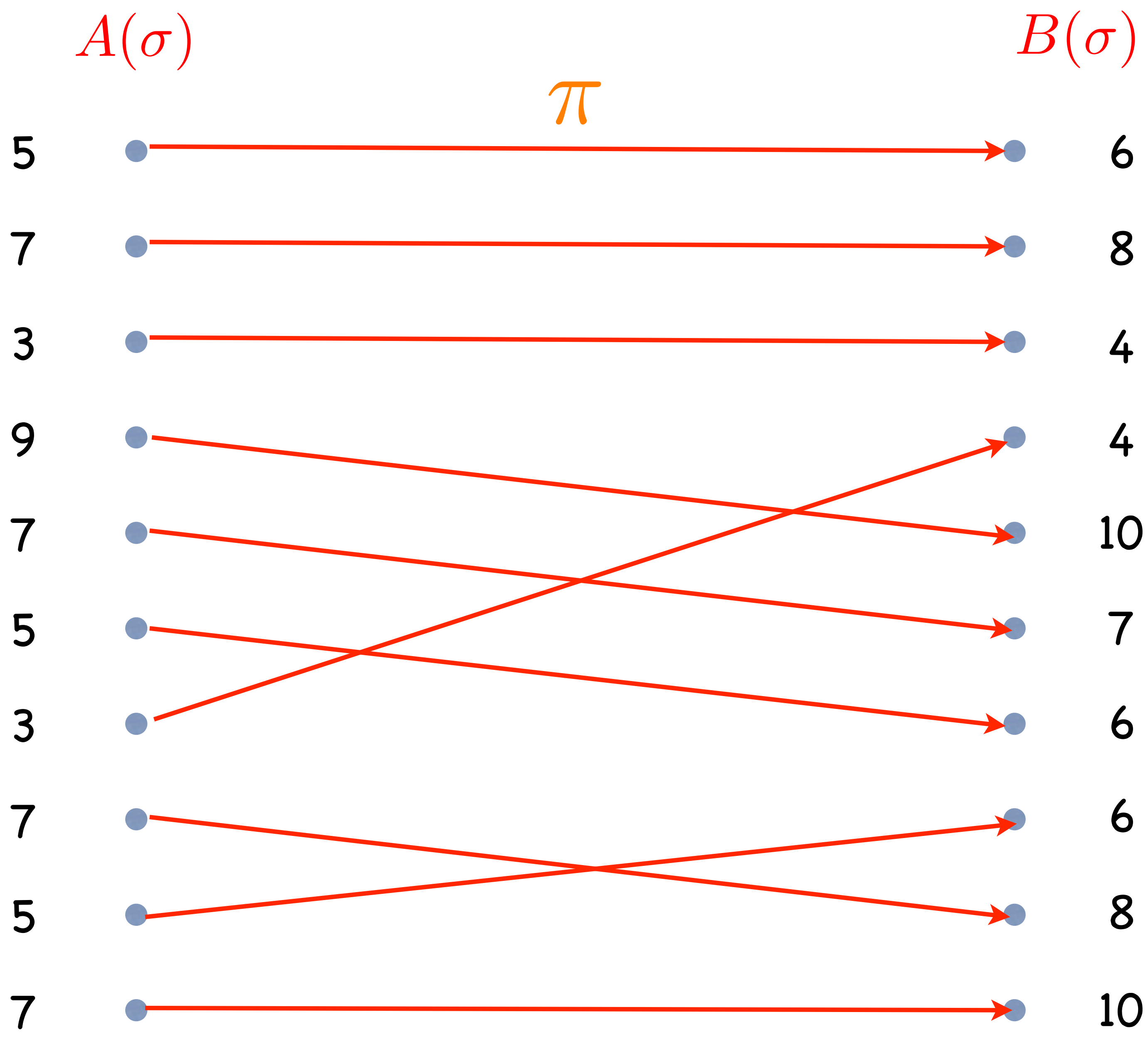
$A(\sigma)$

5
7
3
9
7
5
3
7
5
7

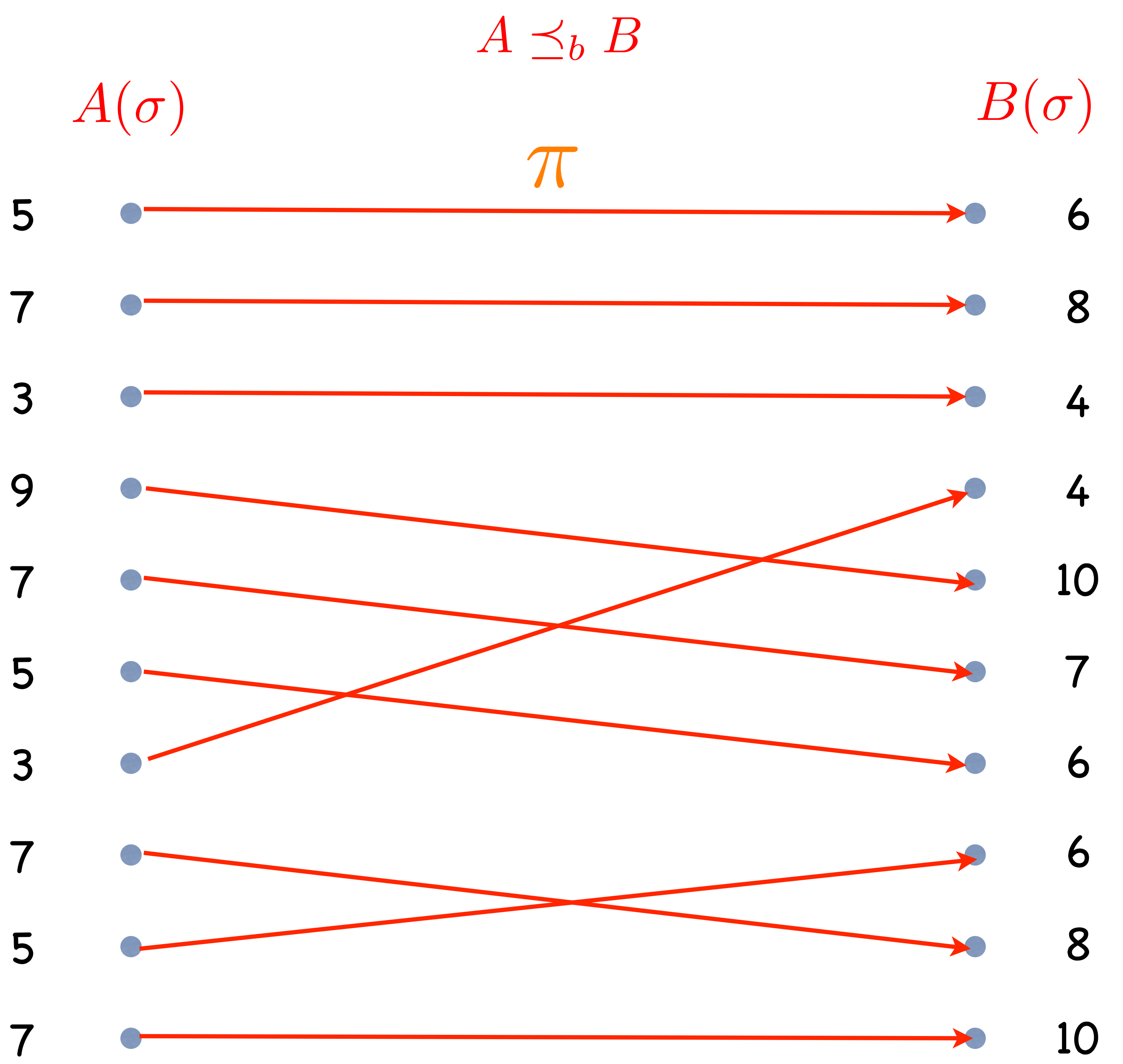
Bijective analysis [A., Dorrigiv, Lopez-Ortiz 07]

$A(\sigma)$	$B(\sigma)$
5	6
7	8
3	4
9	4
7	10
5	7
3	6
7	6
5	8
7	10

Bijective analysis [A., Dorrigiv, Lopez-Ortiz 07]



Bijective analysis [A., Dorrigiv, Lopez-Ortiz 07]



Stochastic dominance

Stochastic dominance

- A random variable X is *stochastically dominated* by a r.v. Y if for all $c \in \mathbb{R}$

$$\mathrm{P}r[X \leq c] \geq \Pr[Y \leq c]$$

Stochastic dominance

- A random variable X is *stochastically dominated* by a r.v. Y if for all $c \in \mathbb{R}$

$$\mathrm{P}r[X \leq c] \geq \Pr[Y \leq c]$$

- It follows that $\mathbb{E}(X) \leq \mathbb{E}(Y)$, and more strongly,

$F(c) \geq G(c)$, where *F* and *G* are the cdf's of X,Y

Stochastic dominance

- A random variable X is *stochastically dominated* by a r.v. Y if for all $c \in \mathbb{R}$

$$\Pr[X \leq c] \geq \Pr[Y \leq c]$$

- It follows that $\mathbb{E}(X) \leq \mathbb{E}(Y)$, and more strongly,

  $F(c) \geq G(c)$, where *F* and *G* are the cdf's of X,Y

- Moreover, for every increasing function *h,*

$$\mathbb{E}(h(X)) \leq \mathbb{E}(h(Y))$$

Stochastic dominance

- A random variable X is *stochastically dominated* by a r.v. Y if for all $c \in \mathbb{R}$

$$\Pr[X \leq c] \geq \Pr[Y \leq c]$$

- It follows that $\mathbb{E}(X) \leq \mathbb{E}(Y)$, and more strongly,

  $F(c) \geq G(c)$, where *F* and *G* are the cdf's of X,Y

- Moreover, for every increasing function *h,*

$$\mathbb{E}(h(X)) \leq \mathbb{E}(h(Y))$$

- Bijective analysis $\equiv$ Stochastic dominance under uniform distribution

Stochastic dominance

- A random variable X is *stochastically dominated* by a r.v. Y if for all $c \in \mathbb{R}$

$$\mathrm{P}r[X \leq c] \geq \Pr[Y \leq c]$$

- It follows that $\mathbb{E}(X) \leq \mathbb{E}(Y)$, and more strongly,

$F(c) \geq G(c)$, where *F* and *G* are the cdf's of X,Y

- Moreover, for every increasing function *h,*

$$\mathbb{E}(h(X)) \leq \mathbb{E}(h(Y))$$

- Bijective analysis $\equiv$ Stochastic dominance under uniform distribution

- Very strong relation, limited applications in online computing [Hiller and Vredeveld 08]

Previous work on bijective analysis / stochastic dominance

Previous work on bijective analysis / stochastic dominance

- Paging and list update [A.,Dorrigiv and Lopez-Ortiz 07, A. Dorrigiv and Lopez-Ortiz 08, A. and Schweitzer '13]

Previous work on bijective analysis / stochastic dominance

- Paging and list update [A.,Dorrigiv and Lopez-Ortiz 07, A. Dorrigiv and Lopez-Ortiz 08, A. and Schweitzer '13]
- Online bin coloring [Hiller and Vredeveld 08]

Previous work on bijective analysis / stochastic dominance

- Paging and list update [A.,Dorrigiv and Lopez-Ortiz 07, A. Dorrigiv and Lopez-Ortiz 08, A. and Schweitzer ’13]
- Online bin coloring [Hiller and Vredeveld 08]
- 2-server problem on a line of 3 points (“*baby server problem*”) [Boyar, Irani, Larsen 09]

Previous work on bijective analysis / stochastic dominance

- Paging and list update [A.,Dorrigiv and Lopez-Ortiz 07, A. Dorrigiv and Lopez-Ortiz 08, A. and Schweitzer ’13]
- Online bin coloring [Hiller and Vredeveld 08]
- 2-server problem on a line of 3 points (“*baby server problem*”) [Boyar, Irani, Larsen 09]
- *Two-headed disk problem* [Hofri 83, Sehari and Rotem 96]

Previous work on bijective analysis / stochastic dominance

- Paging and list update [A.,Dorrigiv and Lopez-Ortiz 07, A. Dorrigiv and Lopez-Ortiz 08, A. and Schweitzer ’13]
- Online bin coloring [Hiller and Vredeveld 08]
- 2-server problem on a line of 3 points (“*baby server problem*”) [Boyar, Irani, Larsen 09]
- *Two-headed disk problem* [Hofri 83, Sehari and Rotem 96]
- Routing in array Mesh networks [Mitzenmacher 96]

Deficiencies of bijective analysis

Deficiencies of bijective analysis

- Previous applications of the analysis were tailored to the problems at hand

Deficiencies of bijective analysis

- Previous applications of the analysis were tailored to the problems at hand
- It’s not easy to come up with a bijection

Deficiencies of bijective analysis

- Previous applications of the analysis were tailored to the problems at hand
- It's not easy to come up with a bijection
- Even worse, a bijection may not exist

Deficiencies of bijective analysis

- Previous applications of the analysis were tailored to the problems at hand
- It’s not easy to come up with a bijection
- Even worse, a bijection may not exist

Our main objectives:

Deficiencies of bijective analysis

- Previous applications of the analysis were tailored to the problems at hand
- It’s not easy to come up with a bijection
- Even worse, a bijection may not exist

Our main objectives:

- Extend the application of exact bijective analysis (when possible)

Deficiencies of bijective analysis

- Previous applications of the analysis were tailored to the problems at hand
- It's not easy to come up with a bijection
- Even worse, a bijection may not exist

Our main objectives:

- Extend the application of exact bijective analysis (when possible)
- Introduce a concept of approximate bijective analysis

More specifically....

More specifically....

- The previous approaches are “all or nothing”

More specifically....

- The previous approaches are “all or nothing”
- What if we suspect that an algorithm is “good” (but not optimal) ?

More specifically....

- The previous approaches are “all or nothing”
- What if we suspect that an algorithm is “good” (but not optimal) ?

Approximate variant of bijective analysis

Bijective ratio

Bijective ratio

Given two online algorithms A and B, we say that the bijective ratio of A against B is ρ if there exists a bijection π (defined over request sequences of a given size n) such that

$$A(\sigma) \leq \rho \cdot B(\pi(\sigma)), \text{ for all } n \geq n_0$$

Bijective ratio

Given two online algorithms A and B, we say that the bijective ratio of A against B is ρ if there exists a bijection π (defined over request sequences of a given size n) such that

$$A(\sigma) \leq \rho \cdot B(\pi(\sigma)), \text{ for all } n \geq n_0$$

The bijective ratio of A is ρ if A has bijective ratio ρ against all online algorithms

Bijective ratio

Given two online algorithms A and B, we say that the bijective ratio of A against B is ρ if there exists a bijection π (defined over request sequences of a given size n) such that

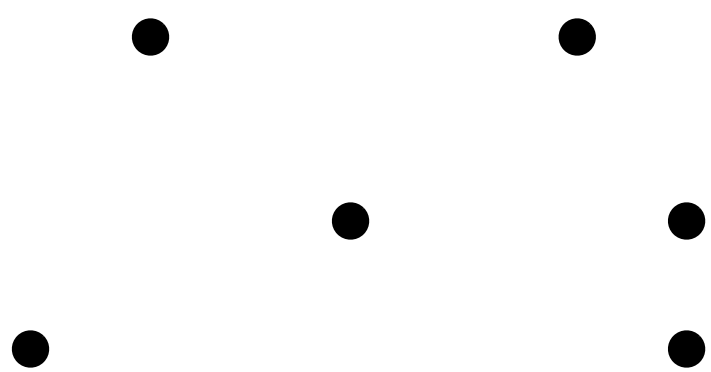
$$A(\sigma) \leq \rho \cdot B(\pi(\sigma)), \text{ for all } n \geq n_0$$

The bijective ratio of A is ρ if A has bijective ratio ρ against all online algorithms

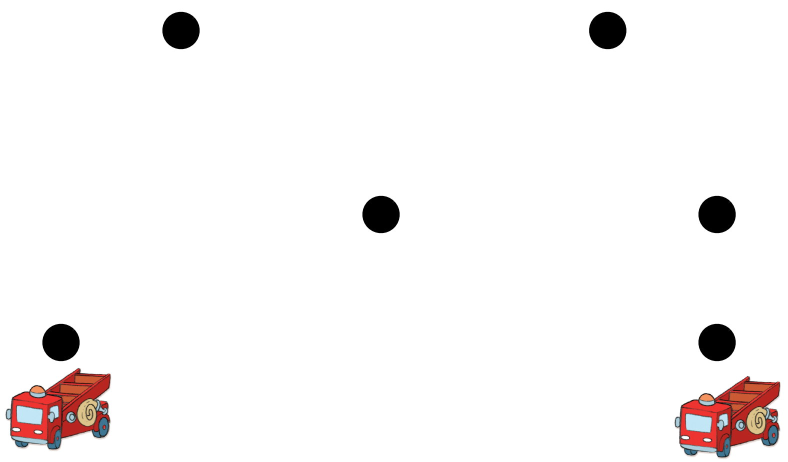
The bijective ratio of an online problem is determined by the best online algorithm

An application: the k-server problem

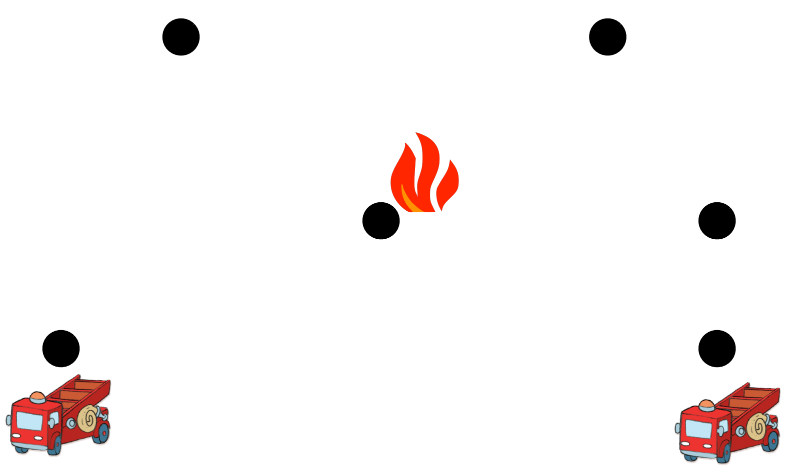
An application: the k-server problem



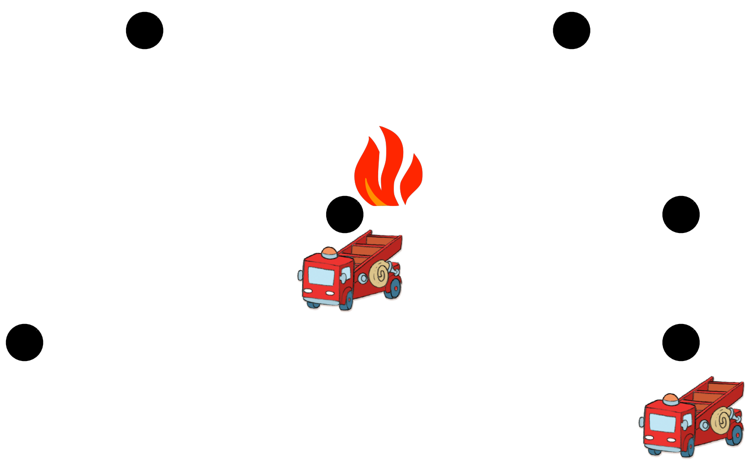
An application: the k-server problem



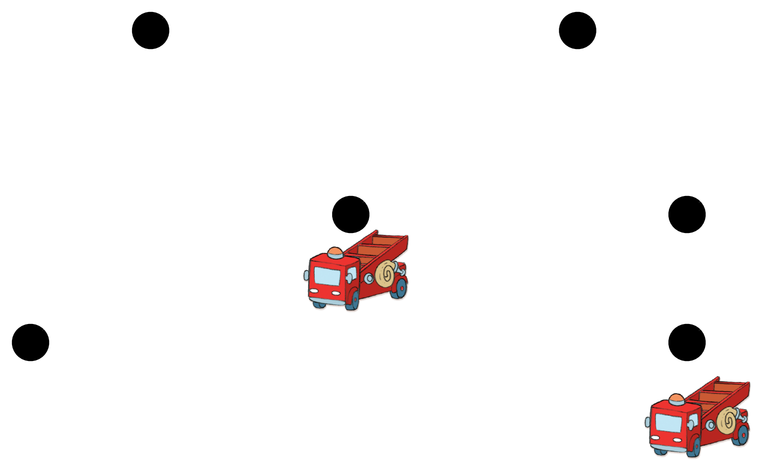
An application: the k-server problem



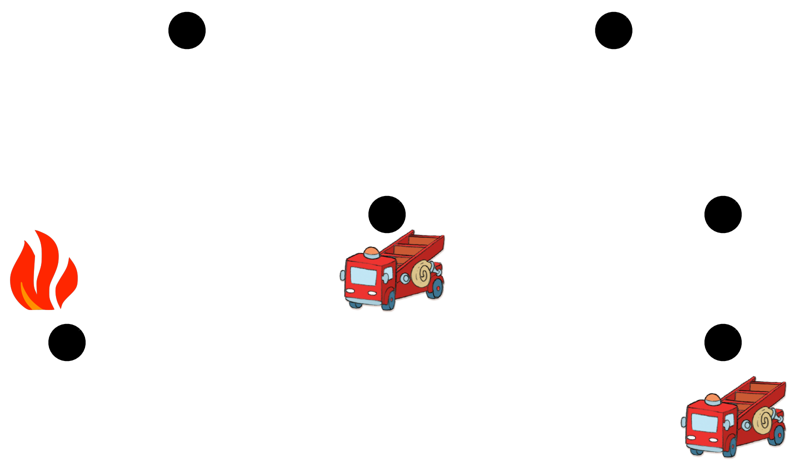
An application: the k-server problem



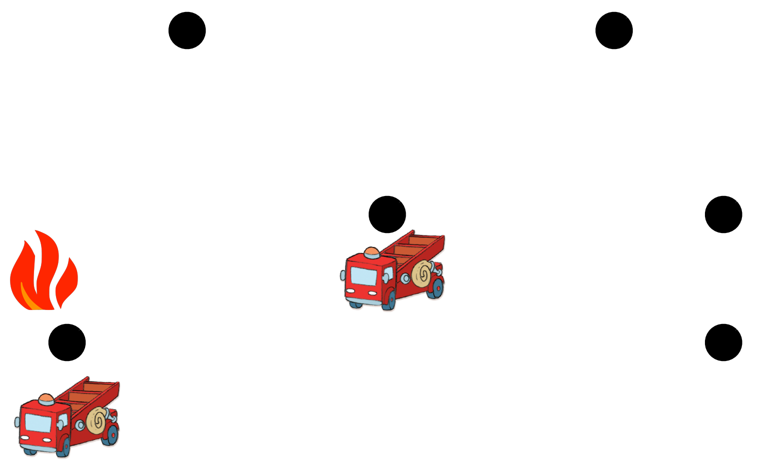
An application: the k-server problem



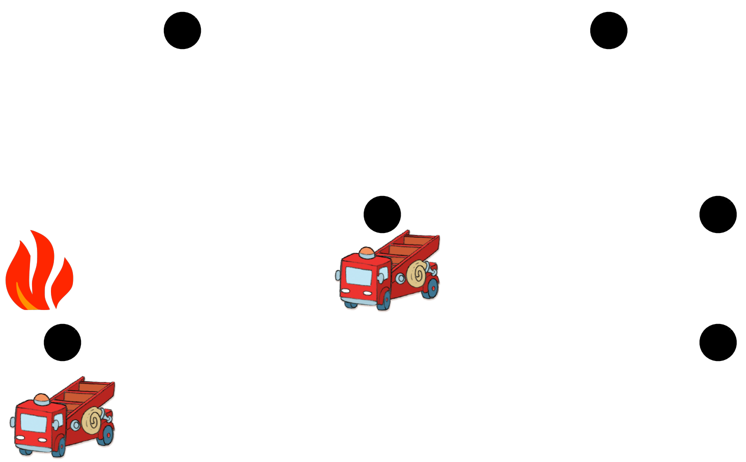
An application: the k-server problem



An application: the k-server problem

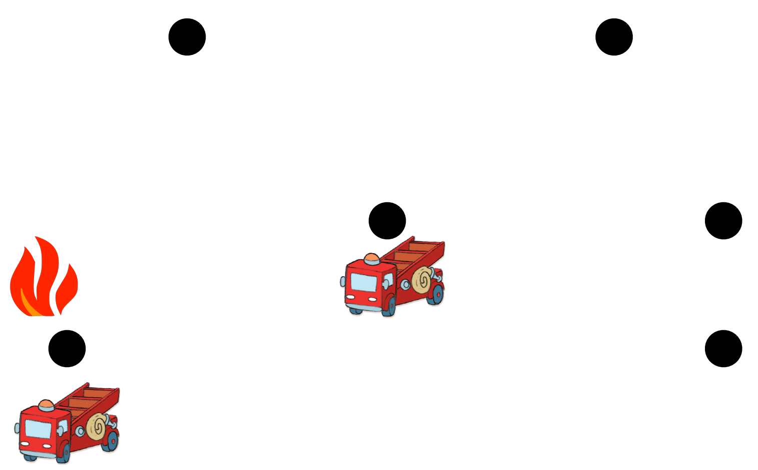


An application: the k-server problem



- request sequence: sequence of points
- cost of online algorithm: total cost for moving the servers
- likewise for opt (who knows the sequence)

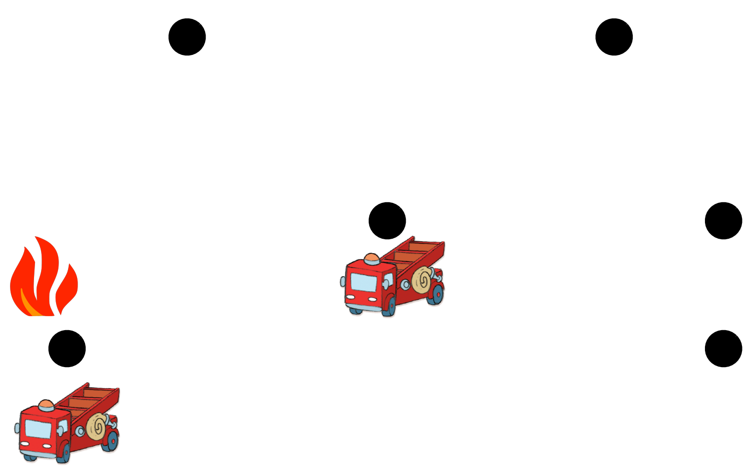
An application: the k-server problem



- request sequence: sequence of points
- cost of online algorithm: total cost for moving the servers
- likewise for opt (who knows the sequence)

Notable results

An application: the k-server problem

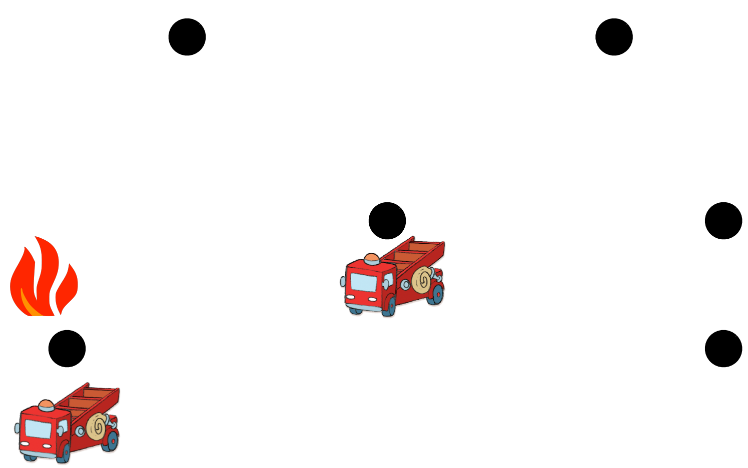


- request sequence: sequence of points
- cost of online algorithm: total cost for moving the servers
- likewise for opt (who knows the sequence)

Notable results

- (2k-1)-competitive algorithm for general metric spaces
  [Koutsoupias & Papadimitriou 95]

An application: the k-server problem

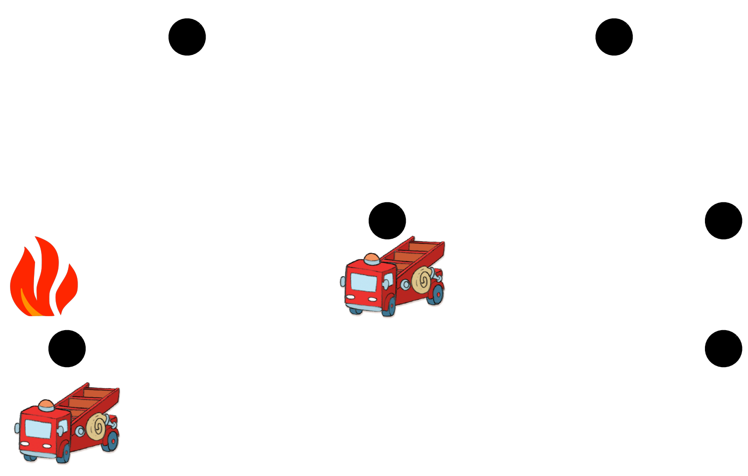


- request sequence: sequence of points
- cost of online algorithm: total cost for moving the servers
- likewise for opt (who knows the sequence)

Notable results

- (2k-1)-competitive algorithm for general metric spaces
  [Koutsoupias & Papadimitriou 95]
- k-competitive algorithms for trees (and thus the line too)
  [Chrobak & Larmore 91]

An application: the k-server problem

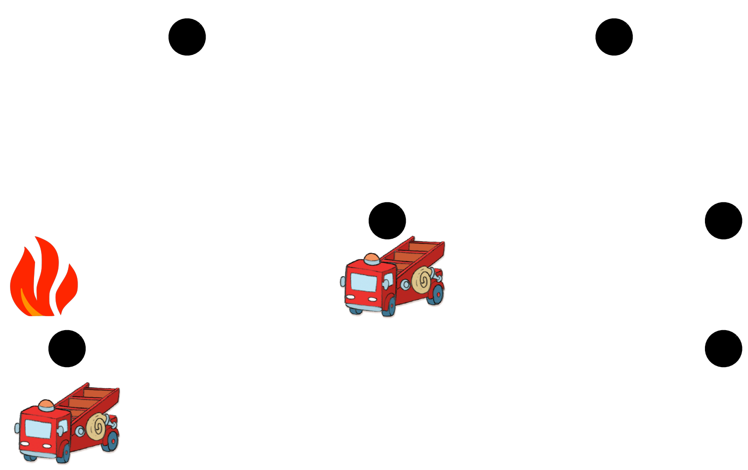


- request sequence: sequence of points
- cost of online algorithm: total cost for moving the servers
- likewise for opt (who knows the sequence)

Notable results

- (2k-1)-competitive algorithm for general metric spaces
  [Koutsoupias & Papadimitriou 95]
- k-competitive algorithms for trees (and thus the line too)
  [Chrobak & Larmore 91]
- Greedy has unbounded competitive ratio…

An application: the k-server problem



- request sequence: sequence of points
- cost of online algorithm: total cost for moving the servers
- likewise for opt (who knows the sequence)

Notable results

- (2k-1)-competitive algorithm for general metric spaces

  [Koutsoupias & Papadimitriou 95]

- k-competitive algorithms for trees (and thus the line too)

  [Chrobak & Larmore 91]

- Greedy has unbounded competitive ratio…
- ….but Greedy seems to be good on average

  [Calderbank et al. 85a, Calderbank et al. 85b, Anagnostopoulos et al. 10, Rudec et al. 10, Rudec et al. 13]

Some results

Some results

- Greedy has bijective ratio on the (continuous) line at most 4k/3

Some results

- Greedy has bijective ratio on the (continuous) line at most 4k/3
- On the circle Greedy has bijective ratio at most k and at least k/3

Some results

- Greedy has bijective ratio on the (continuous) line at most 4k/3
- On the circle Greedy has bijective ratio at most k and at least k/3
- k-Center has bijective ratio 2 for k-server on the line and the circle

Some results

- Greedy has bijective ratio on the (continuous) line at most 4k/3
- On the circle Greedy has bijective ratio at most k and at least k/3
- k-Center has bijective ratio 2 for k-server on the line and the circle
- k-Center has unbounded bijective ratio on the (continuous) star...

Some results

- Greedy has bijective ratio on the (continuous) line at most 4k/3
- On the circle Greedy has bijective ratio at most k and at least k/3
- k-Center has bijective ratio 2 for k-server on the line and the circle
- k-Center has unbounded bijective ratio on the (continuous) star...
- ...but Greedy has bijective ratio at most 4k

Outlook

Outlook

- Many, many open problems:

Outlook

- Many, many open problems: Discrete k-server

Outlook

- Many, many open problems:

Discrete k-server

Continuous k-server on trees: can we go below the competitive ratio?

Outlook

- Many, many open problems:

Discrete k-server

Continuous k-server on trees: can we go below the competitive ratio?

Semi-online algorithms

Outlook

- Many, many open problems:

Discrete k-server

Continuous k-server on trees: can we go below the competitive ratio?

Semi-online algorithms

Online Steiner tree problems

Outlook

- Many, many open problems:

Discrete k-server

Continuous k-server on trees: can we go below the competitive ratio?

Semi-online algorithms

Online Steiner tree problems

Experimental study of the bijective ratio

Outlook

- Many, many open problems:

  Discrete k-server

  Continuous k-server on trees: can we go below the competitive ratio?

  Semi-online algorithms

  Online Steiner tree problems

  Experimental study of the bijective ratio

- The big picture

Outlook

- Many, many open problems:
  - Discrete k-server
  - Continuous k-server on trees: can we go below the competitive ratio?
  - Semi-online algorithms
  - Online Steiner tree problems
  - Experimental study of the bijective ratio
- The big picture
  - New tools from a theoretical point of view ...

Outlook

- Many, many open problems:

  Discrete k-server

  Continuous k-server on trees: can we go below the competitive ratio?

  Semi-online algorithms

  Online Steiner tree problems

  Experimental study of the bijective ratio

- The big picture

  New tools from a theoretical point of view ...

  ... but hopefully also ideas for algorithms that perform better in practice

Outlook

- Many, many open problems:

  Discrete k-server

  Continuous k-server on trees: can we go below the competitive ratio?

  Semi-online algorithms

  Online Steiner tree problems

  Experimental study of the bijective ratio

- The big picture

  New tools from a theoretical point of view ...

  ... but hopefully also ideas for algorithms that perform better in practice

Thank you!

Outlook

Paper available on arXiv

- Many, many open problems:

  Discrete k-server

  Continuous k-server on trees: can we go below the competitive ratio?

  Semi-online algorithms

  Online Steiner tree problems

  Experimental study of the bijective ratio

- The big picture

  New tools from a theoretical point of view ...

  ... but hopefully also ideas for algorithms that perform better in practice

Thank you!