Online Algorithms

a topic in

DM573 - Introduction to Computer Science

Kim Skak Larsen

Department of Mathematics and Computer Science (IMADA)
University of Southern Denmark (SDU)

kslarsen@imada.sdu.dk https://imada.sdu.dk/u/kslarsen/

November 27, 2023

- A highly skilled, successful computer scientist is rewarded a ski vacation until
 her company desperately needs her again, at which point she is called home
 immediately, being notified when she wakes up in the morning.
- At the ski resort, skis cost 10 units in the local currency.
- One can rent skis for 1 unit per day.
- Of course, if she buys, she doesn't have to rent anymore.
- Which algorithm does she employ to minimize her spending?

- A highly skilled, successful computer scientist is rewarded a ski vacation until
 her company desperately needs her again, at which point she is called home
 immediately, being notified when she wakes up in the morning.
- At the ski resort, skis cost 10 units in the local currency.
- One can rent skis for 1 unit per day.
- Of course, if she buys, she doesn't have to rent anymore.
- Which algorithm does she employ to minimize her spending?

• What is a good algorithm?

- A highly skilled, successful computer scientist is rewarded a ski vacation until her company desperately needs her again, at which point she is called home immediately, being notified when she wakes up in the morning.
- At the ski resort, skis cost 10 units in the local currency.
- One can rent skis for 1 unit per day.
- Of course, if she buys, she doesn't have to rent anymore.
- Which algorithm does she employ to minimize her spending?

- What is a good algorithm?
- How do we measure if it's a good algorithm?

Competitive analysis is one way to measure the quality of an online algorithm.

Idea:

Let OPT denote an optimal offline algorithm.

"Offline" means getting the entire input before having to compute.

This corresponds to knowing the future!

We calculate how well we perform compared to OPT.

Competitive analysis is one way to measure the quality of an online algorithm.

Idea:

Let OPT denote an optimal offline algorithm.

"Offline" means getting the entire input before having to compute.

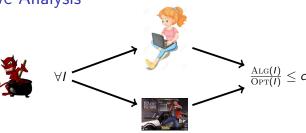
This corresponds to knowing the future!

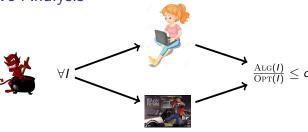
We calculate how well we perform compared to OPT.

Notation:

 $\mathrm{ALG}(I)$ denote the result of running an algorithm ALG on the input sequence I.

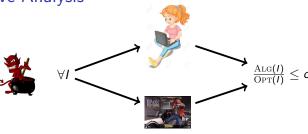
Thus, Opt(I) is the result of running Opt on I.





An algorithm, ALG, is c-competitive if

$$\forall I \colon \frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)} \leq c.$$

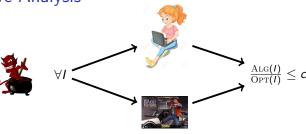


An algorithm, ALG, is *c-competitive* if

$$\forall I \colon \frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)} \leq c.$$

ALG has competitive ratio c if

c is the *best* (smallest) c for which ALG is c-competitive.



An algorithm, ALG, is *c-competitive* if

$$\forall I : \frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)} \leq c.$$

ALG has competitive ratio c if

c is the best (smallest) c for which ALG is c-competitive.

Technically, the definition of being c-competitive is that $\exists b \ \forall I \colon \mathrm{ALG}(I) \leq c \ \mathrm{OPT}(I) + b$, but the additive term, b, does not become relevant for what we consider. Also not relevant today, a best c does not necessarily exist, so the competitive ratio is inf $\{c \mid \mathrm{ALG} \text{ is } c\text{-competitive}\}$.

We have a very simple request sequence with very simple responses:

input sequence decision
it's a new day

input sequence	decision
it's a new day	rent

input sequence	decision
it's a new day	rent
it's a new day	

input sequence	decision
it's a new day	rent
it's a new day	rent

input sequence	decision
it's a new day	rent
it's a new day	rent
÷	:
it's a new day	

input sequence	decision
it's a new day	rent
it's a new day	rent
÷	:
it's a new day	rent

input sequence	decision
it's a new day	rent
it's a new day	rent
÷	:
it's a new day	rent
it's a new day	

input sequence	decision
it's a new day	rent
it's a new day	rent
÷	:
it's a new day	rent
it's a new day	buy

input sequence	decision
it's a new day	rent
it's a new day	rent
÷	:
it's a new day	rent
it's a new day	buy
it's a new day	

input sequence	decision
it's a new day	rent
it's a new day	rent
:	÷.
it's a new day	rent
it's a new day	buy
it's a new day	do nothing

input sequence	decision
it's a new day	rent
it's a new day	rent
÷	:
it's a new day	rent
it's a new day	buy
it's a new day	do nothing
it's a new day	do nothing
:	:

input sequence	decision
it's a new day	rent
it's a new day	rent
÷	:
it's a new day	rent
it's a new day	buy
it's a new day	do nothing
it's a new day	do nothing
÷	:
come home	go home



• So all reasonable algorithms are of the form **Buy on day X** (or never buy).



SDU 🏠

Ski Rental – a simplest online problem

- So all reasonable algorithms are of the form Buy on day X (or never buy).
- Reminder: It costs 10 to buy skis and 1 to rent.

- So all reasonable algorithms are of the form Buy on day X (or never buy).
- Reminder: It costs 10 to buy skis and 1 to rent.
- We choose the algorithm: **Buy on day 10**.

- So all reasonable algorithms are of the form Buy on day X (or never buy).
- Reminder: It costs 10 to buy skis and 1 to rent.
- We choose the algorithm: Buy on day 10.
- The optimal offline algorithm, OPT:

```
d = the number of days we get to stay
```

if d < 10:

rent every day

else: # $d \ge 10$

buy on day 1

- So all reasonable algorithms are of the form Buy on day X (or never buy).
- Reminder: It costs 10 to buy skis and 1 to rent.
- We choose the algorithm: Buy on day 10.
- The optimal offline algorithm, OPT:

```
d = the number of days we get to stay if d < 10:
```

rent every day

else: $\# d \ge 10$ buy on day 1

- ullet If we are called home *before* day 10, we (**Buy on day 10**) are optimal!
 - Our cost: d < 10.
 - Opt's cost: d.

- So all reasonable algorithms are of the form Buy on day X (or never buy).
- Reminder: It costs 10 to buy skis and 1 to rent.
- We choose the algorithm: **Buy on day 10**.
- The optimal offline algorithm, OPT:

```
d= the number of days we get to stay if d<10:
rent every day
```

rent every day else: $\# d \ge 10$

buy on day 1

- ullet If we are called home *before* day 10, we (**Buy on day 10**) are optimal!
 - Our cost: d < 10.
 - Opt's cost: d.
- In the worst case, we are called home at (or any time after) day 10:
 - Our cost: 9 + 10.
 OPT's cost: 10.

- So all reasonable algorithms are of the form Buy on day X (or never buy).
- Reminder: It costs 10 to buy skis and 1 to rent.
- We choose the algorithm: **Buy on day 10**.
- The optimal offline algorithm, OPT: $d = \text{the number of days we get to stay} \\ \text{if } d < 10:$
 - rent every day else: # d > 10
 - buy on day 1
- ullet If we are called home *before* day 10, we (**Buy on day 10**) are optimal!
 - Our cost: *d* < 10.
 - Opt's cost: d.
- In the worst case, we are called home at (or any time after) day 10:
 - Our cost: 9 + 10.
 - Opt's cost: 10.
- Result: we perform at most $\frac{19}{10} < 2$ times worse than OPT.

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

That algorithm never performs worse than $\frac{19}{10}$ times OPT.

Is our algorithm $\frac{19}{10}$ -competitive?

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

That algorithm never performs worse than $\frac{19}{10}$ times $\mathrm{OPT}.$

Is our algorithm $\frac{19}{10}$ -competitive?

Yes

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

That algorithm never performs worse than $\frac{19}{10}$ times $\mathrm{OPT}.$

Is our algorithm $\frac{19}{10}$ -competitive?

Yes

Is our algorithm also 2-competitive?

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

Is our algorithm $\frac{19}{10}$ -competitive?	Yes
Is our algorithm also 2-competitive?	Yes

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

Is our algorithm $\frac{19}{10}$ -competitive?	Yes
Is our algorithm also 2-competitive?	Yes
Is our algorithm also 42-competitive?	

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

Is our algorithm $\frac{19}{10}$ -competitive?	Yes
Is our algorithm also 2-competitive?	Yes
Is our algorithm also 42-competitive?	Yes

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

Is our algorithm $\frac{19}{10}$ -competitive?	Yes
Is our algorithm also 2-competitive?	Yes
Is our algorithm also 42-competitive?	Yes
Is our algorithm 1.5-competitive?	

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

Is our algorithm $\frac{19}{10}$ -competitive?	Yes
Is our algorithm also 2-competitive?	Yes
Is our algorithm also 42-competitive?	Yes
Is our algorithm 1.5-competitive?	No

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

Is our algorithm $\frac{19}{10}$ -competitive?	Yes
Is our algorithm also 2-competitive?	Yes
Is our algorithm also 42-competitive?	Yes
Is our algorithm 1.5-competitive?	No
Does our algorithm have competitive ratio $\frac{19}{10}$?	

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

Is our algorithm $\frac{19}{10}$ -competitive?	Yes
Is our algorithm also 2-competitive?	Yes
Is our algorithm also 42-competitive?	Yes
Is our algorithm 1.5-competitive?	No
Does our algorithm have competitive ratio $\frac{19}{10}$?	Yes

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

Is our algorithm $\frac{19}{10}$ -competitive?	Yes
Is our algorithm also 2-competitive?	Yes
Is our algorithm also 42-competitive?	Yes
Is our algorithm 1.5-competitive?	No
Does our algorithm have competitive ratio $\frac{19}{10}$?	Yes
Does our algorithm have competitive ratio 2?	

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

Is our algorithm $\frac{19}{10}$ -competitive?	Yes
Is our algorithm also 2-competitive?	Yes
Is our algorithm also 42-competitive?	Yes
Is our algorithm 1.5-competitive?	No
Does our algorithm have competitive ratio $\frac{19}{10}$?	Yes
Does our algorithm have competitive ratio 2?	No

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

Is our algorithm $\frac{19}{10}$ -competitive?	Yes
Is our algorithm also 2-competitive?	Yes
Is our algorithm also 42-competitive?	Yes
Is our algorithm 1.5-competitive?	No
Does our algorithm have competitive ratio $\frac{19}{10}$?	Yes
Does our algorithm have competitive ratio 2?	No
Does our algorithm have competitive ratio 1.5?	

In general, we want to *define* good algorithms for online problems and *prove* that they are good.

The ratio $(\frac{19}{10})$ from ski rental is a guarantee:

Is our algorithm $\frac{19}{10}$ -competitive?	Yes
Is our algorithm also 2-competitive?	Yes
Is our algorithm also 42-competitive?	Yes
Is our algorithm 1.5-competitive?	No
Does our algorithm have competitive ratio $\frac{19}{10}$?	Yes
Does our algorithm have competitive ratio 2?	No
Does our algorithm have competitive ratio 1.5?	No

Online Problems

So, what characterizes an online problem?

- Input arrives one request at a time.
- For each request, we have to make an irrevocable decision.
- We want to minimize cost.

Online Problems

So, what characterizes an online problem?

- Input arrives one request at a time.
- For each request, we have to make an irrevocable decision.
- We want to minimize cost.

Sometimes we want to maximize a profit instead of minimizing a cost and there are some technicalities in adjusting definitions to accommodate that possibility.

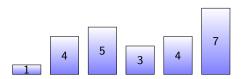
Machine Scheduling

- $m \ge 1$ machines.
- *n* jobs of varying sizes arriving one at a time to be assigned to a machine.
- The goal is to minimize makespan, i.e., finish all jobs as early as possible.

Machine Scheduling

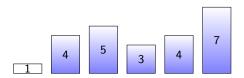
- $m \ge 1$ machines.
- *n* jobs of varying sizes arriving one at a time to be assigned to a machine.
- The goal is to minimize makespan, i.e., finish all jobs as early as possible.
- \bullet Algorithm List Scheduling (Ls): place next job on the least loaded machine.

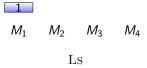




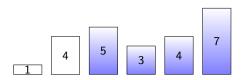
 M_1 M_2 M_3 M_4

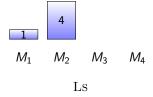




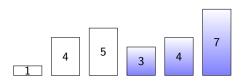


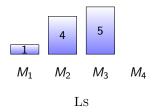




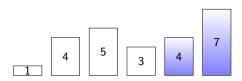


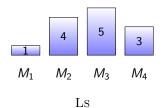




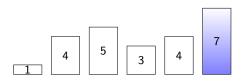


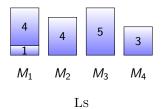




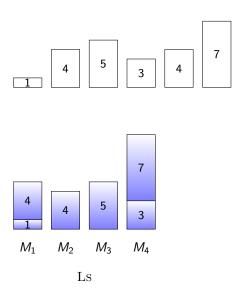




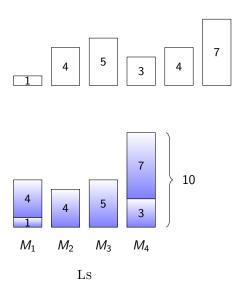




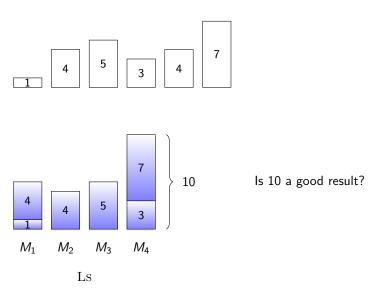












On some sequences, 10 is a good result.

On some sequences, 1.000.000 is a good result.

Sometimes 1000 is a bad result.

It depends on what is possible, i.e., how large or "difficult" the input is.

Competitive Analysis – why OPT?

OPT is not an online algorithm.

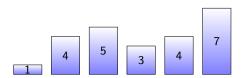
However, it is a natural reference point:

- Our online algorithms cannot perform better than OPT.
- As input sequences I get harder, OPT(I) typically grows, so a comparison to OPT remains somewhat reasonable.

Formally:

- For any ALG and any I, $\frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)} \geq 1$.
- We want to design algorithms with smallest possible competitive ratio.

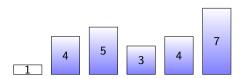


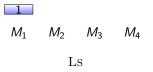


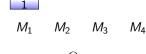
 M_1 M_2 M_3 M_4 Ls

 M_1 M_2 M_3 M_4



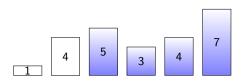


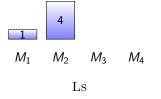


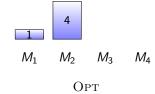


Opt

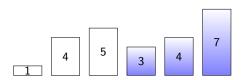




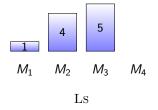


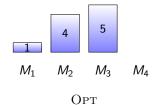




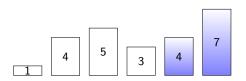


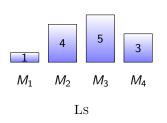
What does OPT do now?

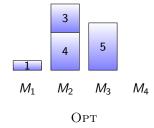




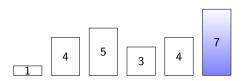


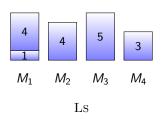


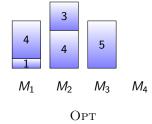




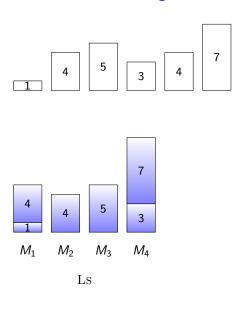


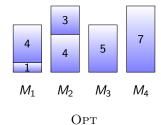




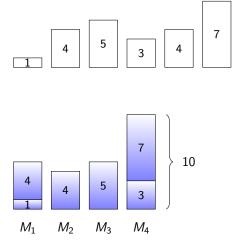


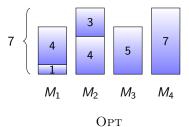












Ls



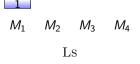


$$M_1$$
 M_2 M_3 M_4

$$M_1$$
 M_2 M_3 M_4



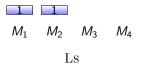




$$M_1$$
 M_2 M_3 M_4 OPT



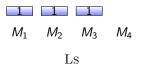








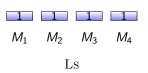


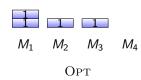






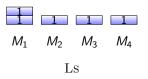


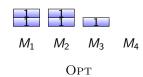






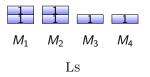


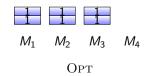






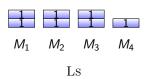


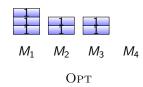






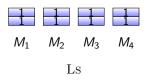


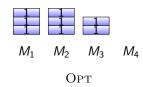






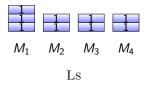


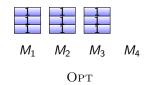






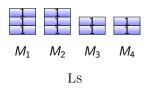


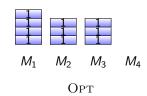






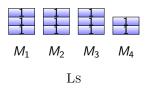


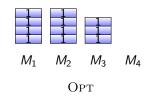






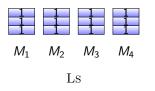


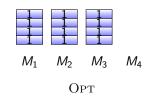






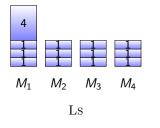


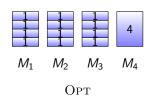






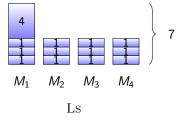








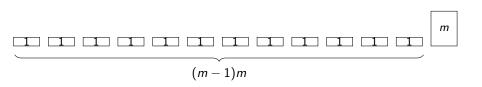


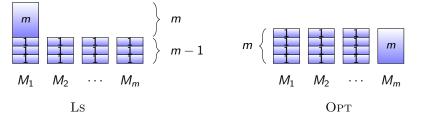


Thus,

$$\frac{\mathrm{Ls}(I)}{\mathrm{Opt}(I)} = \frac{7}{4} = 2 - \frac{1}{4}$$

Machine Scheduling – Ls is at best $(2-\frac{1}{m})$ -competitive





In general,

$$\frac{\operatorname{Ls}(I)}{\operatorname{OPT}(I)} \geq \frac{(m-1)+m}{m} = \frac{2m-1}{m} = 2 - \frac{1}{m}$$

Machine Scheduling - proof techniques

You have just seen a

lower bound proof

In our context, that is relatively easy: Just demonstrate an example (family of examples), where the algorithm performs worse than some ratio.

Machine Scheduling - proof techniques

You have just seen a

lower bound proof

In our context, that is relatively easy: Just demonstrate an example (family of examples), where the algorithm performs worse than some ratio.

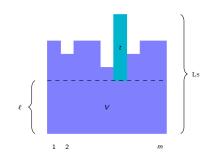
Now we want to show the much harder

upper bound proof

We must show that it is *never* worse than a given ratio for *any* of the (potentially infinitely many) possible input sequences.

Machine Scheduling – Ls is $(2 - \frac{1}{m})$ -competitive





t is the length of the job starting at ℓ and ending at the makespan

T is the total length of all jobs

Define $V = \ell \cdot m$

Now conclude:

 $Opt \ge T/m$ and $Opt \ge t$

 $T \ge V + t$, due to Ls's choice for t

Ls =
$$\ell + t$$

 $\leq \frac{T-t}{m} + t$, since $\ell = \frac{V}{m}$ and $V \leq T - t$
= $\frac{T}{m} + (1 - \frac{1}{m})t$
 $\leq \text{OPT} + (1 - \frac{1}{m}) \text{ OPT}$, from OPT inequalities above
= $(2 - \frac{1}{m}) \text{ OPT}$

Machine Scheduling

Since we have both an upper and lower bound, we have shown the following:

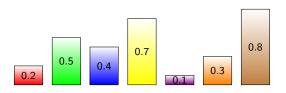
Theorem

The algorithm Ls for minimizing makespan in machine scheduling with $m \ge 1$ machines has competitive ratio $2 - \frac{1}{m}$.

Bin Packing

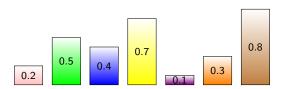
- Unbounded supply of bins of size 1.
- *n* items, each of size strictly between zero and one, to be placed in a bin.
- Obviously, the items placed in any given bin cannot be larger than 1 in total.
- The goal is to minimize the number of bins used.
- ullet Algorithm First-Fit (FF): place next item in the first bin with enough space. The ordering of the bins is determined by the first time they receive an item.

Bin Packing -FF example



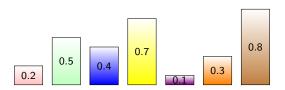


Bin Packing – FF example



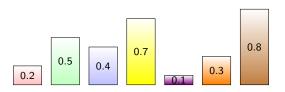


Bin Packing – F_F example



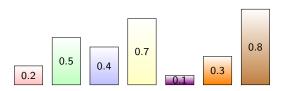


Bin Packing – F_F example



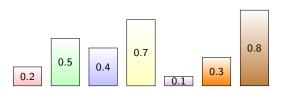


Bin Packing – FF example



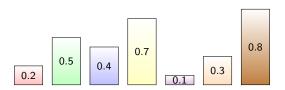


Bin Packing – FF example





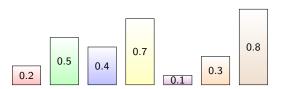
Bin Packing -FF example





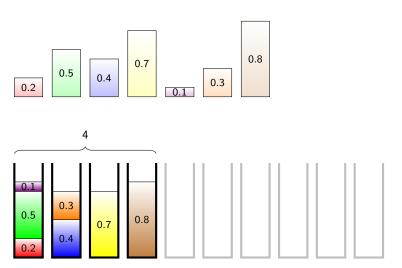
FF

Bin Packing – F_F example





Bin Packing – F_F example



FF



FF

Bin Packing – simple lower bound against ${ m FF}$

1/3 1/3 1/3 1/3 1/2 1/2 1/2 1/2

All items are slightly larger than the indicated fraction, e.g., $\frac{1}{3} + \frac{1}{1000}$.



FF

Bin Packing – simple lower bound against ${ m FF}$

1/3 1/3 1/3 1/3 1/2 1/2 1/2 1/2

All items are slightly larger than the indicated fraction, e.g., $\frac{1}{3} + \frac{1}{1000}$.





 FF

Opt

1/3 1/3 1/3 1/3 1/2 1/2 1/2 1/2

All items are slightly larger than the indicated fraction, e.g., $\frac{1}{3} + \frac{1}{1000}$.



FF OPT

1/3 1/3 1/3



 F_{F}





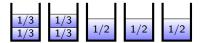
 F_{F}





 F_{F}





 1/2
 1/2
 1/2

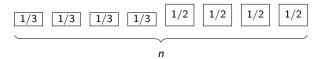
 1/3
 1/3
 1/3

FF OPT

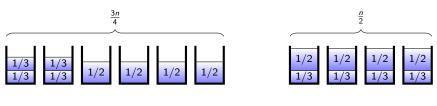


 F_{F}

Bin Packing – simple lower bound against FF

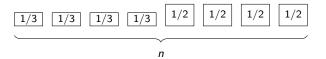


All items are slightly larger than the indicated fraction, e.g., $\frac{1}{3} + \frac{1}{1000}$.

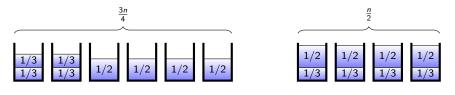


Opt

Bin Packing – simple lower bound against ${\rm FF}$



All items are slightly larger than the indicated fraction, e.g., $\frac{1}{3} + \frac{1}{1000}$.



FF

OPT

Thus,

$$\frac{\mathrm{Fr}(I)}{\mathrm{OPT}(I)} \ge \frac{\frac{n}{4} + \frac{n}{2}}{\frac{n}{2}} = \frac{3}{2}$$

Bin Packing - First-Fit

- FF has been shown to be 1.7-competitive [hard]. Thus, the competitive ratio is *at most* 1.7.
- We have just shown that the competitive ratio is at least 1.5.
- So, the competitive ratio of FF is in the interval [1.5, 1.7].
- We'll approach the precise value in the exercises.

How To Learn More

You'll meet online algorithms in courses, without necessarily being told that they are *online* problems.

A formal treatment of this topic is somewhat abstract and heavy on proofs, and more appropriate for the MS level. At that point, you can take the course

DM860: Online Algorithms

or read one of

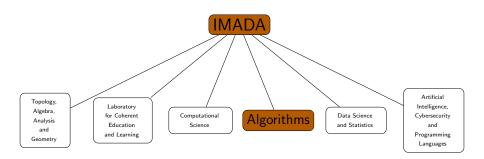




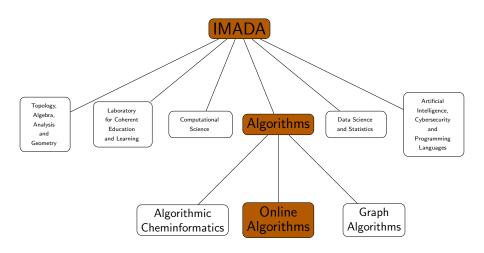
IMADA People in Online Algorithms

Online Algorithms		
	Joan Boyar	Online algorithms, combinatorial optimization, cryptology, computational complexity
	Lene Monrad Favrholdt	Online algorithms, graph algorithms
	Kim Skak Larsen	Online algorithms, algorithms and data structures, database systems, semantics
	Kevin Schewior	Online algorithms, algorithms under uncertainty, approximation algorithms
	Magnus Berg	Online algorithms





IMADA Research Structure



Online Algorithms Group

___ _ Online Algorithms

Research ▼

D===l==

News

Connections

For Students

Contact



Group photo: Revin Schewior, Lene M. Favrholdt, Joan Boyar, Rim S. Larsen.

Online algorithms deal with a particularly challenging branch of the algorithms area. Many real world problems have an online component, some decisions that have to be made immediately, before all related data is available. An example of an online problem is the seat reservation problem, where, as with the IC trains in Demmark, a given train has a fixed number of seats and passengers can make seat reservations before their trip and be given seat numbers immediately. In contrast, if the problem were "offline", the seat reservation system could wait until it had all reservation requests before assigning any seats. The online properly limits how well a system can perform in optimizing for its objective (maximizing the number of passengers able to buy tickets, for example). However, it may be possible for a system to perform quite well, even if the problem is online, as long as the correct online algorithm (method of solving the online problem) is used. The study of online algorithms is aimed at designing and analysing online algorithms.

In analyzing algorithms, it is very useful to have a theoretical tool which can be used in predicting how well algorithms will do in practice. The group has been working extensively on this, concentrating on alternatives to the standard quality measure for online algorithms, the competitive ratio. They have defined and are investigation the relative worst order ratio, which niese results expend to those optained with ormonethive.

Online Algorithms – Subgroup of Algorithms



Research -

People ▼

Subgroups ▼

New

Connections

For Students



Group photo (left to right): Rolf Fagerberg, Yun Wang, Lene M. Favrholdt, Kevin Schewior, Jørgen Bang-Jensen, Joan Boyar, Anders Yeo, Jakob L. Andersen, Simon Erfurth, Kim S. Larsen, Daniel Merkle.

In recent years, the word "algorithms" has become part of everyone's vocabulary. When newscasters explain that some software is really amazing, they state that there is "intelligence" or "algorithms" in the product. The recent popularity of the word "algorithms" is well deserved. In fact, what is often referred to as "intelligence" in systems is really huge amounts of data combined with advanced algorithms for searching and computing hased on this charge.

Algorithms are at the core of Computer Science, so in addition to offering expertise in concrete application areas, a solid background in algorithms makes it easy to enter other subareas of Computer Science; also after

Online Algorithms

a topic in

DM573 - Introduction to Computer Science

Kim Skak Larsen

Department of Mathematics and Computer Science (IMADA)
University of Southern Denmark (SDU)

kslarsen@imada.sdu.dk https://imada.sdu.dk/u/kslarsen/

November 27, 2023