Online Algorithms

a topic in

DM534 - Introduction to Computer Science

Kim Skak Larsen

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

kslarsen@imada.sdu.dk

November 17, 2016

- 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:

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

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

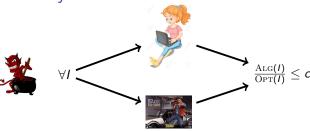


An algorithm, ALG, is *c-competitive* if

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

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

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

ALG has competitive ratio c if

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

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 b-set c does not necessarily exist, so the competitive ratio is inf $\{c \mid \mathrm{ALG} \text{ is } c\text{-competitive}\}$.

4 U 2 4 D2 2 4 E 2 4 E 2 E 3

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

input sequence decision
it's a new day

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

input sequence	decision
it's a new day	rent

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

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

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

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

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	

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

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
:	:
it's a new day	buy do nothing

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

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).

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

- So all reasonable algorithms are of the form Buy on day X (or never buy).
- Reminder: It costs 10 to buy skis.
- 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.
- We choose the algorithm: Buy on day 10.
- The optimal offline algorithm, OPT: d= the day we are called home if d<10: rent every day else: $\#\ d\geq 10$ 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.
- We choose the algorithm: Buy on day 10.
- The optimal offline algorithm, OPT: d= the day we are called home if d<10: rent every day else: $\#\ d\geq 10$

buy on day 1

- If we are called home before day 10, we 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.
- We choose the algorithm: Buy on day 10.
- The optimal offline algorithm, OPT: d= the day we are called home if d<10: rent every day else: $\# d \geq 10$
 - buy on day 1
- If we are called home before day 10, we are optimal!
 - Our cost: d < 10.
 - OPT's cost: d.
- In the worst case, we are called home (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.
- We choose the algorithm: Buy on day 10.
- The optimal offline algorithm, OPT: d= the day we are called home if d<10: rent every day else: $\# d \geq 10$
 - buy on day 1
- If we are called home before day 10, we are optimal!
 - Our cost: *d* < 10.
 - OPT's cost: d.
- In the worst case, we are called home (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:

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

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

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 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 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:

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

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:

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

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:

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

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:

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

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:

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

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*?

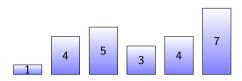
- Input arrives one request at a time.
- For each request, we have to make an irrevocable decision.
- We want to minimize cost.

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

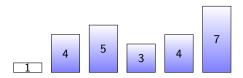
4 D > 4 A > 4 B > 4 B >

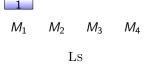
Machine Scheduling

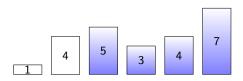
- $m \ge 1$ machines.
- *n* jobs of varying sizes arriving one at a time to be assigned to a machine.
- The goal is to minimize makespan, i.e., finish all jobs as early as possible.
- Algorithm List Scheduling (Ls): place next job on the least loaded machine.

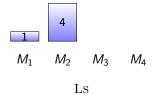


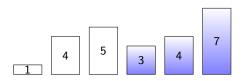
 M_1 M_2 M_3 M_4

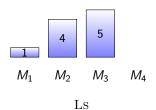


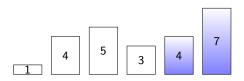


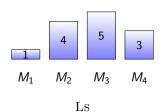


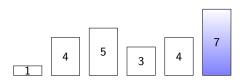


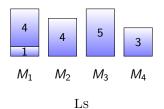


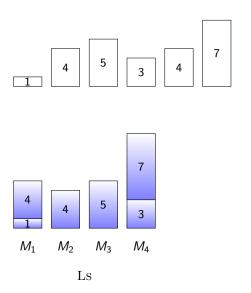


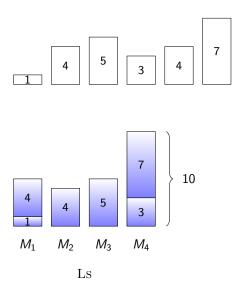


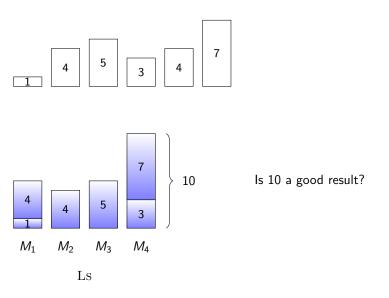












On some sequences, 10 is a good result.

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?

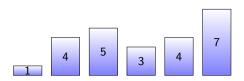
OPT is not an online algorithm.

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:

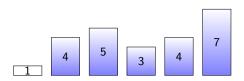
- For any ALG and any I, $\frac{\mathrm{ALG}(I)}{\mathrm{OPT}(I)} \geq 1$.
- We want to design algorithms with smallest possible competitive ratio.

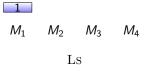


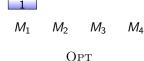
 M_1 M_2 M_3 M_4 Ls

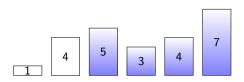
 M_1 M_2 M_3 M_4

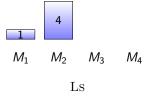
Opt

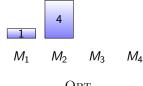




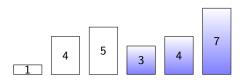




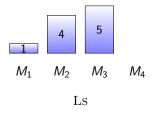


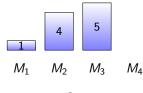


Орт

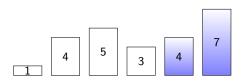


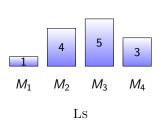
What does OPT do now?

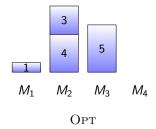


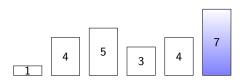


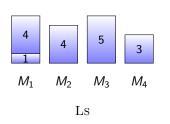
Opt

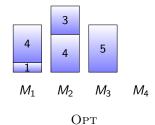


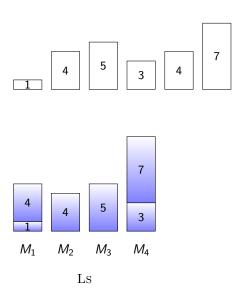


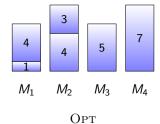


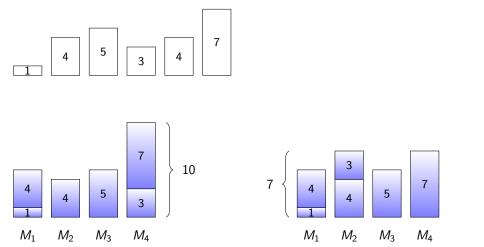












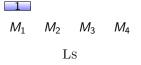
Орт

Ls



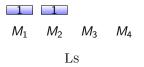
$$M_1$$
 M_2 M_3 M_4 M_1 M_2 M_3 M_4 M_5 Opt





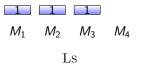
$$M_1$$
 M_2 M_3 M_4 OPT

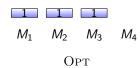




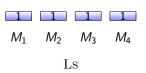


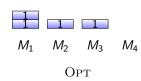




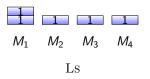


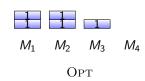




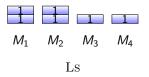


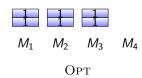




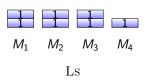


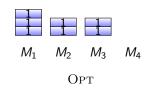






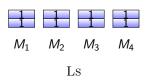


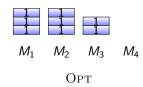




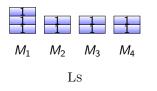
14 / 23

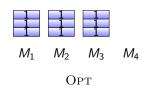




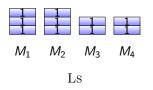


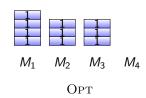




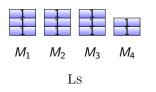


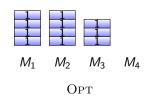




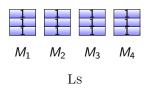


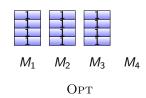




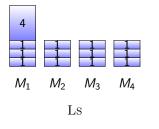


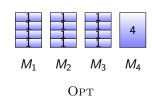




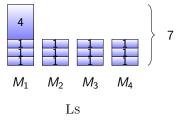






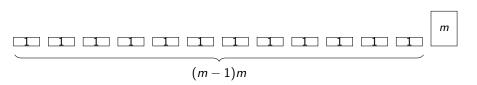


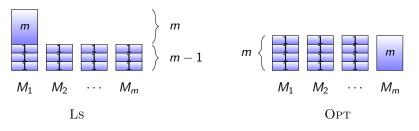




Thus,

$$\frac{\mathrm{ALG}(\mathit{I})}{\mathrm{OPT}(\mathit{I})} = \frac{7}{4} = 2 - \frac{1}{4}$$





In general,

$$\frac{\mathrm{ALG}(\mathit{I})}{\mathrm{OPT}(\mathit{I})} \geq \frac{(m-1)+m}{m} = \frac{2m-1}{m} = 2 - \frac{1}{m}$$

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.

15 / 23

Machine Scheduling - proof techniques

You have just seen a

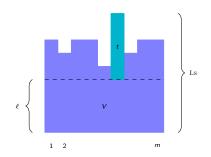
lower bound proof

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.



t ends at the makespan and start at ℓ $t_{\rm max}$ is the length of the longest job T is the total length of all jobs ${
m OPT} \geq T/m$ ${
m OPT} \geq t_{
m max}$ $V = \ell \cdot m$

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

Ls =
$$\ell + t$$

 $\leq \frac{T-t}{m} + t$, since $\ell = \frac{V}{m}$ and $V \leq T - t$
= $\frac{T}{m} + (1 - \frac{1}{m})t$
 $\leq \frac{T}{m} + (1 - \frac{1}{m})t_{\max}$
 $\leq \text{OPT} + (1 - \frac{1}{m})\text{ OPT}$, 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

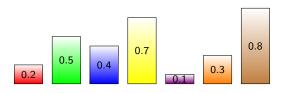
The algorithm Ls for minimizing makespan in machine scheduling with $m \ge 1$ machines has competitive ratio $2 - \frac{1}{m}$.

17 / 23

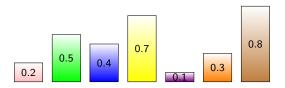
Bin Packing

- 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.
- 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.

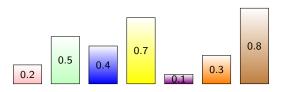
18 / 23



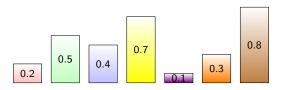




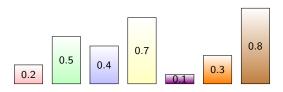




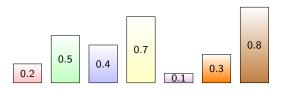




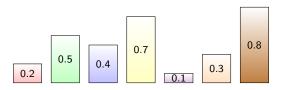




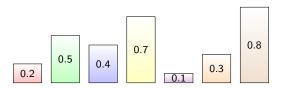






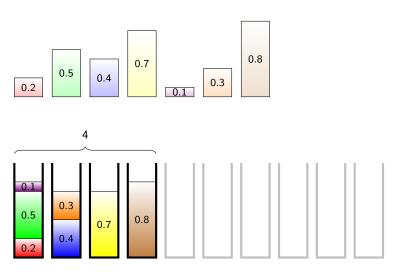








FF



FF

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

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

FF

Орт

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

 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
 1/2
 1/2
 1/2
 1/2

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





 Fr

Opt

20 / 23

Bin Packing – simple lower bound against FF

1/2 1/2

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



FF



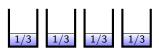
OPT

 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., $\frac{1}{3} + \frac{1}{1000}$.



 F_{F}



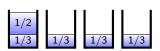
Орт

 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., $\frac{1}{3} + \frac{1}{1000}$.



 F_{F}



Орт

20 / 23

 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., $\frac{1}{3} + \frac{1}{1000}$.



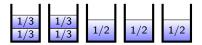
 $F_{\mathbf{F}}$



Орт

 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., $\frac{1}{3} + \frac{1}{1000}$.



Орт

FF

20 / 23

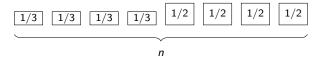
Bin Packing – simple lower bound against ${\rm FF}$

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

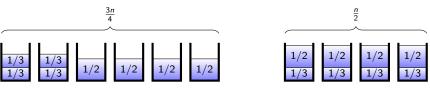
All items are slightly larger than the indicated fraction, e.g., $\frac{1}{3} + \frac{1}{1000}$.



FF OPT

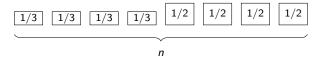


All items are slightly larger than the indicated fraction, e.g., $\frac{1}{3} + \frac{1}{1000}$.

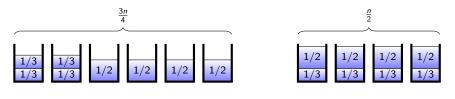


 F_{F}

OPT



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



 F_{F}

OPT

Thus,

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

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.
- ullet So, the competitive ratio of F_F 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



IMADA People in Online Algorithms

Online Algorithms		
	Joan Boyar	Online algorithms, combinatorial optimization, cryptology, data structures, computational complexity
	Lene Monrad Favrholdt	Online algorithms, graph algorithms
	Kim Skak Larsen	Online algorithms, algorithms and data structures, database systems, semantics
	Christian Kudahl	Online algorithms
	Jesper With Mikkelsen	Online algorithms