Approximation Algorithms

On-line Algorithms I

Rasmus Pagh
University of Copenhagen

TODAY

- COMPETITIVE RATIO
- REFRESHER: SCHEDULING JOBS ON PARALLEL MACHINES
- BIN PACKING
  - GREEDY 2-APPROXIMATION
  - COMPETITIVE RATIO 4/3 LOWER BOUND
- ON-LINE STEINER TREE
  - LOWER BOUND
  - UPPER BOUND

COMPETITIVE RATIO

LET OPT DENOTE THE COST OF AN OPTIMAL SOLUTION

DEF.

AN ON-LINE ALGORITHM IS c-COMPETITIVE IF IT ALWAYS FINDS A SOLUTION OF COST $\leq c \cdot OPT$.

IT IS ASYMPTOTICALLY c-COMPETITIVE IF IT ALWAYS FINDS A SOLUTION OF COST $\leq c \cdot OPT + o(OPT)$

BASICALLY SYNONYMOUS WITH APPROXIMATION FACTOR

NEGLIGIBLE WHEN OPT IS LARGE

SCHEDULING JOBS ON MULITIPLE MACHINES

INPUT: NUMBER m OF MACHINES

$p_1, p_2, \ldots, p_n \geq 0$ PROCESSING TIMES,
BEING REVEALED ONE AT A TIME

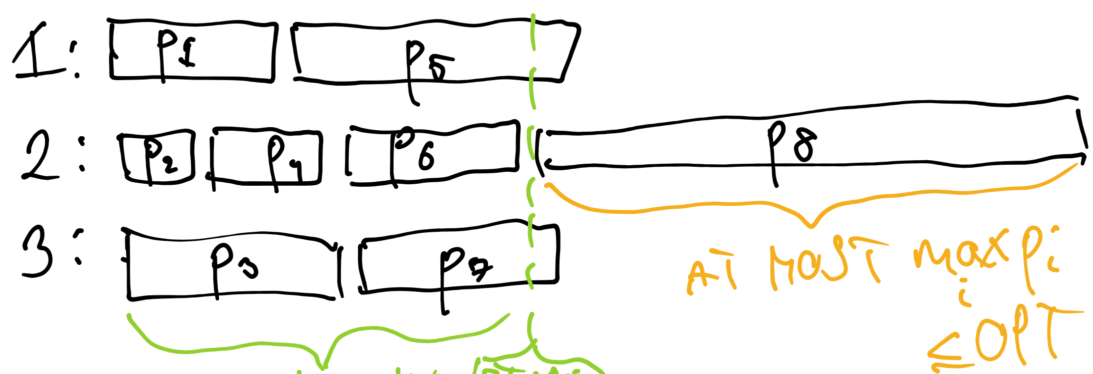
NEW COMPARED TO SETUP IN FIRST LECTURE

GRAHAM'S ALGORITHM:

ASSIGN EACH JOB TO THE MACHINE THAT CURRENTLY HAS THE SMALLEST LOAD
(TIES BROKEN IN ANY WAY)

$\boxed{\text{COMPETITIVE RATIO IS } \leq 2}$ TIGHT

CONFIGURATION $(m=3)$:



AT MOST $\max_i p_i \leq OPT$

NO MACHINE IS IDLE

START OF LAST JOB $\leq \sum_i p_i / m \leq OPT$

TOTAL COST $\leq 2\, OPT$

ON-LINE

BIN-PACKING PROBLEM

INPUT: $a_1, a_2, \ldots, a_n \in (0,1)$, REVEALED ONE AT A TIME

OBJECTIVE: ASSIGN $\{1, \ldots, n\}$ TO m BINS SUCH THAT ELEMENTS IN EACH BIN B_j HAVE $\sum_{i \in B_j} a_i \leq 1$ AND m IS AS SMALL AS POSSIBLE. ONCE ASSIGNED $i \in B_j$, THIS CANNOT BE CHANGED.

"NEXT FIT"

GREEDY ALGORITHM:

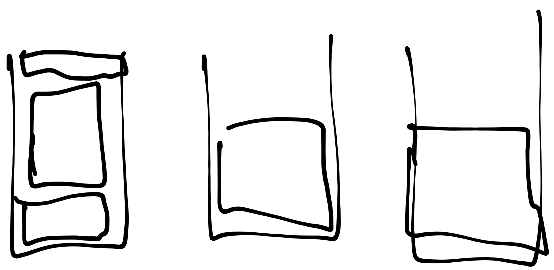
WORK ON ONE BIN j AT A TIME.
IF THERE IS ROOM, PUT a_i IN BIN j, OTHERWISE MOVE ON TO A NEW, EMPTY BIN

PAUSE AND THINK:

SHOW THAT GREEDY HAS COMPETITIVE RATIO EXACTLY 2.

FIRST FIT ALGORITHM

FOR ITEM i, PLACE a_i IN THE FIRST EXISTING BIN WHERE IT FITS, OR OPEN A NEW BIN IF NONE EXISTS



ULLMAN 1971:

FIRST FIT USES $\leq 1.7 OPT + 3$ BINS

AND THE FACTOR 1.7 IS THE BEST ASYMPTOTIC COMPETITIVE RATIO

LOWER BOUND FOR BIN PACKING'S COMPETITIVE RATIO

WANT TO SHOW LIMIT ON <u>ANY</u> ALGORITHM.

CONSIDER, FOR SOME SMALL $\varepsilon > 0$, THE INPUT SEQUENCE

$$\underbrace{\tfrac{1}{2}-\varepsilon, \tfrac{1}{2}-\varepsilon, \tfrac{1}{2}-\varepsilon, \dots, \tfrac{1}{2}-\varepsilon}_{m \text{ TIMES}}, \underbrace{\tfrac{1}{2}+\varepsilon, \tfrac{1}{2}+\varepsilon, \dots, \tfrac{1}{2}+\varepsilon}_{m \text{ TIMES}}$$

(2018) BALOGH ET AL.: COMP. RATIO BETWEEN ~ 1.543 AND 1.578

AFTER FIRST m INPUTS, HAVE $\frac{m}{2}+t$ BINS, OF WHICH $2t$ HAVE 1 ELEMENT AND THE REST HAVE TWO

CAN MATCH $2t$ OF THESE TO BINS WITH ONE ELEMENT, REMAINING $m-2t$ NEED TO BE IN SEPARATE BINS

$$\boxed{\max\left(1+\frac{2t}{m}, \frac{3}{2}-\frac{t}{m}\right) \geq \frac{4}{3} \text{ FOR ALL } t}$$

LOWER BOUND FOR ANY ALGORITHM

COMPETITIVE RATIO $\geq \dfrac{\frac{m}{2}+t}{OPT_m} = \dfrac{\frac{m}{2}+t}{m/2} = 1+\dfrac{2t}{m}$

AT TIME m

COMPETITIVE RATIO $\geq \dfrac{\frac{m}{2}+t+m-2t}{m} = \dfrac{3}{2}-\dfrac{t}{m}$

OPT_{2m}

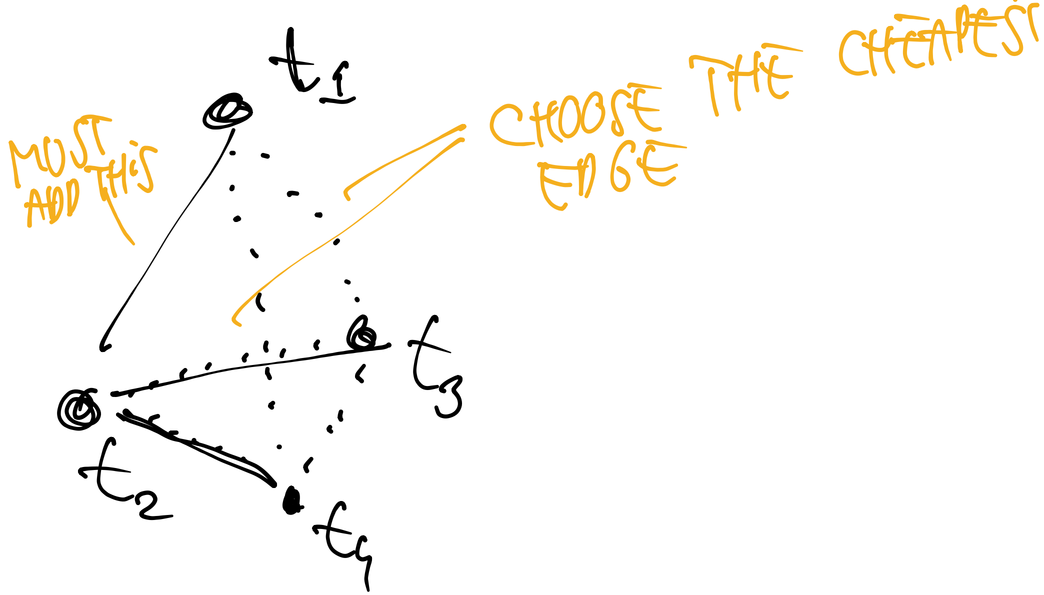
ON-LINE STEINER TREE

FIXED AHEAD OF TIME

INPUT: DISTANCE METRIC GIVEN BY A WEIGHTED GRAPH
SEQUENCE $t_1, \ldots, t_k$ OF "TERMINAL NODES" IN GRAPH
GIVEN ON-LINE ONE AT A TIME

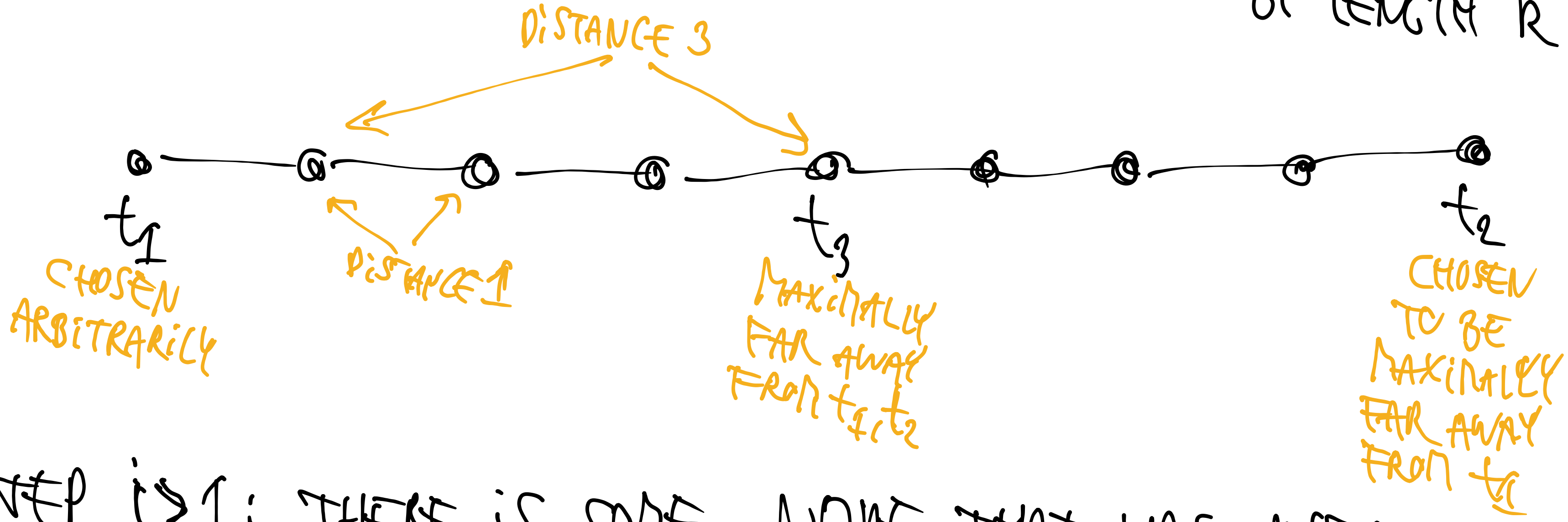
OBJECTIVE: MAINTAIN A STEINER TREE SPANNING THE TERMINAL NODES $t_1, \ldots, t_i$ SEEN SO FAR.
CAN ONLY ADD EDGES, NOT REMOVE ANY

GREEDY ALGORITHM:
IN EACH STEP i, ADD THE CHEAPEST EDGE $\{t_i, t_j\}$ WITH $j<i$



LOWER BOUND FOR GREEDY'S COMPETITIVE RATIO

METRIC: SHORTEST PATH ON GRAPH THAT IS A SIMPLE PATH OF LENGTH k



IN STEP $i > 1$: THERE IS SOME NODE THAT HAS DISTANCE AT LEAST $\frac{k}{i+1}/2$ TO PREVIOUS NODES (SOME INTERVAL OF LENGTH $\frac{k}{i+1}$ IS EMPTY)

TOTAL COST IS: $\sum_{i=2}^{k} \frac{k/2}{i+1} = \Omega(kH_k) = \Omega(k \log k)$

$OPT = k-1$

$\boxed{\text{COMPETITIVE RATIO IS } \Omega(\log k)}$

ANALYSIS OF GREEDY

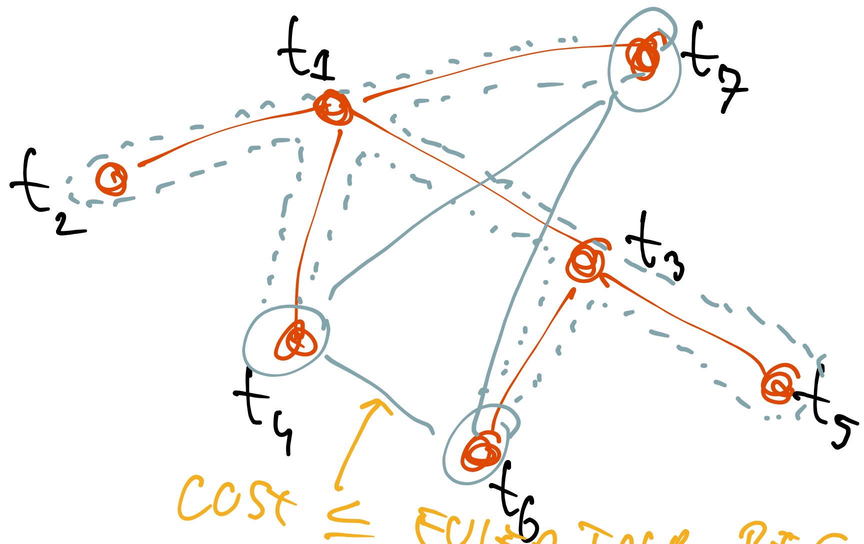
LEMMA FOR $i = 1, \dots, k-1$, THE iTH MOST EXPENSIVE EDGE CHOSEN BY GREEDY HAS COST AT MOST $2\,OPT/i$.

COROLLARY: TOTAL COST IS $\leq \underbrace{2H_{k-1}}$ $\cdot OPT$

BOUND ON COMPETITIVE RATIO MATCHES LOWER BOUND UP TO CONSTANT FACTOR

PROOF OF LEMMA:

CONSIDER THE OPTIMAL STEINER TREE, COST OPT



COST $\leq$ EULER TOUR BECAUSE OF TRIANGLE INEQUALITY

CLAIM

FOR ANY SET OF i TERMINALS, CAN CREATE A CYCLE VISITING THEM ALL OF WEIGHT $\leq 2\,OPT$

$\Rightarrow$ SOME EDGE HAS COST $\leq 2\,OPT/i$

LEMMA FOR $i = 1, \ldots, k-1$, THE iTH MOST EXPENSIVE EDGE CHOSEN BY GREEDY HAS COST AT MOST $2\,OPT/i$.

CLAIM

FOR ANY SET OF i TERMINALS, CAN CREATE A CYCLE VISITING THEM ALL OF WEIGHT $\leq 2\,OPT$

$\Rightarrow$ SOME EDGE HAS COST $\leq 2\,OPT/i$

LAST PART OF PROOF

ORDER THE TERMINALS ACCORDING TO THE COST INCURRED BY GREEDY

$t_{\sigma(1)}, t_{\sigma(2)}, \ldots, t_{\sigma(k)}$

MOST EXPENSIVE TO ADD

CHEAPEST TO ADD

C_i: CYCLE CONTAINING THE i MOST EXPENSIVE TERMINALS

SOME TERMINAL IN C_i COST $\leq 2\,OPT/i$ TO ADD

$\Rightarrow$ THE CHEAPEST IN C_i COSTS AT MOST $2\,OPT/i$.