

Complexity Theory of Polynomial-Time Problems

Lecture 8: (Boolean) Matrix Multiplication

Karl Bringmann

Recall: Boolean Matrix Multiplication

given $n \times n$ matrices A, B with entries in $\{0,1\}$ compute matrix C with $C_{i,j} = \bigvee_{k=1}^{n} A_{i,k} \wedge B_{k,j}$

what we already know about BMM:

BMM is in time $O(n^3/\log n)$ (four Russians)

BMM is equivalent to computing the Transitive Closure of a given graph

BMM can be reduced to APSP $\rightarrow O(n^3/2^{\sqrt{\log n}})$



Exponent of Matrix Multiplication

define ω as the *infimum* over all c such that MM has an $O(n^c)$ algorithm note: MM is in time $O(n^{\omega+\varepsilon})$ for any $\varepsilon > 0$ we will be sloppy and write: MM is in time $O(n^{\omega})$ note: MM is **not** in time $O(n^{\omega-\varepsilon})$ for any $\varepsilon > 0$ $\omega \ge 2$

Thm:	$\omega < 3$	
		$\omega \leq \dots$
	Strassen'69	2.81
	Pan'78	2.79
this is very fast – in theory	Bini et al.'79	2.78
	Schönhage'80	2.52
all these algorithms have	Romani'80	2.52
impractically large constant la	Coppersmith,Winograd'81	2.50
(maybe except Strassen'69)	Strassen'86	2.48
	Coppersmith,Winograd'90	2.376
	Stothers'10	2.374
	Vassilevska-Williams'11	2.37288
max planck institut informatik	l e Gall'14	2.37287

Boolean Matrix Multiplication

Thm:

BMM is in time $O(n^{\omega})$

given $n \times n$ matrices A, B with entries in $\{0,1\}$ compute standard matrix product C' with $C'_{i,j} = \sum_{i=1}^{n} A_{i,k} \cdot B_{k,j}$ define matrix C with $C_{i,j} = [C'_{i,j} > 0]$

then C is the Boolean matrix product of A and B

Hypothesis: BMM is not in time $O(n^{\omega-\varepsilon})$



Combinatorial Algorithms

fast matrix multiplication uses algebraic techniques which are impractical

"combinatorial algorithms": do not use algebraic techniques

not well defined!

Arlazarov,Dinic,Kronrod, Faradzhev'70 (four russians) Bansal,Williams'09 Chan'15 Yu'15 $O(n^3/\log^2 n)$

 $O(n^{3}(\log \log n)^{2}/\log^{9/4}n)$ $O(n^{3}(\log \log n)^{3}/\log^{3}n)$ $O(n^{3} \operatorname{poly} \log \log n / \log^{4}n)$

Hypothesis: BMM has no "combinatorial" algorithm in time $O(n^{3-\varepsilon})$



In this lecture you learn that...

...(B)MM is useful for **designing theoretically fast algorithms**

- Exercise: k-Clique in $O(n^{\omega k/3})$
- Exercise: MaxCut in $O(2^{\omega n/3} \operatorname{poly}(n))$
- Node-Weighted Negative Triangle in $O(n^{\omega})$



- Transitive Closure has no $O(n^{\omega-\varepsilon})$ / combinatorial $O(n^{3-\varepsilon})$ algorithm
- Exercise: pattern matching with 2 patterns
- Sliding Window Hamming Distance
- context-free grammar problems

max planck institut

Outline

1) Relations to Subcubic Equivalences

- 2) Strassen's Algorithm
- 3) Sliding Window Hamming Distance
- 4) Node-Weighted Negative Triangle
- 5) Context-Free Grammars





max planck institut informatik













Triangle Given graph *G* Decide whether there are vertices i, j, k such that i, j, k form a triangle



All-Pairs-Triangle Given graph *G* with vertex set $V = I \cup J \cup K$ Decide for every $i \in I, j \in J$ whether there is a vertex $k \in K$ such that i, j, k form a triangle

Split I, J, K into n/s parts of size s: $I_1, \dots, I_{n/s}, J_1, \dots, J_{n/s}, K_1, \dots, K_{n/s}$

For each of the $(n/s)^3$ triples (I_x, J_y, K_z) : consider graph $G[I_x \cup J_y \cup K_z]$

> max planck institut informatik



Initialize C as $n \times n$ all-zeroes matrix

For each of the $(n/s)^3$ triples of parts (I_x, J_y, K_z) :

While $G[I_x \cup J_y \cup K_z]$ contains a triangle:

Find a triangle (i, j, k) in $G[I_x \cup J_y \cup K_z]$

Set $C[i, j] \coloneqq 1$

Delete edge (i, j)

(i, j) is in no more triangles

- ✓ guaranteed termination: can delete $\leq n^2$ edges
- ✓ correctness:

if (i, j) is in a triangle, we will find one

max planck institut





Find a triangle (i, j, k) in $G[I_x \cup J_y \cup K_z]$

How to **find** a triangle if we can only **decide** whether one exists?

Partition I_x into $I_x^{(1)}, I_x^{(2)}, J_y$ into $J_y^{(1)}, J_y^{(2)}, K_z$ into $K_z^{(1)}, K_z^{(2)}$

Since $G[I_x \cup J_y \cup K_z]$ contains a triangle, at least one of the 2³ subgraphs $G[I_x^{(a)} \cup J_y^{(b)} \cup K_z^{(c)}]$

contains a triangle

Decide for each such subgraph whether it contains a triangle

Recursively find a triangle in one subgraph





 J_y

 K_{z}

Find a triangle (i, j, k) in $G[I_x \cup J_y \cup K_z]$

How to **find** a triangle if we can only **decide** whether one exists?

Partition I_x into $I_x^{(1)}, I_x^{(2)}, J_y$ into $J_y^{(1)}, J_y^{(2)}, K_z$ into $K_z^{(1)}, K_z^{(2)}$

Since $G[I_x \cup J_y \cup K_z]$ contains a triangle, at least one of the 2³ subgraphs $G[I_x^{(a)} \cup J_y^{(b)} \cup K_z^{(c)}]$

contains a triangle

Decide for each such subgraph whether it contains a triangle

Recursively find a triangle in one subgraph





All-Pairs-

Triangle

Running Time: $T_{\text{FindTriangle}}(n) \leq$

 $2^3 \cdot T_{\text{DecideTriangle}}(n)$

 $+ T_{\text{FindTriangle}}(n/2)$

 $= O(T_{\text{DecideTriangle}}(n))$

Initialize C as $n \times n$ all-zeroes matrix

For each of the $(n/s)^3$ triples of parts (I_x, J_y, K_z) :

While $G[I_x \cup J_y \cup K_z]$ contains a triangle:

Find a triangle (i, j, k) in $G[I_x \cup J_y \cup K_z]$

Set $C[i, j] \coloneqq 1$

Delete edge (i, j)

Running Time:

 $(*) = O(T_{\text{FindTriangle}}(s)) = O(T_{\text{DecideTriangle}}(s))$ Total time: $((\#\text{triples}) + (\#\text{triangles found})) \cdot (*)$ $\leq ((n/s)^3 + n^2) \cdot T_{\text{DecideTriangle}}(s)$ Set $s = n^{1/3}$ and assume $T_{\text{DecideTriangle}}(n) = O(n^{3-\varepsilon})$ Total time: $O(n^2 \cdot n^{1-\varepsilon/3}) = O(n^{3-\varepsilon/3})$



If BMM has (combinatorial) $O(n^{3-\varepsilon})$ algorithm then Triangle has (combinatorial) $O(n^{3-\varepsilon})$ algorithm

If Triangle has (combinatorial) $O(n^{3-\varepsilon})$ algorithm then BMM has (combinatorial) $O(n^{3-\varepsilon/3})$ algorithm

→ **subcubic equivalent**, but this mainly makes sense for **combinatorial** algorithms





Outline

1) Relations to Subcubic Equivalences

2) Strassen's Algorithm

- 3) Sliding Window Hamming Distance
- 4) Node-Weighted Negative Triangle
- 5) Context-Free Grammars



Strassen's Algorithm



$$C_{1,1} = A_{1,1} \cdot B_{1,1} + A_{1,2} \cdot B_{2,1}$$

$$C_{1,2} = A_{1,1} \cdot B_{1,2} + A_{1,2} \cdot B_{2,2}$$

$$C_{2,1} = A_{2,1} \cdot B_{1,1} + A_{2,2} \cdot B_{2,1}$$

$$C_{2,2} = A_{2,1} \cdot B_{1,2} + A_{2,2} \cdot B_{2,2}$$

 $T(n) \le 8 T(n/2) + O(n^2)$ $T(n) \le O(n^3)$



Strassen's Algorithm



$$M_{1} = (A_{1,1} + A_{2,2}) \cdot (B_{1,1} + B_{2,2})$$

$$M_{2} = (A_{2,1} + A_{2,2}) \cdot B_{1,1}$$

$$M_{3} = A_{1,1} \cdot (B_{1,2} - B_{2,2})$$

$$M_{4} = A_{2,2} \cdot (B_{2,1} - B_{1,1})$$

$$M_{5} = (A_{1,1} + A_{1,2}) \cdot B_{2,2}$$

$$M_{6} = (A_{2,1} - A_{1,1}) \cdot (B_{1,1} + B_{1,2})$$

$$M_{7} = (A_{1,2} - A_{2,2}) \cdot (B_{2,1} + B_{2,2})$$

max planck institut informatik

$$C_{1,1} = M_1 + M_4 - M_5 + M_7$$
$$C_{1,2} = M_3 + M_5$$
$$C_{2,1} = M_2 + M_4$$
$$C_{2,2} = M_1 - M_2 + M_3 + M_6$$

$$T(n) \le 7T(n/2) + O(n^2)$$
$$T(n) \le O(n^{\log_2 7}) = O(n^{2.8074})$$

2、

Faster Matrix Multiplication



matrix of rank 1: outer product of two vectors matrix of rank r: sum of r rank-1-matrices

matrix rank is in P

tensor of rank 1: outer product of three vectors tensor of rank r: sum of r rank-1-tensors

tensor rank is not known to be in P



Faster Matrix Multiplication



Strassen: rank of MM-tensor for n = 2 is at most 7

max planck institut

any bound on rank of MM-tensor can be transformed into a MM-algorithm

thus search for faster MM-algorithms is a mathematical question

this is **complete**: one can find ω by analyzing tensor rank!

Outline

- 1) Relations to Subcubic Equivalences
- 2) Strassen's Algorithm

3) Sliding Window Hamming Distance

- 4) Node-Weighted Negative Triangle
- 5) Context-Free Grammars



given two strings: text *T* of length *n* and pattern *P* of length m < ncompute for each *i* the Hamming distance of *P* and T[i..i + m - 1]

	а	b	С	b	b	С	а	а	
best known algorithm:	b	b	С	а					
$O(n\sqrt{m} \operatorname{polylog} n)$		b	b	С	а				
$< O(n^{1.5001})$			b	b	С	а			
				b	b	С	а		
					b	b	С	а	

[Indyk,Porat,Clifford'09]

 $\approx 0(n^{1.18})$

max planck institut informatik

Thm: Sliding Window Hamming Distance has no $O(n^{\omega/2-\varepsilon})$ algorithm or combinatorial $O(n^{1.5-\varepsilon})$ algorithm unless the BMM-Hypothesis fails

Open Problem: get rid of "combinatorial" or design improved algorithm using MM

given two strings: text *T* of length *n* and pattern *P* of length m < n compute for each *i* the Hamming distance of *P* and T[i...i + m - 1]



max planck institut

informatik

\$	\$	\$	\$	\$	\$	1	2	у	\$	1	у	3	\$	у	2	3	\$	\$	\$	\$	\$	\$	
						1	x	x	1	2	3	X	2	3									
										1	x	x	1	2	3	x	2	3					
														1	x	x	1	2	3	X	2	3	
			1	x	x	1	2	3	X	2	3												
							1	x	x	1	2	3	x	2	3								
											1	x	X	1	2	3	x	2	3				
1	X	x	1	2	3	X	2	3															
				1	X	x	1	2	3	x	2	3											put a 1 if there is
								1	x	x	1	2	3	x	2	3							at least one mato
					_						_											_	4
	1		0	0								1		1	0								1 1 0
	1		1	1				•				1	()	1				=				1 1 1
	0		1	1								0	-	1	1								1 1 1

```
given Boolean n \times n-matrices A, B
```

we construct text+pattern of length $O(n^2)$ (in time $O(n^2)$)

thus, an $O(n^{\omega/2-\varepsilon})$ algorithm for Sliding Window Hamming Distance would yield an $O(n^{\omega-2\varepsilon})$ algorithm for BMM, contradicting BMM-Hypothesis

and an $O(n^{1.5-\varepsilon})$ combinatorial algorithm for Sliding Window Hamming Dist. would yield an $O(n^{3-2\varepsilon})$ combinatorial algorithm for BMM

Thm: Sliding Window Hamming Distance has no $O(n^{\omega/2-\varepsilon})$ algorithm or combinatorial $O(n^{1.5-\varepsilon})$ algorithm unless the BMM-Hypothesis fails



Outline

- 1) Relations to Subcubic Equivalences
- 2) Strassen's Algorithm
- 3) Sliding Window Hamming Distance
- 4) Node-Weighted Negative Triangle
- 5) Context-Free Grammars



Node-Weighted Negative Triangle



given a directed graph with weights $w_{i,j}$ on **edges**, is there a triangle *i*, *j*, *k*: $w_{j,i} + w_{i,k} + w_{k,j} < 0$?

given a directed graph with weights w_i on **nodes**, is there a triangle *i*, *j*, *k*: $w_i + w_j + w_k < 0$?

given an **unweighted** undirected graph, is there a triangle?



Node-Weighted Negative Triangle



find appropriate **edge** weights that simulate the given **node** weights:

set $w_{i,j} \coloneqq (w_i + w_j)/2$

then for a triangle i, j, k:

$$w_{j,i} + w_{i,k} + w_{k,j} = w_i + w_j + w_k$$



Node-Weighted Negative Triangle







we can assume that the graph is tripartite:



given graph G = (V, E) with $V = I \cup J \cup K$ and node weights w_v ,

compute minimum weight q s.t.

there are $i \in I, j \in J, k \in K$ with $(i, j), (j, k), (k, i) \in E$ and $w_i + w_j + w_k = q$

- assume that *I*, *J*, *K* are sorted by weight
- parameter p (=sufficiently large constant)
- assume $n \coloneqq |I| = |J| = |K| = p^{\ell}$ for some $\ell \in \mathbb{N}$ (add isolated dummy vertices)



given graph G = (V, E) with $V = I \cup J \cup K$ and node weights w_v ,

compute minimum weight q s.t.

there are $i \in I, j \in J, k \in K$ with $(i, j), (j, k), (k, i) \in E$ and $w_i + w_j + w_k = q$

- assume that *I*, *J*, *K* are sorted by weight
- parameter p (=sufficiently large constant)
- assume $n \coloneqq |I| = |J| = |K| = p^{\ell}$ for some $\ell \in \mathbb{N}$ (add isolated dummy vertices)

ALG(G): 0) if n = O(1) then solve in constant time

1) split $I = I_1 \cup \cdots \cup I_p$, $J = J_1 \cup \cdots \cup J_p$, $K = K_1 \cup \cdots \cup K_p$ (in sorted order: $\max w(I_x) \le \min w(I_{x+1})$ and so on)



2) $R \coloneqq \{(x, y, z) \in \{1, \dots, p\}^3 \mid G[I_x \cup J_y \cup K_z] \text{ contains a triangle} \}$

3) for each $(x, y, z) \in R$ s.t. there is no $(x', y', z') \in R$ with x' < x, y' < y, z' < zrun ALG $(G[I_x \cup J_y \cup K_z])$

max planck institut informatik

3) for each (*x*, *y*, *z*) ∈ *R* s.t. there is no (*x*', *y*', *z*') ∈ *R* with *x*' < *x*, *y*' < *y*, *z*' < *z* run ALG(*G*[*I_x* ∪ *J_y* ∪ *K_z*])
 Correctness:

if there is no triangle in $G[I_x \cup J_y \cup K_z]$ then we can ignore it

if
$$(x, y, z) \in R$$
 is dominated by $(x', y', z') \in R$:
let i, j, k be a triangle in $G[I_x \cup J_y \cup K_z]$, and i', j', k' a triangle in $G[I_{x'} \cup J_{y'} \cup K_{z'}]$
then $w_i + w_j + w_k \ge \min w(I_x) + \min w(J_y) + \min w(K_z)$
 $\vee I \qquad \vee I \qquad \vee I$
and $w_{i'} + w_{j'} + w_{k'} \le \max w(I_{x'}) + \max w(J_{y'}) + \max w(K_{z'})$

so we can safely ignore $G[I_x \cup J_y \cup K_z]$



given graph G = (V, E) with $V = I \cup J \cup K$ and node weights w_v ,

compute minimum weight q s.t.

there are $i \in I, j \in J, k \in K$ with $(i, j), (j, k), (k, i) \in E$ and $w_i + w_j + w_k = q$

- assume that *I*, *J*, *K* are sorted by weight
- parameter p (=sufficiently large constant)

max planck institut informatik

- assume $n \coloneqq |I| = |J| = |K| = p^{\ell}$ for some $\ell \in \mathbb{N}$ (add isolated dummy vertices)

ALG(G): 0) if n = O(1) then solve in constant time

1) split $I = I_1 \cup \cdots \cup I_p$, $J = J_1 \cup \cdots \cup J_p$, $K = K_1 \cup \cdots \cup K_p$ (in sorted order: $\max I_x \le \min I_{x+1}$ aso.)

2) $R \coloneqq \{(x, y, z) \in \{1, ..., p\}^3 \mid G[I_x \cup J_y \cup K_z] \text{ contains a triangle}\}$ $O(p^3 n^{\omega})$

3) for each $(x, y, z) \in R$ s.t. there is no $(x', y', z') \in R$ with x' < x, y' < y, z' < zrun ALG $(G[I_x \cup J_y \cup K_z])$ size n/p

how many iterations?

3) for each $(x, y, z) \in R$ s.t. there is no $(x', y', z') \in R$ with x' < x, y' < y, z' < zHow many iterations?

define
$$\Delta_{r,s} \coloneqq \{(x, y, z) \in \{1, ..., p\}^3 \mid x - y = r \text{ and } x - z = s\}$$

= $\{(1, 1 - r, 1 - s), (2, 2 - r, 2 - s), ..., (p, p - r, p - s)\} \cap \{1, ..., p\}^3$
for $r, s \in \{-p, ..., p\}$

the sets $\Delta_{r,s}$ cover $\{1, ..., p\}^3$

max planck institut informatik

line 3) applies to at most one (x, y, z) in $\Delta_{r,s}$ for any r, s!

hence, there are at most $(2p)^2 = 4p^2$ recursive calls to ALG

recursion: $T(n) \le 4p^2 \cdot T(n/p) + O(p^3 n^{\omega})$

given graph G = (V, E) with $V = I \cup J \cup K$ and node weights w_{ν} ,

compute minimum weight q s.t.

there are $i \in I, j \in J, k \in K$ with $(i, j), (j, k), (k, i) \in E$ and $w_i + w_j + w_k = q$

- assume that *I*, *J*, *K* are sorted by weight
- parameter p (=sufficiently large constant)

max planck institut informatik

- assume $n \coloneqq |I| = |J| = |K| = p^{\ell}$ for some $\ell \in \mathbb{N}$ (add isolated dummy vertices)

ALG(G): 0) if n = O(1) then solve in constant time

1) split $I = I_1 \cup \cdots \cup I_p$, $J = J_1 \cup \cdots \cup J_p$, $K = K_1 \cup \cdots \cup K_p$ (in sorted order: $\max I_x \le \min I_{x+1}$ aso.)

2) $R \coloneqq \{(x, y, z) \in \{1, \dots, p\}^3 \mid G[I_x \cup J_y \cup K_z] \text{ contains a triangle}\}$ $O(p^3 n^{\omega})$

3) for each $(x, y, z) \in R$ s.t. there is no $(x', y', z') \in R$ with x' < x, y' < y, z' < zrun ALG($G[I_x \cup J_v \cup K_z]$) size n/p

recursion: $T(n) \leq 4p^2 \cdot T(n/p) + O(p^3 n^{\omega})$

recursion:	T(n)	$\leq 4p^2 \cdot T(n/p)$	$) + O(p^3 n^\omega)$	
	T(n)	$\leq 4p^2 \cdot T(n/p)$	$) + \alpha p^3 n^{\omega}$	for some constant α
want to show:	T(n)	$\leq 2\alpha p^3 n^{\omega}$		
plug in:	T(n)	$\leq 4p^2 \cdot 2\alpha p^3 ($	$n/p)^{\omega} + \alpha p^3 n^{\omega}$	
		$= \alpha p^3 n^\omega (1 +$	$8p^{2-\omega})$	
assume $\omega > 2$:		$\leq 2\alpha p^3 n^{\omega}$	for a sufficien	tly large constant p
so we have:	T(n)	$\leq O(n^{\omega})$		
(if $\omega = 2$: show that	t T(n)	$\leq 2\alpha p^3 n^{\omega + \varepsilon})$		

max planck institut informatik



In this lecture you learned that...

...(B)MM is useful for **designing theoretically fast algorithms**

- Exercise: k-Clique in $O(n^{\omega k/3})$
- Exercise: MaxCut in $O(2^{\omega n/3} \operatorname{poly}(n))$
- Node-Weighted Negative Triangle in $O(n^{\omega})$



- Transitive Closure has no $O(n^{\omega-\varepsilon})$ / combinatorial $O(n^{3-\varepsilon})$ algorithm
- Exercise: pattern matching with 2 patterns
- Sliding Window Hamming Distance
- context-free grammar problems

max planck institut