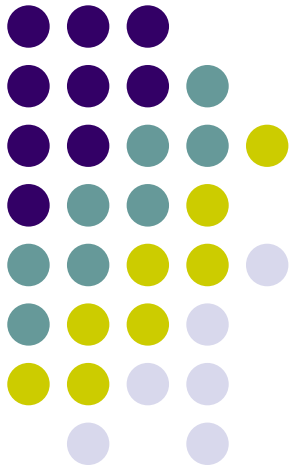
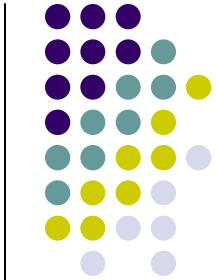
CS256
Applied Theory of Computation

Space-Time Tradeoffs II

John E Savage



Overview



- Space-time tradeoffs using flow properties
- Definition of flow properties of functions
- Flow property of matrix multiplication
- Grigoriev's lower bound
- Application to convolution
- Application to cyclic shift

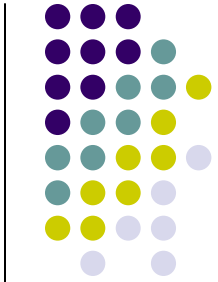
Flow Properties of Functions

- Space-time tradeoffs for straight-line programs (SLPs), operations over $\Omega = \{b : A^r \rightarrow A^s \mid r,s \geq 1\}$.

  Definition Let $f : A^n \rightarrow A^m$ & $h : A^r \rightarrow A^s$, $r \leq n$, $s \leq m$. $h = f|^{Y_1}_{X_1}$ is the subfunction of f obtained by fixing variables in $X_0 = X - X_1$ & discarding outputs in $Y_0 = Y - Y_1$. Thus, h is function of inputs in X_1 & has outputs in Y_1. $f : A^n \rightarrow A^m$ has a $w(u,v)$-**flow** if for all $X_1 \subseteq X$ & $Y_1 \subseteq Y$ satisfying $|X_1| \geq u$ and $|Y_1| \geq v$ there is a sub-function h that has at least $|A|^{w(u,v)}$ points in the image of its domain.

  Note: $w(u,v)$ is an increasing function of u and v.

Flow Properties of Functions

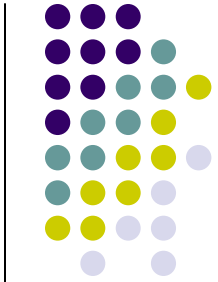


- *f* is *(*α, *n*, *m*, *p)*-**independent** for α ≥1 and *p* ≤ *m* if it has a *w(u,v)*-flow if *w(u,v)* > *(v/*α*) -1* when *n* – *u* + *v* ≤ *p*.

Flow Property of Matrix Multiplication

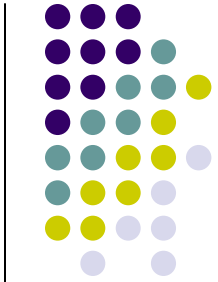
Theorem The $n \times n$ matrix multiplication function C = A × B over ring R is $(1,2n^2,n^2,n)$-independent.

Flow Property of Matrix Multiplication



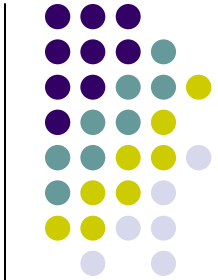
Proof Consider a set X_0 of inputs and a set Y_1 of outputs where $|X_0| + |Y_1| = n$. (Thus, $|X_1| = |X| - |X_0| = 2n^2 - |X_0|$.) Outputs in Y_1 fall into at most $|Y_1|$ columns of C. Inputs in X_0 fall into at most $|X_0|$ columns of A. Thus, at least $n - |X_0| = |Y_1|$ columns of A contain only variables. Make B a permutation matrix that permutes the all-variables columns of A onto the at most $|Y_1|$ columns containing selected outputs. Because the variables in the all-variable columns of A are free to assume any values, the outputs in Y_1, which are in $|Y_1|$ columns, can assume $|R|^y$ different values for $y = |Y_1|$. Thus, the function is $(1,2n^2,n^2,n)$-independent.♥

Grigoriev's Lower Bound



Theorem Let $f : A^n \rightarrow A^m$ and let each output be dependent on each input. Let it have a $w(u,v)$-flow and be realized by SLP with operators over A. Let $b \leq m$. Then every pebbling of SLP DAG requires space S and time T satisfying $T \geq \lfloor m/b \rfloor (n - d)$ where d is the largest integer such that $w(d,b) \leq S$.

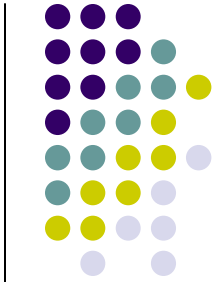
Grigoriev's Lower Bound



Proof Divide steps into intervals during which b outputs are pebbled ($m - b\lfloor m/b \rfloor$ in last interval.) Let Y_1 be outputs pebbled in an interval, $|Y_1| = b$. Let X_0 and X_1 be the inputs pebbled inside and before the interval. Since each output is dependent on each input, $X = X_0 \cup X_1$.

Let $x_0 = |X_0|$ and $x_1 = |X_1|$. By definition there exists assignment of values to inputs in X_0 so that outputs in Y_1 assume at least $|A|^y$ values, $y = w(x_1,b)$. Since all variables in the interval (i.e. X_0) are fixed, the values of the outputs Y_1 are determined by inputs in X_1. Thus, if $w(x_1,b) > S$, outputs Y_1 assume more values than can be carried by $\leq S$ pebbles on a DAG at start of an interval, a contradiction.

Grigoriev's Lower Bound



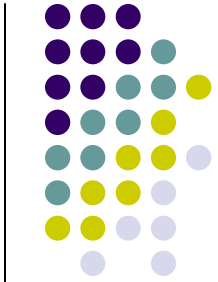
Proof (cont.) Thus, $w(x_1,b) \le S$. Since $w(u,v)$ is increasing in u and v, $x_1 \le d$ where d is the largest integer satisfying $w(d,b) \le S$. Because $x_0 + x_1 = n$ and each output depends on every input, it follows that $x_0 \ge n - d$. Thus for each interval except possibly last, $\ge n - d$ inputs are pebbled in the interval.

If T_I is the number of steps on which inputs are pebbled,

$$T \ge T_I \ge \lfloor m/b \rfloor (n - d) \text{ ♥}$$

Grigoriev established above theorem for $(1, n, m, p)$-independent functions, which we restate as a corollary to the above theorem. It is restated as for (α, n, m, p) -independent functions.

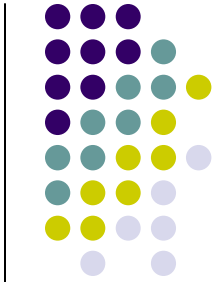
Grigoriev's Lower Bound



Corollary Let $f : A^n \rightarrow A^m$ be (α, n, m, p)-ind. And let it be realized by an SLP over the basis Ω. Every pebbling of every DAG for f requires space S and time T satisfying the inequality

$$\lceil \alpha(S+1) \rceil T \geq mp/4$$

Grigoriev's Lower Bound

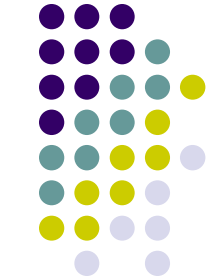


Proof An (α, n, m, p)-independent function on n inputs has $w(u,v)$-flow satisfying $w(u,v) > (v/\alpha) - 1$ for $n-u+v \le p$ where $x_0 = n-u \ge 0$. Since b can be freely chosen, let $b = \lceil \alpha(S+1) \rceil$. Thus, $(b/\alpha) - 1 \ge S$ for $(n-d)+ b \le p$, which contradicts the requirement that $w(d,b) \le S$. It follows that $(n-d)+b \ge p$ or that $(n-d) \ge p - \lceil \alpha(S+1) \rceil$. With the inequality $\lfloor m/x \rfloor \ge (m-x+1)/x$, the following lower bound follows from Theorem.

$$T \ge (m - \lceil \alpha(S+1) \rceil + 1)(p - \lceil \alpha(S+1) \rceil)/\lceil \alpha(S+1) \rceil$$

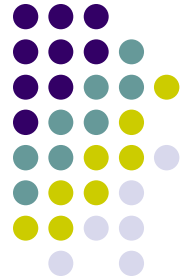
Since $p \le m$, if $\lceil \alpha(S+1) \rceil \le p/2$, the result follows. If $\lceil \alpha(S+1) \rceil > p/2$, $\lceil \alpha(S+1) \rceil T \ge mp/2$, since $T \ge m$.♥

Application to Matrix Multiplication



- We now show that space and time for $n \times n$ matrix multiplication trade as $(S+1)T = \Theta(n^3/4)$.
- That is, the standard algorithm for this problem is optimal.
- Recall that fast algorithms exist for matrix multiplication that have $T = O(n^{2.38})$. It isn't clear if they have near optimal ST.

Space-Time Tradeoff for Matrix Multiplication



Theorem The $n \times n$ matrix multiplication function $C = A \times B$ over ring R is $(1,2n^2,n^2,n)$-independent.

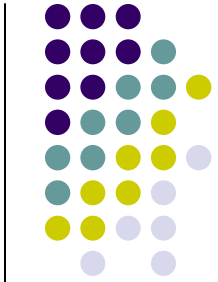
Corollary The time T and space S required to realize the $n \times n$ matrix multiplication function $C = A \times B$ over a ring R using an SLP must satisfy

$$(S+1)T \geq n^3/4$$

Also, $T \leq 2n^3$ when $S = 3$.

Proof Do each inner product as shown below.♥

Convolution

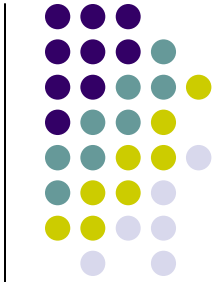


- The wrapped convolution on strings of length *n* over ring R, $f_{\text{wrapped}}^{(n)} : R^{2n} \rightarrow R^n$, can be characterized by following product of circulant matrix with vector.

$$\begin{bmatrix} w_0 \\ w_1 \\ w_2 \\ : \\ w_{n-1} \end{bmatrix} = \begin{bmatrix} u_0 & u_{n-1} & u_{n-2} \ldots & u_1 \\ u_1 & u_0 & u_{n-1} \ldots & u_2 \\ & & : & \\ u_{n-2} & u_{n-3} & \ldots & u_{n-1} \\ u_{n-1} & u_{n-2} & \ldots & u_0 \end{bmatrix} \begin{bmatrix} v_0 \\ v_1 \\ v_2 \\ : \\ v_{n-1} \end{bmatrix}$$

- Note that each row of the matrix M is a cyclic shift of a single row.

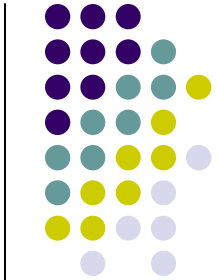
Convolution



Lemma For n even, the function $f_{\text{wrapped}}^{(n)} : R^{2n} \rightarrow R^{n}$ is $(2,2n,n,n/2)$-independent.

Proof Consider $X_0 \subseteq X$ and $Y_1 \subseteq Y$, X and Y inputs and outputs of $f_{\text{wrapped}}^{(n)}$ where $|X_0| + |Y_1| = p = n/2$. For $f_{\text{wrapped}}^{(n)}$ to be $(2,2n,n,n/2)$-independent, there must be assignment to inputs in X0 such that outputs in Y_1 have $> |R|^y$ distinct values for $y = (|Y_1|/2) - 1$ as inputs in $X_1 = X - X_0$ range over all values. Let $e = |X_0 \cap \{u_0, u_1, \ldots, u_{n-1}\}|$. Thus, every row of M contains the same number e of entries from X_0. The |entries in X_1 are free to vary.

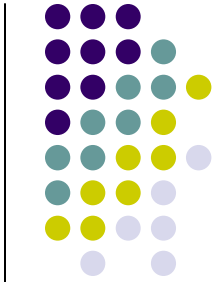
Convolution



Each output in Y_1 corresponds to row of M. Number of instances of input variables in X_1 in these rows is $|Y_1|(n\text{-}e)$. Since these rows have n columns, there is a column, say the tth, containing at least the average $|Y_1|(n\text{-}e)/n \geq |Y_1|/2$ variables from X_1. (Instances of X in a column of M are distinct.)

It follows that by choosing the tth component of $\mathbf{v}$, v_t, to be 1 and the others to be 0, at least $|Y_1|/2$ of the inputs in X_1 are mapped onto outputs in Y_1. Since these inputs (and outputs) can assume $|R|^y$ distinct values for $y = (|Y_1|/2) - 1$ different values, it follows that $f_{\text{wrapped}}^{(n)}$ is $(2,2n,n,n/2)$-independent.♥

Convolution



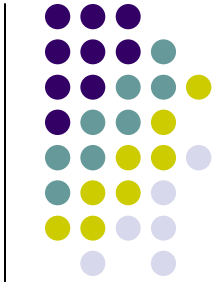
Theorem The time T and space S required to pebble any straight-line program for the standard or wrapped convolution must satisfy the following inequality:

$$(S + 1)T \geq n^2/16$$

This lower bound can be achieved to within a constant multiplicative factor for S = O(1).

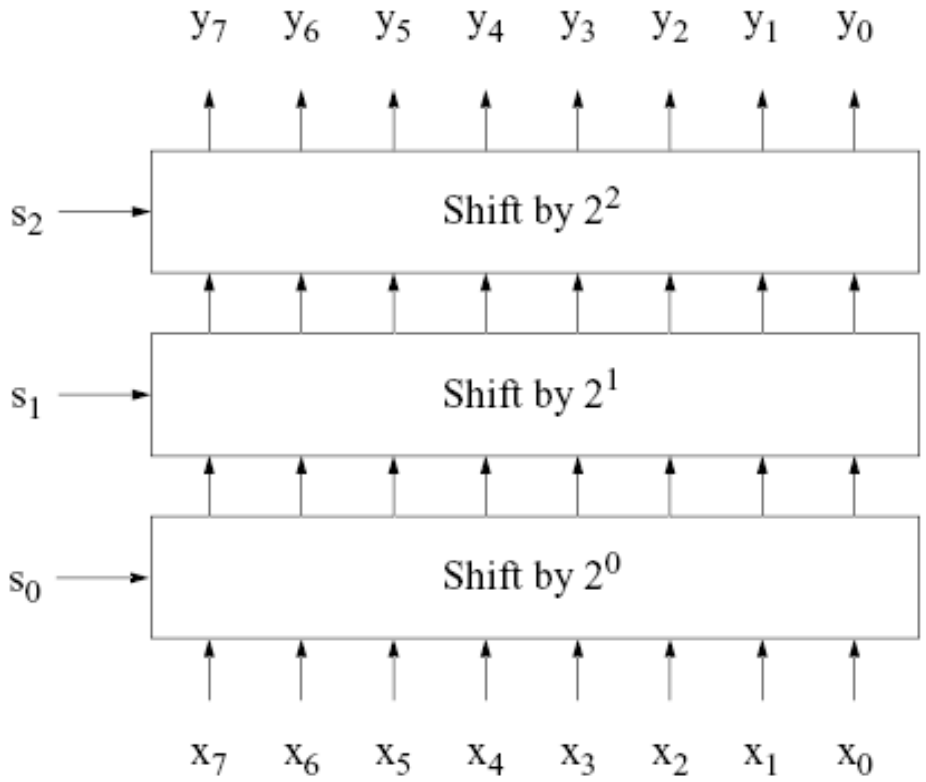
Proof Upper bound follows from standard matrix vector algorithm using the inner product graph shown earlier that can be done with three pebbles.♥

Cyclic Shifting



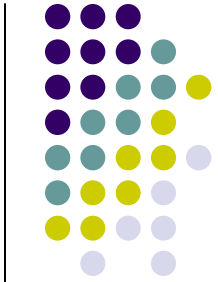
- The cyclic shifting function $f_{\text{cyclic}}^{(n)}: B^{m(n)} \rightarrow B^n$, where $m(n) = n + \lceil \log_2 n \rceil$ is a subfunction of many functions, including integer multiplication and squaring, integer reciprocal, and powers of integers.
- Cyclic shifting is a good example of a problem for which a lower bound on the exchange of space and time exists. We now define $f_{\text{cyclic}}^{(n)}$. Let $k = \lceil \log_2 n \rceil$. Its input variables are $\boldsymbol{x} = (x_{n-1}, \ldots, x_1, x_0)$ and $\boldsymbol{s} = (s_{k-1}, \ldots, s_1, s_0)$
- Here $\boldsymbol{s}$ specifices an amount $|\boldsymbol{s}|$ of the shift of $\boldsymbol{x}$.

Cyclic Shifting

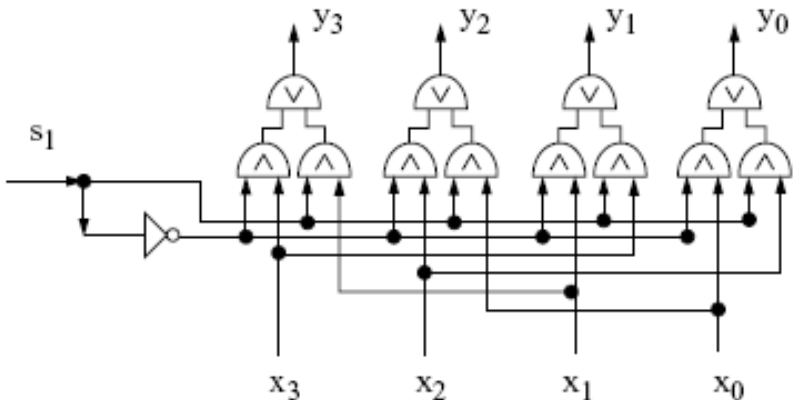


- The above circuit for cyclic shift has size *O(n log n).*

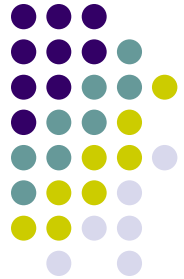
Cyclic Shifting



- Each circuit can be realized as shown below.



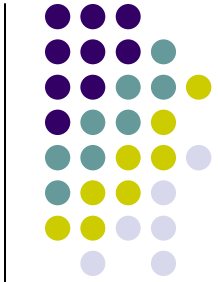
Cyclic Shifting



Lemma $f_{\text{cyclic}}^{(n)} : B^{m(n)} \rightarrow B^n$, $m(n) = n + \lceil \log_2 n \rceil$ is $(2,m(n),n,n/2)$-independent.

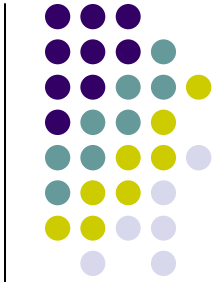
Proof Consider $X_0 \subseteq X$ and $Y_1 \subseteq Y$, X and Y inputs and outputs of $f_{\text{cyclic}}^{(n)}$ where $|X_0| + |Y_1| = p = n/2$. For $f_{\text{cyclic}}^{(n)}$ to be $(2,m(n),n,n/2)$-independent, there must be assignment to inputs in X_0 such that outputs in Y_1 have $> |B|^y$ distinct values for $y = (|Y_1|/2) - 1$ as inputs in $X_1 = X - X_0$ range over all values.

Cyclic Shifting



Proof (cont.) Let X_0 contain e elements from $\boldsymbol{x}$. Let y_i be in Y_1. As we run through all possible shift values, y_i is made equal to every one of the inputs in $\boldsymbol{x}$. For n-e of these shifts y_i is set equal to an input in $X_1 = X - X_0$. (For example, if $n = 6$ and $e = 2$, say with $X_1 = \{x_0, x_3, x_4, x_5\}$ and $Y_1 = \{y_2, y_3, y_5\}$, then as $\boldsymbol{s}$ ranges over all of its values, each of the three y_i's in Y_1 is assigned four different variables in X_1.)

Cyclic Shifting



Proof (cont.) Thus, the number of input variables assigned to outputs, summed over all cyclic shifts, is $|Y_1|(n - e)$. Since there are n cyclic shifts, for some shift the number of variables in X_1 that are matched with outputs in Y_1 is at least the average of this quantity; that is, at least $|Y_1|(1 - e/n) \geq |Y_1|/2$. Thus, some shift matchs at least $|Y_1|/2$ inputs in X_1 with outputs in Y_1. Since these outputs can assume $|B|^y$ distinct values for $y = (|Y_1|/2) - 1$, it follows that $f_{\text{cyclic}}^{(n)}$ is $(2,m(n),n,n/2)$-independent.♥

Cyclic Shifting

Theorem Every pebbling strategy for straight-line programs computing the cyclic shifting function $f_{\text{cyclic}}^{(n)}$ requires S and T satisfying the inequality

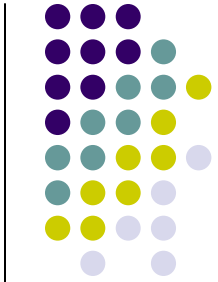
$$(S + 1)T \geq n^2/16$$

A graph and a pebbling strategy exist to compute $f_{\text{cyclic}}^{(n)}$ using space $O(n)$ & time $O(n \log n)$, namely, that satisfies the inequality

$$(S + 1)T = O(n^2 \log n)$$

Proof The lower bound follows from Grigoriev's Theorem. The upper bound follows by pebbling the standard circuit for shifting shown earlier using a total of S = O(n) pebbles.♥

Generalization



- The method used here can be generalized to the class of transitive functions defined by a problem in Chapter 10.