# Saving Space by Algebraization 

## Daniel Lokshtanov and Jesper Nederlof

6 June 2010
"Saving space by algebraization", STOC 2010
Jesper Nederlof
UNIVERSITETET I BERGEN Institutt for informatikk

## Dynamic Programming (DP)

- From the 1950's by Richard Bellman in his book "Bottleneck Problems and Dynamic Programming"; Nowadays a prominent algorithmic technique in designing algorithms.


## Dynamic Programming (DP)

- From the 1950's by Richard Bellman in his book "Bottleneck Problems and Dynamic Programming"; Nowadays a prominent algorithmic technique in designing algorithms.
- Uses a big table of data that has to be stored in the memory.


## Dynamic Programming (DP)

- From the 1950's by Richard Bellman in his book "Bottleneck Problems and Dynamic Programming"; Nowadays a prominent algorithmic technique in designing algorithms.
- Uses a big table of data that has to be stored in the memory.
- A relatively easy procedure computes new table entries using already computed table entries.


## Dynamic Programming (DP)

- From the 1950's by Richard Bellman in his book "Bottleneck Problems and Dynamic Programming"; Nowadays a prominent algorithmic technique in designing algorithms.
- Uses a big table of data that has to be stored in the memory.
- A relatively easy procedure computes new table entries using already computed table entries.
- This easy procedure is often so easy that we just write it down as a single formula, obtaining a recurrence.


## Dynamic Programming (DP)

- From the 1950 's by Richard Bellman in his book "Bottleneck Problems and Dynamic Programming"; Nowadays a prominent algorithmic technique in designing algorithms.
- Uses a big table of data that has to be stored in the memory.
- A relatively easy procedure computes new table entries using already computed table entries.
- This easy procedure is often so easy that we just write it down as a single formula, obtaining a recurrence.
- We are interested in conditions that are sufficient for being able to reduce the space requirement of DP algorithms significantly (without significant loss of speed).


## The approach in a nutshell

- Usually the dependency between table entries is highly unpredictable and interleaved.


## The approach in a nutshell

- Usually the dependency between table entries is highly unpredictable and interleaved.
- The best one can do is keeping track of (almost) the whole table.


## The approach in a nutshell

- Usually the dependency between table entries is highly unpredictable and interleaved.
- The best one can do is keeping track of (almost) the whole table.
- We use a transformation to transform the table into one where the dependency between table entries is very restricted and systematic.


## The approach in a nutshell

- Usually the dependency between table entries is highly unpredictable and interleaved.
- The best one can do is keeping track of (almost) the whole table.
- We use a transformation to transform the table into one where the dependency between table entries is very restricted and systematic.
- This allows us to turn DP algorithms of which the recurrence only uses certain operators in space efficient ones.


## Subset Sum

- Given integers $\left\{e_{1}, \ldots, e_{n}\right\}$ and $t$ in binary representation, count the number of subsets $S \subseteq\{1, \ldots, n\}$ such that $\sum_{i \in S} e_{i}=t$.


## Subset Sum

- Given integers $\left\{e_{1}, \ldots, e_{n}\right\}$ and $t$ in binary representation, count the number of subsets $S \subseteq\{1, \ldots, n\}$ such that $\sum_{i \in S} e_{i}=t$.
- Brute force gives a $\mathcal{O}\left(2^{n}\right)$ time, polynomial space algorithm. DP gives an $\mathcal{O}(n t)$ time and $\mathcal{O}(t)$ space algorithm:


## Subset Sum

- Given integers $\left\{e_{1}, \ldots, e_{n}\right\}$ and $t$ in binary representation, count the number of subsets $S \subseteq\{1, \ldots, n\}$ such that $\sum_{i \in S} e_{i}=t$.
- Brute force gives a $\mathcal{O}\left(2^{n}\right)$ time, polynomial space algorithm. DP gives an $\mathcal{O}(n t)$ time and $\mathcal{O}(t)$ space algorithm:
- Define $A[i, j]$ as the number of subsets $S \subseteq\{1, \ldots, i\}$ such that $\sum_{k \in S} e_{k}=j$. Then


## Subset Sum

- Given integers $\left\{e_{1}, \ldots, e_{n}\right\}$ and $t$ in binary representation, count the number of subsets $S \subseteq\{1, \ldots, n\}$ such that $\sum_{i \in S} e_{i}=t$.
- Brute force gives a $\mathcal{O}\left(2^{n}\right)$ time, polynomial space algorithm. DP gives an $\mathcal{O}(n t)$ time and $\mathcal{O}(t)$ space algorithm:
- Define $A[i, j]$ as the number of subsets $S \subseteq\{1, \ldots, i\}$ such that $\sum_{k \in S} e_{k}=j$. Then

$$
A[i, j]= \begin{cases}0 & \text { if } i=1, e_{1} \neq j, \text { and } e_{1} \neq 0 \\ 1 & \text { if } i=1 \text { and }\left(e_{1}=j \text { or } j=0\right. \\ A[i-1, j]+A\left[i-1, j-e_{i}\right] & \text { if } i>1\end{cases}
$$

## Subset Sum

- Given integers $\left\{e_{1}, \ldots, e_{n}\right\}$ and $t$ in binary representation, count the number of subsets $S \subseteq\{1, \ldots, n\}$ such that $\sum_{i \in S} e_{i}=t$.
- Brute force gives a $\mathcal{O}\left(2^{n}\right)$ time, polynomial space algorithm. DP gives an $\mathcal{O}(n t)$ time and $\mathcal{O}(t)$ space algorithm:
- Define $A[i, j]$ as the number of subsets $S \subseteq\{1, \ldots, i\}$ such that $\sum_{k \in S} e_{k}=j$. Then

$$
A[i, j]= \begin{cases}0 & \text { if } i=1, e_{1} \neq j, \text { and } e_{1} \neq 0 \\ 1 & \text { if } i=1 \text { and }\left(e_{1}=j \text { or } j=0\right) \\ A[i-1, j]+A\left[i-1, j-e_{i}\right] & \text { if } i>1\end{cases}
$$

- The answer of the Subset Sum instance can be read from $A[n, t]$.

Weight
items
1
2
2


$$
A[i, j]= \begin{cases}0 & \text { if } i=1, e_{1} \neq j, \text { and } e_{1} \neq 0 \\ 1 & \text { if } i=1 \text { and }\left(e_{1}=j \text { or } j=0\right) \\ A[i-1, j]+A\left[i-1, j-e_{i}\right] & \text { if } i>1\end{cases}
$$

"Saving space by algebraization", STOC 2010
Jesper Nederlof

Weight
items
1
2
2


$$
A[i, j]= \begin{cases}0 & \text { if } i=1, e_{1} \neq j, \text { and } e_{1} \neq 0 \\ 1 & \text { if } i=1 \text { and }\left(e_{1}=j \text { or } j=0\right) \\ A[i-1, j]+A\left[i-1, j-e_{i}\right] & \text { if } i>1\end{cases}
$$

"Saving space by algebraization", STOC 2010
Jesper Nederlof

Weight
items
1
2
2


$$
A[i, j]= \begin{cases}0 & \text { if } i=1, e_{1} \neq j, \text { and } e_{1} \neq 0 \\ 1 & \text { if } i=1 \text { and }\left(e_{1}=j \text { or } j=0\right) \\ A[i-1, j]+A\left[i-1, j-e_{i}\right] & \text { if } i>1\end{cases}
$$

"Saving space by algebraization", STOC 2010
Jesper Nederlof

Weight

| 1 |  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | $t-6339$ t-3175 t |  |  | $\underline{n t}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|  | 2 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|  | 5 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| items |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 3156 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
|  | 3164 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 65 | 2 | 0 | 0 |
|  | 3175 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 82 | 67 | 2 | 0 |

$$
A[i, j]= \begin{cases}0 & \text { if } i=1, e_{1} \neq j, \text { and } e_{1} \neq 0 \\ 1 & \text { if } i=1 \text { and }\left(e_{1}=j \text { or } j=0\right) \\ A[i-1, j]+A\left[i-1, j-e_{i}\right] & \text { if } i>1\end{cases}
$$

Weight

| 1 |  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |  |  | $t$ | $n t$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|  | 2 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
|  | 5 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 0 | 0 | 0 | 0 |
| items |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 3156 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0 |
|  | 3164 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 65 | 2 | $Q$ | 0 |
|  | , 3175 | 1 | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 82 | 67 | 2 | 0 |

$$
A[i, j]= \begin{cases}0 & \text { if } i=1, e_{1} \neq j, \text { and } e_{1} \neq 0 \\ 1 & \text { if } i=1 \text { and }\left(e_{1}=j \text { or } j=0\right) \\ A[i-1, j]+A\left[i-1, j-e_{i}\right] & \text { if } i>1\end{cases}
$$

Weight


$$
A[i, j]= \begin{cases}0 & \text { if } i=1, e_{1} \neq j, \text { and } e_{1} \neq 0 \\ 1 & \text { if } i=1 \text { and }\left(e_{1}=j \text { or } j=0\right) \\ A[i-1, j]+A\left[i-1, j-e_{i}\right] & \text { if } i>1\end{cases}
$$

Weight


$$
A[i, j]= \begin{cases}0 & \text { if } i=1, e_{1} \neq j, \text { and } e_{1} \neq 0 \\ 1 & \text { if } i=1 \text { and }\left(e_{1}=j \text { or } j=0\right) \\ A[i-1, j]+A\left[i-1, j-e_{i}\right] & \text { if } i>1\end{cases}
$$

## Definition (Convolution operator)

- Define the operator $\otimes$ on column vectors of size $N$ as

$$
\mathbf{a} \otimes \mathbf{b}=\left(\sum_{i+j=k} a_{i} b_{j}\right)_{0 \leq k<N}^{T}
$$

## Definition (Convolution operator)

- Define the operator $\otimes$ on column vectors of size $N$ as

$$
\mathbf{a} \otimes \mathbf{b}=\left(\sum_{i+j=k} a_{i} b_{j}\right)_{0 \leq k<N}^{T}
$$

## Example

$$
\left(\begin{array}{l}
1 \\
3 \\
4 \\
0 \\
0
\end{array}\right) \otimes\left(\begin{array}{l}
2 \\
3 \\
3 \\
0 \\
0
\end{array}\right)=\left(\begin{array}{c}
2 \\
9 \\
20 \\
21 \\
12
\end{array}\right)
$$

## Definition (Convolution operator)

- Define the operator $\otimes$ on column vectors of size $N$ as

$$
\mathbf{a} \otimes \mathbf{b}=\left(\sum_{i+j=k} a_{i} b_{j}\right)_{0 \leq k<N}^{T}
$$

## Example

$$
\left(\begin{array}{l}
1 \\
3 \\
4 \\
0 \\
0
\end{array}\right) \otimes\left(\begin{array}{l}
2 \\
3 \\
3 \\
0 \\
0
\end{array}\right)=\left(\begin{array}{c}
2 \\
9 \\
20 \\
21 \\
12
\end{array}\right)
$$

$$
\left(1+3 x+4 x^{2}\right)\left(2+3 x+3 x^{2}\right)=2+9 x+20 x^{2}+21 x^{3}+12 x^{4}
$$

## Convolution

## Definition (Pointwise multiplication)

- Let • be the point-wise multiplication of two vectors, that is:

$$
\mathbf{a} \cdot \mathbf{b}=\left(\begin{array}{c}
a_{0} \\
a_{1} \\
\vdots \\
a_{N-1}
\end{array}\right) \cdot\left(\begin{array}{c}
b_{0} \\
b_{1} \\
\vdots \\
b_{N-1}
\end{array}\right)=\left(\begin{array}{c}
a_{0} b_{0} \\
a_{1} b_{1} \\
\vdots \\
a_{N-1} b_{N-1}
\end{array}\right)
$$

## Convolution

- Assume we have an invertible matrix $\mathbf{T}$ such that for every $\mathbf{a}, \mathbf{b}$

$$
\mathbf{T}(\mathbf{a} \otimes \mathbf{b})=\mathbf{T a} \cdot \mathbf{T} \mathbf{b}
$$

- and we want to compute $\mathbf{d}_{t}$, where (for example)

$$
\mathbf{d}=(\mathbf{a} \otimes \mathbf{b}) \otimes(\mathbf{c}+\mathbf{a})
$$

- Then we know that

$$
\begin{aligned}
\mathbf{T} \mathbf{d} & =\mathbf{T}((\mathbf{a} \otimes \mathbf{b}) \otimes(\mathbf{c}+\mathbf{a})) \\
& =\mathbf{T}(\mathbf{a} \otimes \mathbf{b}) \cdot \mathbf{T}(\mathbf{c}+\mathbf{a}) \\
& =(\mathbf{T} \mathbf{a} \cdot \mathbf{T} \mathbf{b}) \cdot(\mathbf{T} \mathbf{c}+\mathbf{T a})
\end{aligned}
$$

- And $\mathbf{d}_{t}$ can be obtained using

$$
\mathbf{d}_{t}=\left(\mathbf{T}^{-1} \mathbf{T d}\right)_{t}=\sum_{i=0}^{N-1} \mathbf{T}_{t i}^{-1}\left((\mathbf{T a})_{d}(\mathbf{T b})_{d}\right)\left((\mathbf{T c})_{d}+(\mathbf{T a})_{d}\right)
$$

## Discrete Fourier Transform (DFT)

## Definition

Let $\omega$ be a number s.t. $\omega^{N}=1$ and $\omega^{i} \neq 1$ for $1<i<N$, then the discrete Fourier transform $F$ is defined as:

$$
\mathbf{F}=\left(\begin{array}{cccc}
1 & 1 & \ldots & 1 \\
1 & \omega & \ldots & \omega^{N-1} \\
\vdots & \vdots & \omega^{i j} & \vdots \\
1 & \omega^{N-1} & \ldots & \omega^{(N-1)(N-1)}
\end{array}\right)
$$

## Lemma

$$
\mathbf{F}^{-1}=\frac{1}{N}\left(\omega^{-i j}\right)_{0 \leq i, j<N} \quad \text { and also } \quad \mathbf{F}(\mathbf{a} \otimes \mathbf{b})=\mathbf{F a} \cdot \mathbf{F b}
$$

## Subset Sum revisited

- Now we will use the DFT on the dynamic programming algorithm.
- In order to achieve this we first have to change perspective slightly, and consider the DP table as an array of row vectors.
- For an integer $k$, denote $I_{k}$ as the $k$ 'th column of the identity matrix.


## Example

$$
\left(\begin{array}{l}
1 \\
3 \\
4 \\
0 \\
0
\end{array}\right) \otimes\left(\begin{array}{l}
0 \\
0 \\
1 \\
0 \\
0
\end{array}\right)=\left(\begin{array}{l}
0 \\
0 \\
1 \\
3 \\
4
\end{array}\right)
$$

## Metamorphosis

So we had

$$
A[i, j]= \begin{cases}0 & \text { if } i=1, e_{1} \neq j, \text { and } e_{1} \neq 0 \\ 1 & \text { if } i=1 \text { and }\left(e_{1}=j \text { or } j=0\right) \\ A[i-1, j]+A\left[i-1, j-e_{i}\right] & \text { if } i>1\end{cases}
$$

And we rewrite it as

$$
A[i]= \begin{cases}I_{0}^{T}+I_{e_{1}}^{T} & \text { if } i=1 \\ A[i-1]+A[i-1] \otimes I_{e_{i}}^{T} & \text { if } i>1\end{cases}
$$

- Recall we are interested in $A[n, t]=A[n]_{t}$.


## Metamorphosis

$$
A[i]=\left\{\begin{array}{lll}
I_{0}^{T}+I_{e_{1}}^{T} & \text { if } i=1 \\
A[i-1]+A[i-1] \otimes & I_{e_{i}}^{T} & \text { if } i>1
\end{array}\right.
$$

## Metamorphosis

$$
\mathbf{F}(A[i])=\left\{\begin{array}{lll}
\mathbf{F}\left(I_{0}^{T}+I_{e_{1}}^{T}\right) & & \text { if } i=1 \\
\mathbf{F}(A[i-1]+A[i-1] \otimes & \left.I_{e_{i}}^{T}\right) & \text { if } i>1
\end{array}\right.
$$

## Metamorphosis

$$
\mathbf{F}(A[i])=\left\{\begin{array}{ll}
\mathbf{F} I_{0}^{T}+\mathbf{F} I_{e_{1}}^{T} & \text { if } i=1 \\
\mathbf{F}(A[i-1])+\mathbf{F}(A[i-1] \otimes & \left.I_{e_{i}}^{T}\right)
\end{array} \text { if } i>1 .\right.
$$

## Metamorphosis

$$
\mathbf{F}(A[i])=\left\{\begin{array}{ll}
\mathbf{F} I_{0}^{T}+\mathbf{F} I_{e_{1}}^{T} & \text { if } i=1 \\
\mathbf{F}(A[i-1])+\mathbf{F}(A[i-1] & ) \cdot \mathbf{F}\left(I_{e_{i}}^{T}\right)
\end{array} \text { if } i>1 .\right.
$$

## Metamorphosis

$$
\mathbf{F}(A[i])=\left\{\begin{array}{ll}
\mathbf{F} I_{0}^{T}+\mathbf{F} I_{e_{1}}^{T} & \text { if } i=1 \\
\mathbf{F}(A[i-1])+\mathbf{F}(A[i-1] & ) \cdot \mathbf{F}\left(I_{e_{i}}^{T}\right)
\end{array} \text { if } i>1 .\right.
$$

- And all dependency between different components of vectors is gone, since we only use addition and point wise multiplication.


## The transformed table



## The transformed table

|  |  |  |  |  |  |  |  | ndex | x of | vec |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 6339 | -31 | $t$ | $n t$ |
| 1 | 1 | 5 | 23 | 68 | 79 | 14 | 143 | 87 | 401 | 413 | 154 | 294 | 513 | 94 |
|  | 2 | 41 | 25 | 325 | 83 | 25 | 325 | 6 | 72 | 9 | 97 | 32 | 273 | 26 |
|  | 5 | 43 | 12 | 13 | 91 | 150 | 13 | 267 | 65 | 89 | 256 | 426 | 18 | 103 |
| which |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
|  | 3156 | 6 | 65 | 44 | 12 | 109 | 44 | 325 | 9 | 25 | 169 | 113 | 93 | 284 |
|  | 3164 | 72 | 43 | 98 | 72 | 83 | 98 | 83 | 43 | 78 | 36.5 | 52 | 265 | 185 |
|  | , 3175 | 516 | 12 | 78 | 283 | 12 | 247 | 43 | 102 | 13 | 62 | 77 | 112 | 394 |

- Any component of the bottom row can be computed using $\mathcal{O}(n)$ additions and multiplications!


## The finishing touch

- So we can compute any component of the vector $\mathbf{F}(A[n])$ fast.
- Now we can compute $A[n]_{t}$ according to

$$
A[n]_{t}=\left(\mathbf{F}^{-1}(\mathbf{F}(A[n]))\right)_{t}=
$$

$$
\left(\left(\begin{array}{cccc}
1 & 1 & \cdots & 1 \\
1 & \omega & \cdots & \omega^{-(N-1)} \\
\vdots & \vdots & \omega^{-i j} & \vdots \\
1 & \omega^{-(N-1)} & \cdots & \omega^{-(N-1)(N-1)}
\end{array}\right)\left(\begin{array}{c}
(\mathbf{F}(A[n]))_{1} \\
\vdots \\
(\mathbf{F}(A[n]))_{N-1}
\end{array}\right)\right)_{t}
$$

## The finishing touch

- So we can compute any component of the vector $\mathbf{F}(A[n])$ fast.
- Now we can compute $A[n]_{t}$ according to

$$
A[n]_{t}=\left(\mathbf{F}^{-1}(\mathbf{F}(A[n]))\right)_{t}=
$$

$$
\left(\left(\begin{array}{cccc}
1 & 1 & \cdots & 1 \\
1 & \omega & \cdots & \omega^{-(N-1)} \\
\vdots & \vdots & \omega^{-i j} & \vdots \\
1 & \omega^{-(N-1)} & \cdots & \omega^{-(N-1)(N-1)}
\end{array}\right)\left(\begin{array}{c}
(\mathbf{F}(A[n]))_{1} \\
\vdots \\
(\mathbf{F}(A[n]))_{N-1}
\end{array}\right)\right)_{t}
$$

- and we are done!


## The transformed table



## The transformed table



## The transformed table



## The transformed table



## The transformed table



The new algorithm uses $\tilde{\mathcal{O}}\left(n^{3} t \log t\right)$ time and $\tilde{\mathcal{O}}\left(n^{2}\right)$ space.

## Further remarks

- The algorithm has to work in a field where there exists an $\omega$.
- This can be achieved by using for example complex numbers with finite precision.
- We also used the Möbius transformation to save space for a different type of DP algorithms
- in combination with the DFT, this found applications to among others the Traveling Salesman and Weighted Steiner Tree problems.
- It would be interesting to find more transformations that also are useful to save space for existing DP algorithms.


## Further research

- Can we efficiently save space for "Min-Sum DP algorithms"?
- For the Knapsack problem the approach results in an algorithm much slower then the corresponding DP algorithm.
- Are there space and time efficient algorithms for deciding properties of partial $k$-trees (for example, maximum independent set)?
- Does there exists a positive $\epsilon$ such that
- Subset Sum can be solved in $\mathcal{O}\left((2-\epsilon)^{n}\right)$ time and polynomial space?
- Subset Sum can be solved in $\mathcal{O}\left(n^{c} t^{(1-\epsilon)}\right)$ time?


## Thanks for listening!

