• The Forward-Backward Algorithm
The First Secret of the FBA

• Suppose we are given $N$ elements $g_1, g_2, \ldots, g_N$, in a non-commutative ring and asked to compute the $N$ products

$$ h_1 = g_2 g_3 \cdots g_N $$
$$ h_2 = g_1 g_3 g_4 \cdots g_N $$
$$ \vdots $$
$$ h_N = g_1 g_2 \cdots g_{N-1} $$

• This task appears to require $N(N - 2)$ multiplications, but ...
The First Secret of the FBA

- But if we precompute the $2N$ quantities

\[
\begin{align*}
\alpha_1 &= g_1 \\
\alpha_2 &= (g_1) \cdot g_2 \\
\alpha_3 &= (g_1 g_2) \cdot g_3 \\
& \vdots \\
\alpha_N &= (g_1 \cdots g_{N-1}) \cdot g_N
\end{align*}
\]

\[
\begin{align*}
\beta_1 &= g_1 \cdot (g_2 \cdots g_N) \\
& \vdots \\
\beta_{N-2} &= g_{N-2} \cdot (g_{N-1} g_N) \\
\beta_{N-1} &= g_{N-1} \cdot (g_N) \\
\beta_N &= g_N
\end{align*}
\]

which requires $2(N - 1)$ multiplications, \ldots
The First Secret of the FBA

- Then

\[ h_1 = \beta_2 \]
\[ h_2 = \alpha_1 \cdot \beta_3 \]
\[ h_3 = \alpha_2 \cdot \beta_4 \]

\[ \vdots \]

\[ h_{N-1} = \alpha_{N-2} \cdot \beta_N \]
\[ h_N = \alpha_{N-1} \]

which requires \( N - 2 \) more multiplications, for a total of \( 3N - 4 \) multiplications.
Let $A$ be an $2 \times 2$ matrix. How hard is it to compute $A^n$?

By definition,

$$[A^n]_{i,j} = \sum_{i_2, \ldots, i_{n-1}} a_{i,i_2} a_{i_2,i_3} \cdots a_{i_{n-1},j},$$

which appears to require $2^n$ additions and $2^{n+1}$ multiplications.

But we know it’s not that hard …
The Second Secret of the FBA

• Having computed $A^i$, we can compute $A^{i+1}$ using the associative law:
  
  $$A^{i+1} = A^i \cdot A,$$

  which requires 4 additions and 8 multiplications.

• Therefore we can compute $A^n$ with $4(n-1)$ additions and $8(n-1)$ multiplications.
The Third Secret of the FBA

- The distributive law:

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_i b_j = (a_1 + \cdots + a_n)(b_1 + \cdots + b_n)$$

- LHS = $n^2$ additions and multiplications.
- RHS = $2(n - 1)$ additions and 1 multiplication.
A Weighted Trellis of Rank 5
An Equivalent Set of “Compatible” Matrices

\[
W_1 = A \begin{pmatrix} b & c \\ 1 & 2 \end{pmatrix} \quad W_2 = \begin{pmatrix} d & e & f \\ b & 3 & 0 & 1 \\ c & 2 & 4 & 0 \end{pmatrix}
\]

\[
W_3 = \begin{pmatrix} g & h & j \\ d & 0 & 1 & 2 \end{pmatrix} \quad W_4 = \begin{pmatrix} k & m \\ g & 2 & 0 \end{pmatrix}
\]

\[
W_5 = \begin{pmatrix} B \\ k & 1 \\ m & 3 \end{pmatrix}
\]
Path Weights

- The weight of a path equals the \textit{product} of the component edge weights.
- The \textcolor{red}{red} path has weight 48,
- The flow from $A$ to $B$, denoted $\mu(A, B)$, is the sum of the weights of all paths from $u$ to $v$. 
Vertex-Constrained Flows

• The flow from $u$ to $v$ through $x$, denoted $\mu_x(u,v)$, is the sum of the weights of all paths from $u$ to $v$ which pass through $x$. 
Edge-Constrained Flows

• The flow from $u$ to $v$ through $e$, denoted $\mu_e(u, v)$, is the sum of the weights of all paths from $u$ to $v$ which pass through $x$. 
Problems Solved by the FBA

Problem 1. Compute the flow from $A$ to $B$, i.e.,

$$\mu(A, B) = \sum_{T: A \rightarrow B} w(T).$$

Since there may be as many as $q_1 q_2 \cdots q_{N-1}$ paths from $A$ to $B$, the computation of $\mu(A, B)$ appears to be a formidable task. However, as we shall see, the FBA computes this flow using at most $2(q_0 q_1 + q_1 q_2 + \cdots + q_{N-1} q_N)$ arithmetic operations.
Problems Solved by the FBA

Problem 2. For each vertex $v$, compute the flow from $A$ to $B$ through $v$, i.e.,

$$\mu_v(A, B) = \sum_{T : A \rightarrow B} w(T).$$
Problems Solved by the FBA

**Problem 3.** For each edge $e$, compute the flow from $A$ to $B$ through $e$, i.e.,

$$
\mu_e(A, B) = \sum_{T: A \xrightarrow{e} B} w(T).
$$
Another FBA Secret

Theorem. For all vertices $u$, $v$, and $x$, we have

$$\mu_x(u, v) = \mu(u, x)\mu(x, v).$$

Similarly, if $\text{init}(e) = x$ and $\text{fin}(e) = y$,

$$\mu_e(u, v) = \mu(u, x) \cdot w(e) \cdot \mu(y, v).$$
Suppose there are $m$ paths from $u$ to $x$, say $P_1, \ldots, P_m$, and $n$ paths from $x$ to $v$, say $Q_1, \ldots, Q_n$. Then there are exactly $mn$ paths from $u$ to $v$ through $x$, namely paths of the form $T_i * Q_j$, where "*" denotes concatenation.
Proof, Continued

Then we have

\[ \mu_x(u, v) = \sum_{T: u \mapsto v} w(T) \]

\[ = \sum_{i=1}^{m} \sum_{j=1}^{n} w(T_i \ast Q_j) \]

\[ = \sum_{i=1}^{m} \sum_{j=1}^{n} w(T_i) w(Q_j) \]

\[ = \left( \sum_{i=1}^{m} w(T_i) \right) \left( \sum_{j=1}^{n} w(Q_j) \right) \quad \text{(The distributive law)} \]

\[ = \mu(u, x) \mu(x, v). \]
A Weighted Trellis of Rank 5
Path Weights

- The weight of a path equals the product of the component edge weights.
- The red path has weight 48,
• The flow from $u$ to $v$, denoted $\mu(u,v)$, is the sum of the weights of all paths from $u$ to $v$. 
The Three Problems

Problem 1. Compute the flow $\mu(A, B)$.

Problem 2. For each vertex $v$, compute the flow $\mu_v(A, B)$.

Problem 3. For each edge $e$, compute the flow $\mu_e(A, B)$.
An FBA Secret: The Distributive Law

**Theorem.** For all vertices \( u, v, \) and \( x, \) we have

\[
\mu_x(u, v) = \mu(u, x)\mu(x, v).
\]

Similarly, if \( \text{init}(e) = x \) and \( \text{fin}(e) = y, \)

\[
\mu_e(u, v) = \mu(u, x) \cdot w(e) \cdot \mu(y, v).
\]
The Forward-Backward Recursions

For each \( i = 1, 2, \ldots, N \), define a \( q_{i-1} \times q_i \) matrix \( W_i \), whose rows are indexed by the vertices in \( V_{i-1} \) and columns are indexed by the vertices in \( V_i \):

\[
W_i(u, v) = \sum_{e: \text{init}(e) = u, \text{fin}(e) = v} w(e)
\]
Example

\[ W_1 = A \begin{pmatrix} b & c \\ 1 & 2 \end{pmatrix} \]  \[ W_2 = b \begin{pmatrix} d & e & f \\ 3 & 0 & 1 \\ c & 2 & 4 & 0 \end{pmatrix} \]

\[ W_3 = e \begin{pmatrix} g & h & j \\ d & 0 & 1 & 2 \end{pmatrix} \]  \[ W_4 = h \begin{pmatrix} k & m \end{pmatrix} \]

\[ W_5 = k \begin{pmatrix} B \\ m & \end{pmatrix} \]}
Another Example
The Forward and Backward Recursions

- $\alpha_i$ is a row vector of dimension $q_i$:

  $$\alpha_0 = 1, \quad \alpha_i = \alpha_{i-1} W_i, \quad \text{for } i = 1, \ldots, N.$$  

- $\beta_i$ is a column vector of dimension $q_i$:

  $$\beta_N = 1, \quad \beta_i = W_{i+1} \beta_{i+1}, \quad \text{for } i = N - 1, N - 2, \ldots, 0.$$  

- If $v \in V_i$, we denote the $v$th component of $\alpha_i$ resp. $\beta_i$ by $\alpha_i(v)$ resp. $\beta_i(v)$.  

The Forward and Backward Recursions

- We have by induction

\[
\alpha_i = W_1 W_2 \cdots W_i, \quad \text{for } i = 1, \ldots, N.
\]
\[
\beta_i = W_{i+1} W_{i+2} \cdots W_N, \quad \text{for } i = N - 1, N - 2, \ldots, 0.
\]
$\alpha$ Example

\[ \alpha_0 = \begin{pmatrix} 1 \end{pmatrix} \]
\[ \alpha_1 = \alpha_0 W_1 = \begin{pmatrix} 1 & 2 \end{pmatrix} \]
\[ \alpha_2 = \alpha_1 W_2 = \begin{pmatrix} 7 & 81 & 1 \end{pmatrix} \]
\[ \alpha_3 = \alpha_2 W_3 = \begin{pmatrix} 24 & 16 & 33 \end{pmatrix} \]
\[ \alpha_4 = \alpha_3 W_4 = \begin{pmatrix} 96 & 49 \end{pmatrix} \]
\[ \alpha_5 = \alpha_4 W_5 = \begin{pmatrix} 243 \end{pmatrix} \]
\( \beta \) Example

\[
\begin{align*}
\beta_5 &= (1) \\
\beta_4 &= W_5 \beta_5 = (1 \ 3)^T \\
\beta_3 &= W_4 \beta_4 = (2 \ 6 \ 3)^T \\
\beta_2 &= W_3 \beta_3 = (12 \ 18 \ 15)^T \\
\beta_1 &= W_2 \beta_2 = (51 \ 96)^T \\
\beta_0 &= W_1 \beta_1 = (243)
\end{align*}
\]
An alternative definition of the $\alpha_i$s and $\beta_i$s

\[
\alpha_i(u) = \sum_{\substack{e \in E_{i-1,i}^\text{fin}(e)=u}} \alpha_{i-1}(\text{init}(e)) \cdot w(e).
\]

\[
\beta_i(u) = \sum_{\substack{e \in E_{i,i+1}^\text{init}(e)=u}} w(e) \cdot \beta_{i+1}(\text{fin}(e)).
\]
Example
Complexity of the FBA

These definitions are preferable if the trellis is sparse, and can be used to show that the arithmetic complexity of computing the $\alpha_i$’s and $\beta_i$’s is $O(|E|)$, where $E$ is the total number of edges in the trellis.
The Key Property of the $\alpha$’s and the $\beta$’s

**Theorem.** For any $u \in V_i$, we have

$$\alpha_i(u) = \mu(A, u).$$

Similarly, for any $v \in V_i$, we have

$$\beta_i(v) = \mu(v, B).$$
Proof of Key Property

We use induction on $i$. The case $i = 1$ says that

$$\alpha_1(u) = \mu(A, u).$$

But $\alpha_1 = W_1$, and $W_1(A, u)$ is by definition the sum of the weights of the edges (if any) between $A$ and $u$. So (1) is correct for $i = 1$. 
Proof, Continued

Assuming the truth of (1) for \( i \), we proceed as follows. Let \( u \in V_{i+1} \). Then

\[
\mu(A, u) = \sum_{x \in V_i} \mu_x(A, u) \quad \text{(conservation of flow)}
\]

\[
= \sum_{x \in V_i} \mu(A, x) \mu(x, u) \quad \text{(distributive law)}
\]

\[
= \sum_{x \in V_i} \alpha_i(x) W_{i+1}(x, u) \quad \text{(ind. hyp., def. of } W_{i+1})
\]

\[
= \alpha_{i+1}(u) \quad \text{(def. of } \alpha_{i+1})
\]
Recall the Three Problems

**Problem 1.** Compute the flow $\mu(A, B)$.

**Problem 2.** For each vertex $v$, compute the flow $\mu_v(A, B)$.

**Problem 3.** For each edge $e$, compute the flow $\mu_e(A, B)$. 
The Three Solutions

Solution 1. For any $i = 0, 1, \ldots, N$, we have

$$\mu(A, B) = \alpha_i \beta_i.$$  

In particular, $\mu(A, B) = \alpha_N = \beta_0$. (One-way algorithm will do.)

Corollary. We have

$$\mu(A, B) = W_1 W_2 \cdots W_N.$$  

Thus computing flows is equivalent to matrix multiplication!
Proof of Solution 1

For any index $i$ we have

$$
\mu(A, B) = \sum_{x \in V_i} \mu_x(A, B) \quad \text{(conservation of flow)}
$$

$$
= \sum_{x \in V_i} \mu(A, x) \mu(x, B) \quad \text{(Distributive law)}
$$

$$
\sum_{x \in V_i} \alpha_i(x) \beta_i(x) \quad \text{(Key Property)}
$$

$$
= \alpha_i \beta_i.
$$
Solution to Problem 2

Solution 2. If $v \in V_i$, we have

$$\mu_v(A, B) = \alpha_i(v)\beta_i(v).$$

Proof: We have

$$\mu_v(A, B) = \mu(A, v)\mu(v, B) \quad \text{(Distributive law)}$$

$$= \alpha_i(v)\beta_i(v) \quad \text{(Key Property)}$$
Solution to Problem 3

Solution 3. If $\text{init}(e) = u \in V_{i-1}$ and $\text{fin}(e) = v \in V_i$, we have

$$\mu_e(A, B) = \alpha_{i-1}(u) \cdot w(e) \cdot \beta_i(v).$$

Proof:

$$\mu_e(A, B) = \mu(A, u) \cdot w(e) \cdot \mu(v, B) \quad \text{(Distributive Law)}$$
$$= \alpha_{i-1}(u) \cdot w(e) \cdot \beta_i(v). \quad \text{(Key Property)}$$
Example

\[\begin{align*}
\alpha_0 &= (1) \\
\alpha_1 &= (1 \ 2) \\
\alpha_2 &= (7 \ 81 \ 1) \\
\alpha_3 &= (24 \ 16 \ 33) \\
\alpha_4 &= (96 \ 49) \\
\alpha_5 &= (243) \\
\beta_0 &= (1) \\
\beta_1 &= (1 \ 3)^T \\
\beta_2 &= (2 \ 6 \ 3)^T \\
\beta_3 &= (12 \ 18 \ 15)^T \\
\beta_4 &= (51 \ 96)^T \\
\beta_5 &= (243)
\end{align*}\]

\[\mu(A, B) = \alpha_0 \beta_0 = \alpha_1 \beta_0 = \alpha_2 \beta_2 = \alpha_3 \beta_3 = \alpha_4 \beta_4 = \alpha_5 \beta_5 = 243.\]
Example

Notation: \( \langle x, y \rangle = (x_1y_1, \ldots, x_ny_n) \):

\[
\begin{align*}
\langle \alpha_0, \beta_0 \rangle &= \begin{pmatrix} A \\ 243 \end{pmatrix} & \langle \alpha_1, \beta_1 \rangle &= \begin{pmatrix} b & c \\ 51 & 192 \end{pmatrix} \\
\langle \alpha_2, \beta_2 \rangle &= \begin{pmatrix} d & e & f \\ 84 & 144 & 15 \end{pmatrix} & \langle \alpha_3, \beta_3 \rangle &= \begin{pmatrix} g & h & j \\ 48 & 96 & 99 \end{pmatrix} \\
\langle \alpha_4, \beta_4 \rangle &= \begin{pmatrix} k & m \\ 96 & 147 \end{pmatrix} & \langle \alpha_5, \beta_5 \rangle &= \begin{pmatrix} B \\ 243 \end{pmatrix}
\end{align*}
\]

For example, the flow from \( A \) to \( B \) through vertex \( h \) is 96.
Example

\[
\begin{align*}
\langle \alpha_0 W_1 \beta_1 \rangle &= A \begin{pmatrix} b & c \\ 51 & 192 \end{pmatrix} & \langle \alpha_1 W_2 \beta_2 \rangle &= b \begin{pmatrix} d & e & f \\ 36 & 0 & 15 \\ 48 & 144 & 0 \end{pmatrix} \\
\langle \alpha_2 W_3 \beta_3 \rangle &= e \begin{pmatrix} d & h & j \\ 0 & 42 & 42 \end{pmatrix} & \langle \alpha_3 W_4 \beta_4 \rangle &= g \begin{pmatrix} k & m \\ 48 & 0 \end{pmatrix} \\
\langle \alpha_4 W_5 \beta_5 \rangle &= k \begin{pmatrix} f \\ 0 & 6 & 9 \end{pmatrix} & \langle \alpha_3 W_4 \beta_4 \rangle &= h \begin{pmatrix} 48 & 48 \\ 0 & 99 \end{pmatrix} \\
\end{align*}
\]
For example, the flow from $A$ to $B$ through the edge $e \rightarrow h$ is 48.