L11: Algebraic Path Problems with applications to Internet Routing

Timothy G. Griffin

timothy.griffin@cl.cam.ac.uk
Computer Laboratory
University of Cambridge, UK

Michaelmas Term, 2019
Shortest paths example, \( sp = (\mathbb{N}^\infty, \text{min}, +, \infty, 0) \)

The adjacency matrix

\[
A = \begin{bmatrix}
1 & 2 & 1 & 6 & \infty \\
2 & \infty & 5 & \infty & 4 \\
3 & 1 & 5 & \infty & 4 & 3 \\
4 & 6 & \infty & 4 & \infty & \infty \\
5 & \infty & 4 & 3 & \infty & \infty
\end{bmatrix}
\]
Shortest paths solution

solves this global optimality problem:

\[ A^*(i, j) = \min_{p \in \pi(i, j)} w(p), \]

where \( \pi(i, j) \) is the set of all paths from \( i \) to \( j \).
Widest paths example, \( \text{bw} = (\mathbb{N}^\infty, \text{max}, \min, 0, \infty) \)

A solvers this global optimality problem:

\[
A^*(i, j) = \max_{p \in \pi(i, j)} w(p),
\]

where \( w(p) \) is now the minimal edge weight in \( p \).
Unfamiliar example, \( (2\{a, b, c\}, \cup, \cap, \{\}, \{a, b, c\}) \)

We want \( A^* \) to solve this global optimality problem:

\[
A^*(i, j) = \bigcup_{p \in \pi(i, j)} w(p),
\]

where \( w(p) \) is now the intersection of all edge weights in \( p \).

For \( x \in \{a, b, c\} \), interpret \( x \in A^*(i, j) \) to mean that there is at least one path from \( i \) to \( j \) with \( x \) in every arc weight along the path.

\[
A^*(4, 1) = \{a, b\} \quad A^*(4, 5) = \{b\}
\]
Another unfamiliar example, \((2^{\{a, b, c\}}, \cap, \cup)\)

We want matrix \(R\) to solve this global optimality problem:

\[
A^*(i, j) = \bigcap_{p \in \Pi(i, j)} w(p),
\]

where \(w(p)\) is now the union of all edge weights in \(p\).

For \(x \in \{a, b, c\}\), interpret \(x \in A^*(i, j)\) to mean that every path from \(i\) to \(j\) has at least one arc with weight containing \(x\).

\[
A^*(4, 1) = \{b\} \quad A^*(4, 5) = \{b\} \quad A^*(5, 1) = \{
\]
Semirings (generalise $(\mathbb{R}, +, \times, 0, 1)$)

<table>
<thead>
<tr>
<th>name</th>
<th>$S$</th>
<th>$\oplus$, $\otimes$, $\bar{0}$, $\bar{1}$</th>
<th>possible routing use</th>
</tr>
</thead>
<tbody>
<tr>
<td>sp</td>
<td>$\mathbb{N}^\infty$</td>
<td>$\min$, $+$, $\infty$, $0$</td>
<td>minimum-weight routing</td>
</tr>
<tr>
<td>bw</td>
<td>$\mathbb{N}^\infty$</td>
<td>$\max$, $\min$, $0$, $\infty$</td>
<td>greatest-capacity routing</td>
</tr>
<tr>
<td>rel</td>
<td>$[0, 1]$</td>
<td>$\max$, $\times$, $0$, $1$</td>
<td>most-reliable routing</td>
</tr>
<tr>
<td>use</td>
<td>${0, 1}$</td>
<td>$\max$, $\min$, $0$, $1$</td>
<td>usable-path routing</td>
</tr>
<tr>
<td></td>
<td>$2^W$</td>
<td>$\cup$, $\cap$, ${}$, $W$</td>
<td>shared link attributes?</td>
</tr>
<tr>
<td></td>
<td>$2^W$</td>
<td>$\cap$, $\cup$, ${}$, $W$</td>
<td>shared path attributes?</td>
</tr>
</tbody>
</table>

A wee bit of notation!

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathbb{N}$</td>
<td>Natural numbers (starting with zero)</td>
</tr>
<tr>
<td>$\mathbb{N}^\infty$</td>
<td>Natural numbers, plus infinity</td>
</tr>
<tr>
<td>$\bar{0}$</td>
<td>Identity for $\oplus$</td>
</tr>
<tr>
<td>$\bar{1}$</td>
<td>Identity for $\otimes$</td>
</tr>
</tbody>
</table>
Recommended (on reserve in CL library)
We will look at all of the axioms of semirings, but the most important are

**distributivity**

\[
\begin{align*}
LD & : \quad a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c) \\
RD & : \quad (a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c)
\end{align*}
\]
Distributivity, illustrated

\[ a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c) \]

\( j \) makes the choice \( = \) \( i \) makes the choice
Should distributivity hold in Internet Routing?

- \( j \) prefers long path though one of its customers (not the shorter path through a competitor).
- Given two routes from a provider, \( i \) prefers the one with a shorter path.
- More on inter-domain routing in the Internet later in the term ...
Widest shortest-paths

- Metric of the form \((d, b)\), where \(d\) is distance \((\min, +)\) and \(b\) is capacity \((\max, \min)\).
- Metrics are compared lexicographically, with distance considered first.
- Such things are found in the vast literature on Quality-of-Service (QoS) metrics for Internet routing.
Widest shortest-paths

```
0  (1,10)  1  (1,100)  2
    |        |        |
    |        |        |
(1,5)  3  (1,100)  (1,100)
```
Weights are globally optimal (we have a semiring)

Widest shortest-path weights computed by Dijkstra and Bellman-Ford

\[ R = \begin{bmatrix}
0 & (0, \infty) & (1, 10) & (3, 10) & (2, 5) & (2, 10) \\
1 & (1, 10) & (0, \infty) & (2, 100) & (1, 5) & (1, 100) \\
2 & (3, 10) & (2, 100) & (0, \infty) & (1, 100) & (1, 100) \\
3 & (2, 5) & (1, 5) & (1, 100) & (0, \infty) & (2, 100) \\
4 & (2, 10) & (1, 100) & (1, 100) & (2, 100) & (0, \infty)
\end{bmatrix} \]
But what about the paths themselves?

Four optimal paths of weight $(3, 10)$.

$$P_{\text{optimal}}(0, 2) = \{(0, 1, 2), (0, 1, 4, 2)\}$$
$$P_{\text{optimal}}(2, 0) = \{(2, 1, 0), (2, 4, 1, 0)\}$$

There are standard ways to extend Bellman-Ford and Dijkstra to compute paths (or the associated next hops).

Do these extended algorithms find all optimal paths?
Surprise!

### Four optimal paths of weight (3, 10)

\[
\begin{align*}
P_{\text{optimal}}(0, 2) &= \{ (0, 1, 2), (0, 1, 4, 2) \} \\
P_{\text{optimal}}(2, 0) &= \{ (2, 1, 0), (2, 4, 1, 0) \}
\end{align*}
\]

### Paths computed by (extended) Dijkstra

\[
\begin{align*}
P_{\text{Dijkstra}}(0, 2) &= \{ (0, 1, 2), (0, 1, 4, 2) \} \\
P_{\text{Dijkstra}}(2, 0) &= \{ (2, 4, 1, 0) \}
\end{align*}
\]

Notice that 0’s paths cannot both be implemented with next-hop forwarding since \(P_{\text{Dijkstra}}(1, 2) = \{ (1, 4, 2) \} \).

### Paths computed by distributed Bellman-Ford

\[
\begin{align*}
P_{\text{Bellman}}(0, 2) &= \{ (0, 1, 4, 2) \} \\
P_{\text{Bellman}}(2, 0) &= \{ (2, 1, 0), (2, 4, 1, 0) \}
\end{align*}
\]
Optimal paths from 0 to 2. Computed by Dijkstra but not by Bellman-Ford.
Optimal paths from 2 to 1. Computed by Bellman-Ford but not by Dijkstra
How can we understand this (algebraically)?

The Algorithm to Algebra (A2A) method

\[
\begin{pmatrix}
\text{original metric} \\
\text{complex algorithm}
\end{pmatrix} 
\overset{+}{\rightarrow} 
\begin{pmatrix}
\text{modified metric} \\
\text{matrix equations (generic algorithm)}
\end{pmatrix}
\]

Preview

- We can add paths explicitly to the widest shortest-path semiring to obtain a new algebra.
- We will see that distributivity does not hold for this algebra.
- Why? We will see that it is because min is not cancellative! 
  \((a \min b = a \min c \text{ does not imply that } b = c)\)
Towards a non-classical theory of algebraic path finding

We need theory that can accept algebras that violate distributivity.

Global optimality

\[ A^*(i, j) = \bigoplus_{p \in P(i, j)} w(p), \]

Left local optimality (distributed Bellman-Ford)

\[ L = (A \otimes L) \oplus I. \]

Right local optimality (Dijkstra’s Algorithm)

\[ R = (R \otimes A) \oplus I. \]

Embrace the fact that all three notions can be distinct.
Lectures 2, 3

- Semigroups
- A few important semigroup properties
- Semigroup and partial orders
A semigroup \((S, \bullet)\) is a non-empty set \(S\) with a binary operation such that

\[ a \bullet (b \bullet c) = (a \bullet b) \bullet c \]

Important Assumption — We will ignore trivial semigroups
We will implicitly assume that \(2 \leq |S|\).

Note
Many useful binary operations are not semigroup operations. For example, \((\mathbb{R}, \bullet)\), where \(a \bullet b \equiv (a + b)/2\).
Some Important Semigroup Properties

<table>
<thead>
<tr>
<th>ID</th>
<th>identity</th>
<th>$\equiv$ $\exists \alpha \in S, \forall a \in S, a = \alpha \cdot a = a \cdot \alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>AN</td>
<td>annihilator</td>
<td>$\equiv$ $\exists \omega \in S, \forall a \in S, \omega = \omega \cdot a = a \cdot \omega$</td>
</tr>
<tr>
<td>CM</td>
<td>commutative</td>
<td>$\equiv$ $\forall a, b \in S, a \cdot b = b \cdot a$</td>
</tr>
<tr>
<td>SL</td>
<td>selective</td>
<td>$\equiv$ $\forall a, b \in S, a \cdot b \in {a, b}$</td>
</tr>
<tr>
<td>IP</td>
<td>idempotent</td>
<td>$\equiv$ $\forall a \in S, a \cdot a = a$</td>
</tr>
</tbody>
</table>

A semigroup with an identity is called a monoid.

Note that

$$\text{SL}(S, \cdot) \implies \text{IP}(S, \cdot)$$
A few concrete semigroups

<table>
<thead>
<tr>
<th>$S$</th>
<th>•</th>
<th>description</th>
<th>$\alpha$</th>
<th>$\omega$</th>
<th>CM</th>
<th>SL</th>
<th>IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$S$</td>
<td>left</td>
<td>$x \leftarrow y = x$</td>
<td>$\epsilon$</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>$S$</td>
<td>right</td>
<td>$x \rightarrow y = y$</td>
<td></td>
<td>$\epsilon$</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>$S^*$</td>
<td>.</td>
<td>concatenation</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>$S^+$</td>
<td>.</td>
<td>concatenation</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>${t, f}$</td>
<td>$\wedge$</td>
<td>conjunction</td>
<td>$t$</td>
<td>$f$</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>${t, f}$</td>
<td>$\vee$</td>
<td>disjunction</td>
<td>$f$</td>
<td>$t$</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>$\mathbb{N}$</td>
<td>$\min$</td>
<td>minimum</td>
<td>$0$</td>
<td>$0$</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>$\mathbb{N}$</td>
<td>$\max$</td>
<td>maximum</td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>$2^W$</td>
<td>$\cup$</td>
<td>union</td>
<td>$\emptyset$</td>
<td>$W$</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$2^W$</td>
<td>$\cap$</td>
<td>intersection</td>
<td>$W$</td>
<td>$\emptyset$</td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{fin}(2^U)$</td>
<td>$\cup$</td>
<td>union</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\text{fin}(2^U)$</td>
<td>$\cap$</td>
<td>intersection</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\mathbb{N}$</td>
<td>+</td>
<td>addition</td>
<td>$0$</td>
<td>$0$</td>
<td>*</td>
<td></td>
<td>*</td>
</tr>
<tr>
<td>$\mathbb{N}$</td>
<td>$\times$</td>
<td>multiplication</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$W$ a finite set, $U$ an infinite set. For set $Y$, $\text{fin}(Y) \equiv \{ X \in Y \mid X \text{ is finite} \}$
A few abstract semigroups

<table>
<thead>
<tr>
<th>$S$</th>
<th>$\bullet$</th>
<th>description</th>
<th>$\alpha$</th>
<th>$\omega$</th>
<th>CM</th>
<th>SL</th>
<th>IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>$2^U$</td>
<td>$\cup$</td>
<td>union</td>
<td>${}$</td>
<td>$U$</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>$2^U$</td>
<td>$\cap$</td>
<td>intersection</td>
<td>$U$</td>
<td>${}$</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>$2^{U \times U}$</td>
<td>$\times$</td>
<td>relational join</td>
<td>$\mathcal{I}_U$</td>
<td>${}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$X \rightarrow X$</td>
<td>$\circ$</td>
<td>composition</td>
<td>$\lambda x. x$</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

$U$ an infinite set

$X \times Y \equiv \{(x, z) \in U \times U | \exists y \in U, (x, y) \in X \land (y, z) \in Y\}$

$\mathcal{I}_U \equiv \{(u, u) | u \in U\}$

**subsemigroup**

Suppose $(S, \bullet)$ is a semigroup and $T \subseteq S$. If $T$ is closed w.r.t $\bullet$ (that is, $\forall x, y \in T, x \bullet y \in T$), then $(T, \bullet)$ is a subsemigroup of $S$. 
Order Relations

We are interested in order relations \( \leq \subseteq S \times S \)

**Definition (Important Order Properties)**

- **Reflexive** \( RX \): \( a \leq a \)
- **Transitive** \( TR \): \( a \leq b \land b \leq c \rightarrow a \leq c \)
- **Antisymmetric** \( AY \): \( a \leq b \land b \leq a \rightarrow a = b \)
- **Total** \( TO \): \( a \leq b \lor b \leq a \)

<table>
<thead>
<tr>
<th></th>
<th>pre-order</th>
<th>partial order</th>
<th>preference order</th>
<th>total order</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RX</strong></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td><strong>TR</strong></td>
<td>*</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td><strong>AY</strong></td>
<td>*</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td><strong>TO</strong></td>
<td></td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>
**Definition (Canonical pre-orders)**

\[
\begin{align*}
  a \leq^R b & \iff \exists c \in S : b = a \cdot c \\
  a \leq^L b & \iff \exists c \in S : a = b \cdot c
\end{align*}
\]

**Lemma ( Sanity check)**

*Associativity of \( \cdot \) implies that these relations are transitive.*

**Proof.**

Note that \( a \leq^R b \) means \( \exists c_1 \in S : b = a \cdot c_1 \), and \( b \leq^R c \) means \( \exists c_2 \in S : c = b \cdot c_2 \). Letting \( c_3 = c_1 \cdot c_2 \) we have \( c = b \cdot c_2 = (a \cdot c_1) \cdot c_2 = a \cdot (c_1 \cdot c_2) = a \cdot c_3 \). That is, \( \exists c_3 \in S : c = a \cdot c_3 \), so \( a \leq^R c \). The proof for \( \leq^L \) is similar.
Canonically Ordered Semigroup

Definition (Canonically Ordered Semigroup)
A commutative semigroup \((S, \bullet)\) is canonically ordered when \(a \subseteq^R c\) and \(a \subseteq^L c\) are partial orders.

Definition (Groups)
A monoid is a group if for every \(a \in S\) there exists a \(a^{-1} \in S\) such that \(a \bullet a^{-1} = a^{-1} \bullet a = \alpha\).
Lemma (THE BIG DIVIDE)

Only a trivial group is canonically ordered.

Proof.
If $a, b \in S$, then $a = \alpha \cdot a = (b \cdot b^{-1}) \cdot a = b \cdot (b^{-1} \cdot a) = b \cdot c$, for $c = b^{-1} \cdot a$, so $a \trianglelefteq_L b$. In a similar way, $b \trianglelefteq_R a$. Therefore $a = b$. \qed
Natural Orders

Definition (Natural orders)

Let \((S, \bullet)\) be a semigroup.

\[ a \preceq^L b \iff a = a \bullet b \]
\[ a \preceq^R b \iff b = a \bullet b \]

Lemma

If \(\bullet\) is commutative and idempotent, then \(a \preceq^D b \iff a \preceq^D b\), for \(D \in \{R, L\}\).

Proof.

\[ a \preceq^R b \iff b = a \bullet c = (a \bullet a) \bullet c = a \bullet (a \bullet c) = a \bullet b \iff a \preceq^R b \]
\[ a \preceq^L b \iff a = b \bullet c = (b \bullet b) \bullet c = b \bullet (b \bullet c) = b \bullet a = a \bullet b \iff a \preceq^L b \]
**Lemma (Natural Bounds)**

- If $\alpha$ exists, then for all $a$, $a \leq^{\text{L}} \alpha$ and $\alpha \leq^{\text{R}} a$
- If $\omega$ exists, then for all $a$, $\omega \leq^{\text{L}} a$ and $a \leq^{\text{R}} \omega$
- If $\alpha$ and $\omega$ exist, then $S$ is bounded.

\[
\omega \leq^{\text{L}} a \leq^{\text{L}} \alpha \\
\alpha \leq^{\text{R}} a \leq^{\text{R}} \omega
\]

**Remark (Thanks to Iljitsch van Beijnum)**

Note that this means for $(\min, +)$ we have

\[
0 \leq^{\text{L}_{\min}} a \leq^{\text{L}_{\min}} \infty \\
\infty \leq^{\text{R}_{\min}} a \leq^{\text{R}_{\min}} 0
\]

and still say that this is bounded, even though one might argue with the terminology!
Examples of special elements

<table>
<thead>
<tr>
<th></th>
<th>S</th>
<th>( \bullet )</th>
<th>( \alpha )</th>
<th>( \omega )</th>
<th>( \leq^L )</th>
<th>( \leq^R )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \mathbb{N}^\infty )</td>
<td>min</td>
<td>( \infty )</td>
<td>0</td>
<td>( \leq )</td>
<td>( \geq )</td>
<td></td>
</tr>
<tr>
<td>( \mathbb{N}^{-\infty} )</td>
<td>max</td>
<td>0</td>
<td>( -\infty )</td>
<td>( \geq )</td>
<td>( \leq )</td>
<td></td>
</tr>
<tr>
<td>( \mathcal{P}(W) )</td>
<td>( \cup )</td>
<td>( { } )</td>
<td>( W )</td>
<td>( \subseteq )</td>
<td>( \supseteq )</td>
<td></td>
</tr>
<tr>
<td>( \mathcal{P}(W) )</td>
<td>( \cap )</td>
<td>( W )</td>
<td>( { } )</td>
<td>( \subseteq )</td>
<td>( \supseteq )</td>
<td></td>
</tr>
</tbody>
</table>
Lemma

Let $D \in \{R, L\}$.

1. $IP(S, \bullet) \iff RX(S, \preceq^D)$
2. $CM(S, \bullet) \implies AY(S, \preceq^D)$
3. $AS(S, \bullet) \implies TR(S, \preceq^D)$
4. $CM(S, \bullet) \implies (SL(S, \bullet) \iff TO(S, \preceq^D))$

Proof.

1. $a \preceq^D a \iff a = a \cdot a$
2. $a \preceq^L b \land b \preceq^L a \iff a = a \cdot b \land b = b \cdot a \implies a = b$
3. $a \preceq^L b \land b \preceq^L c \iff a = a \cdot b \land b = b \cdot c \implies a = a \cdot (b \cdot c) = (a \cdot b) \cdot c = a \cdot c \implies a \preceq^L c$
4. $a = a \cdot b \lor b = a \cdot b \iff a \preceq^L b \lor b \preceq^L a$
Bounds

Suppose \((S, \leq)\) is a partially ordered set.

**greatest lower bound**

For \(a, b \in S\), the element \(c \in S\) is the greatest lower bound of \(a\) and \(b\), written \(c = a \land b\), if it is a lower bound (\(c \leq a\) and \(c \leq b\)), and for every \(d \in S\) with \(d \leq a\) and \(d \leq b\), we have \(d \leq c\).

**least upper bound**

For \(a, b \in S\), the element \(c \in S\) is the least upper bound of \(a\) and \(b\), written \(c = a \lor b\), if it is an upper bound (\(a \leq c\) and \(b \leq c\)), and for every \(d \in S\) with \(a \leq d\) and \(b \leq d\), we have \(c \leq d\).
Semi-lattices

Suppose \((S, \leq)\) is a partially ordered set.

**meet-semilattice**

S is a **meet-semilattice** if \(\text{glb } a, b\) exists for each \(a, b \in S\).

**join-semilattice**

S is a **join-semilattice** if \(\text{lub } a, b\) exists for each \(a, b \in S\).
Fun Facts

**Fact 1**
Suppose \((S, \bullet)\) is a commutative and idempotent semigroup.
- \((S, \leq^L)\) is a meet-semilattice with \(a \mathsf{glb} b = a \bullet b\).
- \((S, \leq^R)\) is a join-semilattice with \(a \mathsf{lub} b = a \bullet b\).

**Fact 2**
Suppose \((S, \leq)\) is a partially ordered set.
- If \((S, \leq)\) is a meet-semilattice, then \((S, \mathsf{glb})\) is a commutative and idempotent semigroup.
- If \((S, \leq)\) is a join-semilattice, then \((S, \mathsf{lub})\) is a commutative and idempotent semigroup.

That is, semi-lattices represent the same class of structures as commutative and idempotent semigroups.
Semirings
Matrix semirings
Shortest paths
Minimax
## Bi-semigroups and Pre-Semirings

$(S, \oplus, \otimes)$ is a **bi-semigroup** when

- $(S, \oplus)$ is a semigroup
- $(S, \otimes)$ is a semigroup

$(S, \oplus, \otimes)$ is a **pre-semiring** when

- $(S, \oplus, \otimes)$ is a bi-semigroup
- $\oplus$ is commutative

and left- and right-distributivity hold,

\[
\begin{align*}
LD : \quad a \otimes (b \oplus c) &= (a \otimes b) \oplus (a \otimes c) \\
RD : \quad (a \oplus b) \otimes c &= (a \otimes c) \oplus (b \otimes c)
\end{align*}
\]
Semirings

\((S, \oplus, \otimes, 0, 1)\) is a semiring when

- \((S, \oplus, \otimes)\) is a pre-semiring
- \((S, \oplus, 0)\) is a (commutative) monoid
- \((S, \otimes, 1)\) is a monoid
- \(0\) is an annihilator for \(\otimes\)
### Examples

#### Pre-semirings

<table>
<thead>
<tr>
<th>name</th>
<th>$S$</th>
<th>$\oplus$</th>
<th>$\otimes$</th>
<th>$0$</th>
<th>$1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>min_plus</td>
<td>$\mathbb{N}$</td>
<td>min</td>
<td>+</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>max_min</td>
<td>$\mathbb{N}$</td>
<td>max</td>
<td>min</td>
<td>0</td>
<td></td>
</tr>
</tbody>
</table>

#### Semirings

<table>
<thead>
<tr>
<th>name</th>
<th>$S$</th>
<th>$\oplus$</th>
<th>$\otimes$</th>
<th>$0$</th>
<th>$1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>sp</td>
<td>$\mathbb{N}^\infty$</td>
<td>min</td>
<td>$\infty$</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>bw</td>
<td>$\mathbb{N}^\infty$</td>
<td>max</td>
<td>min</td>
<td>0</td>
<td>$\infty$</td>
</tr>
</tbody>
</table>

Note the sloppiness — the symbols $+$, max, and min in the two tables represent different functions....
How about $(\max, +)$?

<table>
<thead>
<tr>
<th>Pre-semiring</th>
<th>name</th>
<th>$S$</th>
<th>$\oplus$, $\otimes$</th>
<th>$\overline{0}$, $\overline{1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>max_plus</td>
<td>$\mathbb{N}$</td>
<td>max</td>
<td>+</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

- What about "$\overline{0}$ is an annihilator for $\otimes$"? No!

Fix that ...

<table>
<thead>
<tr>
<th>Pre-semiring</th>
<th>name</th>
<th>$S$</th>
<th>$\oplus$, $\otimes$</th>
<th>$\overline{0}$, $\overline{1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>max_plus</td>
<td>$\mathbb{N} \cup {-\infty}$</td>
<td>max</td>
<td>+ $-\infty$</td>
<td>0</td>
</tr>
</tbody>
</table>
Matrix Semirings

- \((S, \oplus, \otimes, \overline{0}, \overline{1})\) a semiring
- Define the semiring of \(n \times n\)-matrices over \(S\): \((\mathbb{M}_n(S), \oplus, \otimes, J, I)\)

\(\oplus\) and \(\otimes\)

\[
(A \oplus B)(i, j) = A(i, j) \oplus B(i, j)
\]

\[
(A \otimes B)(i, j) = \bigoplus_{1 \leq q \leq n} A(i, q) \otimes B(q, j)
\]

\(J\) and \(I\)

\[
J(i, j) = \overline{0}
\]

\[
I(i, j) = \begin{cases} 
\overline{1} & \text{(if } i = j) \\
\overline{0} & \text{(otherwise)}
\end{cases}
\]
Associativity

\[ A \otimes (B \otimes C) = (A \otimes B) \otimes C \]

\[
(A \otimes (B \otimes C))(i, j) \\
= \bigoplus_{1 \leq u \leq n} A(i, u) \otimes (B \otimes C)(u, j) \quad \text{(def } \rightarrow ) \\
= \bigoplus_{1 \leq u \leq n} \bigoplus_{1 \leq v \leq n} A(i, u) \otimes (B(u, v) \otimes C(v, j)) \quad \text{(LD)} \\
= \bigoplus_{1 \leq v \leq n} \bigoplus_{1 \leq u \leq n} (A(i, u) \otimes B(u, v)) \otimes C(v, j) \quad \text{(AS, CM)} \\
= \bigoplus_{1 \leq v \leq n} (A \otimes B)(i, v) \otimes C(v, j) \quad \text{(def } \leftarrow ) \\
= ((A \otimes B) \otimes C)(i, j) \quad \text{(def } \leftarrow )
\]
Left Distributivity

\[ A \otimes (B \oplus C) = (A \otimes B) \oplus (A \otimes C) \]

\[
\begin{align*}
(A \otimes (B \oplus C))(i, j) \\
= & \bigoplus_{1 \leq q \leq n} A(i, q) \otimes (B \oplus C)(q, j) & \text{(def \rightarrow)} \\
= & \bigoplus_{1 \leq q \leq n} A(i, q) \otimes (B(q, j) \oplus C(q, j)) & \text{(def \rightarrow)} \\
= & \bigoplus_{1 \leq q \leq n} (A(i, q) \otimes B(q, j)) \oplus (A(i, q) \otimes C(q, j)) & \text{(LD)} \\
= & (\bigoplus_{1 \leq q \leq n} A(i, q) \otimes B(q, j)) \oplus (\bigoplus_{1 \leq q \leq n} A(i, q) \otimes C(q, j)) & \text{(AS, CM)} \\
= & ((A \otimes B) \oplus (A \otimes C))(i, j) & \text{(def \leftarrow)}
\end{align*}
\]
Matrix encoding path problems

- $(S, \oplus, \otimes, 0, 1)$ a semiring
- $G = (V, E)$ a directed graph
- $w \in E \rightarrow S$ a weight function

Path weight

The weight of a path $p = i_1, i_2, i_3, \ldots, i_k$ is

$$w(p) = w(i_1, i_2) \otimes w(i_2, i_3) \otimes \cdots \otimes w(i_{k-1}, i_k).$$

The empty path is given the weight $\overline{1}$.

Adjacency matrix $A$

$$A(i, j) = \begin{cases} 
    w(i, j) & \text{if } (i, j) \in E, \\
    0 & \text{otherwise}
\end{cases}$$
The general problem of finding globally optimal path weights

Given an adjacency matrix $A$, find $A^*$ such that for all $i, j \in V$

$$A^*(i, j) = \bigoplus_{p \in \pi(i, j)} w(p)$$

where $\pi(i, j)$ represents the set of all paths from $i$ to $j$.

How can we solve this problem?
Stability

- \((S, \oplus, \otimes, 0, 1)\) a semiring

\(a \in S\), define powers \(a^k\)

\[
\begin{align*}
    a^0 &= 1 \\
    a^{k+1} &= a \otimes a^k
\end{align*}
\]

Closure, \(a^*\)

\[
\begin{align*}
    a^{(k)} &= a^0 \oplus a^1 \oplus a^2 \oplus \cdots \oplus a^k \\
    a^* &= a^0 \oplus a^1 \oplus a^2 \oplus \cdots \oplus a^k \oplus \cdots
\end{align*}
\]

Definition (q stability)

If there exists a \(q\) such that \(a^{(q)} = a^{(q+1)}\), then \(a\) is \(q\)-stable. By induction: \(\forall t, 0 \leq t, a^{(q+t)} = a^{(q)}\). Therefore, \(a^* = a^{(q)}\).
Matrix methods

Matrix powers, $A^k$

$$A^0 = I$$
$$A^{k+1} = A \otimes A^k$$

Closure, $A^*$

$$A^{(k)} = I \oplus A^1 \oplus A^2 \oplus \cdots \oplus A^k$$
$$A^* = I \oplus A^1 \oplus A^2 \oplus \cdots \oplus A^k \oplus \cdots$$

Note: $A^*$ might not exist. Why?
Matrix methods can compute optimal path weights

- Let $\pi(i, j)$ be the set of paths from $i$ to $j$.
- Let $\pi^k(i, j)$ be the set of paths from $i$ to $j$ with exactly $k$ arcs.
- Let $\pi^{(k)}(i, j)$ be the set of paths from $i$ to $j$ with at most $k$ arcs.

Theorem

\begin{align*}
(1) \quad A^k(i, j) &= \bigoplus_{p \in \pi^k(i, j)} w(p) \\
(2) \quad A^{(k)}(i, j) &= \bigoplus_{p \in \pi^{(k)}(i, j)} w(p) \\
(3) \quad A^*(i, j) &= \bigoplus_{p \in \pi(i, j)} w(p)
\end{align*}

Warning again: for some semirings the expression $A^*(i, j)$ might not be well-defined. Why?
Proof of (1)

By induction on $k$. Base Case: $k = 0$.

$$\pi^0(i, i) = \{\epsilon\},$$

so $A^0(i, i) = I(i, i) = 1 = w(\epsilon)$.

And $i \neq j$ implies $\pi^0(i, j) = \emptyset$. By convention

$$\bigoplus_{p \in \emptyset} w(p) = 0 = I(i, j).$$
Proof of (1)

Induction step.

\[ A^{k+1}(i, j) = (A \otimes A^k)(i, j) \]

\[ = \bigoplus_{1 \leq q \leq n} A(i, q) \otimes A^k(q, j) \]

\[ = \bigoplus_{1 \leq q \leq n} A(i, q) \otimes \left( \bigoplus_{p \in \pi^k(q, j)} w(p) \right) \]

\[ = \bigoplus_{1 \leq q \leq n} \bigoplus_{p \in \pi^k(q, j)} A(i, q) \otimes w(p) \]

\[ = \bigoplus_{(i, q) \in E} \bigoplus_{p \in \pi^k(q, j)} w(i, q) \otimes w(p) \]

\[ = \bigoplus_{p \in \pi^{k+1}(i, j)} w(p) \]
Fun Facts

Fact 3
If $\overline{1}$ is an annihilator for $\oplus$, then every $a \in S$ is 0-stable!

Fact 4
If $S$ is 0-stable, then $\mathbb{M}_n(S)$ is $(n - 1)$-stable. That is,

$$A^* = A^{(n-1)} = I \oplus A^1 \oplus A^2 \oplus \cdots \oplus A^{n-1}$$

Why? Because we can ignore paths with loops.

$$(a \otimes c \otimes b) \oplus (a \otimes b) = a \otimes (\overline{1} \oplus c) \otimes b = a \otimes \overline{1} \otimes b = a \otimes b$$

Think of $c$ as the weight of a loop in a path with weight $a \otimes b$. 
Shortest paths example, \((\mathbb{N}^\infty, \min, +)\)

The adjacency matrix

\[
A = \begin{bmatrix}
0 & 1 & 2 & 3 & 4 \\
0 & \infty & 2 & 1 & 6 & \infty \\
1 & 2 & \infty & 5 & \infty & 4 \\
2 & 1 & 5 & \infty & 4 & 3 \\
3 & 6 & \infty & 4 & \infty & \infty \\
4 & \infty & 4 & 3 & \infty & \infty
\end{bmatrix}
\]

Note that the longest shortest path is \((1, 0, 2, 3)\) of length 3 and weight 7.
(min, +) example

Our theorem tells us that $A^* = A^{(n-1)} = A^{(4)}$

$$A^* = A^{(4)} = I \min A \min A^2 \min A^3 \min A^4 = \begin{bmatrix}
0 & 1 & 2 & 3 & 4 \\
0 & 2 & 1 & 5 & 4 \\
2 & 0 & 3 & 7 & 4 \\
1 & 3 & 0 & 4 & 3 \\
5 & 7 & 4 & 0 & 7 \\
4 & 4 & 3 & 7 & 0 \\
\end{bmatrix}$$
(min, +) example

\[
A = \begin{bmatrix}
0 & 1 & 2 & 3 & 4 \\
0 & \infty & 2 & 1 & 6 & \infty \\
1 & 2 & \infty & 5 & \infty & 4 \\
2 & 1 & 5 & \infty & 4 & 3 \\
3 & 6 & \infty & 4 & \infty & \infty \\
4 & \infty & 4 & 3 & \infty & \infty
\end{bmatrix}
\]

\[
A^2 = \begin{bmatrix}
0 & 1 & 2 & 3 & 4 \\
0 & 2 & 6 & 7 & \underline{5} & 4 \\
1 & 6 & 4 & 3 & 8 & 8 \\
2 & 7 & \underline{3} & 2 & 7 & 9 \\
3 & \underline{5} & 8 & 7 & 8 & 7 \\
4 & 4 & 8 & 9 & 7 & 6
\end{bmatrix}
\]

\[
A^3 = \begin{bmatrix}
0 & 1 & 2 & 3 & 4 \\
0 & 8 & 4 & 3 & 8 & 10 \\
1 & 4 & 8 & 7 & \underline{7} & 6 \\
2 & 3 & 7 & 8 & 6 & 5 \\
3 & 8 & 7 & 6 & 11 & 10 \\
4 & 10 & 6 & 5 & 10 & 12
\end{bmatrix}
\]

\[
A^4 = \begin{bmatrix}
0 & 1 & 2 & 3 & 4 \\
0 & 4 & 8 & 9 & 7 & 6 \\
1 & 8 & 6 & 5 & 10 & 10 \\
2 & 9 & 5 & 4 & 9 & 11 \\
3 & 7 & 10 & 9 & 10 & 9 \\
4 & 6 & 10 & 11 & 9 & 8
\end{bmatrix}
\]

First appearance of final value is in red and underlined. Remember: we are looking at all paths of a given length, even those with cycles!
Lemma

If $\oplus$ is idempotent, then

$$(A \oplus I)^k = A^{(k)}.$$  

Proof. Base case: When $k = 0$ both expressions are $I$.
Assume $(A \oplus I)^k = A^{(k)}$. Then

$$(A \oplus I)^{k+1} = (A \oplus I)(A \oplus I)^k$$
$$= (A \oplus I)A^{(k)}$$
$$= AA^{(k)} \oplus A^{(k)}$$
$$= A(I \oplus A \oplus \cdots \oplus A^k) \oplus A^{(k)}$$
$$= A \oplus A^2 \oplus \cdots \oplus A^{k+1} \oplus A^{(k)}$$
$$= A^{k+1} \oplus A^{(k)}$$
$$= A^{(k+1)}$$
back to \((\text{min, +})\) example

\[
(A \oplus I)^1 = \begin{bmatrix}
0 & 1 & 2 & 3 & 4 \\
0 & 2 & 1 & 6 & \infty \\
2 & 5 & 0 & 4 & 3 \\
6 & \infty & 4 & 0 & \infty \\
\infty & 4 & 3 & \infty & 0
\end{bmatrix}
\]

\[
(A \oplus I)^2 = \begin{bmatrix}
0 & 1 & 2 & 3 & 4 \\
0 & 2 & 1 & 5 & 4 \\
2 & 3 & 0 & 4 & 3 \\
5 & 8 & 4 & 0 & 7 \\
4 & 4 & 3 & 7 & 0
\end{bmatrix}
\]

\[
(A \oplus I)^3 = \begin{bmatrix}
0 & 1 & 2 & 3 & 4 \\
0 & 2 & 1 & 5 & 4 \\
2 & 3 & 0 & 4 & 3 \\
5 & 7 & 4 & 0 & 7 \\
4 & 4 & 3 & 7 & 0
\end{bmatrix}
\]
Semigroup properties (so far)

\[
\begin{align*}
\text{AS}(S, \bullet) & \iff \forall a, b, c \in S, \ a \bullet (b \bullet c) = (a \bullet b) \bullet c \\
\text{IID}(S, \bullet, \alpha) & \iff \forall a \in S, \ a = \alpha \bullet a = a \bullet \alpha \\
\text{ID}(S, \bullet) & \iff \exists \alpha \in S, \ \text{IID}(S, \bullet, \alpha) \\
\text{IAN}(S, \bullet, \omega) & \iff \forall a \in S, \ \omega = \omega \bullet a = a \bullet \omega \\
\text{AN}(S, \bullet) & \iff \exists \omega \in S, \ \text{IAN}(S, \bullet, \omega) \\
\text{CM}(S, \bullet) & \iff \forall a, b \in S, \ a \bullet b = b \bullet a \\
\text{SL}(S, \bullet) & \iff \forall a, b \in S, \ a \bullet b \in \{a, b\} \\
\text{IP}(S, \bullet) & \iff \forall a \in S, \ a \bullet a = a \\
\text{IR}(S, \bullet) & \iff \forall s, t \in S, \ s \bullet t = t \\
\text{IL}(S, \bullet) & \iff \forall s, t \in S, \ s \bullet t = s
\end{align*}
\]

Recall that is right (IR) and is left (IL) are forced on us by wanting an \(\Leftrightarrow\)-rule for \(\text{SL}((S, \bullet) \times (T, \diamond))\).
### Bisemigroup properties (so far)

<table>
<thead>
<tr>
<th>Property</th>
<th>Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>AAS(S, ⊕, ⊗)</td>
<td>$AAS(S, \oplus) \equiv AS(S, \oplus)$</td>
</tr>
<tr>
<td>AID(S, ⊕, ⊗)</td>
<td>$AID(S, \oplus, \otimes) \equiv ID(S, \oplus)$</td>
</tr>
<tr>
<td>ACM(S, ⊕, ⊗)</td>
<td>$ACM(S, \oplus, \otimes) \equiv CM(S, \oplus)$</td>
</tr>
<tr>
<td>MAS(S, ⊕, ⊗)</td>
<td>$MAS(S, \oplus, \otimes) \equiv AS(S, \otimes)$</td>
</tr>
<tr>
<td>MID(S, ⊕, ⊗)</td>
<td>$MID(S, \oplus, \otimes) \equiv ID(S, \otimes)$</td>
</tr>
<tr>
<td>LD(S, ⊕, ⊗)</td>
<td>$LD(S, \oplus, \otimes) \equiv \forall a, b, c \in S, a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c)$</td>
</tr>
<tr>
<td>RD(S, ⊕, ⊗)</td>
<td>$RD(S, \oplus, \otimes) \equiv \forall a, b, c \in S, (a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c)$</td>
</tr>
<tr>
<td>ZA(S, ⊕, ⊗)</td>
<td>$ZA(S, \oplus, \otimes) \equiv \exists \bar{0} \in S, IID(S, \oplus, \bar{0}) \land IAN(S, \otimes, \bar{0})$</td>
</tr>
<tr>
<td>OA(S, ⊕, ⊗)</td>
<td>$OA(S, \oplus, \otimes) \equiv \exists \bar{1} \in S, IID(S, \otimes, \bar{1}) \land IAN(S, \oplus, \bar{1})$</td>
</tr>
<tr>
<td>ASL(S, ⊕, ⊗)</td>
<td>$ASL(S, \oplus, \otimes) \equiv SL(S, \oplus)$</td>
</tr>
<tr>
<td>AIP(S, ⊕, ⊗)</td>
<td>$AIP(S, \oplus, \otimes) \equiv IP(S, \oplus)$</td>
</tr>
</tbody>
</table>
A Minimax Semiring

\[
\text{minimax} \equiv (\mathbb{N}^{\infty}, \min, \max, \infty, 0)
\]

\[
17 \min \infty = 17
\]

\[
17 \max \infty = \infty
\]

How can we interpret this?

\[
A^*(i, j) = \min_{p \in \pi(i, j)} \max_{(u, v) \in p} A(u, v),
\]
One possible interpretation of Minimax

- Given an adjacency matrix $A$ over minimax,
- suppose that $A(i, j) = 0 \iff i = j$,
- suppose that $A$ is symmetric ($A(i, j) = A(j, i)$),
- interpret $A(i, j)$ as measured dissimilarity of $i$ and $j$,
- interpret $A^*(i, j)$ as inferred dissimilarity of $i$ and $j$,

Many uses

- Hierarchical clustering of large data sets
- Classification in Machine Learning
- Computational phylogenetics
- ...
Dendrograms from *Hierarchical Clustering With Prototypes via Minimax Linkage*, Bien and Tibshirani, 2011.
A minimax graph
The solution \textbf{A*} drawn as a dendrogram
Hierarchical clustering? Why?

Suppose \((Y, \leq, +)\) is a totally ordered with least element 0.

**Metric**

A **metric** for set \(X\) over \((Y, \leq, +)\) is a function \(d \in X \times X \rightarrow Y\) such that

- \(\forall x, y \in X, \ d(x, y) = 0 \iff x = y\)
- \(\forall x, y \in X, \ d(x, y) = d(y, x)\)
- \(\forall x, y, z \in X, \ d(x, y) \leq d(x, z) + d(z, y)\)

**Ultrametric**

An **ultrametric** for set \(X\) over \((Y, \leq)\) is a function \(d \in X \times X \rightarrow Y\) such that

- \(\forall x \in X, \ d(x, x) = 0\)
- \(\forall x, y \in X, \ d(x, y) = d(y, x)\)
- \(\forall x, y, z \in X, \ d(x, y) \leq d(x, z) \max d(z, y)\)
Fun Facts

Fact 5
If $A$ is an $n \times n$ symmetric minimax adjacency matrix, then $A^*$ is a finite ultrametric for $\{0, 1, \ldots, n-1\}$ over $(\mathbb{N}^\infty, \leq)$.

Fact 6
Suppose each arc weight is unique. Then the set of arcs
\[
\{(i, j) \in E \mid A(i, j) = A^*(i, j)\}
\]
is a minimum spanning tree.
A spanning tree derived from $A$ and $A^*$