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

The adjacency matrix

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

\[
\begin{bmatrix}
1 & 2 & 3 & 4 & 5 \\
1 & 0 & 2 & 1 & 5 & 4 \\
2 & 2 & 0 & 3 & 7 & 4 \\
3 & 1 & 3 & 0 & 4 & 3 \\
4 & 5 & 7 & 4 & 0 & 7 \\
5 & 4 & 4 & 3 & 7 & 0
\end{bmatrix}
\]

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, \( bw = (\mathbb{N}^\infty, \max, \min, 0, \infty) \)

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

solves 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, $\langle a, b, c \rangle$, $\cup$, $\cap$, $\emptyset$, $\{a, b, c\}$

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, $\langle a, b, c \rangle$, $\cap$, $\cup$

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 $x$.

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

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

**Recommended (on reserve in CL library)**
Semiring axioms ...

We will look at all of the axioms of semirings, but the most important are

**Distributivity**

- **LD**: \( a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c) \)
- **RD**: \( (a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c) \)

Distributivity, illustrated

\[
\begin{align*}
a \otimes (b \oplus c) &= (a \otimes b) \oplus (a \otimes c) \\
j \text{ makes the choice} &= i \text{ makes the choice}
\end{align*}
\]
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

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 \((3, 10)\).

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

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 \((3, 10)\)

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

Paths computed by (extended) **Dijkstra**

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

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

Paths computed by **distributed Bellman-Ford**

\[
\begin{align*}
\mathbf{P}_{\text{Bellman}}(0, 2) &= \{(0, 1, 4, 2)\} \\
\mathbf{P}_{\text{Bellman}}(2, 0) &= \{(2, 1, 0), (2, 4, 1, 0)\}
\end{align*}
\]
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

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

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! \((a \min b = a \min c)\) 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.
Semigroups

A **semigroup** \((S, \bullet)\) is a non-empty set \(S\) with a binary operation such that

\[
\text{AS associative } \equiv \forall a, b, c \in S, \ a \bullet (b \bullet c) = (a \bullet b) \bullet c
\]

**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, \((\mathbb{R}, \cdot)\), where \(a \cdot b \equiv (a + b)/2\).
Some Important Semigroup Properties

<table>
<thead>
<tr>
<th>ID</th>
<th>identity</th>
<th>( \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>( \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>( \forall a, b \in S, a \cdot b = b \cdot a )</td>
</tr>
<tr>
<td>SL</td>
<td>selective</td>
<td>( \forall a, b \in S, a \cdot b \in {a, b} )</td>
</tr>
<tr>
<td>IP</td>
<td>idempotent</td>
<td>( \forall a \in S, a \cdot a = a )</td>
</tr>
</tbody>
</table>

A semigroup with an identity is called a **monoid**.

Note that

\[ SL(S, \cdot) \implies IP(S, \cdot) \]

A few concrete semigroups

<table>
<thead>
<tr>
<th></th>
<th>( \cdot )</th>
<th>description</th>
<th>( \alpha )</th>
<th>( \omega )</th>
<th>CM</th>
<th>SL</th>
<th>IP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>S</td>
<td>left y = x</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>S</td>
<td>right y = y</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>S*</td>
<td>concatenation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>S+</td>
<td>concatenation</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>{t, f}</td>
<td>conjunction</td>
<td>t</td>
<td>f</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>{t, f}</td>
<td>disjunction</td>
<td>f</td>
<td>t</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>min</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>max</td>
<td></td>
<td>0</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>2(W)</td>
<td>union</td>
<td></td>
<td>{}</td>
<td></td>
<td>W</td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>2(W)</td>
<td>intersection</td>
<td>{}</td>
<td>W</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>fin(2(U))</td>
<td>union</td>
<td></td>
<td>{}</td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>fin(2(U))</td>
<td>intersection</td>
<td>{}</td>
<td></td>
<td></td>
<td>*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>addition</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>*</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>multiplication</td>
<td>1</td>
<td>0</td>
<td></td>
<td></td>
<td>*</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>S</th>
<th>•</th>
<th>description</th>
<th>α</th>
<th>ω</th>
<th>CM</th>
<th>SL</th>
<th>IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>(2^U)</td>
<td></td>
<td>union</td>
<td>{}</td>
<td>(U)</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>(2^U\times U)</td>
<td></td>
<td>intersection</td>
<td>{}</td>
<td>(U)</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>(X \times X)</td>
<td></td>
<td>relational join</td>
<td>{}</td>
<td>(U)</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>(X \rightarrow X)</td>
<td></td>
<td>composition</td>
<td>{}</td>
<td>(U)</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>

\(U\) an infinite set

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

\(I_U \equiv \{ (u, u) \mid u \in U \} \)

**subsemigroup**

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

**Order Relations**

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

**Definition (Important Order Properties)**

- **RX** (reflexive) \(\equiv a \leq a\)
- **TR** (transitive) \(\equiv a \leq b \land b \leq c \rightarrow a \leq c\)
- **AY** (antisymmetric) \(\equiv a \leq b \land b \leq a \rightarrow a = b\)
- **TO** (total) \(\equiv a \leq b \lor b \leq a\)

<table>
<thead>
<tr>
<th></th>
<th>partial</th>
<th>preference</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>RX</strong></td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td><strong>TR</strong></td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td><strong>AY</strong></td>
<td>*</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TO</strong></td>
<td></td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
Canonical Pre-order of a Commutative Semigroup

**Definition (Canonical pre-orders)**

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

**Lemma (Sanity check)**

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

**Proof.**

Note that \(a \preceq^R b\) means \(\exists c_1 \in S : b = a \cdot c_1\), and \(b \preceq^R c\) 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 \(a \preceq^R c\). The proof for \(\preceq^L\) is similar.

Canonically Ordered Semigroup

**Definition (Canonically Ordered Semigroup)**

A commutative semigroup \((S, \cdot)\) is canonically ordered when \(a \preceq^R c\) and \(a \preceq^L c\) 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 \(a \cdot a^{-1} = a^{-1} \cdot a = \alpha\).
Canonically Ordered Semigroups vs. Groups

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 \( a \triangleleft^L b \). In a similar way, \( b \triangleleft^R a \). Therefore \( a = b \).

Natural Orders

Definition (Natural orders)

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

\[
\begin{align*}
   a \triangleleft^L b & \iff a = a \cdot b \\
   a \triangleleft^R b & \iff b = a \cdot b
\end{align*}
\]

Lemma

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

Proof.

\[
\begin{align*}
   a \triangleleft^R b & \iff b = a \cdot c = (a \cdot a) \cdot c = a \cdot (a \cdot c) \\
   & = a \cdot b \iff a \triangleleft^R b \\
   a \triangleleft^L b & \iff a = b \cdot c = (b \cdot b) \cdot c = b \cdot (b \cdot c) \\
   & = b \cdot a = a \cdot b \iff a \triangleleft^L b
\end{align*}
\]
Special elements and natural orders

Lemma (Natural Bounds)

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

\[
\omega \leq^L a \leq^L \alpha \\
\alpha \leq^R a \leq^R \omega
\]

Remark (Thanks to Iljitsch van Beijnum)

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

\[
0 \leq^L \min a \leq^L \min \infty \leq^L \infty \\
\infty \leq^R \min a \leq^R \min 0
\]

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

Examples of special elements

<table>
<thead>
<tr>
<th>$S$</th>
<th>$\cdot$</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>
</tr>
<tr>
<td>$\mathbb{N}^{-\infty}$</td>
<td>$\max$</td>
<td>$0$</td>
<td>$-\infty$</td>
<td>$\geq$</td>
<td>$\leq$</td>
</tr>
<tr>
<td>$\mathcal{P}(W)$</td>
<td>$\cup$</td>
<td>$\emptyset$</td>
<td>$W$</td>
<td>$\subseteq$</td>
<td>$\supseteq$</td>
</tr>
<tr>
<td>$\mathcal{P}(W)$</td>
<td>$\cap$</td>
<td>$W$</td>
<td>$\emptyset$</td>
<td>$\subseteq$</td>
<td>$\supseteq$</td>
</tr>
</tbody>
</table>


Property Management

Lemma

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

1. $\text{IP}(S, \cdot) \iff \text{RX}(S, \leq^D)\$
2. $\text{CM}(S, \cdot) \implies \text{AY}(S, \leq^D)\$
3. $\text{AS}(S, \cdot) \implies \text{TR}(S, \leq^D)\$
4. $\text{CM}(S, \cdot) \implies (\text{SL}(S, \cdot) \iff \text{TO}(S, \leq^D))$

Proof.

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

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 $c = a \text{ glb } b$, 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 $c = a \text{ lub } b$, 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 \(a \text{ glb } b\) exists for each \(a, b \in S\).

**Join-semilattice**

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

Fun Facts

**Fact 1**

Suppose \((S, \cdot)\) is a commutative and idempotent semigroup.

- \((S, \leq^L)\) is a meet-semilattice with \(a \text{ glb } b = a \cdot b\).
- \((S, \leq^R)\) is a join-semilattice with \(a \text{ lub } b = a \cdot b\).

**Fact 2**

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

- If \((S, \leq)\) is a meet-semilattice, then \((S, \text{ glb})\) is a commutative and idempotent semigroup.
- If \((S, \leq)\) is a join-semilattice, then \((S, \text{ lub})\) is a commutative and idempotent semigroup.

That is, semi-lattices represent the same class of structures as commutative and idempotent semigroups.
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,

\[
\begin{align*}
\text{LD} : \quad a \otimes (b \oplus c) &= (a \otimes b) \oplus (a \otimes c) \\
\text{RD} : \quad (a \oplus b) \otimes c &= (a \otimes c) \oplus (b \otimes c)
\end{align*}
\]
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>+</td>
<td>$\infty$</td>
<td>0</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 \((\max, +)\)?

### Pre-semiring

<table>
<thead>
<tr>
<th>name</th>
<th>( S )</th>
<th>( \oplus )</th>
<th>( \otimes )</th>
<th>( \bar{0} )</th>
<th>( \bar{1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>max_plus</td>
<td>( \mathbb{N} )</td>
<td>( \max )</td>
<td>+</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

- What about “\( \bar{0} \) is an annihilator for \( \otimes \)”? No!

### Fix that ...

<table>
<thead>
<tr>
<th>name</th>
<th>( S )</th>
<th>( \oplus )</th>
<th>( \otimes )</th>
<th>( \bar{0} )</th>
<th>( \bar{1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>max_plus</td>
<td>( \mathbb{N} \cup { -\infty } )</td>
<td>( \max )</td>
<td>+</td>
<td>(-\infty)</td>
<td>0</td>
</tr>
</tbody>
</table>

---

### Matrix Semirings

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

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

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

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

#### \(\mathbf{J}\) and \(\mathbf{I}\)

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

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

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

\[
(A \otimes (B \otimes C))(i, j) = \bigoplus_{1 \leq u \leq n} A(i, u) \otimes (B \otimes C)(u, j) \quad \text{(def \rightarrow)}
\]

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

\[
= \bigoplus_{1 \leq u \leq n} \bigoplus_{1 \leq v \leq n} (A(i, u) \otimes B(u, v)) \otimes C(v, j) \quad \text{(LD)}
\]

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

\[
= \bigoplus_{1 \leq v \leq n} (A \otimes B)(i, v) \otimes C(v, j) \quad \text{(def \leftarrow)}
\]

\[
= ((A \otimes B) \otimes C)(i, j) \quad \text{(def \leftarrow)}
\]

**Left Distributivity**

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

\[
(A \otimes (B \oplus C))(i, j) = \bigoplus_{1 \leq q \leq n} A(i, q) \otimes (B \oplus C)(q, j) \quad \text{(def \rightarrow)}
\]

\[
= \bigoplus_{1 \leq q \leq n} A(i, q) \otimes (B(q, j) \oplus C(q, j)) \quad \text{(def \rightarrow)}
\]

\[
= \bigoplus_{1 \leq q \leq n} (A(i, q) \otimes B(q, j)) \oplus (A(i, q) \otimes C(q, j)) \quad \text{(LD)}
\]

\[
= \bigoplus_{1 \leq q \leq n} (A(i, q) \otimes B(q, j)) \oplus (A(i, q) \otimes C(q, j)) \quad \text{(AS, CM)}
\]

\[
= ((A \otimes B) \oplus (A \otimes C))(i, j) \quad \text{(def \leftarrow)}
\]
Matrix encoding path problems

- \((S, \oplus, \otimes, \bar{0}, \overline{1})\) 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 \(A\)

\[
A(i, j) = \begin{cases} 
  w(i, j) & \text{if } (i, j) \in E, \\
  \bar{0} & \text{otherwise}
\end{cases}
\]

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

- \((S, \oplus, \otimes, \bar{0}, \bar{T})\) a semiring

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

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

Closure, \(a^*\)

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

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, \(a^* = a^{(q)}\).

Note: \(A^*\) might not exist. Why?

Matrix methods

Matrix powers, \(A^k\)

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

Closure, \(A^*\)

\[
\begin{align*}
A^{(k)} &= I \oplus A^1 \oplus A^2 \oplus \ldots \oplus A^k \\
A^* &= I \oplus A^1 \oplus A^2 \oplus \ldots \oplus A^k \oplus \ldots
\end{align*}
\]
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**

1. $A^k(i, j) = \bigoplus_{p \in \pi^k(i, j)} w(p)$
2. $A^{(k)}(i, j) = \bigoplus_{p \in \pi^{(k)}(i, j)} w(p)
3. $A^*(i, j) = \bigoplus_{p \in \pi(i, j)} w(p)$

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.

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

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

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

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

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

Fun Facts

Fact 3
If \( \bar{T} \) 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,

\[
A^* = A^{(n-1)} = I \oplus A^1 \oplus A^2 \oplus \ldots \oplus A^{n-1}
\]

Why? Because we can ignore paths with loops.

\[
(a \otimes c \otimes b) \oplus (a \otimes b) = a \otimes (\bar{T} \oplus c) \otimes b = a \otimes \bar{T} \otimes b = a \otimes b
\]

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

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

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

(min, +) example

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

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

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

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

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

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

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!

A vs A ⊕ I

**Lemma**

If ⊕ is idempotent, then

$$(A ⊕ I)^k = A^{(k)}.$$ 

**Proof.** Base case: When $k = 0$ both expressions are I.

Assume $(A ⊕ I)^k = A^{(k)}$. Then

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

\[
\begin{pmatrix}
\infty & 0 & 2 & 1 & 6 \\
0 & \infty & 4 & 0 & 3 \\
2 & 1 & 0 & 5 & 4 \\
3 & 6 & \infty & 4 & 0 \\
4 & \infty & 4 & 3 & \infty
\end{pmatrix}^1 = \begin{pmatrix}
0 & 1 & 2 & 3 & 4 \\
0 & 2 & 1 & 5 & 4 \\
2 & 1 & 3 & 0 & 4 \\
3 & 5 & 7 & 4 & 0 \\
4 & 4 & 3 & 7 & 0
\end{pmatrix}
\]

\[
\begin{pmatrix}
\infty & 0 & 2 & 1 & 5 \\
0 & \infty & 4 & 0 & 7 \\
1 & 2 & 0 & 3 & 8 \\
2 & 1 & 3 & 0 & 4 \\
3 & 5 & 8 & 4 & 0
\end{pmatrix}^2 = \begin{pmatrix}
0 & 1 & 2 & 3 & 4 \\
0 & 2 & 1 & 5 & 4 \\
2 & 1 & 3 & 0 & 4 \\
3 & 5 & 7 & 4 & 0 \\
4 & 4 & 3 & 7 & 0
\end{pmatrix}
\]

Semigroup properties (so far)

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

Recall that right (IR) and is left (IL) are forced on us by wanting an \(\Leftrightarrow\)-rule for \SL($(S, \cdot) \times (T, \circ)$).
Bisemigroup properties (so far)

\[
\begin{align*}
AAS(S, \oplus, \otimes) &\equiv AS(S, \oplus) \\
AID(S, \oplus, \otimes) &\equiv ID(S, \oplus) \\
ACM(S, \oplus, \otimes) &\equiv CM(S, \oplus) \\
MAS(S, \oplus, \otimes) &\equiv AS(S, \otimes) \\
MID(S, \oplus, \otimes) &\equiv ID(S, \otimes) \\
LD(S, \oplus, \otimes) &\equiv \forall a, b, c \in S, \ a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c) \\
RD(S, \oplus, \otimes) &\equiv \forall a, b, c \in S, \ (a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c) \\
ZA(S, \oplus, \otimes) &\equiv \exists 0 \in S, \ \text{IID}(S, \oplus, 0) \land \text{IAN}(S, \otimes, 0) \\
OA(S, \oplus, \otimes) &\equiv \exists 1 \in S, \ \text{IID}(S, \oplus, 1) \land \text{IAN}(S, \otimes, 1) \\
ASL(S, \oplus, \otimes) &\equiv SL(S, \oplus) \\
AIP(S, \oplus, \otimes) &\equiv IP(S, \oplus)
\end{align*}
\]

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 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 \(d \in X \times X \to Y\) such that

1. \(\forall x, y \in X, \ d(x, y) = 0 \iff x = y\)
2. \(\forall x, y \in X, \ d(x, y) = d(y, x)\)
3. \(\forall x, y, z \in X, \ d(x, y) \leq d(x, z) + d(z, y)\)

**Ultrametric**

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

1. \(\forall x \in X, \ d(x, x) = 0\)
2. \(\forall x, y \in X, \ d(x, y) = d(y, x)\)
3. \(\forall x, y, z \in X, \ d(x, y) \leq \max(d(x, z), d(z, y))\)

**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 $A^*$