

L11: Algebraic Path Problems with applications to Internet Routing

Lecture 3

Timothy G. Griffin

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

Michaelmas Term, 2018

Lecture 3

- Semirings
- Matrix semirings
- Matrix encoding of a path problem
- \mathbf{A}^* solves a path problem
- Computing \mathbf{A}^*
- \mathbf{A}^* as a solution to certain matrix equations

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,

$$\begin{array}{lcl} \text{LD} & : & a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c) \\ \text{RD} & : & (a \oplus b) \otimes c = (a \otimes c) \oplus (b \otimes c) \end{array}$$

Semirings

$(S, \oplus, \otimes, \bar{0}, \bar{1})$ is a **semiring** when

- (S, \oplus, \otimes) is a pre-semiring
- $(S, \oplus, \bar{0})$ is a (commutative) monoid
- $(S, \otimes, \bar{1})$ is a monoid
- $\bar{0}$ is an annihilator for \otimes

Examples

Pre-semirings

name	S	$\oplus,$	\otimes	$\bar{0}$	$\bar{1}$
min_plus	\mathbb{N}	min	+	0	
max_min	\mathbb{N}	max	min	0	

Semirings

name	S	$\oplus,$	\otimes	$\bar{0}$	$\bar{1}$
sp	\mathbb{N}^∞	min	+	∞	0
bw	\mathbb{N}^∞	max	min	0	∞

Note the sloppiness — the symbols +, max, and min in the two tables represent different functions....

How about (max, +)?

Pre-semiring

name	S	$\oplus,$	\otimes	$\bar{0}$	$\bar{1}$
max_plus	\mathbb{N}	max	+	0	0

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

Fix that ...

name	S	\oplus	\otimes	$\bar{0}$	$\bar{1}$
max_plus $^{-\infty}$	$\mathbb{N} \uplus \{-\infty\}$	max	+	$-\infty$	0

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

$$(\mathbf{A} \oplus \mathbf{B})(i, j) = \mathbf{A}(i, j) \oplus \mathbf{B}(i, j)$$

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

\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

$$\mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}) = (\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C}$$

$$\begin{aligned} & (\mathbf{A} \otimes (\mathbf{B} \otimes \mathbf{C}))(i, j) \\ &= \bigoplus_{1 \leq u \leq n} \mathbf{A}(i, u) \otimes (\mathbf{B} \otimes \mathbf{C})(u, j) && (\text{def } \rightarrow) \\ &= \bigoplus_{1 \leq u \leq n} \mathbf{A}(i, u) \otimes \left(\bigoplus_{1 \leq v \leq n} \mathbf{B}(u, v) \otimes \mathbf{C}(v, j) \right) && (\text{def } \rightarrow) \\ &= \bigoplus_{1 \leq u \leq n} \bigoplus_{1 \leq v \leq n} \mathbf{A}(i, u) \otimes (\mathbf{B}(u, v) \otimes \mathbf{C}(v, j)) && (\text{LID}) \\ &= \bigoplus_{1 \leq v \leq n} \bigoplus_{1 \leq u \leq n} (\mathbf{A}(i, u) \otimes \mathbf{B}(u, v)) \otimes \mathbf{C}(v, j) && (\text{AS, CM}) \\ &= \bigoplus_{1 \leq v \leq n} \left(\bigoplus_{1 \leq u \leq n} \mathbf{A}(i, u) \otimes \mathbf{B}(u, v) \right) \otimes \mathbf{C}(v, j) && (\text{RID}) \\ &= \bigoplus_{1 \leq v \leq n} (\mathbf{A} \otimes \mathbf{B})(i, v) \otimes \mathbf{C}(v, j) && (\text{def } \leftarrow) \\ &= ((\mathbf{A} \otimes \mathbf{B}) \otimes \mathbf{C})(i, j) && (\text{def } \leftarrow) \end{aligned}$$

Left Distributivity

$$\mathbf{A} \otimes (\mathbf{B} \oplus \mathbf{C}) = (\mathbf{A} \otimes \mathbf{B}) \oplus (\mathbf{A} \otimes \mathbf{C})$$

$$\begin{aligned} & (\mathbf{A} \otimes (\mathbf{B} \oplus \mathbf{C}))(i, j) \\ = & \bigoplus_{1 \leq q \leq n} \mathbf{A}(i, q) \otimes (\mathbf{B} \oplus \mathbf{C})(q, j) & (\text{def } \rightarrow) \\ = & \bigoplus_{1 \leq q \leq n} \mathbf{A}(i, q) \otimes (\mathbf{B}(q, j) \oplus \mathbf{C}(q, j)) & (\text{def } \rightarrow) \\ = & \bigoplus_{1 \leq q \leq n} (\mathbf{A}(i, q) \otimes \mathbf{B}(q, j)) \oplus (\mathbf{A}(i, q) \otimes \mathbf{C}(q, j)) & (\text{LD}) \\ = & \left(\bigoplus_{1 \leq q \leq n} \mathbf{A}(i, q) \otimes \mathbf{B}(q, j) \right) \oplus \left(\bigoplus_{1 \leq q \leq n} \mathbf{A}(i, q) \otimes \mathbf{C}(q, j) \right) & (\text{AS, CM}) \\ = & ((\mathbf{A} \otimes \mathbf{B}) \oplus (\mathbf{A} \otimes \mathbf{C}))(i, j) & (\text{def } \leftarrow) \end{aligned}$$

Matrix encoding path problems

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

Path weight

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

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

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

Adjacency matrix \mathbf{A}

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

The general problem of finding globally optimal path weights

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

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

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

How can we solve this problem?

Stability

- $(S, \oplus, \otimes, \bar{0}, \bar{1})$ a semiring

$a \in S$, define powers a^k

$$\begin{aligned} a^0 &= \bar{1} \\ a^{k+1} &= a \otimes a^k \end{aligned}$$

Closure, a^*

$$\begin{aligned} a^{(k)} &= a^0 \oplus a^1 \oplus a^2 \oplus \cdots \oplus a^k \\ a^* &= a^0 \oplus a^1 \oplus a^2 \oplus \cdots \oplus a^k \oplus \cdots \end{aligned}$$

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, \mathbf{A}^k

$$\mathbf{A}^0 = \mathbf{I}$$

$$\mathbf{A}^{k+1} = \mathbf{A} \otimes \mathbf{A}^k$$

Closure, \mathbf{A}^*

$$\mathbf{A}^{(k)} = \mathbf{I} \oplus \mathbf{A}^1 \oplus \mathbf{A}^2 \oplus \dots \oplus \mathbf{A}^k$$

$$\mathbf{A}^* = \mathbf{I} \oplus \mathbf{A}^1 \oplus \mathbf{A}^2 \oplus \dots \oplus \mathbf{A}^k \oplus \dots$$

Note: \mathbf{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

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

Warning again: for some semirings the expression $\mathbf{A}^*(i, j)$ might not be well-defined. Why?

Proof of (1)

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

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

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

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

$$\bigoplus_{p \in \{\}} w(p) = \bar{0} = \mathbf{I}(i, j).$$

Proof of (1)

Induction step.

$$\begin{aligned} \mathbf{A}^{k+1}(i, j) &= (\mathbf{A} \otimes \mathbf{A}^k)(i, j) \\ &= \bigoplus_{1 \leq q \leq n} \mathbf{A}(i, q) \otimes \mathbf{A}^k(q, j) \\ &= \bigoplus_{1 \leq q \leq n} \mathbf{A}(i, q) \otimes \left(\bigoplus_{p \in \pi^k(q, j)} w(p) \right) \\ &= \bigoplus_{1 \leq q \leq n} \bigoplus_{p \in \pi^k(q, j)} \mathbf{A}(i, q) \otimes w(p) \\ &= \bigoplus_{(i, q) \in E} \bigoplus_{p \in \pi^k(q, j)} w(i, q) \otimes w(p) \\ &= \bigoplus_{p \in \pi^{k+1}(i, j)} w(p) \end{aligned}$$

Fun Facts

Fact 3

If $\bar{1}$ 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,

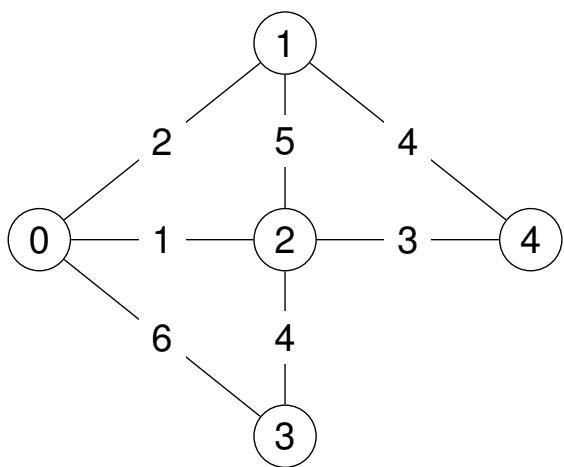
$$\mathbf{A}^* = \mathbf{A}^{(n-1)} = \mathbf{I} \oplus \mathbf{A}^1 \oplus \mathbf{A}^2 \oplus \cdots \oplus \mathbf{A}^{n-1}$$

Why? Because we can ignore paths with loops.

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

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

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

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

(min, +) example

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

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

(min, +) example

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

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

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 \oplus I

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

$$\begin{aligned} (A \oplus I)^{k+1} &= (A \oplus I)(A \oplus I)^k \\ &= (A \oplus I)A^{(k)} \\ &= AA^{(k)} \oplus A^{(k)} \\ &= A(I \oplus A \oplus \dots \oplus A^k) \oplus A^{(k)} \\ &= A \oplus A^2 \oplus \dots \oplus A^{k+1} \oplus A^{(k)} \\ &= A^{k+1} \oplus A^{(k)} \\ &= A^{(k+1)} \end{aligned}$$

back to (min, +) example

$$(A \oplus I)^1 = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 1 & 6 & \infty \\ 1 & 2 & 0 & 5 & \infty & 4 \\ 2 & 1 & 5 & 0 & 4 & 3 \\ 3 & 6 & \infty & 4 & 0 & \infty \\ 4 & \infty & 4 & 3 & \infty & 0 \end{bmatrix} \quad (A \oplus I)^3 = \begin{bmatrix} 0 & 1 & 2 & 3 & 4 \\ 0 & 0 & 2 & 1 & 5 & 4 \\ 1 & 2 & 0 & 3 & 7 & 4 \\ 2 & 1 & 3 & 0 & 4 & 3 \\ 3 & 5 & 7 & 4 & 0 & 7 \\ 4 & 4 & 4 & 3 & 7 & 0 \end{bmatrix}$$

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

Solving (some) equations

Theorem 6.1

If \mathbf{A} is q -stable, then \mathbf{A}^* solves the equations

$$\mathbf{L} = \mathbf{AL} \oplus \mathbf{I}$$

and

$$\mathbf{R} = \mathbf{RA} \oplus \mathbf{I}.$$

For example, to show $\mathbf{L} = \mathbf{A}^*$ solves the first equation:

$$\begin{aligned}\mathbf{A}^* &= \mathbf{A}^{(q)} \\ &= \mathbf{A}^{(q+1)} \\ &= \mathbf{A}^{q+1} \oplus \mathbf{A}^q \oplus \dots \oplus \mathbf{A}^2 \oplus \mathbf{A} \oplus \mathbf{I} \\ &= \mathbf{A}(\mathbf{A}^q \oplus \mathbf{A}^{q-1} \oplus \dots \oplus \mathbf{A} \oplus \mathbf{I}) \oplus \mathbf{I} \\ &= \mathbf{AA}^{(q)} \oplus \mathbf{I} \\ &= \mathbf{AA}^* \oplus \mathbf{I}\end{aligned}$$

Note that if we replace the assumption “ \mathbf{A} is q -stable” with “ \mathbf{A}^* exists,” then we require that \otimes distributes over infinite sums.

A more general result

Theorem Left-Right

If \mathbf{A} is q -stable, then $\mathbf{L} = \mathbf{A}^* \mathbf{B}$ solves the equation

$$\mathbf{L} = \mathbf{AL} \oplus \mathbf{B}$$

and $\mathbf{R} = \mathbf{BA}^*$ solves

$$\mathbf{R} = \mathbf{RA} \oplus \mathbf{B}.$$

For the first equation:

$$\begin{aligned}\mathbf{A}^* \mathbf{B} &= \mathbf{A}^{(q)} \mathbf{B} \\ &= \mathbf{A}^{(q+1)} \mathbf{B} \\ &= (\mathbf{A}^{q+1} \oplus \mathbf{A}^q \oplus \dots \oplus \mathbf{A}^2 \oplus \mathbf{A} \oplus \mathbf{I}) \mathbf{B} \\ &= (\mathbf{A}^{q+1} \oplus \mathbf{A}^q \oplus \dots \oplus \mathbf{A}^2 \oplus \mathbf{A}) \mathbf{B} \oplus \mathbf{B} \\ &= \mathbf{A}(\mathbf{A}^q \oplus \mathbf{A}^{q-1} \oplus \dots \oplus \mathbf{A} \oplus \mathbf{I}) \mathbf{B} \oplus \mathbf{B} \\ &= \mathbf{A}(\mathbf{A}^{(q)} \mathbf{B}) \oplus \mathbf{B} \\ &= \mathbf{A}(\mathbf{A}^* \mathbf{B}) \oplus \mathbf{B}\end{aligned}$$

The “best” solution

Suppose \mathbf{Y} is a matrix such that

$$\mathbf{Y} = \mathbf{AY} \oplus \mathbf{I}$$

$$\begin{aligned}\mathbf{Y} &= \mathbf{AY} \oplus \mathbf{I} \\ &= \mathbf{A}^1 \mathbf{Y} \oplus \mathbf{A}^{(0)} \\ &= \mathbf{A}((\mathbf{AY} \oplus \mathbf{I})) \oplus \mathbf{I} \\ &= \mathbf{A}^2 \mathbf{Y} \oplus \mathbf{A} \oplus \mathbf{I} \\ &= \mathbf{A}^2 \mathbf{Y} \oplus \mathbf{A}^{(1)} \\ &\vdots \quad \vdots \quad \vdots \\ &= \mathbf{A}^{k+1} \mathbf{Y} \oplus \mathbf{A}^{(k)}\end{aligned}$$

If \mathbf{A} is q -stable and $q < k$, then

$$\mathbf{Y} = \mathbf{A}^k \mathbf{Y} \oplus \mathbf{A}^*$$

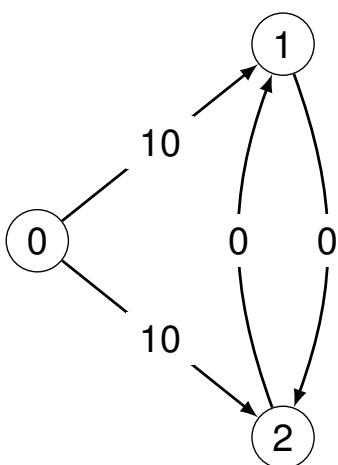
$$\mathbf{Y} \trianglelefteq_{\oplus}^L \mathbf{A}^*$$

and if \oplus is idempotent, then

$$\mathbf{Y} \leq_{\oplus}^L \mathbf{A}^*$$

So \mathbf{A}^* is the largest solution. What does this mean in terms of the sp semiring?

Example with zero weighted cycles using sp semiring



\mathbf{A}^* ($= \mathbf{A} \oplus \mathbf{I}$ in this case) solves

$$\mathbf{X} = \mathbf{XA} \oplus \mathbf{I}.$$

But so does this (dishonest) matrix!

$$\mathbf{F} = \begin{bmatrix} 0 & 1 & 2 \\ 0 & 9 & 9 \\ 1 & 0 & 0 \\ 2 & 0 & 0 \end{bmatrix}$$

For example :

$$\begin{aligned}\mathbf{A} &= \begin{bmatrix} 0 & 1 & 2 \\ 0 & \infty & 10 & 10 \\ 1 & \infty & \infty & 0 \\ 2 & \infty & 0 & \infty \end{bmatrix} & & \begin{aligned} & (\mathbf{FA} \oplus \mathbf{I})(0, 1) \\ &= \min_{q \in \{0, 1, 2\}} \mathbf{F}(0, q) + \mathbf{A}(q, 1) \\ &= \min(0 + 10, 9 + \infty, 9 + 0) \\ &= 9 \\ &= \mathbf{F}(0, 1)\end{aligned}\end{aligned}$$