Simon Wachter: Assignment 3

siwachte@ethz.ch, 19-920-198

05/01/2023 - 09:44h

Question 1:

a) We show that under Definition 1, the following holds:

$$\begin{aligned}
a^* &\stackrel{!}{=} 1 \oplus a \otimes a^* && (1)\\
&= 1 \oplus a \otimes \bigotimes_{n=0}^{\infty} a^{\otimes n} && (2)\\
&= 1 \oplus \bigotimes_{n=0}^{\infty} a^{\otimes n+1} && (3)\\
&= 1 \oplus \bigotimes_{n=1}^{\infty} a^{\otimes n} && (4)\\
&= \bigotimes_{n=0}^{\infty} a^{\otimes n} && (5)\\
&= a^* && (6)
\end{aligned}$$

b) First we show that $\log a \oplus a = \log(2) + a$:

$$\begin{aligned}
a \oplus \log a &= \log(e^a + e^a) && (7)\\
&= \log(2e^a) && (8)\\
&= \log(2) + a && (9)
\end{aligned}$$

Then we calculate the Kleene start:

$$\begin{aligned}
\bigoplus_{\log\ n=0}^{\infty} a^{\oplus n} &= a^{\oplus 0} \oplus_{\log} \left(\bigoplus_{\log\ n=1}^{\infty} a^{\oplus n} \right) && (10)\\
&= 0 \oplus_{\log} \left(\bigoplus_{\log\ n=1}^{\infty} a^{\oplus n} \right) && (11)\\
&= 0 \oplus_{\log} a \oplus_{\log} 2a \oplus_{\log} 3a \oplus_{\log} \cdots && (12)\\
&= \log\left(e^0 + e^a\right) \oplus_{\log} 2a \oplus_{\log} 3a \oplus_{\log} \cdots && (13)\\
&= \log\left(e^{\log(e^0+e^a)} + e^{2a}\right) \oplus_{\log} 3a \oplus_{\log} \cdots && (14)\\
&= \log\left(e^0 + e^a + e^{2a}\right) \oplus_{\log} 3a \oplus_{\log} \cdots && (15)\\
&= \log\left(\sum_{n=0}^{\infty} e^{a^{\oplus n}}\right) && (16)
\end{aligned}$$

We have two cases here, either $a > 0$ or $a \leq 0$. In the first case, the sum diverges and we get ∞. In the second case, the sum converges:

$$\sum_{n=0}^{\infty} e^{a^{\oplus n}} = \frac{1}{1-e^a} \qquad \text{limit geometric series} \tag{17}$$

Therefore, we get:

$$\begin{aligned}
\log\left(\sum_{n=0}^{\infty} e^{a^{\oplus n}}\right) &= \log\left(\frac{1}{1-e^a}\right) && (18)\\
&= \log(1) - \log(1-e^a) && (19)\\
&= \log(1-e^a) && (20)
\end{aligned}$$

c) First we derive a closed form solution for $(x, y)^{\oplus n}$:

$$\begin{aligned}
\langle x, y\rangle^{\oplus 1} &= \langle x, y\rangle && (21)\\
\langle x, y\rangle^{\oplus 2} &= \langle x^2, 2xy\rangle && (22)\\
\langle x, y\rangle^{\oplus 3} &= \langle x^3, 3x^2 y\rangle && (23)\\
\langle x, y\rangle^{\oplus 4} &= \langle x^4, 4x^3 y\rangle && (24)\\
\langle x, y\rangle^{\oplus n} &= \langle x^n, nx^{n-1} y\rangle && (25)
\end{aligned}$$

Then we derive a closed form for a^*:

$$\begin{aligned}
\bigoplus_{n=0}^{\infty} \langle x, y\rangle^{\oplus n} &= \bigoplus_{n=0}^{\infty} \langle x^n, nx^{n-1} y\rangle && \text{eq. (25)} \quad (26)\\
&= \left\langle \sum_{n=0}^{\infty} x^n, \sum_{n=0}^{\infty} nx^{n-1} y \right\rangle && (27)
\end{aligned}$$

Both parts only converge for $|x| < 1$. The left parts is a geometric series and has limit $\frac{1}{1-x}$. And the right parts, which is a power series:

$$\begin{aligned}
\sum_{n=0}^{\infty} nx^{n-1} y &= y \sum_{n=0}^{\infty} nx^{n-1} && (28)\\
&= y\left(0 + \sum_{n=1}^{\infty} nx^{n-1}\right) && (29)\\
&= y \sum_{n=0}^{\infty} (n+1)x^n && (30)\\
&= y \frac{1}{(x-1)^2} && \text{limit power series} \quad (31)\\
&= \frac{y}{(x-1)^2} && (32)
\end{aligned}$$

Hence our closed form with inserted limits is given by:

$$\left\langle \sum_{n=0}^{\infty} x^n, \sum_{n=0}^{\infty} nx^{n-1} y \right\rangle = \left\langle \frac{1}{1-x}, \frac{y}{(x-1)^2} \right\rangle \qquad |x| \leq 1 \tag{33}$$

If $|x| > 1$ both parts diverge.

d)

$$\mathcal{W}_{\text{lang}} = \left\langle 2^{\Sigma^*}, \bigcup, \otimes, \{\}, \{\epsilon\} \right\rangle \tag{34}$$

We first show that $\mathcal{W}_{\text{lang}}$ is a semiring:

- $(2^{\Sigma^*}, \oplus, \mathbf{0})$ must be a commutative monoid with identity element $\mathbf{0}$:

$$\begin{align} (x \oplus y) \oplus z &= (x \cup y) \cup z \tag{35}\\ &= \{x, y\} \cup z \tag{36}\\ &= \{x, y, z\} \tag{37}\\ &= x \oplus \{y, z\} \tag{38}\\ &= x \oplus (y \oplus z) \tag{39} \end{align}$$

$$\begin{align} \mathbf{0} \oplus x &= \{\} \oplus x \tag{40}\\ &= \{\} \cup x \tag{41}\\ &= \oplus \tag{42} \end{align}$$

$$\begin{align} x \oplus y &= x \cup y \tag{43}\\ &= \{x, y\} \tag{44}\\ &= y \cup x \tag{45}\\ &= y + x \tag{46} \end{align}$$

- $(2^{\Sigma^*}, \otimes, \mathbf{1})$ must be a monoid with identity element $\mathbf{1}$:

$$\begin{align} (x \otimes y) \otimes z &= xy \otimes z \tag{47}\\ &= xyz \tag{48}\\ &= x \otimes yz \tag{49}\\ &= x \otimes (y \otimes z) \tag{50} \end{align}$$

$$\begin{align} \mathbf{1} \otimes x &= \{\epsilon\} \otimes x \tag{51}\\ &= x \tag{52}\\ &= x \otimes \{\epsilon\} \tag{53}\\ &= x \otimes \mathbf{1} \tag{54} \end{align}$$

- Multiplication left and right distributes over addition:

$$\begin{align} x \otimes (y \oplus z) &= x \otimes \{y, z\} \tag{55}\\ &= \{xy, xz\} \tag{56}\\ &= \{xy\} \cup \{xz\} \tag{57}\\ &= (x \otimes y) \oplus (x \otimes z) \tag{58} \end{align}$$

$$\begin{align} (x \oplus y) \otimes z &= \{x, y\} \otimes z \tag{59}\\ &= \{xz, yz\} \tag{60}\\ &= \{xz\} \cup \{yz\} \tag{61}\\ &= (x \otimes z) \oplus (y \otimes z) \tag{62} \end{align}$$

- Multiplication by $\mathbf{0}$ annihilates $\mathbb{R} \times \mathbb{R}$:

$$\begin{aligned} \mathbf{0} \otimes x &= \{a \circ b \,|\, a \in A, b \in \{\}\} && (63)\\ &= \{\} && \text{by definition of } \circ \text{, because no } b \text{ exists} \quad (64)\\ &= x \otimes \mathbf{0} && (65)\\ & && (66) \end{aligned}$$

The Kleene start for $\mathcal{W}_{\text{lang}}$ given by:

$$A^{\otimes n} = \{\bigotimes_{i=0}^{n} a_i \mid a_i \in A\} \tag{67}$$

$$\begin{aligned} A^* &= \bigoplus_{n=0}^{\infty} A^{\otimes n} && (68)\\ &= \bigoplus_{n=0}^{\infty} \left\{ \bigotimes_{i=0}^{n} a_i \mid a_i \in A \right\} && (69)\\ &= \left\{ \bigotimes_{i=0}^{n} a_i \mid a_i \in A, n \in \mathbb{Z} \right\} && (70) \end{aligned}$$

Question 2:

a) Tropical semiring is 0-closed:

$$\begin{aligned} a \oplus \mathbf{0} &= \min(a, \mathbf{0}) && (71)\\ &= \min(a, 0) && (72)\\ &= 0 && \text{because } a \in \mathbb{R}_{\geq 0} \quad (73) \end{aligned}$$

Arctic semiring is 0-closed:

$$\begin{aligned} a \oplus \mathbf{0} &= \max(a, \mathbf{0}) && (74)\\ &= \max(a, 0) && (75)\\ &= 0 && \text{because } a \in \mathbb{R}_{\leq 0} \quad (76) \end{aligned}$$

b) Proof by induction:
Base case $i = 1$:

$$\begin{aligned} M^1 &= M && (77)\\ M_{ij} &= w_{ij} && \text{by def of } M \quad (78) \end{aligned}$$

w_{ij} is exactly the semiring-sum over all paths from i to j of length 1. This holds because there is only one path of length 1 from i to j.
Induction hypothesis: M^i_{ij} is the semiring-sum over all paths from i to j of length i.
Induction step $i \to i+1$:

$$\begin{aligned} M^{i+1} &= M^i \otimes M && (79)\\ M^{i+1}_{kj} &= \sum_{l=0}^{n} M^i_{kl} \otimes M_{lj} && \text{def. matrix mult.} \quad (80) \end{aligned}$$

In eq. (80) we sum over the product of all possible paths of length i from k to another node l and all possible paths of length 1 from nodes l to j. This sum is exactly the semiring-sum over all possible paths of length $i+1$ from k to j.

c) If we have a graph G with N vertices, then a path with length $l \geq N$ must visit at least one vertex v twice. From this follows, that we can just remove this cycle at vertex v and reduce the path to a path of length at most $N-1$. If there are multiple cycles, we can remove all of them until we arrive at a path of length at most $N-1$. Let us now assume that the longest path in G has length $l \geq N-1$. The weight of this path is the following:

$$v_0 \overset{w_0}{\rightarrow} v_1 \rightarrow \cdots \rightarrow v_k \rightarrow \cdots \rightarrow v_k \rightarrow \ldots \overset{v_{l-2}}{\rightarrow} v_{l-1} \overset{w_{l-1}}{\rightarrow} v_l \tag{81}$$
$$w_0 \otimes \cdots \otimes w_{k-1} \otimes w_k \otimes \cdots \otimes w_k' \otimes w_{k+1} \otimes \cdots \otimes w_{l-1} \tag{82}$$

Notice the cycle in the middle, which we know exists given our reasoning before. We can remove this cycle and arrive at the following path with new weight:

$$v_0 \overset{w_0}{\rightarrow} v_1 \rightarrow \cdots \rightarrow v_k \rightarrow \ldots \overset{v_{l-2}}{\rightarrow} v_{l-1} \overset{w_{l-1}}{\rightarrow} v_l \tag{83}$$
$$w_0 \otimes \cdots \otimes w_{k-1} \otimes w_{k+1} \otimes \cdots \otimes w_{l-1} \tag{84}$$

We first define weights for subpaths:

$$s_0 = w_0 \otimes \cdots \otimes w_{k-1} \qquad \text{path to } k \tag{85}$$
$$s_1 = w_k \otimes \cdots \otimes w_k' \qquad \text{cycle at } k\text{-th vertex} \tag{86}$$
$$s_2 = w_{k+1} \otimes \cdots \otimes w_{l-1} \qquad \text{path from } k \text{ to } l \tag{87}$$
$$\tag{88}$$

We now rewright out path weights and add them over our semiring:

$$(s_0 \otimes s_1 \otimes s_2) \oplus (s_0 \otimes s_2) \tag{89}$$
$$(s_0) \otimes ((s_1 \otimes s_2) \oplus s_2) \tag{90}$$
$$(s_0) \otimes (s_2 \otimes (s_1 \oplus 1)) \tag{91}$$
$$(s_0) \otimes (s_2 \otimes 1) \qquad \text{def. 0-closed} \tag{92}$$
$$s_0 \otimes s_2 \qquad \text{def. } \mathbf{1} \tag{93}$$

We see that eq. (93) is exactly the weight of the path without the cycle. Since this holds for every path of length $l \geq N$ we can conclude that the longest path in G has length at most $N-1$.

d) Per definition of the Kleene start, we have:

$$M^* = \bigoplus_{i=0}^{\infty} M^{\otimes i} \tag{94}$$

In b) we have shown that $M^{\otimes i}$ is the semiring-sum over all paths of length i and in c) we have shown that the shortest path depends only on paths of length $l \leq N-1$. We also showed that under the $\bigoplus$ operation, only paths of length $l \leq N-1$ are considered. Therefore, we know that M^* depends only on:

$$M^* = \bigoplus_{i=0}^{N-1} M^{\otimes i} \tag{95}$$

e) We define a simple algorithm:

Algorithm 1: Matrix multiplication for Kleene star

M;
$M' \leftarrow M$;
$M^* \leftarrow \mathbf{0}$;
for $i = 0$ *to* $N-1$ **do**
    for $j = 0$ *to len(M)* **do**
        for $k = 0$ *to len(M)* **do**
            for $l = 0$ *to len(M)* **do**
                $M'_{j,k} \leftarrow M'_{j,k} \oplus (M_{j,l} \otimes M_{l,k})$
            $M^*_{j,k} \leftarrow M^*_{j,k} \oplus M'_{j,k}$
    $M \leftarrow M'$;

The inner for loops calculate the matrix multiplication, $M^{\otimes n}$, and the outer loop iterates $N-1$ times to calculate the Kleene star. Since each loop has $\mathcal{O}(N)$ iterations, the algorithm has a runtime of $\mathcal{O}(N^4)$.

f)

$$\begin{aligned} a \oplus a &= a \otimes (\mathbf{1} \oplus \mathbf{1}) && (96)\\ &= a \otimes \mathbf{1} && \text{def. 0-closed} \quad (97)\\ &= a && \text{def. } \mathbf{1} \quad (98) \end{aligned}$$

g) We show given equality with an induction proof: Base case $n = 0$:

$$\begin{aligned} \bigoplus_{i=0}^{0} M^i &= M^0 && (99)\\ &= (I \oplus M)^0 && (100) \end{aligned}$$

Where we assumed that:

$$I \oplus M = M \tag{101}$$

Our induction hypothesis is:

$$\bigoplus_{i=0}^{n} M^i = (I \oplus M)^i \tag{102}$$

Now for the inductive step we have $n \to n+1$:

$$\bigoplus_{i=0}^{n+1} M^i = \bigoplus_{i=0}^{n} M^i \oplus M^{n+1} \tag{103}$$

$$= M^0 \oplus \bigoplus_{i=0}^{n} M^i \oplus M^{n+1} \tag{104}$$

$$= M^0 \oplus \bigoplus_{i=0}^{n} \left(M^i \oplus M^i\right) \oplus M^{n+1} \qquad \text{(def. Idempotent)} \tag{105}$$

$$= \bigoplus_{i=0}^{n} M^i \oplus \bigoplus_{i=1}^{n+1} M^i \tag{106}$$

$$= \bigoplus_{i=0}^{n} M^i \otimes (\mathbf{I} \oplus M) \qquad \text{(def. distributive)} \tag{107}$$

$$= (\mathbf{I} \oplus M)^n \otimes (\mathbf{I} \oplus M) \qquad \text{(def. I.H.)} \tag{108}$$

$$= (\mathbf{I} \oplus M)^{n+1} \tag{109}$$

h) With the log factor, we immediately think about binary representation. We use the product of power rule to rewrite our left side of the equation:

$$\bigotimes_{k=0}^{\lfloor \log_2 n \rfloor} M^{\alpha_k 2^k} = M^{\sum_{k=0}^{\lfloor \log_2 n \rfloor} \alpha_k 2^k} \tag{110}$$

We now analyze the exponent of M more closely. If we choose α_k to represent the k-th bit in the binary representation of n, we can see that $\sum_{k=0}^{\lfloor \log_2 n \rfloor} \alpha_k 2^k = n$. Therefore, we can rewrite the equation as:

$$\bigotimes_{k=0}^{\lfloor \log_2 n \rfloor} M^{\alpha_k 2^k} = M^{\sum_{k=0}^{\lfloor \log_2 n \rfloor} \alpha_k 2^k} \tag{111}$$

$$= M^n \tag{112}$$

With this insight, we can rewrite the algorithm to calculate M^*:

Algorithm 2: Matrix multiplication for Kleene star

M;
$M' \leftarrow M$;
$M^* \leftarrow \mathbf{0}$;
for $i = 0$ *to* $\log_2(N-1)$ **do**
    if $a_i == 0$ **then**
        for $j = 0$ *to len(M)* **do**
            for $k = 0$ *to len(M)* **do**
                for $l = 0$ *to len(M)* **do**
                    $M'_{j,k} \leftarrow M'_{j,k} \oplus (M_{j,l} \otimes M_{l,k})$
                $M^*_{j,k} \leftarrow M^*_{j,k} \oplus M'_{j,k}$
        $M \leftarrow M'$;

As we have just remove some iterations from the outer most loop, our runtime changes to $\mathcal{O}(n^3 \log n)$

i) We derive the SVD of A:

$$
\begin{aligned}
A &= U\Sigma V^T && (113)\\
||A||_2 &= ||U\Sigma V^T||_2 = ||\Sigma||_2 && (114)
\end{aligned}
$$

Further we can rewrite the operator norm by w.l.o.g choosing x such that $||x||_2 = 1$:

$$
\begin{aligned}
||A||_2 &= \sup_{x \neq 0} \frac{||Ax||_2}{||x||_2} && (115)\\
&= \sup_{x \neq 0, ||x||_2 = 1} ||Ax||_2 && (116)\\
& && (117)
\end{aligned}
$$

Combining these two insight we get:

$$
\begin{aligned}
||A||_2 &= \sup_{x \neq 0, ||x||_2 = 1} ||\Sigma x||_2 && (118)\\
&= \sigma_{\max}(A) && \text{(min-max theorem)} \quad (119)
\end{aligned}
$$

j)

$$
\begin{aligned}
||A^* - \sum_{n=0}^{K} A^n||_2 &= ||\sum_{n=0}^{\infty} A^n - \sum_{n=0}^{K} A^n||_2 && (120)\\
&= ||\sum_{n=K+1}^{\infty} A^n||_2 && (121)\\
&= \sigma_{\max}\left(\sum_{n=K+1}^{\infty} A^n\right) && \text{(ex. i))} \quad (122)\\
&\leq \sum_{n=K+1}^{\infty} \sigma_{\max}\left(A^n\right) && \text{(singular value inequalities)} \quad (123)\\
&\leq \sum_{n=K+1}^{\infty} \sigma_{\max}\left(A\right)^n && \text{(singular value inequalities)} \quad (124)\\
&= \frac{\sigma_{\max}(A)^{K+1}}{1 - \sigma_{\max}(A)} && \text{(geom. series if } \sigma_{\max}(A) < 1\text{))} \quad (125)\\
& && (126)
\end{aligned}
$$

eq. (125) shows the closed form solution if $\sigma_{\max}(A) < 1$. If $\sigma_{\max}(A) \geq 1$, the closed form solution is not defined as the geometric series diverges.
The condition on $\sigma_{\max}(A)$ for the truncation error to converge to A^* therefore is $\sigma_{\max}(A) < 1$:

$$
\lim_{K \to \infty} \frac{\sigma_{\max}(A)^{K+1}}{1 - \sigma_{\max}(A)} = 0 \qquad (127)
$$

k) If $\sigma_{\max}(A) \geq 1$ the truncation error is unbounded and the truncation is not a good approximation to asteration.
In the case where $\sigma_{\max}(A) < 1$ the truncation error in $\mathcal{O}$ is given by:

$$\mathcal{O}\left(\frac{\sigma_{\max}(A)^{K+1}}{1-\sigma_{\max}(A)}\right) = \mathcal{O}\left(\sigma_{\max}(A)^{K}\right) \tag{128}$$

Hence the error decays exponentially, which is then generally deemed to be acceptable for an error and so a truncation is a good approximation to asteration.