Prof. Ryan Cotterell

Yannick Wattenberg: Assignment 02

ywattenberg@inf.ethz.ch, 19-947-464.

12/11/2022 - 09:44h

1 Question: Entropy of a Conditional Random Field

a)

Prove that the **expectation semiring** satisfies the semiring axioms:
Let $x_1, y_1, x_2, y_2, x_3, y_3 \in \mathbb{R}$ for this subsection.

Axiom 1

$$(\mathbb{R} \times \mathbb{R}, \oplus, \langle 0, 0 \rangle) \text{ is a commutative monoid with identity element } \langle 0, 0 \rangle$$

Associativity and Commutativity of $\oplus$:

$$\begin{aligned}
(\langle x_1, y_1 \rangle \oplus \langle x_2, y_2 \rangle) \oplus \langle x_3, y_3 \rangle &\stackrel{!}{=} \langle x_1, y_1 \rangle \oplus (\langle x_3, y_3 \rangle \oplus \langle x_2, y_2 \rangle) && (1)\\
(\langle x_1, y_1 \rangle \oplus \langle x_2, y_2 \rangle) \oplus \langle x_3, y_3 \rangle &= \langle x_1 + x_2, y_1 + y_2 \rangle \oplus \langle x_3, y_3 \rangle \quad \text{(def. } \oplus) && (2)\\
&= \langle (x_1 + x_2) + x_3, (y_1 + y_2) + y_3 \rangle \quad \text{(def. } \oplus) && (3)\\
&= \langle x_1 + (x_2 + x_3), y_1 + (y_2 + y_3) \rangle \quad \text{(ass. of } +) && (4)\\
&= \langle x_1, y_1 \rangle \oplus \langle x_2 + x_3, y_2 + y_3 \rangle \quad \text{(def. } \oplus) && (5)\\
&= \langle x_1, y_1 \rangle \oplus \langle x_3 + x_2, y_3 + y_2 \rangle \quad \text{(comm. of } +) && (6)\\
&= \langle x_1, y_1 \rangle \oplus (\langle x_3, y_3 \rangle \oplus \langle x_2, y_2 \rangle) \quad \text{(def. } \oplus) && (7)
\end{aligned}$$

$\langle 0, 0 \rangle$ is the identity element:

$$\begin{aligned}
\langle 0, 0 \rangle \oplus \langle x_1, y_1 \rangle &\stackrel{!}{=} \langle x_1, y_1 \rangle && (8)\\
\langle 0, 0 \rangle \oplus \langle x_1, y_1 \rangle &= \langle x_1 + 0, y_1 + 0 \rangle \quad \text{(def. } \oplus) && (9)\\
&= \langle x_1, y_1 \rangle \quad = \langle x_1 + 0, y_1 + 0 \rangle = \langle x_1, y_1 \rangle \oplus \langle 0, 0 \rangle \quad \text{(def. } \oplus) && (10)
\end{aligned}$$

Axiom 2

$$(\mathbb{R} \times \mathbb{R}, \otimes, \langle 1, 0\rangle) \text{ is a monoid with identity element } \langle 1, 0\rangle$$

Associativity of $\otimes$:

$$
\begin{aligned}
&(\langle x_1, y_1\rangle \otimes \langle x_2, y_2\rangle) \otimes \langle x_3, y_3\rangle \stackrel{!}{=} \langle x_1, y_1\rangle \otimes (\langle x_2, y_2\rangle \otimes \langle x_3, y_3\rangle) && && (11)\\
&(\langle x_1, y_1\rangle \otimes \langle x_2, y_2\rangle) \otimes \langle x_3, y_3\rangle = \langle x_1 \cdot x_2, x_1 \cdot y_2 + y_1 \cdot x_2\rangle \otimes \langle x_3, y_3\rangle && \text{(def. } \otimes) && (12)\\
=&\langle (x_1 \cdot x_2) \cdot x_3, (x_1 \cdot x_2) \cdot y_3 + (x_1 \cdot y_2 + y_1 \cdot x_2) \cdot x_3\rangle && \text{(def. } \otimes) && (13)\\
=&\langle (x_1 \cdot x_2) \cdot x_3, x_1 \cdot x_2 \cdot y_3 + x_1 \cdot y_2 \cdot x_3 + y_1 \cdot x_2 \cdot x_3\rangle && \text{(diss. } \cdot) && (14)\\
=&\langle (x_1 \cdot x_2) \cdot x_3, x_1 \cdot (x_2 \cdot y_3 + y_2 \cdot x_3) + y_1 \cdot x_2 \cdot x_3\rangle && \text{(diss. } \cdot) && (15)\\
=&\langle x_1 \cdot (x_2 \cdot x_3), x_1 \cdot (x_2 \cdot y_3 + y_2 \cdot x_3) + y_1 \cdot (x_2 \cdot x_3)\rangle && \text{(ass. } \cdot) && (16)\\
=&\langle x_1, y_1\rangle \otimes \langle x_2 \cdot x_3, x_2 \cdot y_3 + y_2 \cdot x_3\rangle && \text{(def. } \otimes) && (17)\\
=&\langle x_1, y_1\rangle \otimes (\langle x_2, y_2\rangle \otimes \langle x_3, y_3\rangle) && \text{(def. } \otimes) && (18)
\end{aligned}
$$

$\langle 1, 0\rangle$ is the identity element:

$$
\begin{aligned}
\langle x_1, y_1\rangle \otimes \langle 1, 0\rangle &\stackrel{!}{=} \langle x_1, y_1\rangle && && (19)\\
\langle x_1, y_1\rangle \otimes \langle 1, 0\rangle &= \langle x_1 \cdot 1, x_1 \cdot 0 + y_1 \cdot 1\rangle && \text{(def. } \otimes) && (20)\\
&= \langle x_1, y_1\rangle \qquad = \langle 1 \cdot x_1, 1 \cdot y_1 + 0 \cdot x_1\rangle && && (21)\\
&= \langle 1, 0\rangle \otimes \langle x_1, y_1\rangle && \text{(def. } \otimes) && (22)
\end{aligned}
$$

Axiom 3

$\otimes$ distributes left and right over $\oplus$:

$$\langle x_1, y_1\rangle \otimes (\langle x_2, y_2\rangle \oplus \langle x_3, y_3\rangle) \stackrel{!}{=} (\langle x_1, y_1\rangle \otimes \langle x_2, y_2\rangle) \oplus (\langle x_1, y_1\rangle \otimes \langle x_3, y_3\rangle) \tag{23}$$

$$
\begin{aligned}
&\langle x_1, y_1\rangle \otimes (\langle x_2, y_2\rangle \oplus \langle x_3, y_3\rangle) = \langle x_1, y_1\rangle \otimes \langle x_2 + x_3, y_2 + y_3\rangle && \text{(def. } \oplus) && (24)\\
=&\langle x_1 \cdot (x_2 + x_3), x_1 \cdot (x_2 + x_3) + y_1 \cdot (y_2 + y_3)\rangle && \text{(def. } \otimes) && (25)\\
=&\langle (x_1 \cdot x_2) + (x_1 \cdot x_3), (x_1 \cdot y_2) + (y_1 \cdot x_2) + (x_1 \cdot y_3) + (y_1 \cdot x_3)\rangle && \text{(diss. } \cdot) && (26)\\
=&\langle x_1 \cdot x_2, (x_1 \cdot y_2) + (y_1 \cdot x_2)\rangle \oplus \langle x_1 \cdot x_3, (x_1 \cdot y_3) + (y_1 \cdot x_3)\rangle && \text{(def. } \oplus) && (27)\\
=&(\langle x_1, y_1\rangle \otimes \langle x_2, y_2\rangle) \oplus (\langle x_1, y_1\rangle \otimes \langle x_3, y_3\rangle) && \text{(def. } \otimes) && (28)
\end{aligned}
$$

$$(\langle x_2, y_2\rangle \oplus \langle x_3, y_3\rangle) \otimes \langle x_1, y_1\rangle \stackrel{!}{=} (\langle x_2, y_2\rangle \otimes \langle x_1, y_1\rangle) \oplus (\langle x_3, y_3\rangle \otimes \langle x_1, y_1\rangle) \tag{29}$$

$$
\begin{aligned}
&(\langle x_2, y_2\rangle \oplus \langle x_3, y_3\rangle) \otimes \langle x_1, y_1\rangle = (\langle x_2 + x_3, y_2 + y_3\rangle) \otimes \langle x_1, y_1\rangle && \text{(def. } \oplus) && (30)\\
=&\langle (x_2 + x_3) \cdot x_1, (x_2 + x_3) \cdot y_1 + (y_2 + y_3) \cdot x_1\rangle && \text{(def. } \otimes) && (31)\\
=&\langle (x_2 \cdot x_1) \cdot (x_3 \cdot x_1), (x_2 \cdot y_1) + (x_3 \cdot y_1) + (y_2 \cdot x_1) + (y_3 \cdot x_1)\rangle && \text{(diss. } \cdot) && (32)\\
=&\langle x_2 \cdot x_1, x_2 \cdot y_1 + y_2 \cdot x_1\rangle \oplus \langle x_3 \cdot x_1, x_3 \cdot y_1 + y_3 \cdot x_1\rangle && \text{(def. } \oplus) && (33)\\
=&(\langle x_2, y_2\rangle \otimes \langle x_1, y_1\rangle) \oplus (\langle x_3, y_3\rangle \otimes \langle x_1, y_1\rangle) && \text{(def. } \otimes) && (34)
\end{aligned}
$$

Axiom 4

$$\langle 0,0\rangle \otimes \langle x_1, y_1\rangle \stackrel{!}{=} \langle 0,0\rangle \stackrel{!}{=} \langle x_1, y_1\rangle \otimes \langle 0,0\rangle \tag{35}$$

$$\langle 0,0\rangle \otimes \langle x_1, y_1\rangle = \langle 0 \cdot x_1, 0 \cdot y_1 + 0 \cdot x_1\rangle \qquad \text{(def. } \otimes) \tag{36}$$

$$= \langle 0,0\rangle \tag{37}$$

$$= \langle x_1 \cdot 0, x_1 \cdot 0 + y_1 \cdot 0\rangle \tag{38}$$

$$= \langle x_1, y_1\rangle \otimes \langle 0,0\rangle \qquad \text{(def. } \otimes) \tag{39}$$

With this we have proven that the **expectation semiring** is a semiring.

b)

The definition of the forward algorithm is as follows:

Algorithm 1: Forward algorithm

$\beta(w, t_0) = \nVdash$
for $i = 1$ **to** N **do**
    $v_t = \oplus_{t_i \in T} v_i \otimes$
end

c)