Rationalizing neural predictions

Andrej Švec a Tao Lei (MIT)

Review

the beer was n't what i expected, and i'm not sure it's "true to style", but i thought it was delicious. **a very pleasant ruby red-amber color** with a relatively brilliant finish, but a limited amount of carbonation, from the look of it. aroma is what i think an amber ale should be - a nice blend of caramel and happiness bound together.

Ratings

Look: 5 stars

Smell: 4 stars

this beer pours ridiculously clear with tons of carbonation that forms a rather impressive rocky head that settles slowly into a fairly dense layer of foam. this is a real good lookin' beer, unfortunately it gets worse from here ... first, the aroma is kind of bubblegum-like and grainy. next, the taste is sweet and grainy with an unpleasant bitterness in the finish. ... ... overall, the fat weasel is good for a fairly cheap buzz, but only if you like your beer grainy and bitter .

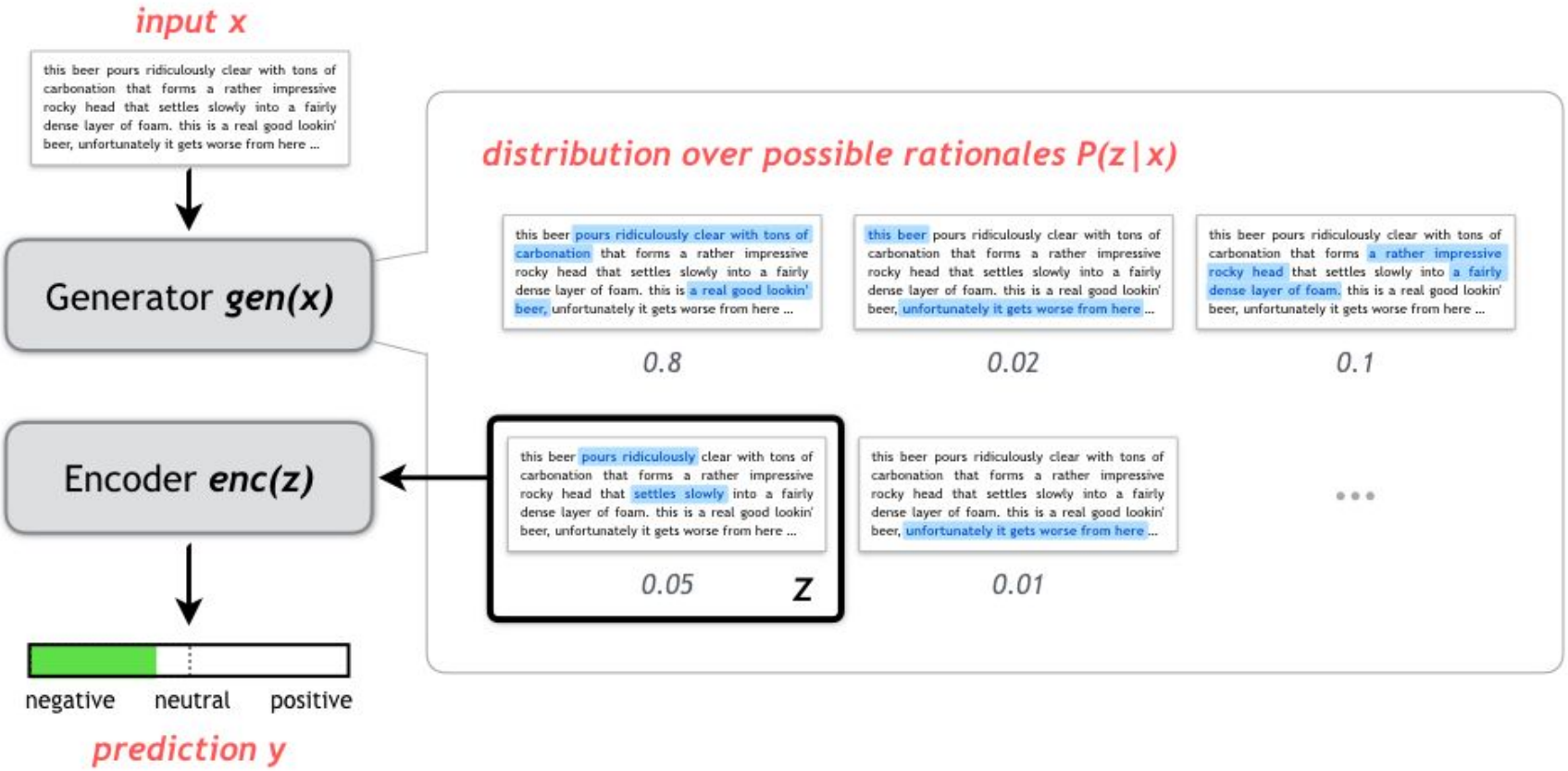
Ratings

Look: 5 stars

Aroma: 2 stars

multi-aspect sentiment analysis

Generator and encoder



Encoder

- input x (words) $\mathbf{x} = \{x_t\}_{t=1}^{l}$
- output y (sentiment) $\mathbf{y} \in [0, 1]^m$
- loss $\mathcal{L}(\mathbf{x}, \mathbf{y}) = \|\tilde{\mathbf{y}} - \mathbf{y}\|_2^2 = \|\mathbf{enc}(\mathbf{x}) - \mathbf{y}\|_2^2$

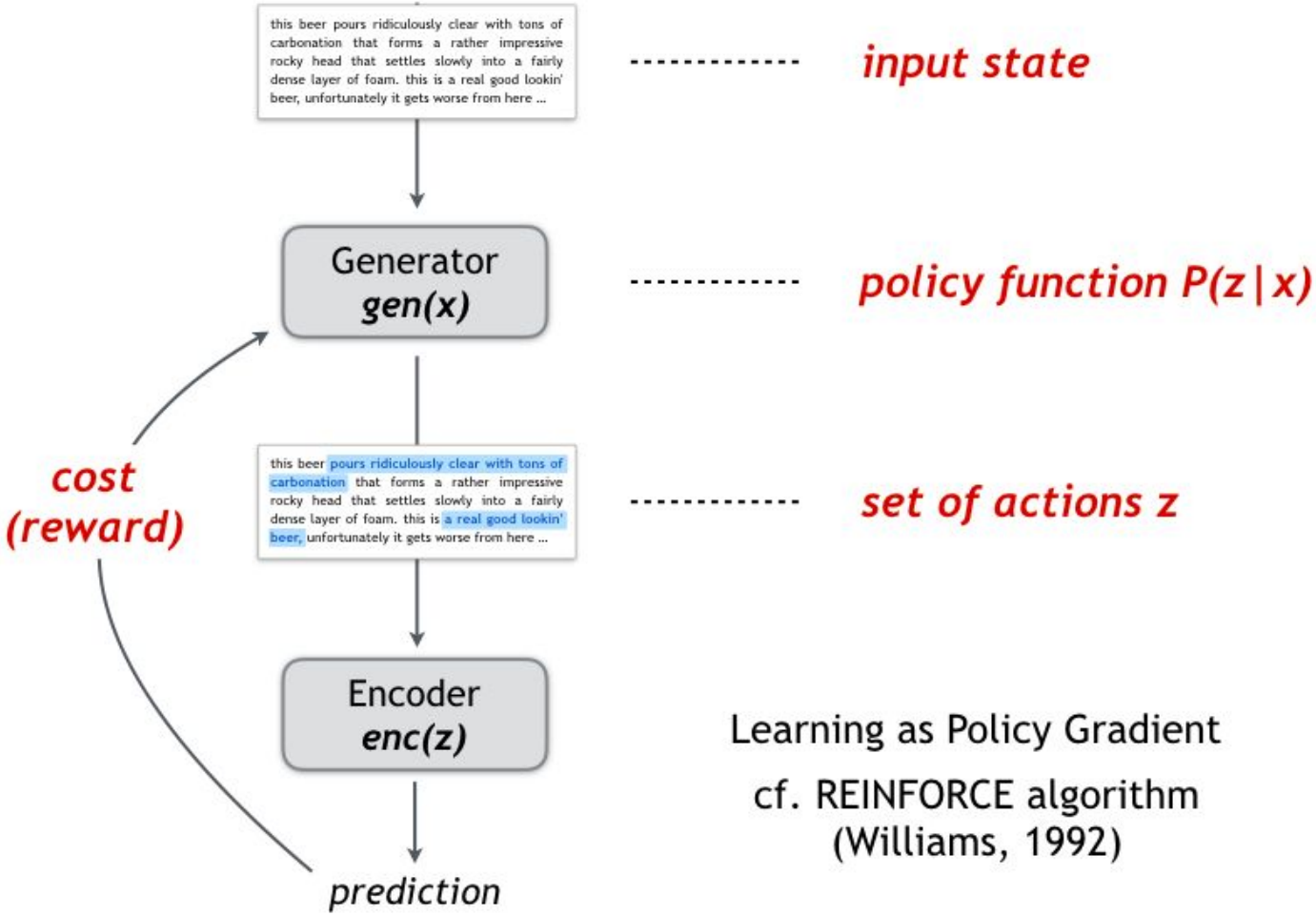
Generator

- input (words) $\mathbf{x} = \{x_t\}_{t=1}^{l}$
- output (probability word is selected) $\mathbf{z} \sim \mathbf{gen}(\mathbf{x}) \equiv p(\mathbf{z}|\mathbf{x})$
- loss $\Omega(\mathbf{z}) = \lambda_1 \|\mathbf{z}\| + \lambda_2 \sum_t |\mathbf{z}_t - \mathbf{z}_{t-1}|$

Learning process

$$\text{cost}(\mathbf{z}, \mathbf{x}, \mathbf{y}) = \mathcal{L}(\mathbf{z}, \mathbf{x}, \mathbf{y}) + \Omega(\mathbf{z})$$

$$\min_{\theta_e, \theta_g} \sum_{(\mathbf{x},\mathbf{y}) \in D} \mathbb{E}_{\mathbf{z} \sim \mathbf{gen}(\mathbf{x})} \left[\text{cost}(\mathbf{z}, \mathbf{x}, \mathbf{y}) \right]$$



Learning as Policy Gradient

cf. REINFORCE algorithm
(Williams, 1992)