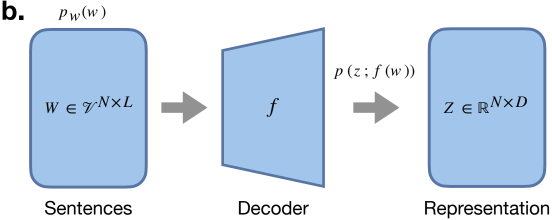
a.

```
def construct_z():
    # 1. Describe $Z$ using a compressed code $W$
    def p_w(w):  # $K(p_W)$ = length of this function
        ... # returns a probability
    w_compressed = [...]  # $K(W \mid p_W) = \sum -\log p_W(w_n)$
    w = decode_algo(w_compressed, p_w)  # $K$ = small constant

    # 2. Decode $Z$ from $W$
    def f(w):  # $K(f)$ = length of this function
        ... # returns mean and std of a normal distribution
    def p_normal(mu, std):  # $K$ = small constant
        ... # returns a probability
    z_mu, z_std = f(w)  # $K$ = small constant
    z_correction = [...]  # $K(Z \mid W, f) = \sum -\log \mathcal{N}(z_n; z_n^{\mu}, z_n^{\sigma})$
    z = decode_algo(z_correction, p_normal, z_mu, z_std)  # $K$ = small constant

    return z
```


b.
$p_W(w)$
$W \in \mathscr{V}^{N \times L}$
Sentences
f
Decoder
$p(z; f(w))$
$Z \in \mathbb{R}^{N \times D}$
Representation

c.

$$K(Z) \approx \min_{p_W, W, f} K(p_W) + K(W \mid p_W) + K(f) + K(Z \mid W, f)$$