

### Examples

1. block *650* sent
2. network interface is *up*

1. [[EMBEDDING], block, 650, sent]
2. [[EMBEDDING], network, interface, is, up]

### Dictionary

[EMBEDDING] : 0,  
block : 1, 650: 2,  
sent : 3,  
network : 4,  
interface : 5, is: 6,  
up: 7

### Indexing

1. [0, 1, 2, 3]
2. [0, 4, 5, 6, 7]

### Masking

1. [0, mask, 2, 3], 1
2. [0, 4, mask, 6, 7], 5

### Tokenization and masking

Transformer encoder  
with multi-head dot-product self-attention

$\mathbf{z}$  = Numerical vector of the [EMBEDDING] token  
(summarizes the log message)  
[0.23, 0.48, 0.78, ..., 0.34]

- Log template extraction (parsing)
1. block <\*> sent
  2. Network interface is <\*>