

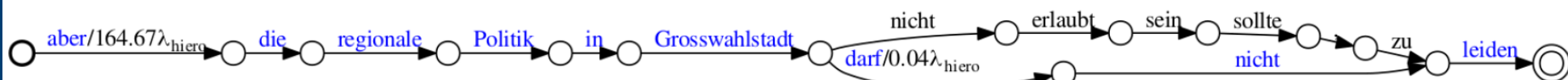


Introduction

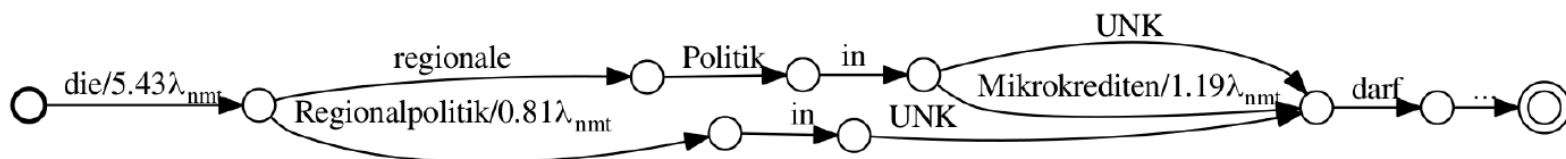
- We investigate the combination of neural and syntactical machine translation
- Lattice rescoring is very effective for single NMT, but lattices are too narrow for NMT ensembles
- We propose a loose coupling scheme based on the edit distance using standard FST operations

Loose Coupling of Hiero and NMT via Edit Distance

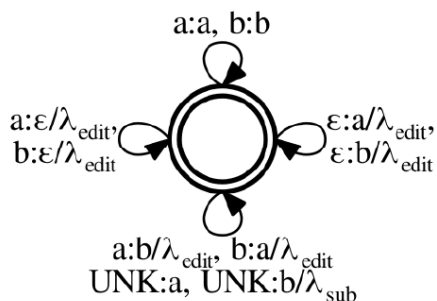
Hiero lattice H :



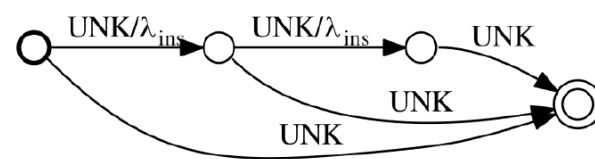
NMT lattice N :



Modified edit distance transducer E :

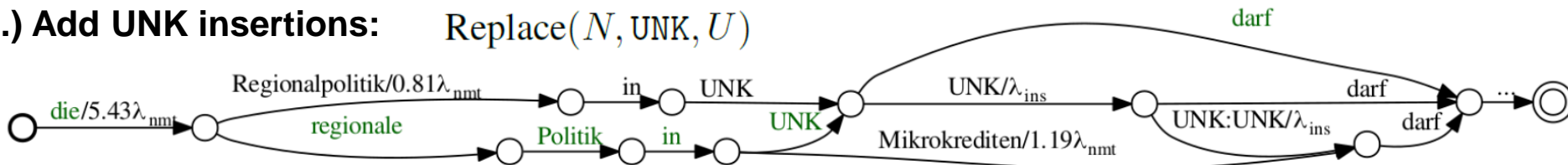


UNK insertion transducer U :



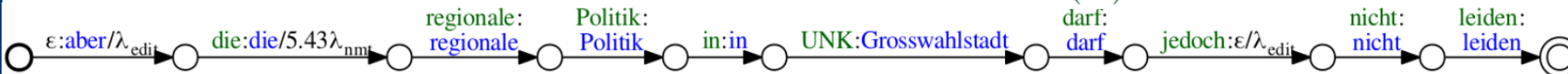
1.) Add UNK insertions:

$\text{Replace}(N, \text{UNK}, U)$



2.) Composition with edit distance transducer, shortest path:

$C := \text{Replace}(N, \text{UNK}, U) \circ E \circ H$
ShortestPath(C)

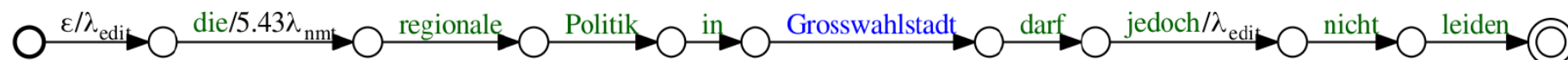


3.) Projection:

$\hat{t}_{comb} = \Pi_{UNK}(\text{ShortestPath}(C))$

$$\hat{t}_{NMT}, \hat{t}_{Hiero} = \operatorname{argmax}_{(t_N, t_H) \in N \times H} \left(e^{-d_{edit}(t_N, t_H)} \cdot P(t_N, t_H | s) \right)$$

$$P(t_N, t_H | s) := P_N(t_N | s)^{\lambda_{nmt}} \cdot P_H(t_H | s)^{\lambda_{hiero}}$$



8 NMT + 1 Hiero

	WMT15	WMT16
Best in competition	25.2	34.8
Hiero baseline	21.2	26.0
NMT baseline (8-ens.)	21.7	25.4
SGNMT (Stahlberg et al., 2016)	24.6	29.7
This work	25.7	31.3

Lattices vs. n-best lists

Depth of \hat{t}_{Hiero} in Hiero n-best list

