

HotpotQA (Transformer_{BASE})

Codelength per sample (bits)

10^1

9×10^0

8×10^0

7×10^0

6×10^0

- No SubQ
- PseudoD
- Seq2Seq
- ONUS
- DLM
- Oracle

10^2

10^3

10^4

Number of training samples

