

A Cross-Task Analysis of Text Span Representations

Shubham Toshniwal

Haoyue Shi

Bowen Shi

Lingyu Gao

Karen Livescu

Kevin Gimpel



github.com/shtoshni92/span-rep

Sentence vs Span

- Lot of prior work in evaluating fixed-dimensional sentence representations.

Sentence vs Span

- Lot of prior work in evaluating fixed-dimensional sentence representations.
- Fixed-dimensional span representations used by QA, coreference resolution, etc.
- **Lack of a comprehensive evaluation for span representations!**

Probing Setup

Tasks

Edge probing setup of Tenney et al. 2019:

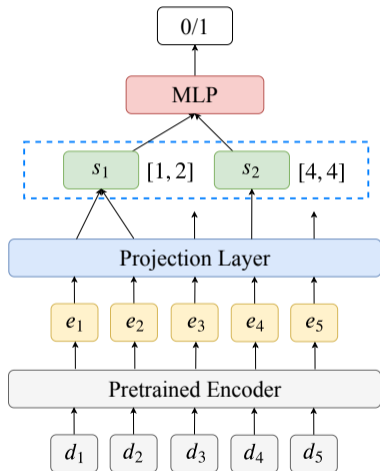
- Constituent labeling
- Named entity labeling (NEL)
- Semantic role labeling (SRL)
- Coreference arc prediction

Tasks

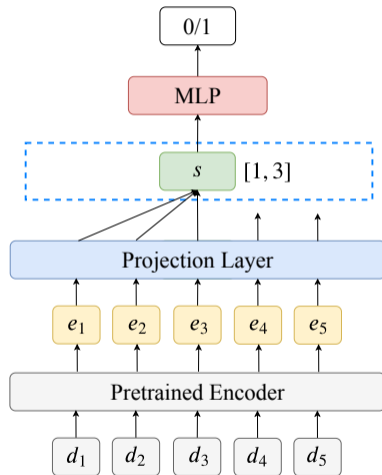
Edge probing setup of Tenney et al. 2019:

- Constituent labeling
- *Constituent detection*
- Named entity labeling (NEL)
- Semantic role labeling (SRL)
- *Mention detection*
- Coreference arc prediction

Probing Model



Labels
Classifier
Span
Representations



Contextual
Vectors

Input
Tokens

Span Representations

- Average pooling
 - Attention pooling
 - Max pooling
- } Entire-span
- Endpoint
 - Diff-Sum
 - Coherent
- } Boundary-based

Results

Frozen Encoder

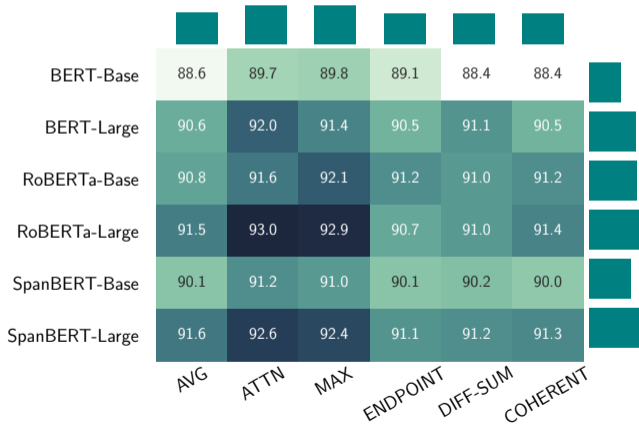
Constituent Labeling

	AVG	ATTN	MAX	ENDPOINT	DIFF-SUM	COHERENT
BERT-Base	85.0	93.0	96.0	96.4	96.5	96.3
BERT-Large	84.2	92.2	95.9	96.6	96.4	96.3
RoBERTa-Base	83.9	92.6	95.5	96.2	96.6	96.1
RoBERTa-Large	82.6	90.4	94.5	96.0	96.2	95.8
SpanBERT-Base	83.3	92.5	95.5	96.5	96.7	96.4
SpanBERT-Large	85.3	92.1	95.8	96.6	96.5	96.5

Boundary-based
representations preferred!

Frozen Encoder

Coreference Arc Prediction



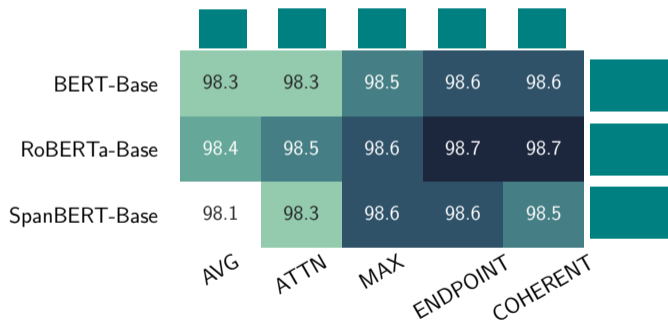
Entire-span
representations preferred!

Takeaways from Frozen Encoder Experiments

- *Syntactic tasks*: Boundary-based span representations outperform entire-span representations.
- *Semantic tasks*: Entire-span representations outperform boundary-based representations.

Fine-tuned Encoder

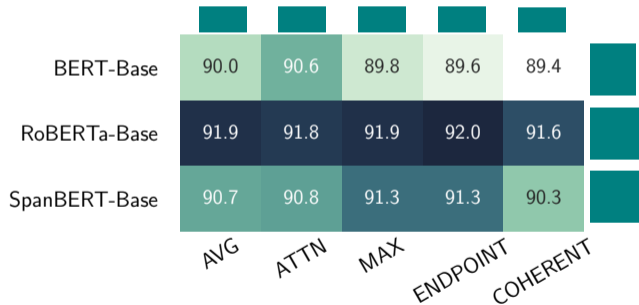
Constituent Labeling



Impact of span representation choice less pronounced!

Fine-tuned Encoder

Coreference Arc Prediction



Impact of span representation choice less pronounced!

Conclusion

- Optimal span representation varies for tasks.
- Choice of span representation is crucial for frozen encoders but less important for fine-tuned encoders.