

Topological Sort for Sentence Ordering

**Shrimai Prabhume, Ruslan Salakhutdinov,
Alan W Black**



Carnegie Mellon University

Language Technologies Institute

Sentence Ordering Task

Sentence Ordering Task

Before

Ironman steals the Infinity Stones back from Thanos and uses them to disintegrate Thanos and his army, at the cost of his life.

Hulk travels to New York City in 2012 and convinces the Ancient One to give him the Time Stone.

Thor decapitates Thanos.

Five years later, AntMan escapes from the quantum realm.

Ironman builds a time machine to save the world.

Sentence Ordering Task

Before

Ironman steals the Infinity Stones back from Thanos and uses them to disintegrate Thanos and his army, at the cost of his life.

Hulk travels to New York City in 2012 and convinces the Ancient One to give him the Time Stone.

Thor decapitates Thanos.

Five years later, AntMan escapes from the quantum realm.

Ironman builds a time machine to save the world.



After

Thor decapitates Thanos.

Five years later, AntMan escapes from the quantum realm.

Ironman builds a time machine to save the world.

Hulk travels to New York City in 2012 and convinces the Ancient One to give him the Time Stone.

Ironman steals the Infinity Stones back from Thanos and uses them to disintegrate Thanos and his army, at the cost of his life.

Methodology

- Constraint Solving Problem vs Sequence Prediction Task
- For a document with n sentences ($\{s_1 \dots s_n\}$)
 - $|\mathcal{C}| = \binom{n}{2}$ constraints
 - Predicted constraints of the form $s_1 < s_2$
 - 4 sentences in a document then 6 constraints
 - $\{s_1 < s_2, s_1 < s_3, s_1 < s_4, s_2 < s_3, s_2 < s_4, s_3 < s_4\}$
 - Topological sort to find an order given \mathcal{C}
 - Graph: $s_1 \rightarrow s_2$ if $s_1 < s_2$

Constraint Learning

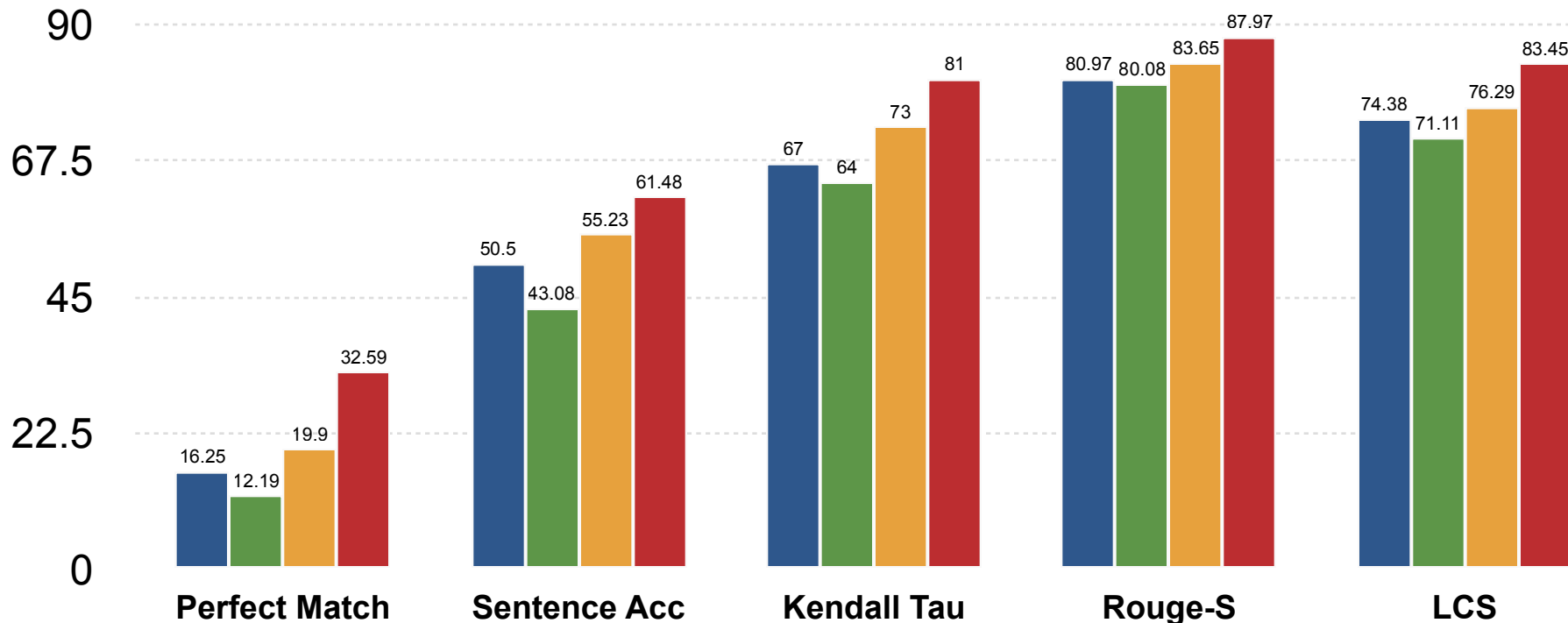
- ***BERT based Representation (B-TSort)***
 - Next Sentence Prediction
 - $\text{MLP}(\text{BERT}(s_1[\text{SEP}]s_2))$
- ***LSTM based Representation (L-TSort)***
 - $\mathbf{h}_1 = \text{LSTM}(s_1); \mathbf{h}_2 = \text{LSTM}(s_2)$
 - $\text{MLP}([\mathbf{h}_1; \mathbf{h}_2])$

Baselines

- ***Attention Order Network (AON)***
 - LSTM: sentence representation
 - Transformer: document representation
 - LSTM decoder: generate order
- ***BERT Attention Order Network (B-AON)***
 - BERT: sentence representation

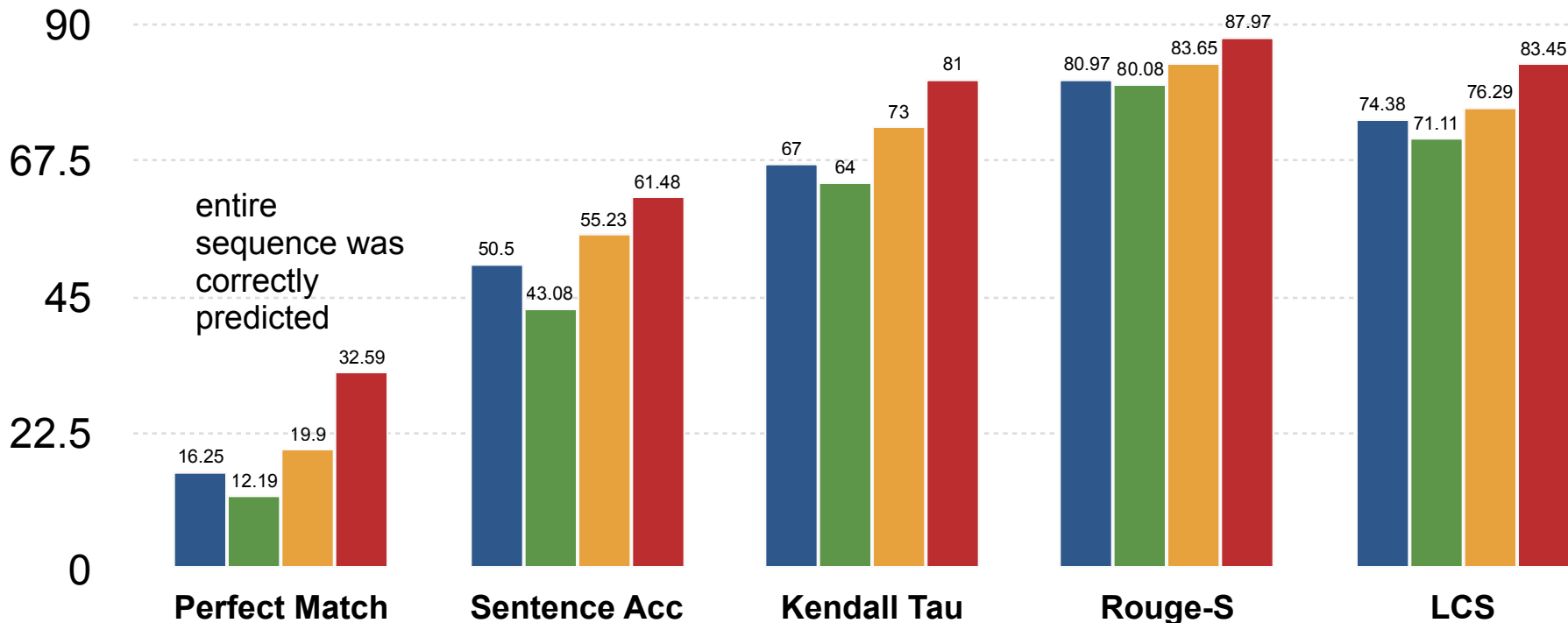
Results for NIPS abstracts

AON L-TSort B-AON B-TSort



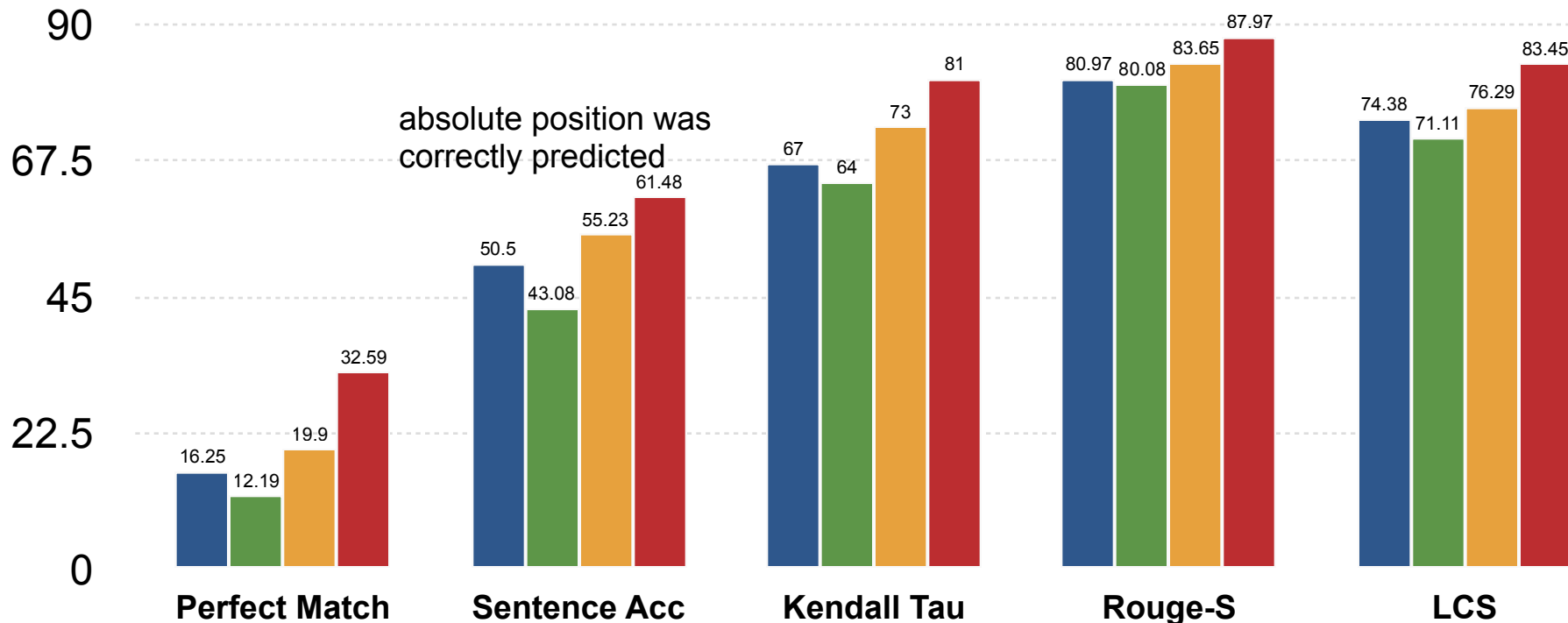
Results for NIPS abstracts

AON L-TSort B-AON B-TSort



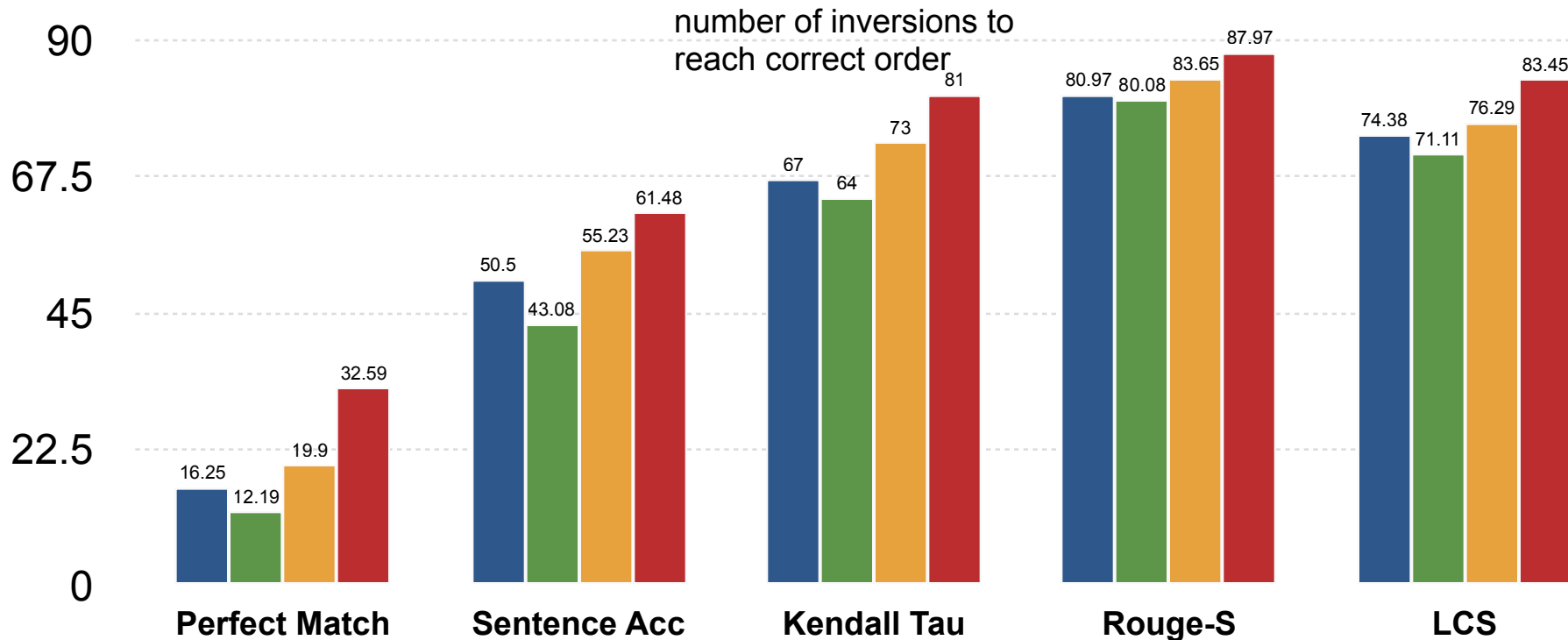
Results for NIPS abstracts

AON L-TSort B-AON B-TSort



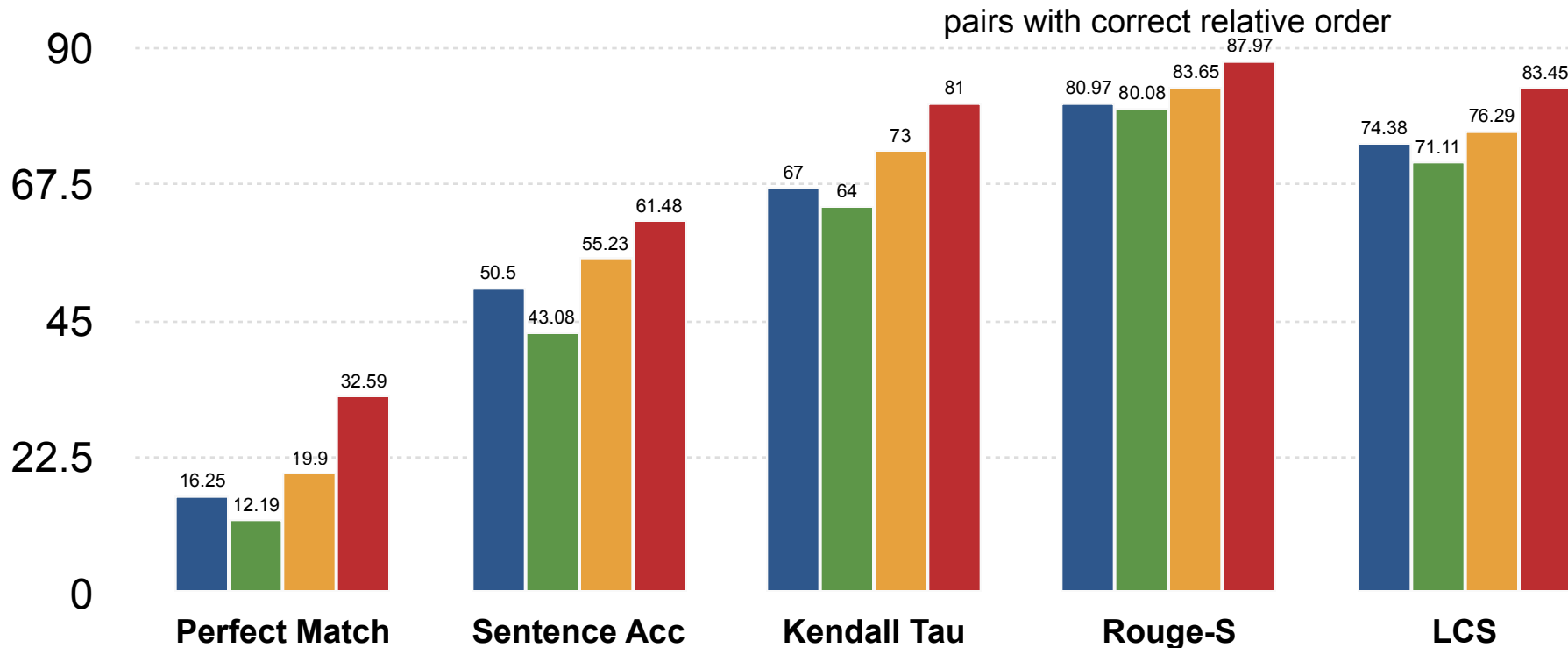
Results for NIPS abstracts

AON L-TSort B-AON B-TSort



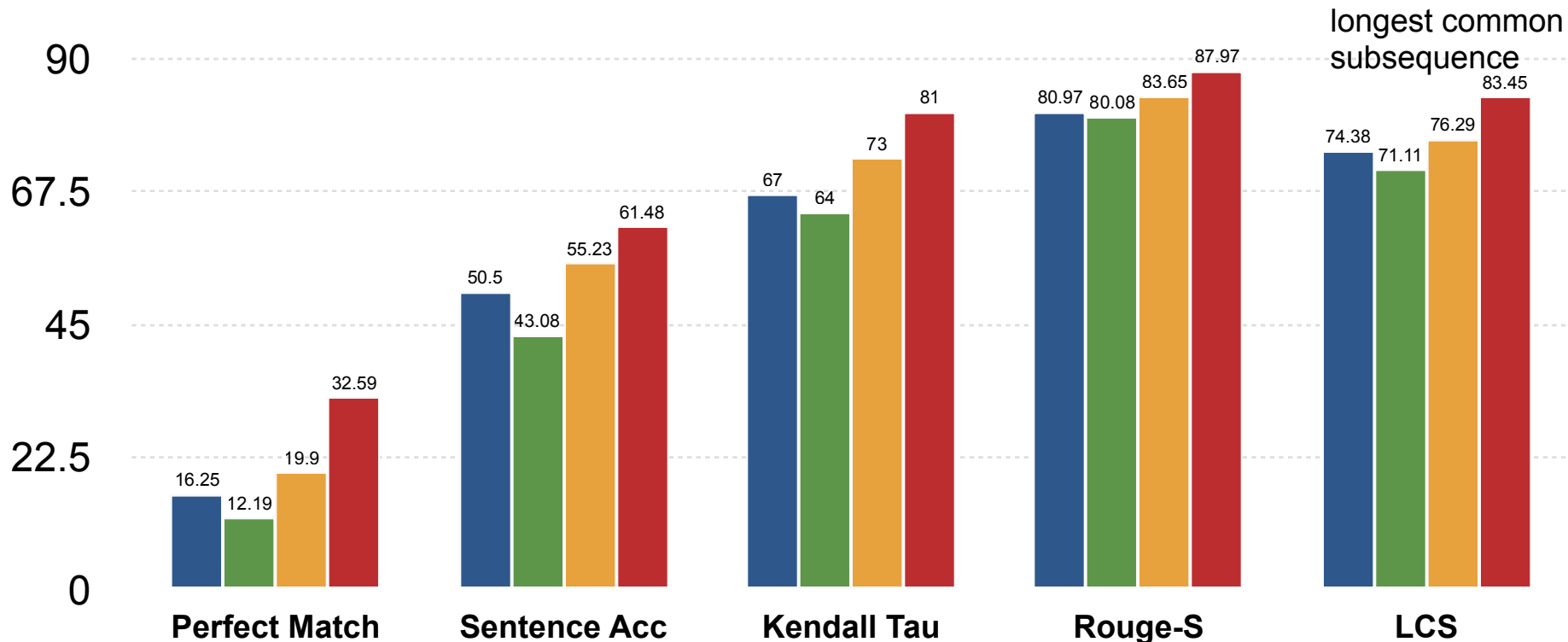
Results for NIPS abstracts

AON L-TSort B-AON B-TSort



Results for NIPS abstracts

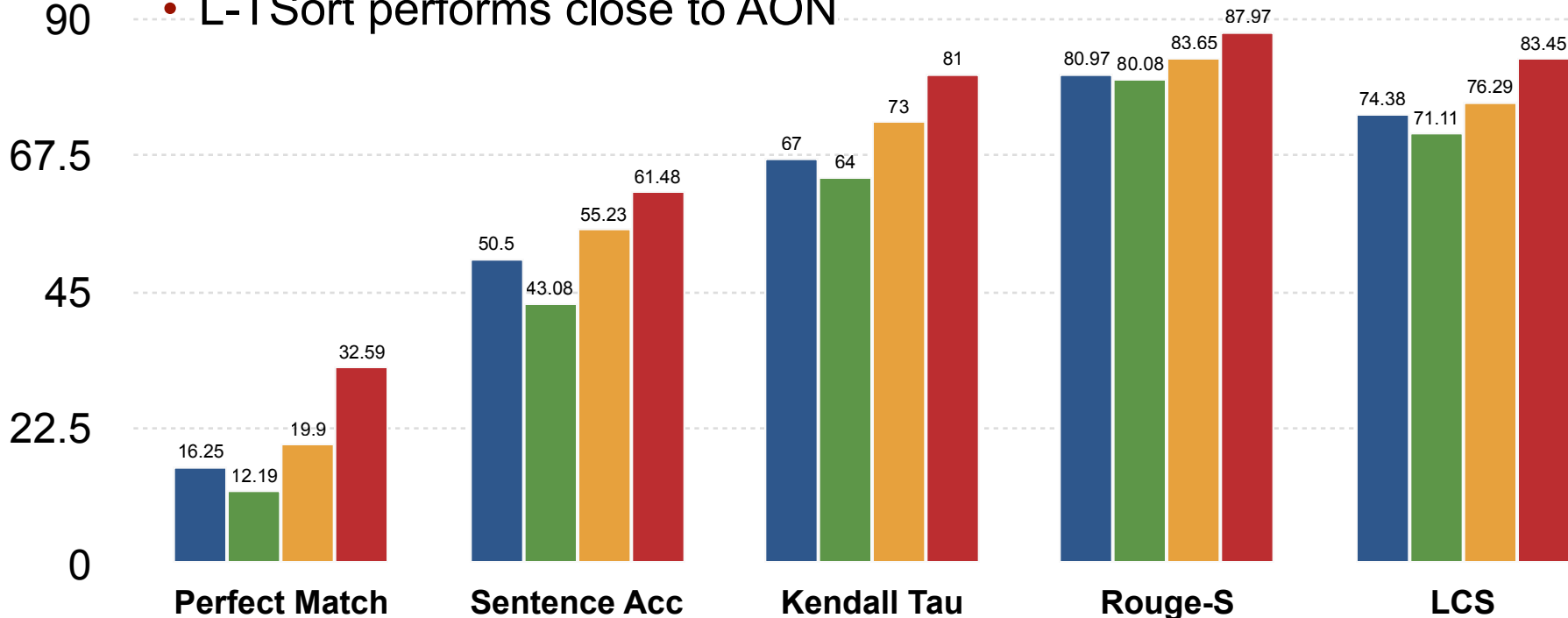
AON L-TSort B-AON B-TSort



Results for NIPS abstracts

■ AON ■ L-TSort ■ B-AON ■ B-TSort

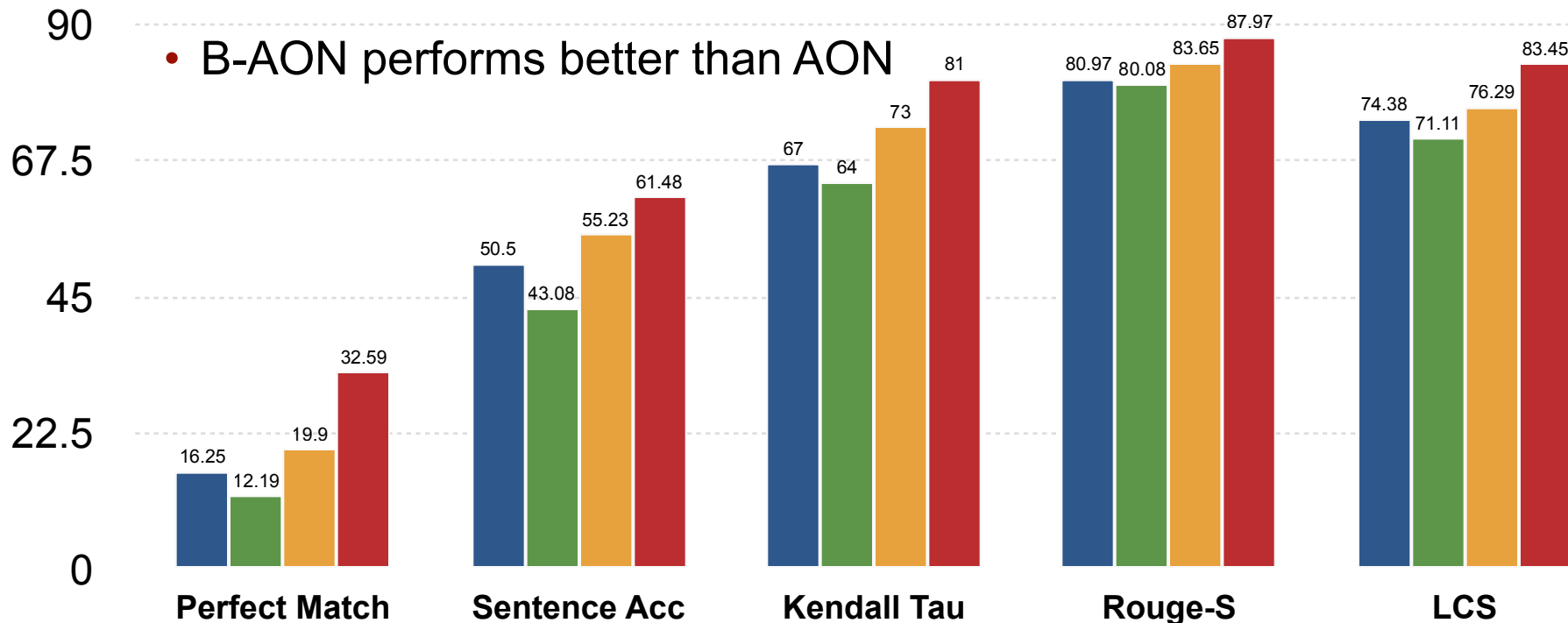
• L-TSort performs close to AON



Results for NIPS abstracts

AON L-TSort B-AON B-TSort

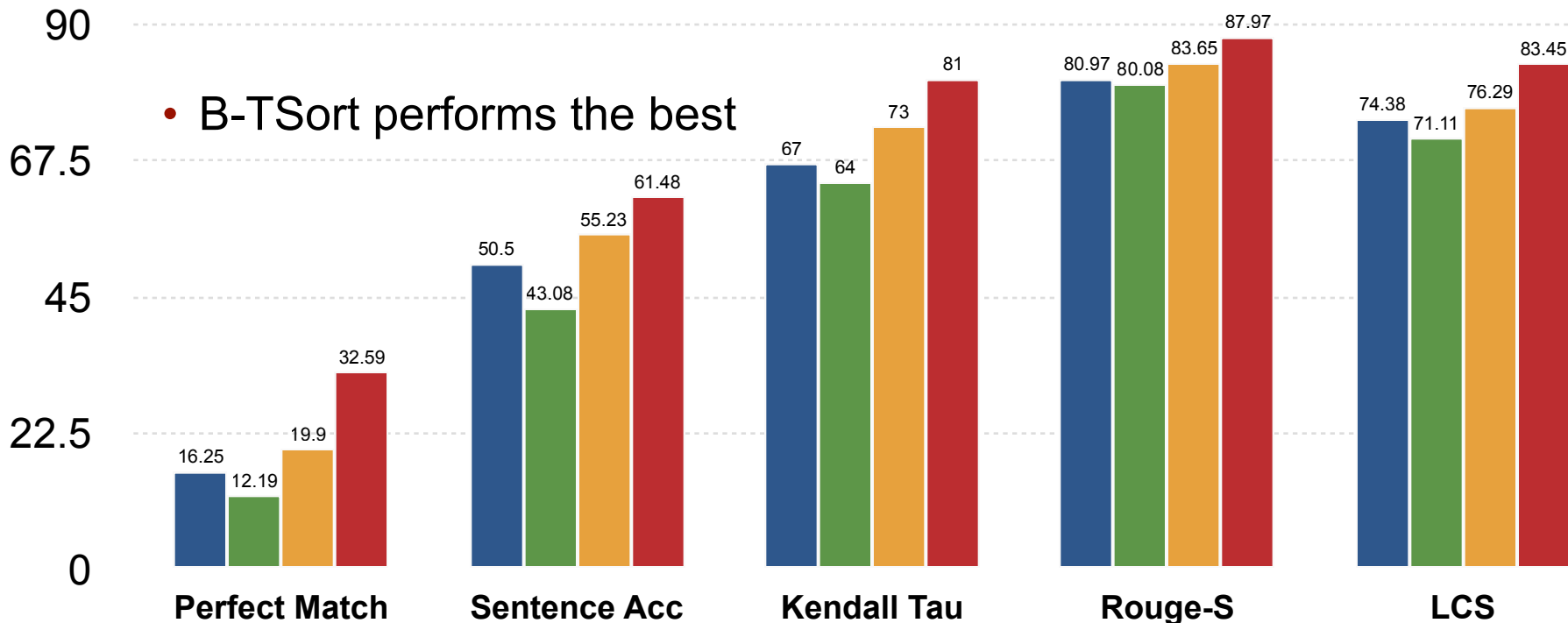
• B-AON performs better than AON



Results for NIPS abstracts

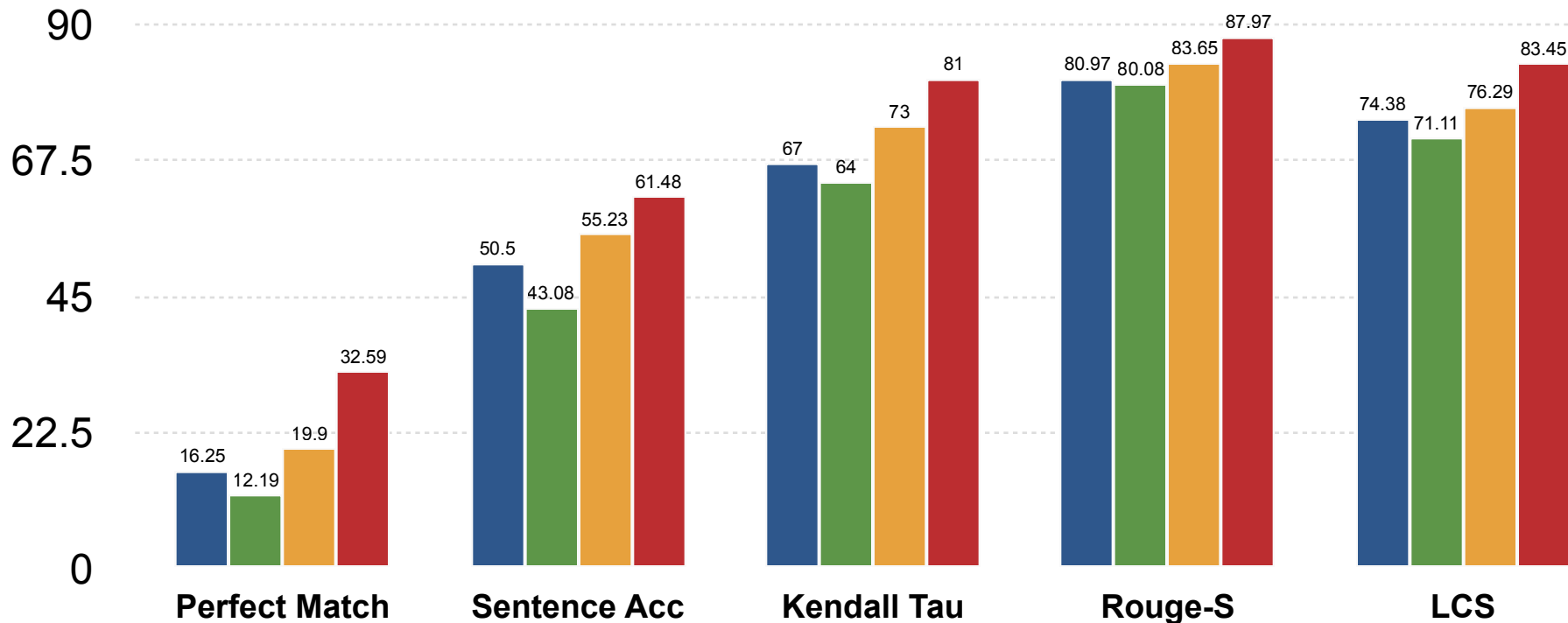
■ AON ■ L-TSort ■ B-AON ■ B-TSort

• B-TSort performs the best



Results for NIPS abstracts

AON L-TSort B-AON B-TSort



Results for Human Evaluation

■ B-TSort ■ No Preference ■ B-AON ■ Gold

B-TSort vs B-AON



B-TSort vs Gold



B-AON vs Gold



0

25

50

75

100

Results

- B-TSort performs best in all metrics for SIND captions, NSF abstract, AAN abstract datasets
- Analysis of *displaced sentences*
 - distance between the predicted position and the actual position of a sentence.
 - B-TSort displaces less sentences.
- Analysis of documents with more than 10 sentences
 - B-TSort also performs better for longer sentences
- Percentage of *mismatch* in input and output for AON
 - Order generation can have a mismatch between the input sentences and the generated orders.
 - This problem does not arise for our models.