CROWN
Conversational Passage Ranking
by Reasoning over Word Networks

Magdalena Kaiser, Rishiraj Saha Roy and Gerhard Weikum
Max Planck Institute for Informatics, Germany
SAMPLE CONVERSATION
Turn 1: What flowering plants work for cold climates?
Turn 2: Which of them survive frost?
Turn 3: How much cold can pansies tolerate?
Turn 4: What’s the UK hardiness rating?
….  
https://commons.wikimedia.org/w/index.php?curid=1954190
Method Overview

★ CROWN is an *unsupervised* method for passage ranking

★ Pseudo-relevant passages obtained with any standard retrieval system (e.g. Indri) using an *expanded* conversational query

★ CROWN models passage relevance as combination of *similarity* and *coherence*

★ Creates *Word Proximity Network (WPN)* from large corpus as backbone for passage scoring
Method Overview

★ The WPN stores co-occurrences of words, within a context window, as measured by Normalized Pointwise Mutual Information (NPMI)

★ Similarity between query and passage terms measured in terms of embedding vectors (node weights)

★ Coherence measured using proximities of significant pairs of passage terms (edge weights)
Word Proximity Network

★ Estimate term relationships from a large corpus
★ Store term co-occurrence in WPN
★ Graph as a very flexible structure with extension possibilities

“The hardiness rating of palm trees is …”, “Here are the RHS hardiness ratings …”, “A hardiness rating gives an estimate on …”
SAMPLE CONVERSATION

Turn 1: What *flowering plants* work for *cold climates*?

... 

Turn T-2: ...

Turn T-1: How much *cold* can *pansies* *tolerate*?

Turn T: What’s the *UK hardiness rating*?
Step 1: Query Expansion

SAMPLE CONVERSATION
Turn 1: What *flowering plants* work for *cold climates*?
... 
Turn T-2: ...
Turn T-1: How much *cold* can *pansies tolerate*?
Turn T: What’s the *UK hardiness rating*?

Conversational Query:

```
combine ( w₁ * (flowering plants cold climates) )
```

```
... 

w_{T-1} * (pansies tolerate)
```

```
w_T * (UK hardiness rating) 
```
Step 2: Candidate passage retrieval

Issue conversational query to standard retrieval system (e.g. Indri)
Step 2: Candidate passage retrieval

Issue conversational query to standard retrieval system (e.g. Indri)

```
Retrieval System
```

```
P_1
```
```
P_2
```
```
P_3
```
```
...
```

“indri-scores”
(1 / rank):
1.0
0.5
0.3

Step 3: Modeling Similarity

Conversational Query:
\[ w_1^* \text{ (flowering plants cold climates)} \quad w_{T-1}^* \text{ (pansies tolerate)} \quad w_T^* \text{ (UK hardiness rating)} \]

Candidate Passage:
P2: “Winter pansies have a UK hardiness rating of H5. In cold climate, pansies can ...”
Step 3: Modeling Similarity

Conversational Query:

\( w_1^* \) (flowering plants cold climates) \( w_{T-1}^* \) (pansies tolerate) \( w_T^* \) (UK hardiness rating)
Step 3: Modeling Similarity

Conversational Query:

\( w_1^* \) (flowering plants cold climates) \( w_{T-1}^* \) (pansies tolerate) \( w_T^* \) (UK hardiness rating)
Step 3: Modeling Similarity

Conversational Query:

$w_1^* (\text{flowering plants cold climates}) \ w_{T-1}^* (\text{pansies tolerate}) \ w_T^* (\text{UK hardiness rating})$
Step 3: Modeling Similarity

Conversational Query:

$w_1^*$ (flowering plants cold climates) $w_{T-1}^*$ (pansies tolerate) $w_T^*$ (UK hardiness rating)
Step 3: Modeling Similarity

Conversational Query:

\( w_1^* \) (flowering plants cold climates) \( w_{T-1}^* \) (pansies tolerate) \( w_T^* \) (UK hardiness rating)
Step 3: Modeling Similarity

Conversational Query:

\(w_1^* \) (flowering plants cold climates) \(w_{T-1}^* \) (pansies tolerate) \(w_T^* \) (UK hardiness rating)

threshold:
node-weight \(>= 0.7\)
Step 3: Modeling Similarity

Conversational Query:
\[ w_1^* \text{ (flowering plants cold climates)} \ w_{T-1}^* \text{ (pansies tolerate)} \ w_T^* \text{ (UK hardiness rating)} \]
Step 3: Modeling Similarity

Conversational Query:

\[ \mathbf{w}_1^* \text{ (flowering plants cold climates)} \mathbf{w}_{T-1}^* \text{ (pansies tolerate)} \mathbf{w}_T^* \text{ (UK hardiness rating)} \]

\[
\text{node-score}(P_2) = \mathbf{w}_1^* (0.99 + 1.0) + \mathbf{w}_2^* 1.0 + \mathbf{w}_3^* (1.0 + 1.0 + 1.0)
\]
Step 3: Modeling Similarity

$w_1^*$ (flowering plants cold climates) $w_{T-1}^*$ (pansies tolerate) $w_T^*$ (UK hardiness rating)
Step 4: Modeling Coherence

P2: “Winter pansies have a UK hardiness rating of H5. In cold climate, pansies can ...”
P2: “Winter pansies have a UK hardiness rating of H5. In cold climate, pansies can ...”
Step 4: Modeling Coherence

**P2:** “Winter pansies have a UK hardiness rating of H5. In cold climate, pansies can ...”

**Pairs within context window 3:**
(winter, pansies), (pansies, UK), (pansies, hardiness), (pansies, rating), (winter, UK), (UK hardiness), (UK rating), (hardiness, rating), (UK, H5), (hardiness, H5), (rating, H5), (H5, cold), (H5, climate), (H5, pansies), (cold, climate), (cold, pansies), (climate, pansies)
Step 4: Modeling Coherence

\[ P_2 \rightarrow \text{WPN} \]
Step 4: Modeling Coherence

Relevant pairs:
(winter, pansies), (UK rating), (hardiness, rating), (cold, climate)

threshold: edge-weights $\geq 0.3$
Step 4: Modeling Coherence

Relevant pairs:
(winter, pansies), (UK rating), (hardiness, rating), (cold, climate)
Step 4: Modeling Coherence

Relevant pairs:
(winter, pansies), (UK rating), (hardiness, rating), (cold, climate)
Step 4: Modeling Coherence

Relevant pairs:
(winter, pansies), (UK rating), (hardiness, rating), (cold, climate)
Step 4: Modeling Coherence

Relevant pairs:
(winter, pansies), (UK rating), (hardiness, rating), (cold, climate)

\[ \text{edge-score}(P_2) = 0.3 + 0.7 + 0.9 \]
WPN adapted to current query

\[ w_1^* \text{(flowering plants cold climates)} \] \[ w_{T-1}^* \text{(pansies tolerate)} \] \[ w_T^* \text{(UK hardiness rating)} \]
Step 5: Putting it together

\[
\text{score}(P_i) = \text{indri-score}(P_i) + \text{node-score}(P_i) + \text{edge-score}(P_i)
\]
Step 5: Putting it together

\[ \text{score}(P_i) = h_1 \cdot \text{indri-score}(P_i) + h_2 \cdot \text{node-score}(P_i) + h_3 \cdot \text{edge-score}(P_i) \]

with hyperparameters \( h_1, h_2 \) and \( h_3 \)
Results

- **Submitted four runs** that explored some variations of our CROWN method
- **Three out of four** runs achieved **above median performance** over all submitted runs (AP@5 and nDCG@1000)
Conclusion

★ CROWN is an unsupervised method for passage ranking
★ Word proximity network as underlying structure
★ Similarity as node weights and coherence as edge weights
★ In future work:
  ○ Consider term position
  ○ Better normalization
  ○ More parameter tuning

Contact: mkaiser@mpi-inf.mpg.de