≥ ir_axioms

Intuitive Axiomatic Retrieval Experimentation.

A. Bondarenko M. Fröbe J. H. Reimer B. Stein M. Völske M. Hagen

Axioms

What they promise 😄

- Axioms are (pairwise) **preferences**
- ≈ "Rules" what documents to rank first
- **Easy** to understand

Use-cases

- Explain ranker decisions
- Evaluate ranking errors
- Re-rank

Working with axioms… 😐

in the past

PROX1 [17] "Prefer the document with shorter total distance between query term pairs."
Given $|Q| > 1, \forall_{q \in Q} q \in D_1 \wedge q \in D_2, \quad M(D,q) = \{i : t_i \in D \wedge t_i = q\}$
$\delta(D,q_1,q_2) = \frac{1}{M(D,q_1)\cdot M(D,q_2)} \sum_{(i,j)\in M(D,q_1)\times M(D,q_2)} |i-j|$
$\sum_{(q_i,q_j)\in Q\times Q} \delta(D_1,q_i,q_j) < \sum_{(q_i,q_j)} \delta(D_2,q_i,q_j) \quad \Rightarrow \quad D_1 >_{\text{PROX1}} D_2$

PROX2 [17] "Prefer documents where query terms occur earlier."
Given $|Q| > 1, \forall_{q \in Q} q \in D_1 \wedge q \in D_2, \quad \text{first}(q,D) = \min\{i : t_i \in D \wedge t_i = q\}$
$\sum_{q\in Q} \text{first}(q,D_1) < \sum_{q\in Q} \text{first}(q,D_2) \quad \Rightarrow \quad D_1 >_{\text{PROX2}} D_2$

PROX3 [17] "Prefer documents where the query occurs earlier as a phrase."
Given $Q = \{q_1,\ldots,q_l\}, \forall_{q\in Q} q \in D_1 \wedge q \in D_2,$
$\tau(Q,D) = \min\{i : t_i \in D \wedge t_i = q_1,\ldots,t_{i+l} = q_l\} \cup \{\infty\}$
$\tau(D_1,Q) < \tau(D_2,Q) \quad \Rightarrow \quad D_1 >_{\text{PROX3}} D_2$

PROX4 [17] "Prefer documents that cover all query terms in a shorter sub-string."
Given $|Q| > 1, \forall_{q\in Q} q \in D_1 \wedge q \in D_2,$
$\omega(D,Q) = \min\{j - i : i < j \wedge t_i \in D \wedge t_j \in D \wedge \forall_{q\in Q} q \in D_{|i\ldots j}\}$
$\omega(D_1,Q) < \omega(D_2,Q) \quad \Rightarrow \quad D_1 >_{\text{PROX4}} D_2$

PROX5 [17] "Prefer documents where query terms are closer together on average."
Given $|Q| > 1, \forall_{q\in Q} q \in D_1 \wedge q \in D_2, \quad M(D,Q) = \{i : t_i \in D \wedge t_i \in Q\}$
$s(D,Q,i) = \min\{k - j : j \le i \wedge k \ge i \wedge \forall_{q\in Q} q \in D_{|j\ldots k}\}$
$\frac{\sum_{i\in M(D_1,Q)} s(D_1,Q,i)}{|M(D_1,Q)|} < \frac{\sum_{i\in M(D_2,Q)} s(D_2,Q,i)}{|M(D_2,Q)|} \quad \Rightarrow \quad D_1 >_{\text{PROX5}} D_2$

Problems

- Not always "easy to understand" ™
- Many implementation caveats (uff!)
- Hard to maintain etc.…

`ir_axioms` to the rescue!

webis-de/ir_axioms `pip install ir_axioms`

- Reference implementations for **25 common axioms**
- **Define and combine** axioms declaratively
- Tightly integrates with **PyTerrier & Pyserini**

Experiments

```
experiment = AxiomaticExperiment(
  [bm25, monot5, ...],
  dataset.get_topics(),
  dataset.get_qrels(),
  index,
  axioms=[ArgUC(), QTArg(), QTPArg(), ...]
)

experiment.preferences
experiment.preference_distribution
experiment.preference_consistency
experiment.inconsistent_pairs
```

Re-ranking

```
bm25 = BatchRetrieve(index, "BM25")

axiom = (ArgUC() & QTArg() & QTPArg()) | ORIG()

# Re-rank top-20 documents with KwikSort.
kwiksort = bm25 % 20 >> \
  KwikSortReranker(axiom, index)

pipeline = kwiksort ^ bm25
```

Thank you, stay tuned for more!