Thread-Level Information for Comment Classification in Community Question Answering

Alberto Barrón-Cedeño, Simone Filice, Giovanni Da San Martino, Shafiq Joty, Lluís Màrquez, Preslav Nakov, Alessandro Moschitti
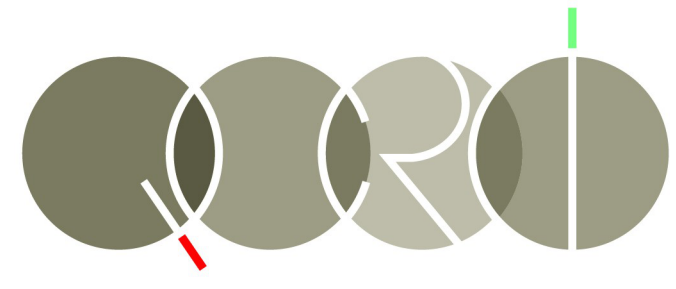
Qatar Computing Research Institute, Hamad Bin Khalifa University

{albarron, sfilice, gmartino, sjoty, lmarquez, pnakov, amoschitti}@qf.org.qa

1. Answer Selection

Question   GOOD*
POTENTIAL*
BAD*

SemEval-2015 Task 3:
Answer Selection in cQA

2. The Issues

Good Scissor
Dears,
Anyone have an idea where to
find a good scissor for hair
and beard trimming please???
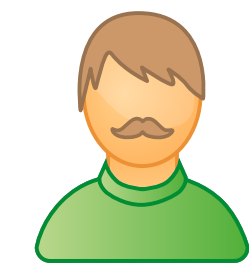
Visit Family food center
Visit Family food center   B→G
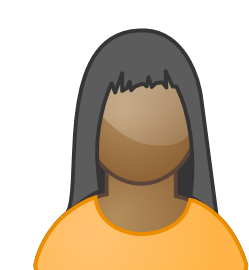
Al rawnaq airport road...U'll   B→G
Al rawnaq airport road...U'll
find all types of scissors there

Thank you all
Thanks for all... I will try that.   B→B
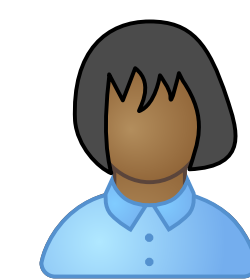
Gymnastic world cup
Does anyone know what time
the competition starts today?

sorry - is this being held
sorry - is this being held here   B→B
in Doha? If so, I'd love to go.
Expat Sueo
P.S. Is that a labradoodle in
your avatar?

**No actually a Cockapoo! Yes
the comp. runs from today**   B→B
until Wednesday at Aspire.

Thanks for the info - maybe
Thanks for the info - maybe I'll   G→B
turn up after the TableTop
Sale is done and dusted!

3. The Baseline

Similarity Features
• greedy string tiling
• longest common subseq.
• Jaccard coefficient
• word containment
• idf-weighted intersection
• cosine similarity
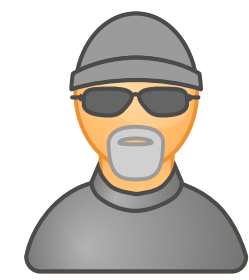• PTK similarity between
  syntactic trees

Heuristic Features
The comment (c) includes...
• URL or email
• "yes", "sure", "no", "@", etc.
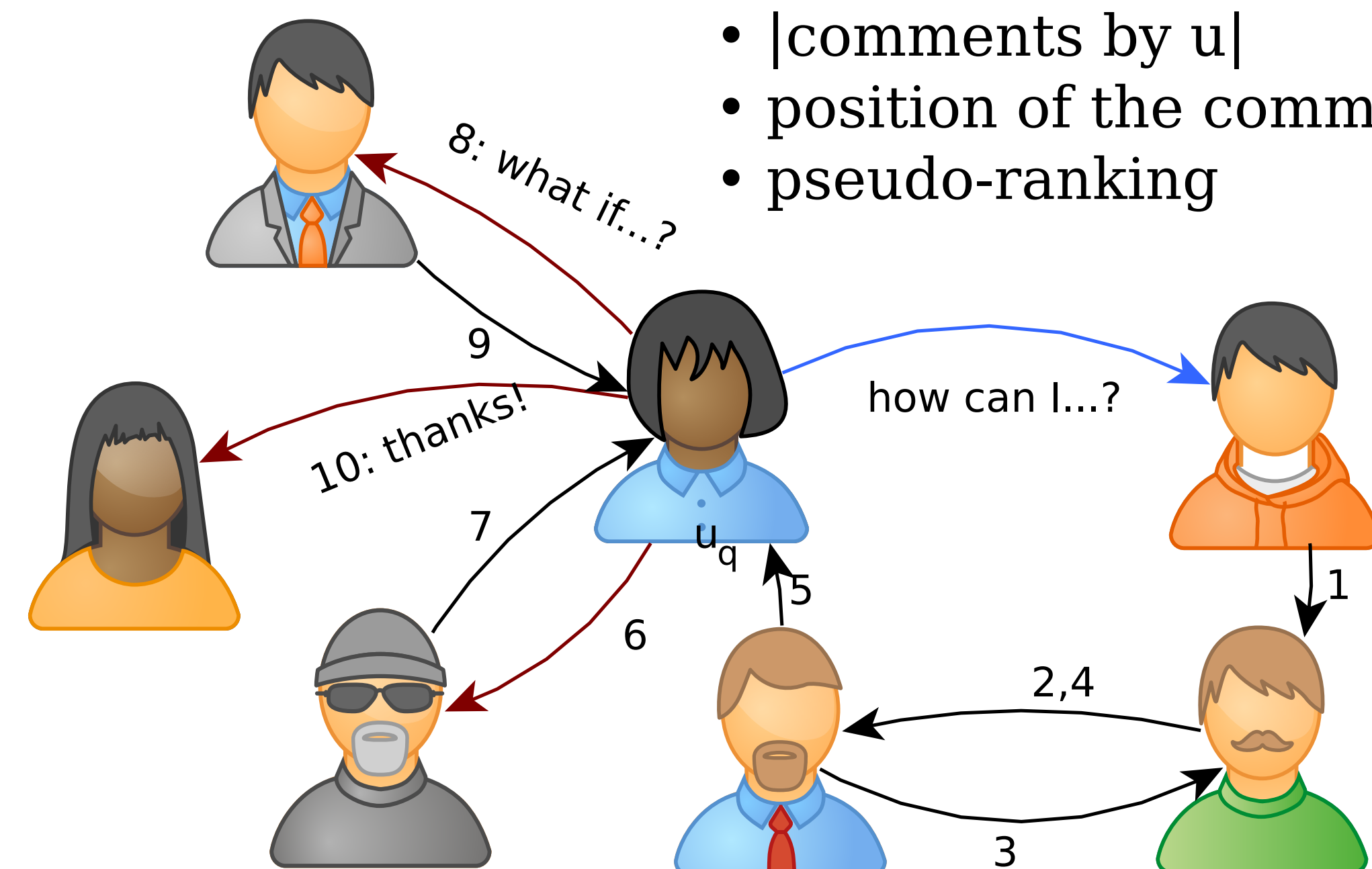• 3 or more repeated chars.
• longer than expected words
• thank* ack*

• c is posted by the same user
  who asked the question (u_q)
• the category of the question
• length of the comment

4. The Alternatives

Structured Prediction Models

SVM → SVM^{hmm}

Logistic regression → CRF

Thread-Level Global Features

c appears before a comment by u_q which
• contains an aknowledgement
• contains no acknowledgements
• contains a question
• contains a question (after c)

• Identification of dialogues
• |comments by u|
• position of the comment
• pseudo-ranking

5. The Outcome

Comment- and Thread-Level Evaluation: Test | Cross-validation

	F_1	A	$F_{1,avg}$	A_{avg}	F_1	A	$F_{1,avg}$	A_{avg}
Baseline Features								
SVM	76.89	74.39	66.52	76.13	74.98±0.7	72.90±1.0	64.56±1.0	75.32±0.4
SVM^{hmm}	78.49	76.37	68.55	77.58	75.28±1.0	73.75±1.6	65.25±1.2	74.68±1.0
LogReg	75.96	70.85	68.84	74.79	74.42±0.8	69.99±0.9	66.00±1.3	73.04±1.0
CRF$_{map}$	78.94	76.67	67.17	76.55	74.42±1.3	72.66±1.7	63.90±1.7	73.51±0.7
CRF$_{mpm}$	77.38	75.15	66.54	75.42	74.93±1.2	73.20±1.8	64.53±1.4	74.32±0.9
Baseline+Thread-based Features								
SVM	79.96	78.44	67.65	76.02	77.59±0.9	76.23±1.4	66.41±1.3	76.23±0.4
SVM^{hmm}	78.82	77.43	66.61	77.06	77.21±1.2	76.20±1.8	65.33±1.1	76.43±0.9
LogReg	79.39	77.33	68.10	75.57	77.22±1.1	75.43±1.5	66.57±1.5	75.05±0.7
CRF$_{map}$	79.31	77.43	66.37	76.08	76.75±1.3	75.14±1.7	65.36±1.4	75.61±0.6
CRF$_{mpm}$	79.51	77.78	67.36	76.63	77.00±1.3	75.43±1.8	65.57±1.5	75.71±0.7

6. Remarks

• Answer dependencies are helpful
  to improve in this task.

• The dependencies can be exploited
  with thread-level features, causing
  the context of a comment to affect
  quality estimation.

• If opting for an ad-hoc learning
  algorithm (e.g., CRF or SVM^{hmm}),
  the class of a comment depends on
  the class assigned to its
  neighbours as well.

• Robust context-aware features are
  the best way to push up the perfor-
  mance of the classifier.