Predicting Thread Discourse Structure over Technical Web Forums

Li Wang,[♠♡] Marco Lui,[♠♡] Su Nam Kim,[♠♡] Joakim Nivre[♢] and Timothy Baldwin[♠♡]

♠ Dept. of Computer Science and Software Engineering, University of Melbourne
♡ NICTA Victoria Research Laboratory
♢ Dept. of Linguistics and Philology, Uppsala University

July 27, 2011

Introduction

Example Thread

HTML Input Code - CNET Coding & scripting Forums

User A Post 1	**HTML Input Code** ...Please can someone tell me how to create an input box that asks the user to enter their ID, and then allows them to press go. It will then redirect to the page ...
User B Post 2	**Re: html input code** Part 1: create a form with a text field. See ... Part 2: give it a Javascript action
User C Post 3	**asp.net c\# video** I've prepared for you video.link click ...
User A Post 4	**Thank You!** Thanks a lot for that ... I have Microsoft Visual Studio 6, what program should I do this in? Lastly, how do I actually include this in my site? ...
User D Post 5	**A little more help** ... You would simply do it this way: ... You could also just ... An example of this is ...

Resources: http://forums.cnet.com/

Example Thread

HTML Input Code - CNET Coding & scripting Forums

User A Post 1	**HTML Input Code** ...Please can someone tell me how to create an input box that asks the user to enter their ID, and then allows them to press go. It will then redirect to the page ...
User B Post 2	**Re: html input code** Part 1: create a form with a text field. See ... Part 2: give it a Javascript action
User C Post 3	**asp.net c\# video** I've prepared for you video.link click ...
User A Post 4	**Thank You!** Thanks a lot for that ... I have Microsoft Visual Studio 6, what program should I do this in? Lastly, how do I actually include this in my site? ...
User D Post 5	**A little more help** ... You would simply do it this way: ... You could also just ... An example of this is ...

External Link

External Video

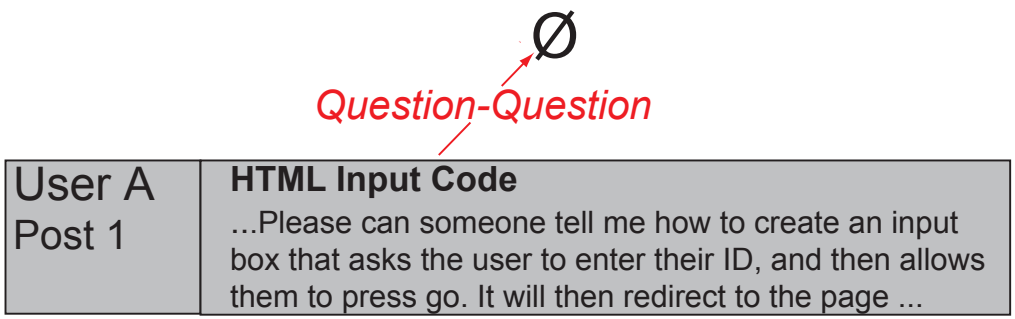
500 words in total

Resources: http://forums.cnet.com/

Discourse Structure of Forum Threads

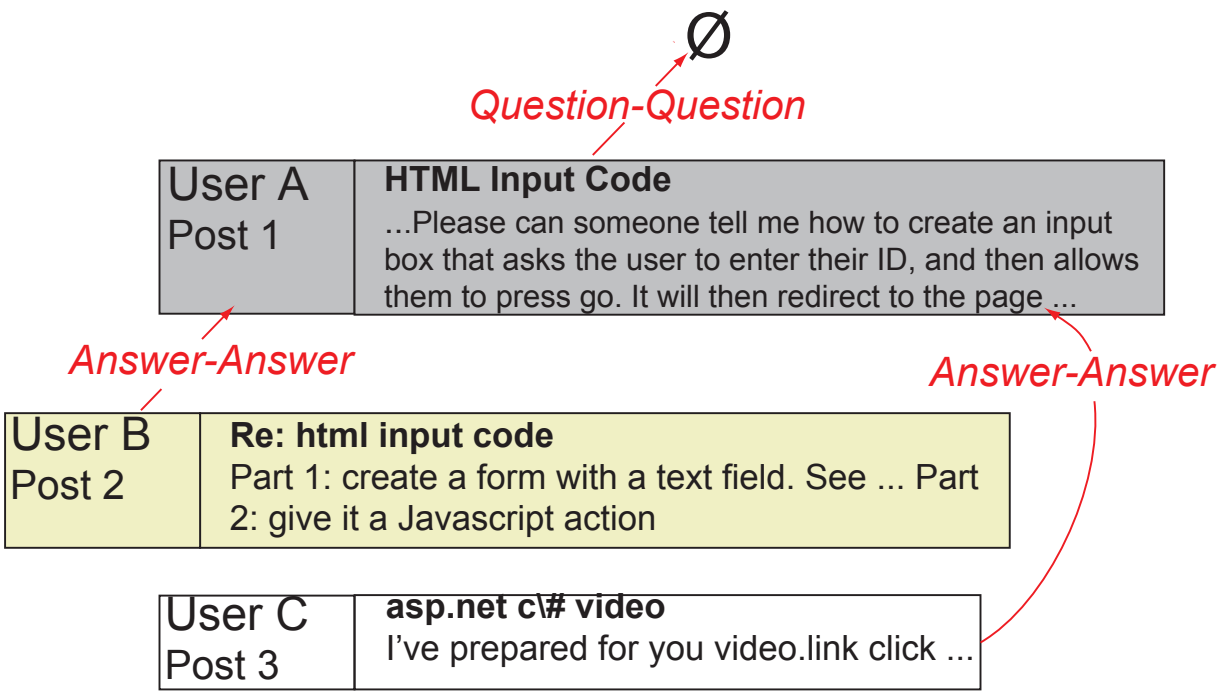
Reference: Kim et al., 2010

Discourse Structure of Forum Threads



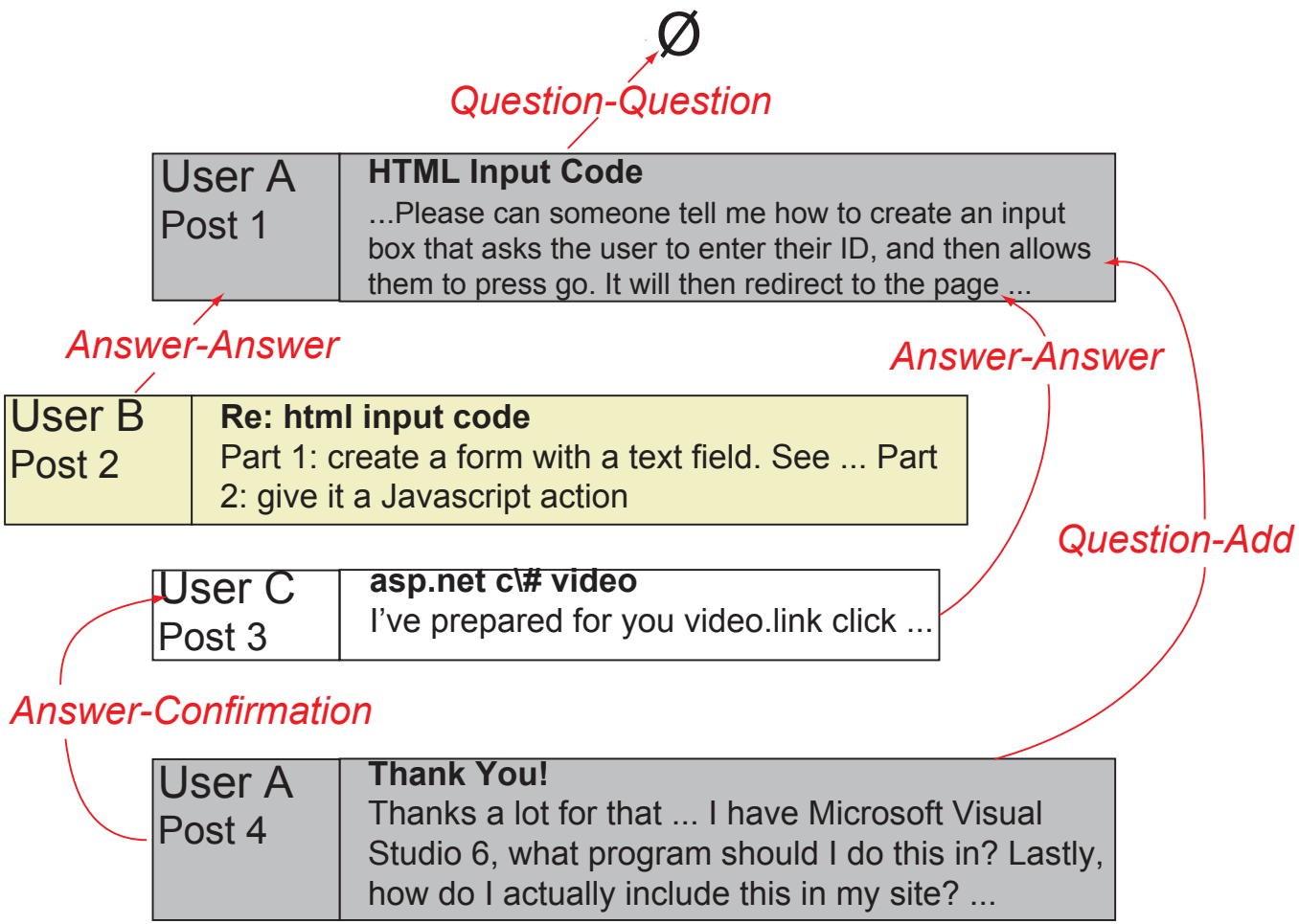
Reference: Kim et al., 2010

Discourse Structure of Forum Threads



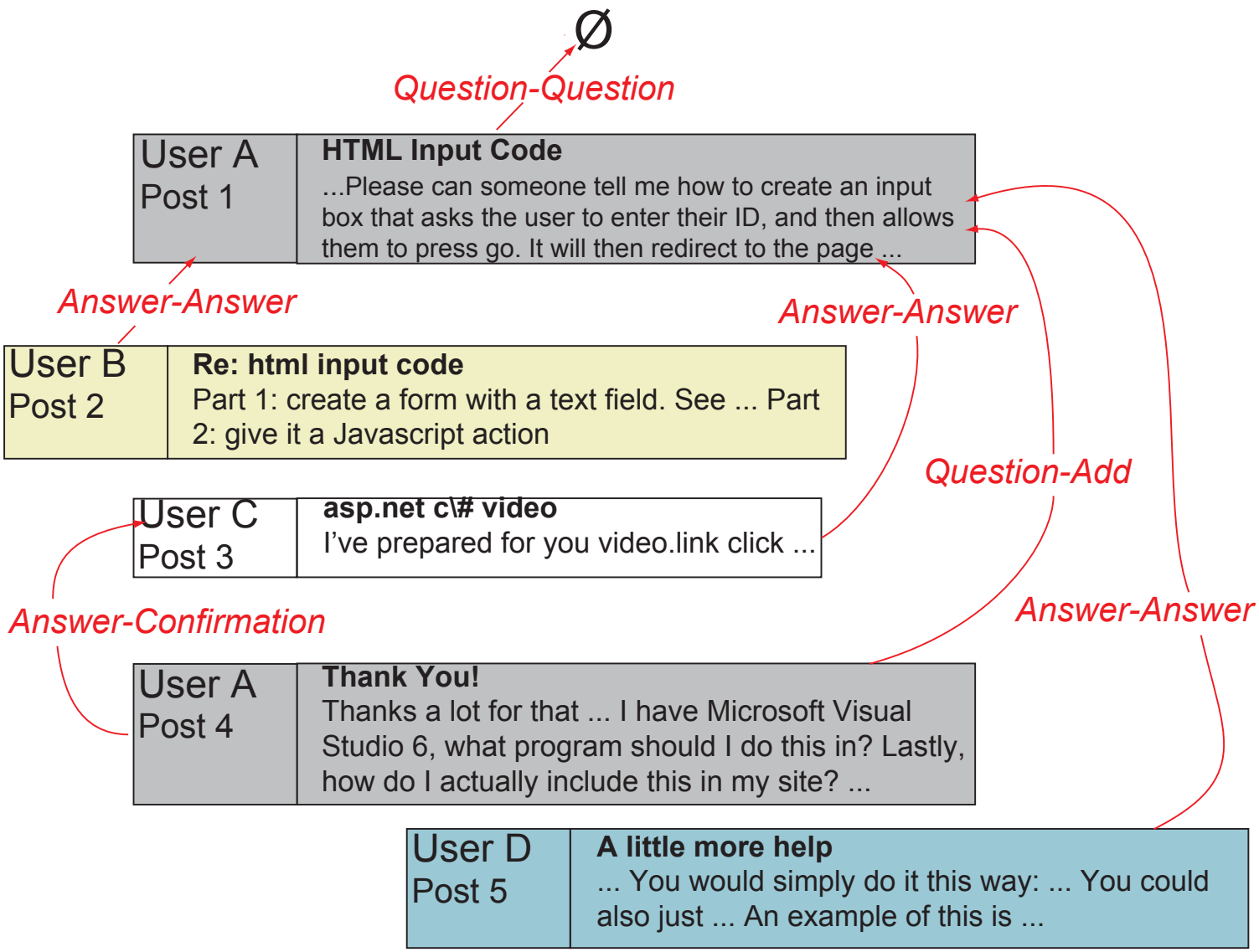
Reference: Kim et al., 2010

Discourse Structure of Forum Threads



Reference: Kim et al., 2010

Discourse Structure of Forum Threads



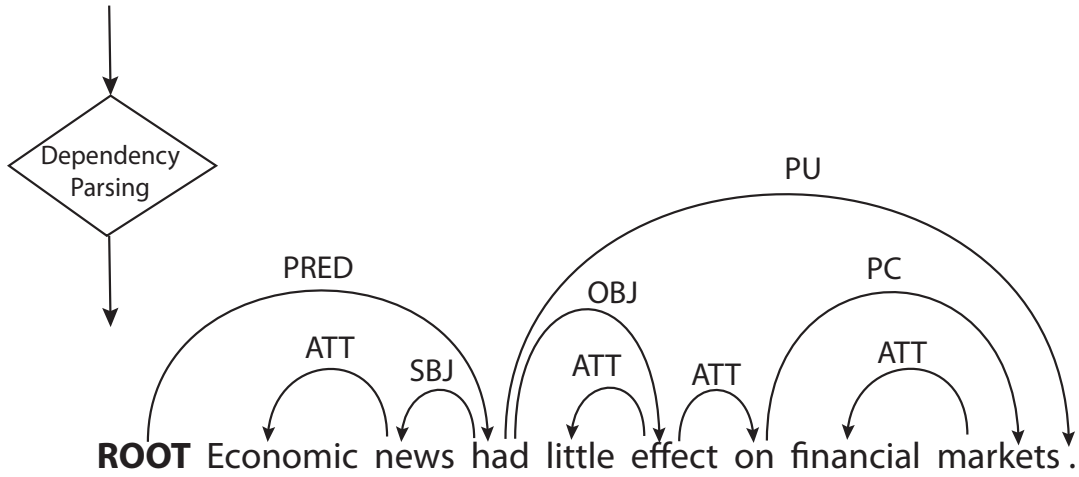
Reference: Kim et al., 2010

Research Aim and Contributions

- Aim:
  - jointly classify the discourse structure of forum threads
- Contributions:
  - apply structural learning and dependency parsing
  - in situ classification analysis

Dependency Parsing of Forum Threads

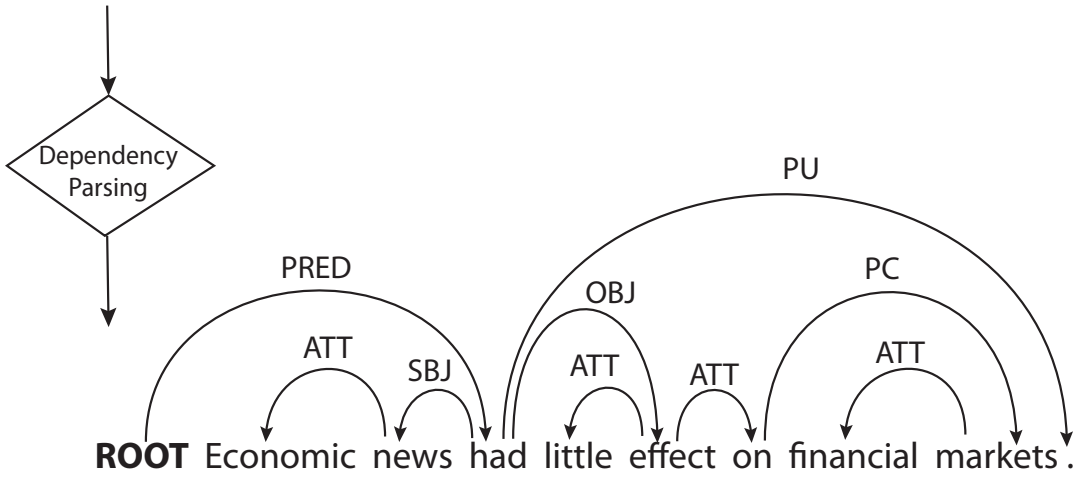
Economic news had little effect on financial markets.



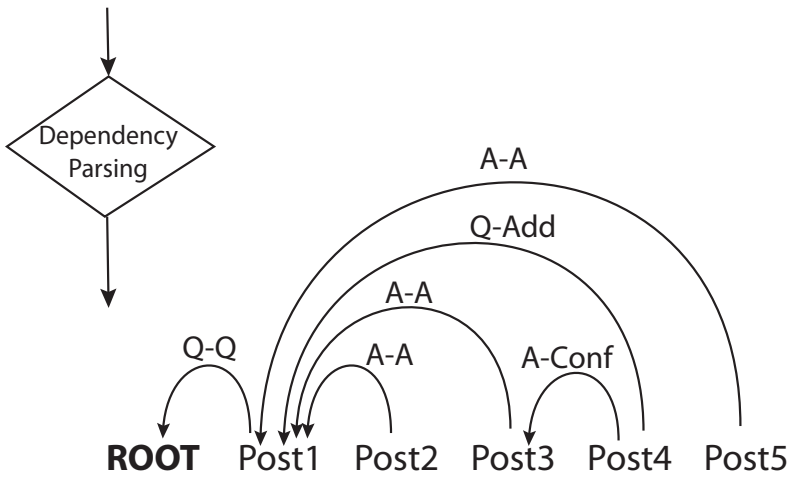
Reference: Kübler et al., 2009

Dependency Parsing of Forum Threads

Economic news had little effect on financial markets.



Post1 Post2 Post3 Post4 Post5



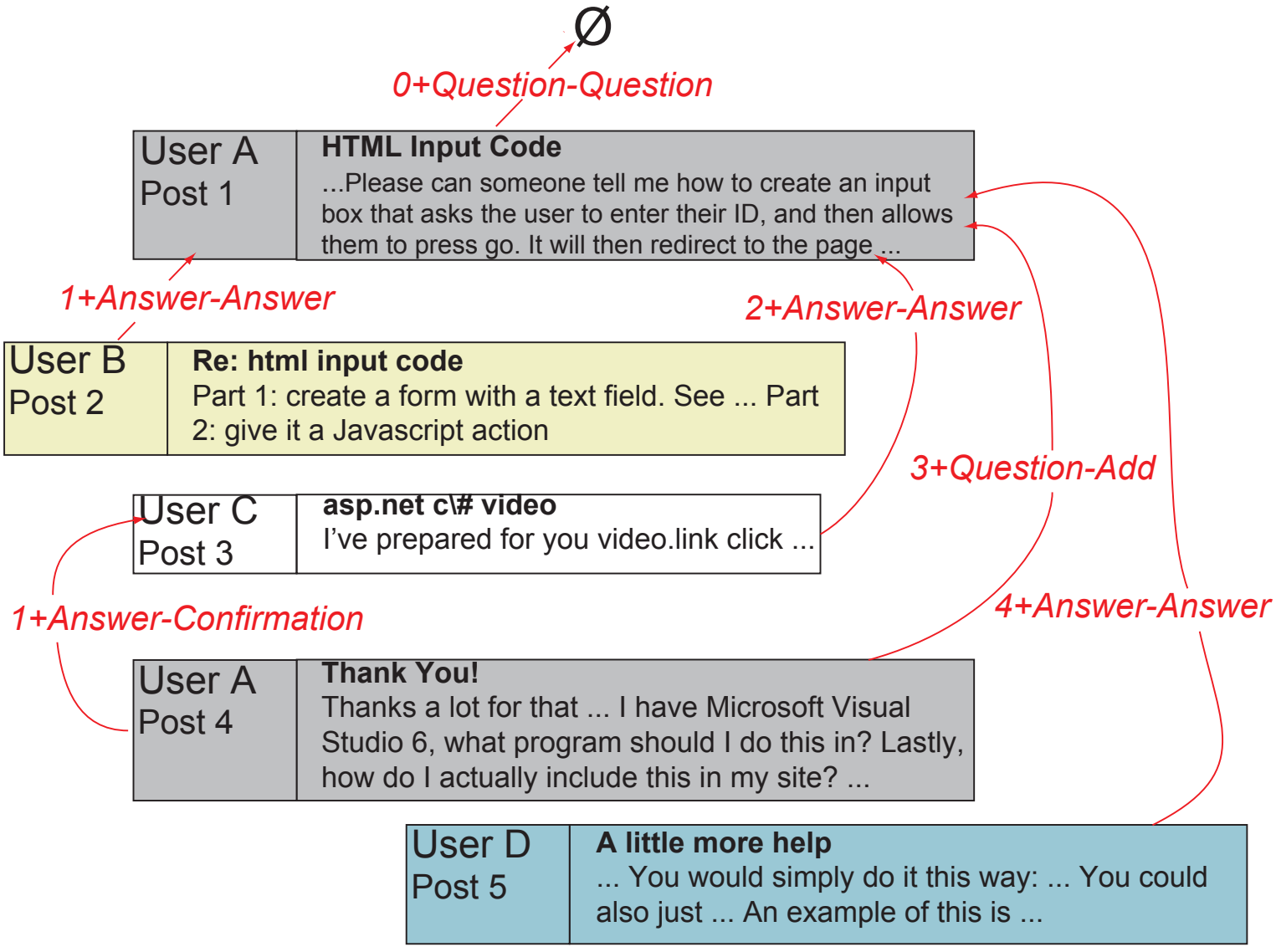
Reference: Kübler et al., 2009

Experimental Setup

Dataset

- From Kim et al. [2010], 1332 posts spanning 315 threads from CNET
- Each post is labelled with one or more links, each link is labelled with a dialogue act
  - Question
    - Question, Add, Correction, Confirmation
  - Answer
    - Answer, Add, Objection, Confirmation
  - Resolution
  - Reproduction
  - Other
- Most common label: 1+Answer-Answer (28.4%)

Recap


Ø
0+Question-Question
User A
Post 1
HTML Input Code
...Please can someone tell me how to create an input box that asks the user to enter their ID, and then allows them to press go. It will then redirect to the page ...
1+Answer-Answer
2+Answer-Answer
User B
Post 2
Re: html input code
Part 1: create a form with a text field. See ... Part 2: give it a Javascript action
3+Question-Add
User C
Post 3
asp.net c\# video
I've prepared for you video.link click ...
1+Answer-Confirmation
4+Answer-Answer
User A
Post 4
Thank You!
Thanks a lot for that ... I have Microsoft Visual Studio 6, what program should I do this in? Lastly, how do I actually include this in my site? ...
User D
Post 5
A little more help
... You would simply do it this way: ... You could also just ... An example of this is ...

Task Description

- **Main task:** joint classification of inter-post links (Link) and dialogue acts (DA)
- Explore two different learning approaches to the task
  - a linear-chain CRF (CRFSGD)
  - a dependency parser (MaltParser)
- The task is a natural fit for dependency parsing, with some special properties:
  - ⊕ strict reverse-chronological directionality (100%)
  - ⊖ non-projective dependencies (2%)
  - ⊖ multi-headedness (6%)
  - ⊖ disconnected sub-graphs (2%)

Features

- Structural features:
  - **Initiator:** binary feature indicating whether the current post's author is the thread initiator
  - **Position:** relative position of the current post
- Semantic features:
  - **TitSim:** relative location of the post which has the most similar title to the current post.
  - **PostSim:** relative location of the post which has the most similar content to the current post.
  - **Punct:** number of question marks (QusCount), exclamation marks (ExcCount) and URLs (UrlCount) in the current post.
  - **UserProf:** class distribution of the current post's author

Reference: Kim et al., 2010

An Example of Feature Representation

- The feature representation of the third post in a thread of length 8:

Feature	Value	Explanation
Initiator	True	post from the initiator
ExcCount	4	4 exclamation marks
QusCount	0	0 question marks
UrlCount	0	0 URLs
Position	0.25	$\frac{i-1}{n} = \frac{3-1}{8}$
PostSim	2	most similar to post 1
TitSim	2	most similar to post 1
UserProf	$\vec{x}$	counts for posts of each class from the same author in the training data

Classification Methodology

Evaluation Metrics

- Stratified (at the thread level) 10-fold cross-validation
- Primarily use post-level micro-averaged F-score
- Also use thread-level F-score/classification accuracy
- Significance test: randomised estimation with $p < 0.05$

Joint Classification

- Joint classification with CRF (CRFSGD)
  - **Composition:** classify the Link and DA separately, and compose the predictions to form the joint classification
  - **Combine:** combine the Link and DA labels into a single class, and apply the learner over the combined class
- Joint classification with dependency parsing (MaltParser)
  - naturally handles the combination of Link and DA

Experiments and Analysis

Post/thread-level Joint Classification F-scores

Method	CRFSGD post/thread	MaltParser post/thread
Heuristic	.515/.311	
NoFeatures	.508/.394	.533/.356
Composition	.728/.553	—
Joint +ALL	.756/.578	.738/.578
−Initiator	.745/.569	.708/.534
−Position	.750/.565	.736/.568
−PostSim	.753/.578	.737/.568
−TitSim	.760/.587	.734/.571
−Punct	.745/.571	.735/.578
−UserProf	.672/.527	.701/.536

Post/thread-level Joint Classification F-scores

Method	CRFSGD post/thread	MaltParser post/thread
Heuristic	**.515/.311**	
NoFeatures	.508/.394	.533/.356
Composition	.728/.553	—
Joint +ALL	.756/.578	.738/.578
−Initiator	.745/.569	.708/.534
−Position	.750/.565	.736/.568
−PostSim	.753/.578	.737/.568
−TitSim	.760/.587	.734/.571
−Punct	.745/.571	.735/.578
−UserProf	.672/.527	.701/.536

Post/thread-level Joint Classification F-scores

Method	CRFSGD post/thread	MaltParser post/thread
Heuristic	.515/.311	
NoFeatures	**.508**/**.394**	**.533**/**.356**
Composition	.728/.553	—
Joint +ALL	.756/.578	.738/.578
−Initiator	.745/.569	.708/.534
−Position	.750/.565	.736/.568
−PostSim	.753/.578	.737/.568
−TitSim	.760/.587	.734/.571
−Punct	.745/.571	.735/.578
−UserProf	.672/.527	.701/.536

Post/thread-level Joint Classification F-scores

Method	CRFSGD post/thread	MaltParser post/thread
Heuristic	.515/.311	
NoFeatures	.508/.394	.533/.356
Composition	**.728/.553**	—
Joint +ALL	.756/.578	.738/.578
−Initiator	.745/.569	.708/.534
−Position	.750/.565	.736/.568
−PostSim	.753/.578	.737/.568
−TitSim	.760/.587	.734/.571
−Punct	.745/.571	.735/.578
−UserProf	.672/.527	.701/.536

Post/thread-level Joint Classification F-scores

Method	CRFSGD post/thread	MaltParser post/thread
Heuristic	.515/.311	
NoFeatures	.508/.394	.533/.356
Composition	.728/.553	—
Joint +ALL	**.756/.578**	**.738/.578**
−Initiator	.745/.569	.708/.534
−Position	.750/.565	.736/.568
−PostSim	.753/.578	.737/.568
−TitSim	.760/.587	.734/.571
−Punct	.745/.571	.735/.578
−UserProf	.672/.527	.701/.536

Post/thread-level Joint Classification F-scores

Method	CRFSGD post/thread	MaltParser post/thread
Heuristic	.515/.311	
NoFeatures	.508/.394	.533/.356
Composition	.728/.553	—
Joint +ALL	.756/.578	.738/.578
−Initiator	.745/.569	.708/.534
−Position	.750/.565	.736/.568
−PostSim	.753/.578	.737/.568
−TitSim	.760/.587	.734/.571
−Punct	.745/.571	.735/.578
−UserProf	.672/.527	.701/.536

Post/thread-level Joint Classification F-scores

Method	CRFSGD **post**/thread	MaltParser **post**/thread
Heuristic	**.515**/.311	
NoFeatures	**.508**/.394	**.533**/.356
Composition	**.728**/.553	—
Joint +ALL	**.756**/.578	**.738**/.578
−Initiator	**.745**/.569	**.708**/.534
−Position	**.750**/.565	**.736**/.568
−PostSim	**.753**/.578	**.737**/.568
−TitSim	**.760**/.587	**.734**/.571
−Punct	**.745**/.571	**.735**/.578
−UserProf	**.672**/.527	**.701**/.536

Post/thread-level Joint Classification F-scores

Method	CRFSGD post/**thread**	MaltParser post/**thread**
Heuristic	.515/**.311**	
NoFeatures	.508/**.394**	.533/**.356**
Composition	.728/**.553**	—
Joint +ALL	.756/**.578**	.738/**.578**
−Initiator	.745/**.569**	.708/**.534**
−Position	.750/**.565**	.736/**.568**
−PostSim	.753/**.578**	.737/**.568**
−TitSim	.760/**.587**	.734/**.571**
−Punct	.745/**.571**	.735/**.578**
−UserProf	.672/**.527**	.701/**.536**

Post/thread-level Joint Classification F-scores

Method	CRFSGD post/thread	MaltParser post/thread
Heuristic	.515/.311	
NoFeatures	.508/.394	.533/.356
Composition	.728/.553	—
Joint +ALL	.756/.578	.738/.578
−Initiator	.745/.569	.708/.534
−Position	.750/.565	.736/.568
−PostSim	.753/.578	.737/.568
−TitSim	.760/.587	.734/.571
−Punct	.745/.571	.735/.578
−UserProf	.672/.527	.701/.536

Post/thread-level Joint Classification F-scores

Method	CRFSGD post/thread	MaltParser post/thread
Heuristic	.515/.311	
NoFeatures	.508/.394	.533/.356
Composition	.728/.553	—
Joint +ALL	**.756**/.578	**.738**/.578
−Initiator	.745/.569	.708/.534
−Position	.750/.565	.736/.568
−PostSim	.753/.578	.737/.568
−TitSim	.760/.587	.734/.571
−Punct	.745/.571	.735/.578
−UserProf	**.672**/.527	**.701**/.536

- Post-level analysis
- ⋆ UserProf has the greatest impact

Post/thread-level Joint Classification F-scores

Method	CRFSGD post/thread	MaltParser post/thread
Heuristic	.515/.311	
NoFeatures	.508/.394	.533/.356
Composition	.728/.553	—
Joint +ALL	**.756**/.578	**.738**/.578
−Initiator	**.745**/.569	**.708**/.534
−Position	.750/.565	.736/.568
−PostSim	.753/.578	.737/.568
−TitSim	.760/.587	.734/.571
−Punct	.745/.571	.735/.578
−UserProf	**.672**/.527	**.701**/.536

- Post-level analysis
- ⋆ UserProf has the greatest impact
- ⋆ Initiator affects MaltParser significantly

Post/thread-level Joint Classification F-scores

Method	CRFSGD post/thread	MaltParser post/thread
Heuristic	.515/.311	
NoFeatures	.508/.394	.533/.356
Composition	.728/.553	—
Joint +ALL	.756/.578	.738/.578
−Initiator	.745/.569	.708/.534
−Position	.750/.565	.736/.568
−PostSim	.753/.578	.737/.568
−TitSim	.760/.587	.734/.571
−Punct	.745/.571	.735/.578
−UserProf	.672/.527	.701/.536

Post/thread-level Joint Classification F-scores

Method	CRFSGD post/thread	MaltParser post/thread
Heuristic	.515/.311	
NoFeatures	.508/.394	.533/.356
Composition	.728/.553	—
Joint +ALL	.756/.578	.738/**.578**
−Initiator	.745/.569	.708/.534
−Position	.750/.565	.736/.568
−PostSim	.753/.578	.737/.568
−TitSim	.760/**.587**	.734/.571
−Punct	.745/.571	.735/**.578**
−UserProf	.672/.527	.701/.536

- Thread-level analysis
- ⋆ the best thread-level F-scores from the two learners are not significantly different

Post/thread-level Joint Classification F-scores

Method	CRFSGD post/thread	MaltParser post/thread
Heuristic	.515/.311	
NoFeatures	.508/.394	.533/.356
Composition	.728/.553	—
Joint +ALL	.756/.578	**.738**/**.578**
−Initiator	.745/.569	.708/.534
−Position	.750/.565	.736/.568
−PostSim	.753/.578	.737/.568
−TitSim	**.760**/**.587**	.734/.571
−Punct	.745/.571	.735/**.578**
−UserProf	.672/.527	.701/.536

- Thread-level analysis
- ⋆ the best thread-level F-scores from the two learners are not significantly different

User Profile Feature Analysis

- The user profile feature (UserProf) is the most effective feature for both CRFSGD and MaltParser
- To gain a deeper insight into the behaviour of the feature:
  - use *uscore* to measure the average training–test post ratio per user in cross-validation:

User Profile Feature Analysis

- The user profile feature (UserProf) is the most effective feature for both CRFSGD and MaltParser
- To gain a deeper insight into the behaviour of the feature:
  - use *uscore* to measure the average training–test post ratio per user in cross-validation:

Bin	*uscore*	Posts per user	Total users	Total posts
High	224.6	251	1	251
Medium	1∼41.7	4∼48	45	395
Low	0	2∼4	157	377
Very Low	0	1	309	309

User Profile Feature Analysis

- The user profile feature (UserProf) is the most effective feature for both CRFSGD and MaltParser
- To gain a deeper insight into the behaviour of the feature:
  - use *uscore* to measure the average training–test post ratio per user in cross-validation:

Bin	*uscore*	Posts per user	Total users	Total posts
High	224.6	251	1	251
Medium	1∼41.7	4∼48	45	395
Low	**0**	2∼4	157	377
Very Low	**0**	1	309	309

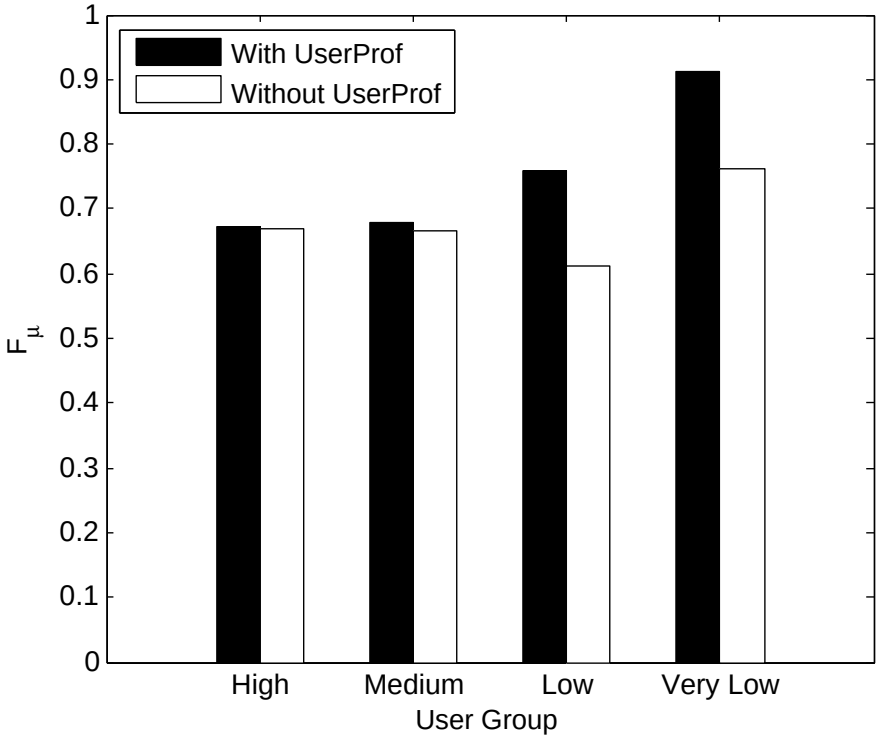
User Profile Feature Analysis

- The user profile feature (UserProf) is the most effective feature for both CRFSGD and MaltParser
- To gain a deeper insight into the behaviour of the feature:
  - use *uscore* to measure the average training–test post ratio per user in cross-validation:

Bin	*uscore*	Posts per user	Total users	Total posts
High	**224.6**	**251**	**1**	**251**
Medium	1∼41.7	4∼48	45	395
Low	0	2∼4	157	377
Very Low	0	1	309	309

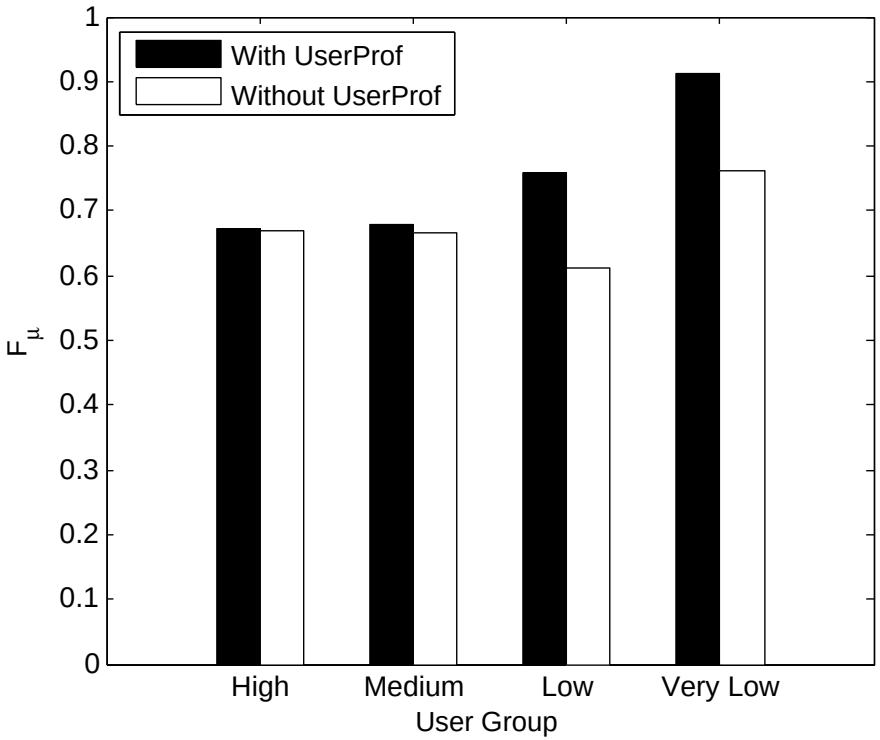
User Profile Feature Analysis

- Post-level joint classification results for users binned by *uscore*, based on CRFSGD with and without UserProf features:



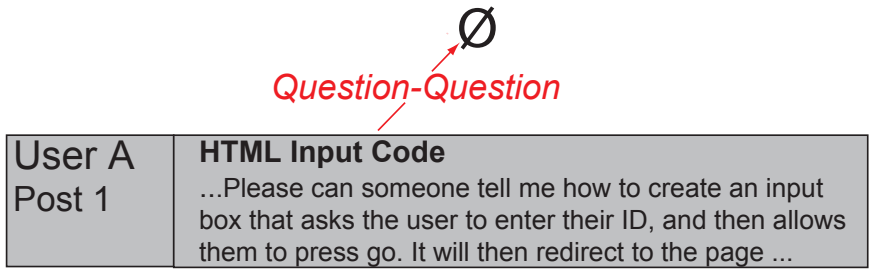
User Profile Feature Analysis

- Post-level joint classification results for users binned by *uscore*, based on CRFSGD with and without UserProf features:

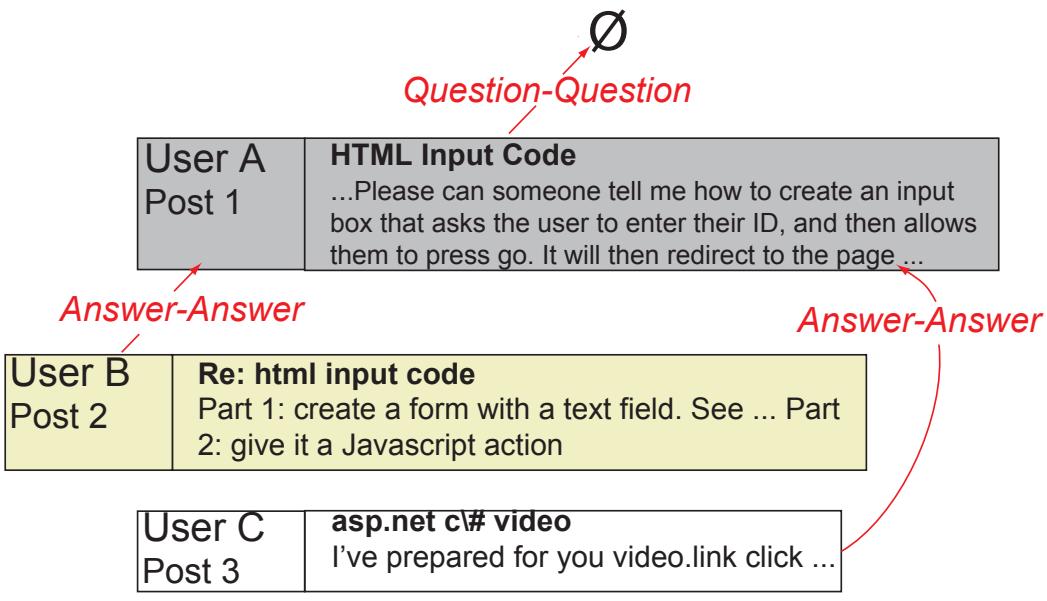


⋆ UserProf has the greatest impact for users with fewer posts.

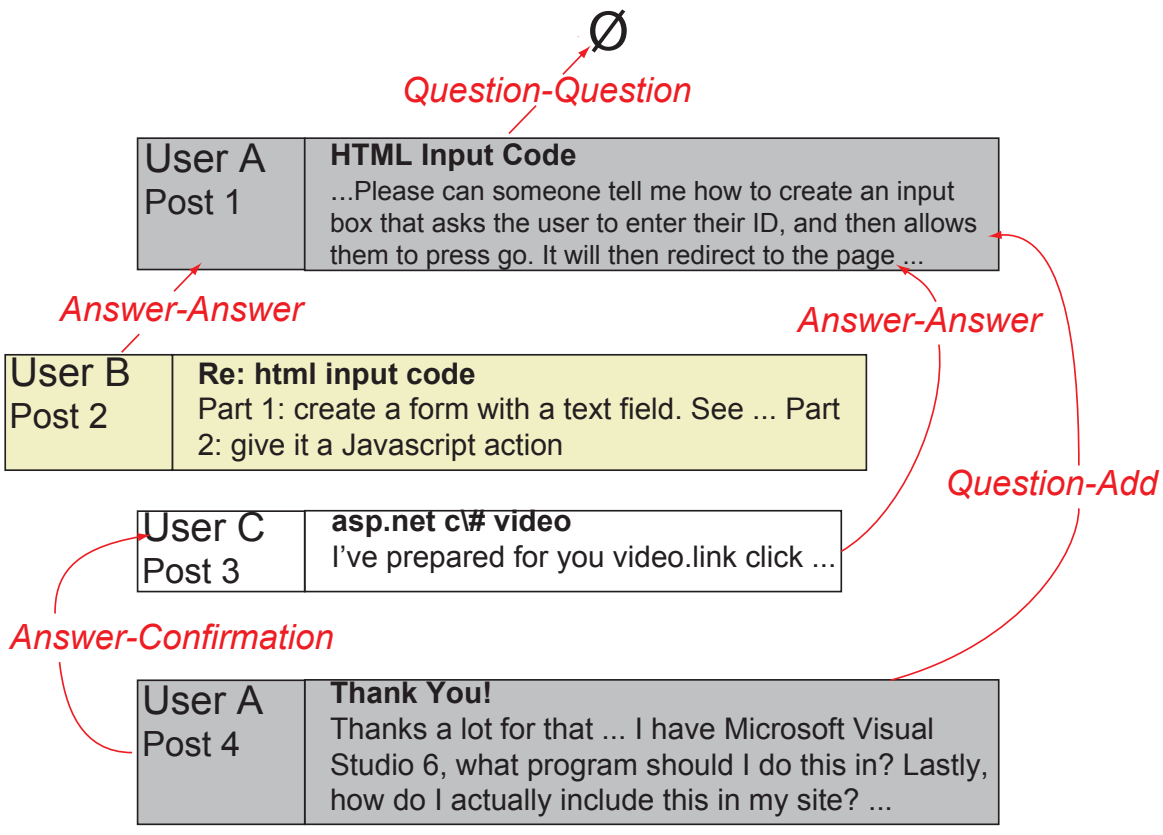
Threads Evolve Over Time



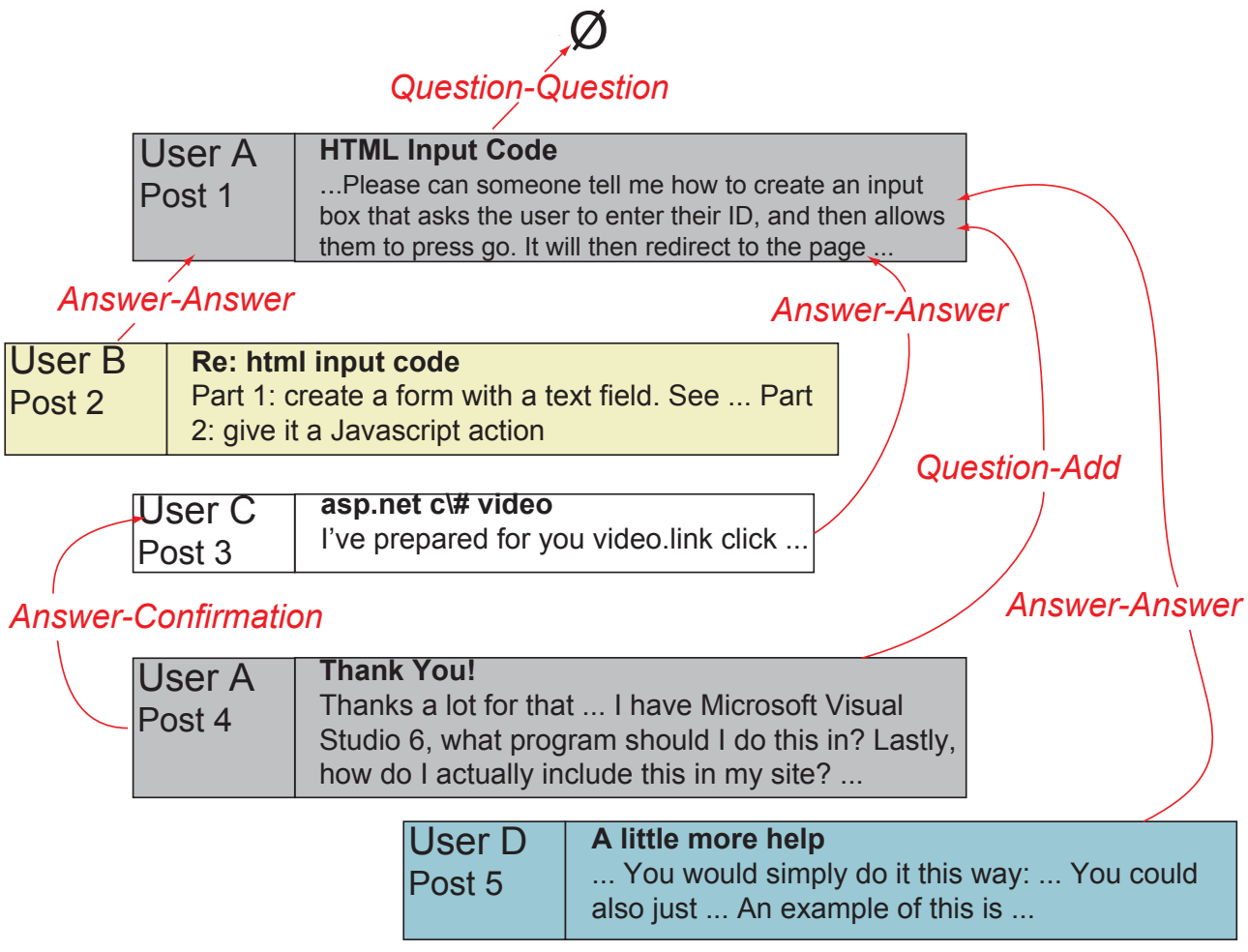
Threads Evolve Over Time



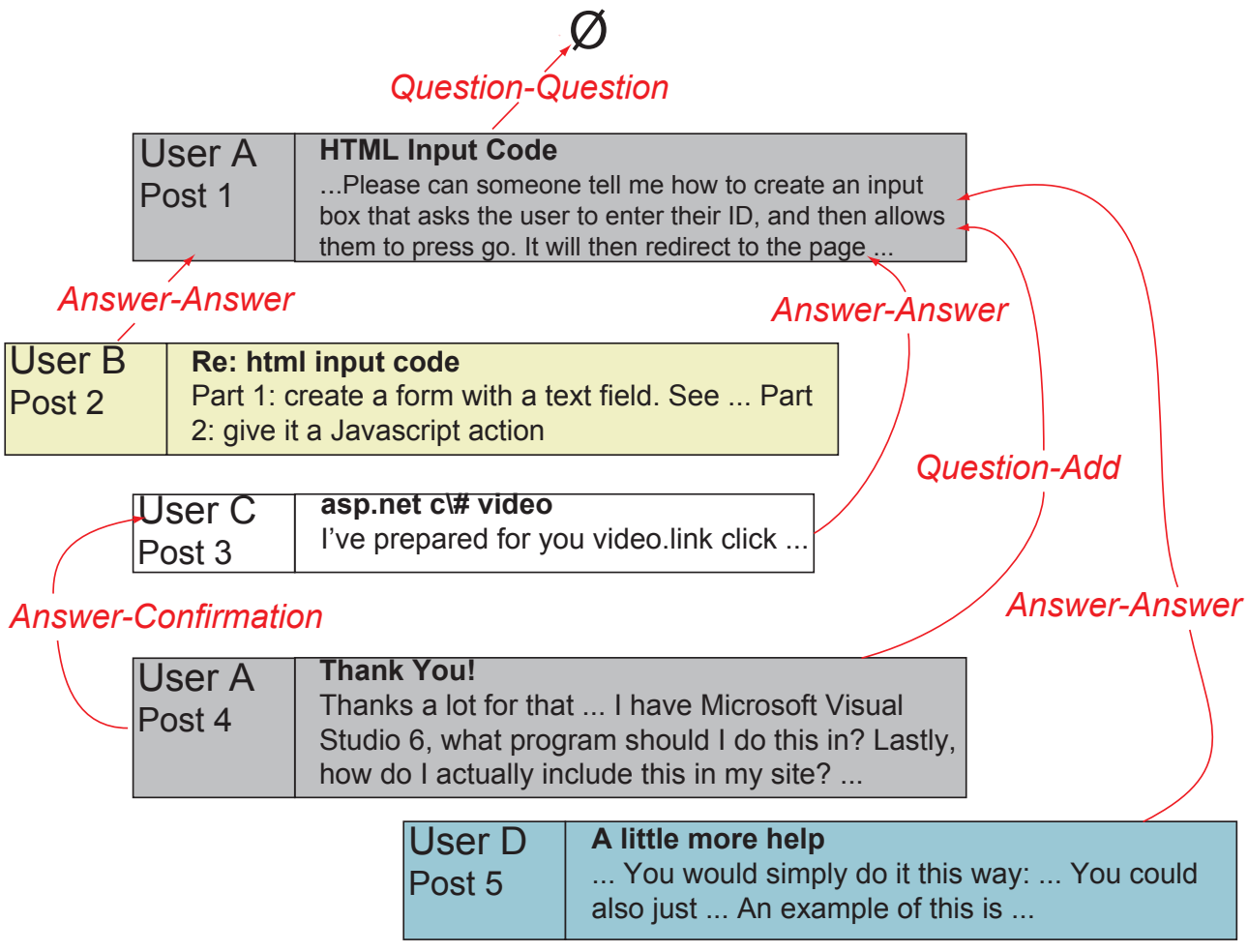
Threads Evolve Over Time



Threads Evolve Over Time



Threads Evolve Over Time



- In situ classification — compare the accuracy of different models when applied to partial threads vs. complete threads.

Classify the "Evolving Threads"

Classify the "Evolving Threads"

HTML Input Code - CNET Coding & scripting Forums

User A Post 1	**HTML Input Code** ...Please can someone tell me how to create an input box that asks the user to enter their ID, and then allows them to press go. It will then redirect to the page ...
User B Post 2	**Re: html input code** Part 1: create a form with a text field. See ... Part 2: give it a Javascript action

Classify first 2 posts

Classify the "Evolving Threads"

HTML Input Code - CNET Coding & scripting Forums

User A Post 1	**HTML Input Code** ...Please can someone tell me how to create an input box that asks the user to enter their ID, and then allows them to press go. It will then redirect to the page ...
User B Post 2	**Re: html input code** Part 1: create a form with a text field. See ... Part 2: give it a Javascript action
User C Post 3	**asp.net c\# video** I've prepared for you video.link click ...
User A Post 4	**Thank You!** Thanks a lot for that ... I have Microsoft Visual Studio 6, what program should I do this in? Lastly, how do I actually include this in my site? ...

Classify first 4 posts

Classify the "Evolving Threads"

HTML Input Code - CNET Coding & scripting Forums

User A Post 1	**HTML Input Code** ...Please can someone tell me how to create an input box that asks the user to enter their ID, and then allows them to press go. It will then redirect to the page ...
User B Post 2	**Re: html input code** Part 1: create a form with a text field. See ... Part 2: give it a Javascript action
User C Post 3	**asp.net c\# video** I've prepared for you video.link click ...
User A Post 4	**Thank You!** Thanks a lot for that ... I have Microsoft Visual Studio 6, what program should I do this in? Lastly, how do I actually include this in my site? ...
User D Post 5	**A little more help** ... You would simply do it this way: ... You could also just ... An example of this is ...

Classify all posts

Evaluation of In situ Classification

User A Post 1	**HTML Input Code** ...Please can someone tell me how to create an input box that asks the user to enter their ID, and then allows them to press go. It will then redirect to the page ...
User B Post 2	**Re: html input code** Part 1: create a form with a text field. See ... Part 2: give it a Javascript action

User A Post 1	**HTML Input Code** ...Please can someone tell me how to create an input box that asks the user to enter their ID, and then allows them to press go. It will then redirect to the page ...
User B Post 2	**Re: html input code** Part 1: create a form with a text field. See ... Part 2: give it a Javascript action
User C Post 3	**asp.net c\# video** I've prepared for you video.link click ...
User A Post 4	**Thank You!** Thanks a lot for that ... I have Microsoft Visual Studio 6, what program should I do this in? Lastly, how do I actually include this in my site? ...

Evaluate first 2 posts

User A Post 1	**HTML Input Code** ...Please can someone tell me how to create an input box that asks the user to enter their ID, and then allows them to press go. It will then redirect to the page ...
User B Post 2	**Re: html input code** Part 1: create a form with a text field. See ... Part 2: give it a Javascript action
User C Post 3	**asp.net c\# video** I've prepared for you video.link click ...
User A Post 4	**Thank You!** Thanks a lot for that ... I have Microsoft Visual Studio 6, what program should I do this in? Lastly, how do I actually include this in my site? ...
User D Post 5	**A little more help** ... You would simply do it this way: ... You could also just ... An example of this is ...

Evaluation of In situ Classification

User A Post 1	**HTML Input Code** …Please can someone tell me how to create an input box that asks the user to enter their ID, and then allows them to press go. It will then redirect to the page …
User B Post 2	**Re: html input code** Part 1: create a form with a text field. See … Part 2: give it a Javascript action

User A Post 1	**HTML Input Code** …Please can someone tell me how to create an input box that asks the user to enter their ID, and then allows them to press go. It will then redirect to the page …
User B Post 2	**Re: html input code** Part 1: create a form with a text field. See … Part 2: give it a Javascript action
User C Post 3	**asp.net c\# video** I've prepared for you video.link click …
User A Post 4	**Thank You!** Thanks a lot for that … I have Microsoft Visual Studio 6, what program should I do this in? Lastly, how do I actually include this in my site? …

Evaluate first 4 posts

User A Post 1	**HTML Input Code** …Please can someone tell me how to create an input box that asks the user to enter their ID, and then allows them to press go. It will then redirect to the page …
User B Post 2	**Re: html input code** Part 1: create a form with a text field. See … Part 2: give it a Javascript action
User C Post 3	**asp.net c\# video** I've prepared for you video.link click …
User A Post 4	**Thank You!** Thanks a lot for that … I have Microsoft Visual Studio 6, what program should I do this in? Lastly, how do I actually include this in my site? …
User D Post 5	**A little more help** … You would simply do it this way: … You could also just … An example of this is …

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	$[1,2]$	$[1,4]$	$[1,6]$	$[1,8]$	$[All]$
$[1,2]$	.947/.947	—	—	—	—
$[1,4]$	.946/.947	.836/.841	—	—	—
$[1,6]$	.946/.947	.840/.841	.800/.794	—	—
$[1,8]$	.946/.947	.840/.841	.800/.794	.780/.769	—
$[All]$	.946/.946	.840/.838	.800/.791	.776/.767	.756/.738

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	[1, 2]	[1, 4]	[1, 6]	[1, 8]	[*All*]
[1, 2]	.947/.947	—	—	—	—
[1, 4]	.946/.947	.836/.841	—	—	—
[1, 6]	.946/.947	.840/.841	.800/.794	—	—
[1, 8]	.946/.947	.840/.841	.800/.794	.780/.769	—
[*All*]	.946/.946	.840/.838	.800/.791	.776/.767	.756/.738

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	$[1,2]$	$[1,4]$	$[1,6]$	$[1,8]$	$[All]$
$[1,2]$	.947/.947	—	—	—	—
$[1,4]$	.946/.947	.836/.841	—	—	—
$[1,6]$	.946/.947	.840/.841	.800/.794	—	—
$[1,8]$	.946/.947	.840/.841	.800/.794	.780/.769	—
$[All]$	.946/.946	.840/.838	.800/.791	.776/.767	.756/.738

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	[1, 2]	[1, 4]	[1, 6]	[1, 8]	[*All*]
[1, 2]	.947/.947	—	—	—	—
[1, 4]	.946/.947	.836/.841	—	—	—
[1, 6]	.946/.947	.840/.841	.800/.794	—	—
[1, 8]	.946/.947	.840/.841	.800/.794	.780/.769	—
[*All*]	.946/.946	.840/.838	.800/.791	.776/.767	.756/.738

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	$[1,2]$	$[1,4]$	$[1,6]$	$[1,8]$	$[All]$
$[1,2]$	.947/.947	—	—	—	—
$[1,4]$	.946/.947	.836/.841	—	—	—
$[1,6]$	.946/.947	.840/.841	.800/.794	—	—
$[1,8]$	.946/.947	.840/.841	.800/.794	.780/.769	—
$[All]$	.946/.946	.840/.838	.800/.791	.776/.767	.756/.738

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	[1, 2]	[1, 4]	[1, 6]	[1, 8]	[*All*]
[1, 2]	.947/.947	—	—	—	—
[1, 4]	.946/.947	.836/.841	—	—	—
[1, 6]	.946/.947	.840/.841	.800/.794	—	—
[1, 8]	.946/.947	.840/.841	.800/.794	.780/.769	—
[*All*]	.946/.946	.840/.838	.800/.791	.776/.767	.756/.738

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	$[1,2]$	$[1,4]$	$[1,6]$	$[1,8]$	$[All]$
$[1,2]$	.947/.947	—	—	—	—
$[1,4]$	.946/.947	.836/.841	—	—	—
$[1,6]$	.946/.947	.840/.841	.800/.794	—	—
$[1,8]$	.946/.947	.840/.841	.800/.794	.780/.769	—
$[All]$	.946/.946	.840/.838	.800/.791	.776/.767	.756/.738

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	$[1,2]$	$[1,4]$	$[1,6]$	$[1,8]$	$[All]$
$[1,2]$	.947/.947	—	—	—	—
$[1,4]$	.946/.947	.836/.841	—	—	—
$[1,6]$	.946/.947	.840/.841	.800/.794	—	—
$[1,8]$	.946/.947	.840/.841	.800/.794	.780/.769	—
$[All]$	.946/.946	.840/.838	.800/.791	.776/.767	.756/.738

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	$[1,2]$	$[1,4]$	$[1,6]$	$[1,8]$	$[All]$
$[1,2]$	.947/.947	—	—	—	—
$[1,4]$	.946/.947	.836/.841	—	—	—
$[1,6]$	.946/.947	.840/.841	.800/.794	—	—
$[1,8]$	.946/.947	.840/.841	.800/.794	.780/.769	—
$[All]$	.946/.946	.840/.838	.800/.791	.776/.767	.756/.738

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	$[1,2]$	$[1,4]$	$[1,6]$	$[1,8]$	$[All]$
$[1,2]$	.947/.947	—	—	—	—
$[1,4]$	.946/.947	.836/.841	—	—	—
$[1,6]$	.946/.947	.840/.841	.800/.794	—	—
$[1,8]$	.946/.947	.840/.841	.800/.794	.780/.769	—
$[All]$	.946/.946	.840/.838	.800/.791	.776/.767	.756/.738

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	[1, 2]	[1, 4]	[1, 6]	[1, 8]	[*All*]
[1, 2]	.947/.947	—	—	—	—
[1, 4]	.946/.947	.836/.841	—	—	—
[1, 6]	.946/.947	.840/.841	.800/.794	—	—
[1, 8]	.946/.947	.840/.841	.800/.794	.780/.769	—
[*All*]	.946/.946	.840/.838	.800/.791	.776/.767	.756/.738

⋆ no evaluation of [1,4] for sub-thread [1,2]

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	[1, 2]	[1, 4]	[1, 6]	[1, 8]	[*All*]
[1, 2]	.947/.947	—	—	—	—
[1, 4]	.946/.947	.836/.841	—	—	—
[1, 6]	.946/.947	.840/.841	.800/.794	—	—
[1, 8]	.946/.947	.840/.841	.800/.794	.780/.769	—
[*All*]	.946/.946	.840/.838	.800/.791	.776/.767	.756/.738

⋆ no evaluation of [1,4] for sub-thread [1,2]

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	$[1,2]$	$[1,4]$	$[1,6]$	$[1,8]$	$[All]$
$[1,2]$	.947/.947	—	—	—	—
$[1,4]$	.946/.947	.836/.841	—	—	—
$[1,6]$	.946/.947	.840/.841	.800/.794	—	—
$[1,8]$	.946/.947	.840/.841	.800/.794	.780/.769	—
$[All]$	.946/.946	.840/.838	.800/.791	.776/.767	.756/.738

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	$[1,2]$	$[1,4]$	$[1,6]$	$[1,8]$	$[All]$
$[1,2]$	**.947**/**.947**	—	—	—	—
$[1,4]$	**.946**/**.947**	.836/.841	—	—	—
$[1,6]$	**.946**/**.947**	.840/.841	.800/.794	—	—
$[1,8]$	**.946**/**.947**	.840/.841	.800/.794	.780/.769	—
$[All]$	**.946**/**.946**	.840/.838	.800/.791	.776/.767	.756/.738

⋆ both learners are very robust over partial threads

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	[1, 2]	[1, 4]	[1, 6]	[1, 8]	[*All*]
[1, 2]	.947/.947	—	—	—	—
[1, 4]	.946/.947	**.836**/**.841**	—	—	—
[1, 6]	.946/.947	**.840**/**.841**	.800/.794	—	—
[1, 8]	.946/.947	**.840**/**.841**	.800/.794	.780/.769	—
[*All*]	.946/.946	**.840**/**.838**	.800/.791	.776/.767	.756/.738

⋆ both learners are very robust over partial threads

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	$[1,2]$	$[1,4]$	$[1,6]$	$[1,8]$	$[All]$
$[1,2]$	.947/.947	—	—	—	—
$[1,4]$	.946/.947	.836/.841	—	—	—
$[1,6]$	.946/.947	.840/.841	**.800**/**.794**	—	—
$[1,8]$	.946/.947	.840/.841	**.800**/**.794**	.780/.769	—
$[All]$	.946/.946	.840/.838	**.800**/**.791**	.776/.767	.756/.738

⋆ both learners are very robust over partial threads

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	[1, 2]	[1, 4]	[1, 6]	[1, 8]	[*All*]
[1, 2]	.947/.947	—	—	—	—
[1, 4]	.946/.947	.836/.841	—	—	—
[1, 6]	.946/.947	.840/.841	.800/.794	—	—
[1, 8]	.946/.947	.840/.841	.800/.794	**.780**/**.769**	—
[*All*]	.946/.946	.840/.838	.800/.791	**.776**/**.767**	.756/.738

⋆ both learners are very robust over partial threads

In Situ Classification

- Link-DA F-score for CRFSGD/MaltParser for in situ classification over sub-threads of different lengths, broken down over different post extents

Test \ B/down	$[1,2]$	$[1,4]$	$[1,6]$	$[1,8]$	$[All]$
$[1,2]$	.947/.947	—	—	—	—
$[1,4]$	.946/.947	.836/.841	—	—	—
$[1,6]$	.946/.947	.840/.841	.800/.794	—	—
$[1,8]$	.946/.947	.840/.841	.800/.794	.780/.769	—
$[All]$	.946/.946	.840/.838	.800/.791	.776/.767	.756/.738

⋆ therefore, our method can be robustly applied to real-time analysis of dynamically evolving threads.

Summary

- Joint classification of web user forum thread discourse structure
- Conclusion:
  - joint classification: achieve state-of-the-art results
  - in situ classification: our method is robust over dynamically evolving threads
- Future work:
  - multi-headedness and disconnected subgraphs in dependency parsing
  - meta-classification
  - unsupervised user-level features

Questions?

References I

Timothy Baldwin, David Martinez, and Richard B. Penman. Automatic thread classification for Linux user forum information access. In *Proceedings of the 12th Australasian Document Computing Symposium (ADCS 2007)*, pages 72–79, Melbourne, Australia, 2007.

Jonathan L. Elsas and Jaime G. Carbonell. It pays to be picky: An evaluation of thread retrieval in online forums. In *Proc. SIGIR'09*, pages 714–715, 2009.

Su Nam Kim, Li Wang, and Timothy Baldwin. Tagging and linking web forum posts. In *Proceedings of the 14th Conference on Computational Natural Language Learning (CoNLL-2010)*, pages 192–202, Uppsala, Sweden, 2010.

Sandra Kübler, Ryan McDonald, and Joakim Nivre. Dependency parsing. *Synthesis Lectures on Human Language Technologies*, 2(1):1–127, 2009.

Marco Lui and Timothy Baldwin. You are what you post: User-level features in threaded discourse. In *Proceedings of the 14th Australasian Document Computing Symposium (ADCS 2009)*, Sydney, Australia, 2009.

Marco Lui and Timothy Baldwin. Classifying user forum participants: Separating the gurus from the hacks, and other tales of the internet. In *Proceedings of the 2010 Australasian Language Technology Workshop (ALTW 2010)*, Melbourne, Australia, 2010.

Jangwon Seo, W. Bruce Croft, and David A. Smith. Online community search using thread structure. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management (CIKM 2009)*, pages 1907–1910, Hong Kong, China, 2009.

Dataset Statistics

Thread len	Count
2	105
3	59
4	57
5	25
6	18
7	10
8	14
9	7
10	4
11	4
12	3
13	7
14	2

Link	Count
0	321
1	801
2	151
3	49
4	27
5	17
6	11
7	11
8	4
9	2
10	2
11	1

Dialogue Act	Count
Question-question	316
Question-add	157
Question-correction	3
Question-confirmation	54
Answer-answer	560
Answer-add	108
Answer-objection	29
Answer-confirmation	14
Resolution	118
Reproduction	20
Other	18

Dataset Statistics

- Of the 1332 posts, 65 posts have multiple labels, 22 posts link to two different links, 43 posts have one head with multiple labels.
- 5 threads contain non-projective dependencies

Component-wise Classification

- One approach to joint classification with CRFSGD is to firstly conduct component-wise classification over Link and DA separately, and compose the predictions
- Post/thread-level component-wise classification F-scores for Link and DA classes:

Method	Link	DA
Kim et al. [2010]	.863 / .676	.751 / .543
CRFSGD	.891 / .727	.795 / .609

Component-wise Classification

- One approach to joint classification with CRFSGD is to firstly conduct component-wise classification over Link and DA separately, and compose the predictions
- Post/thread-level component-wise classification F-scores for Link and DA classes:

Method	Link	DA
Kim et al. [2010]	.863 / .676	.751 / .543
CRFSGD	.891 / .727	.795 / .609

  - ⋆ at the component-wise tasks, our method is superior to Kim et al. [2010], based on a different learner and slightly different feature set.

Joint Classification Decomposition

- Post/thread-level Link and DA F-scores from component-wise classification, and from Link-DA classification decomposition ("*" signifies a significantly *worse* result than the **best** result in that column)

Approaches	Link	DA
Component-wise	.891 / .727*	**.795 / .609**
CRFSGD decomp	**.893 / .749**	.785 / .603
MaltParser decomp	.870* / .730*	.766* / .571*

Joint Classification Decomposition

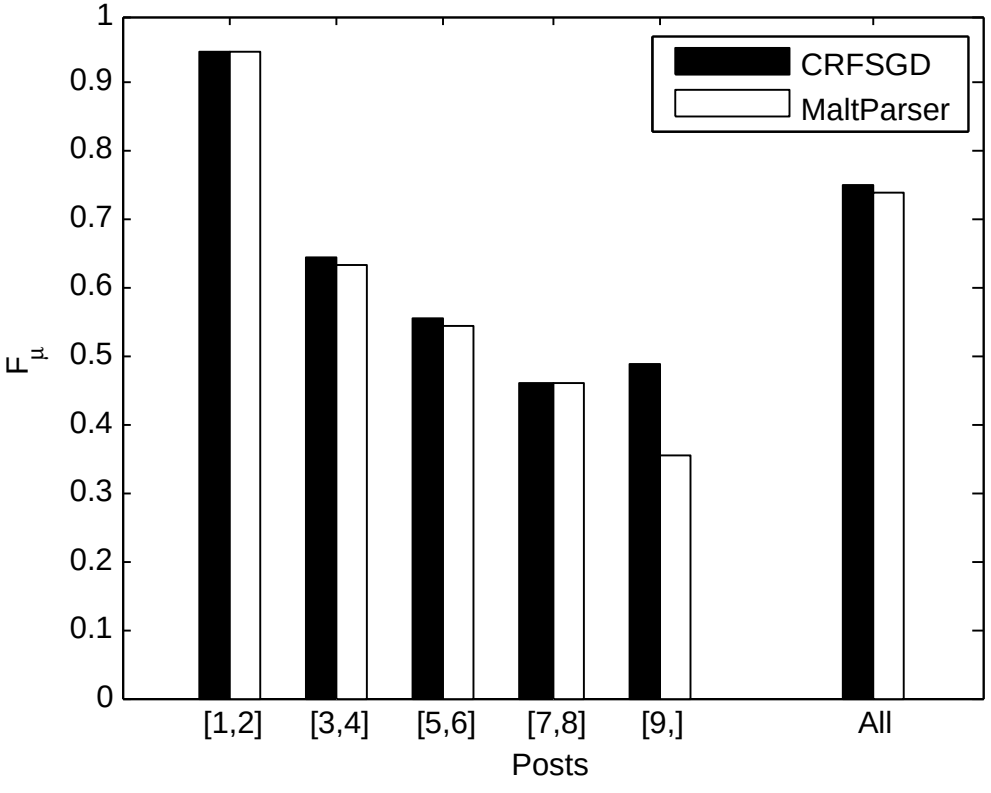
- Post/thread-level Link and DA F-scores from component-wise classification, and from Link-DA classification decomposition ("*" signifies a significantly *worse* result than the **best** result in that column)

Approaches	Link	DA
Component-wise	.891 / .727*	**.795 / .609**
CRFSGD decomp	**.893 / .749**	.785 / .603
MaltParser decomp	.870* / .730*	.766* / .571*

  - ⋆ the decomposed predictions are mostly slightly worse than the results for the component-wise classification, despite achieving higher F-score for the joint classification task
  - ⋆ simply due to the combined method tending to get both labels correct or both labels wrong, for a given post

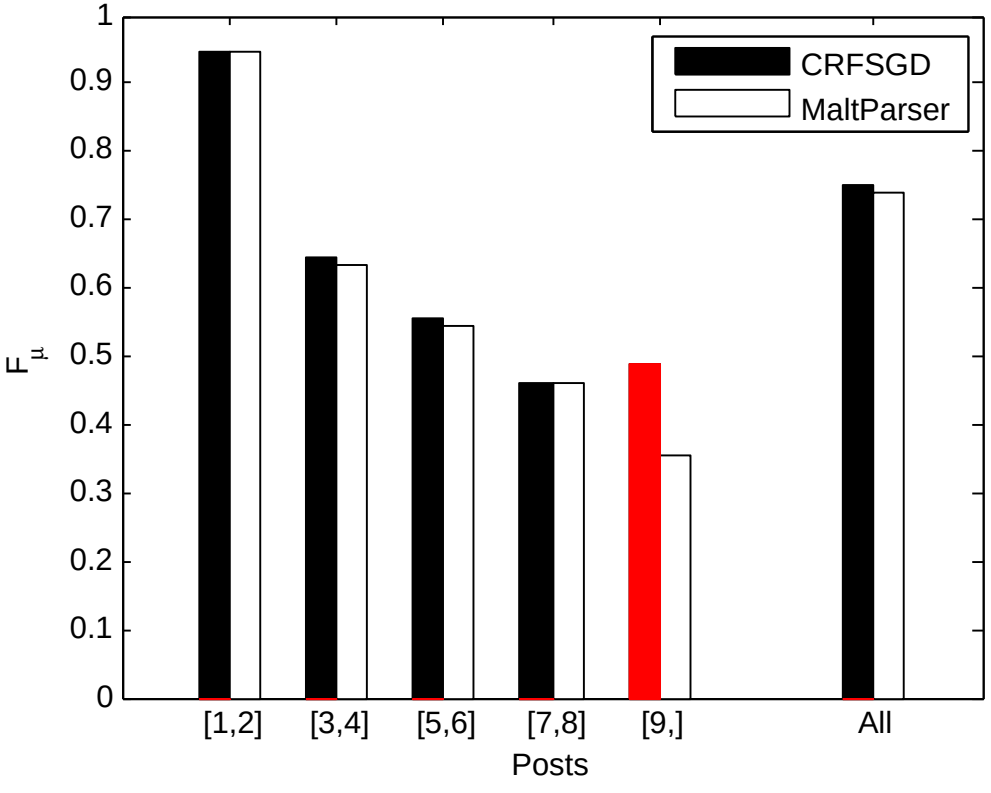
Post Position-based Result Breakdown

- How accurate are the predictions at different depths?
- Breakdown of post-level Link-DA results for CRFSGD and MaltParser based on post position:



Post Position-based Result Breakdown

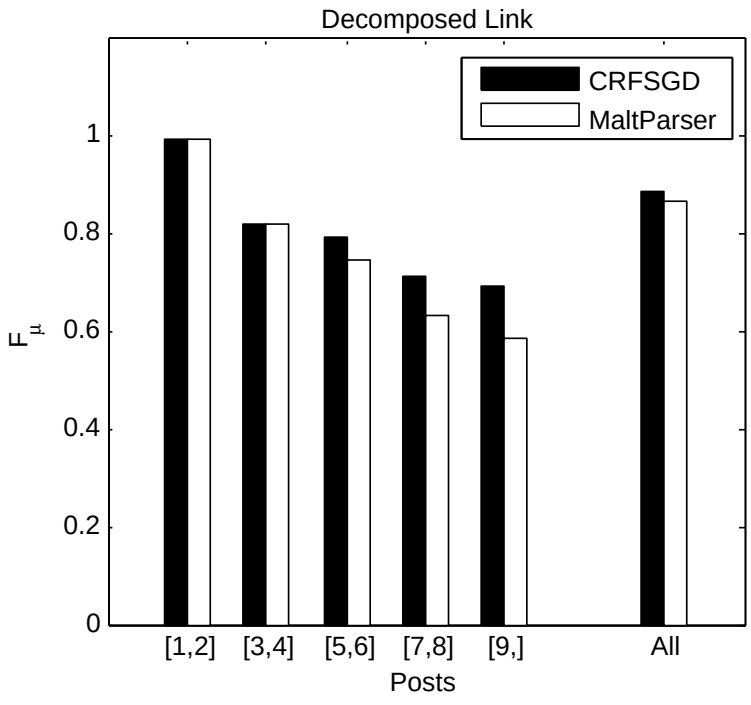
- How accurate are the predictions at different depths?
- Breakdown of post-level Link-DA results for CRFSGD and MaltParser based on post position:

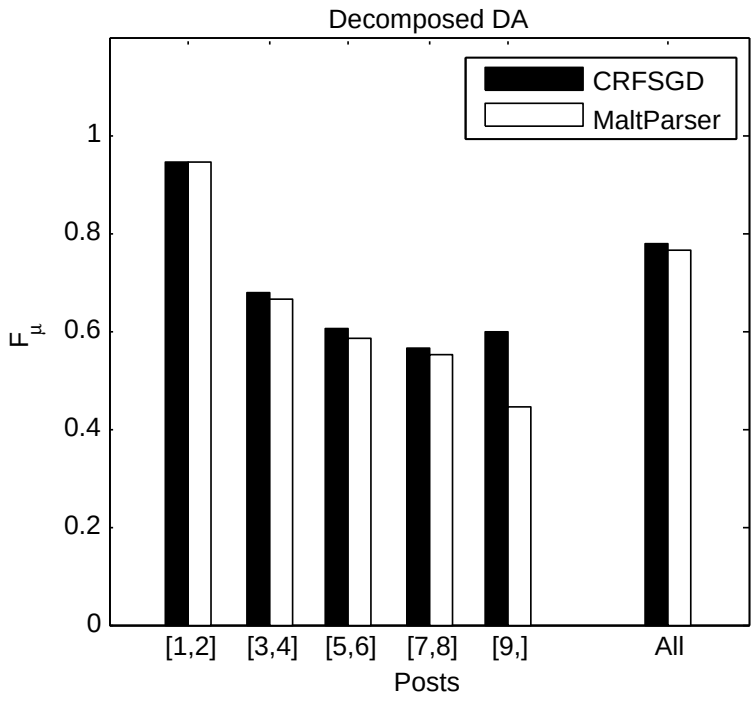


★ the results for CRFSGD improve for later posts

Post Position-based Result Breakdown

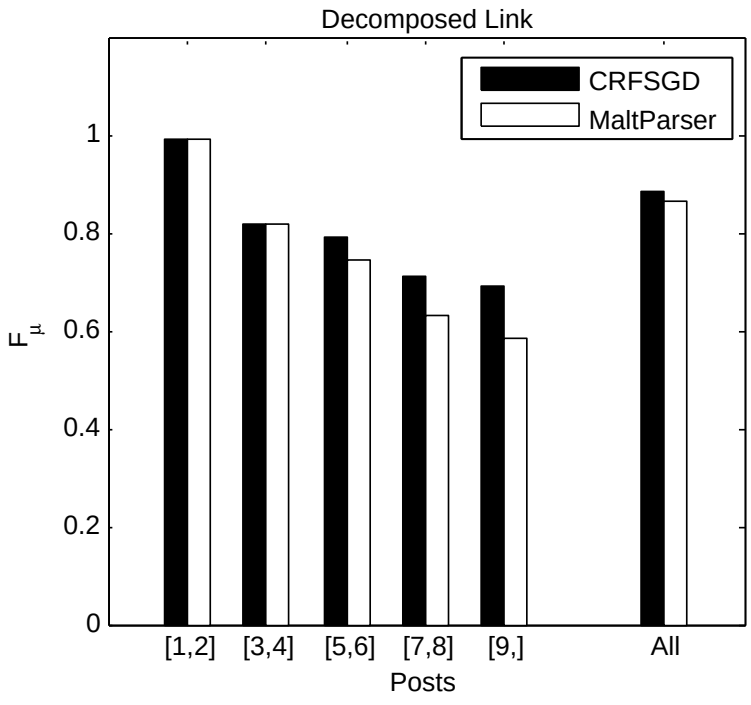
- Breakdown of post-level Link and DA F-score based on the decomposition of CRFSGD and MaltParser classifications:

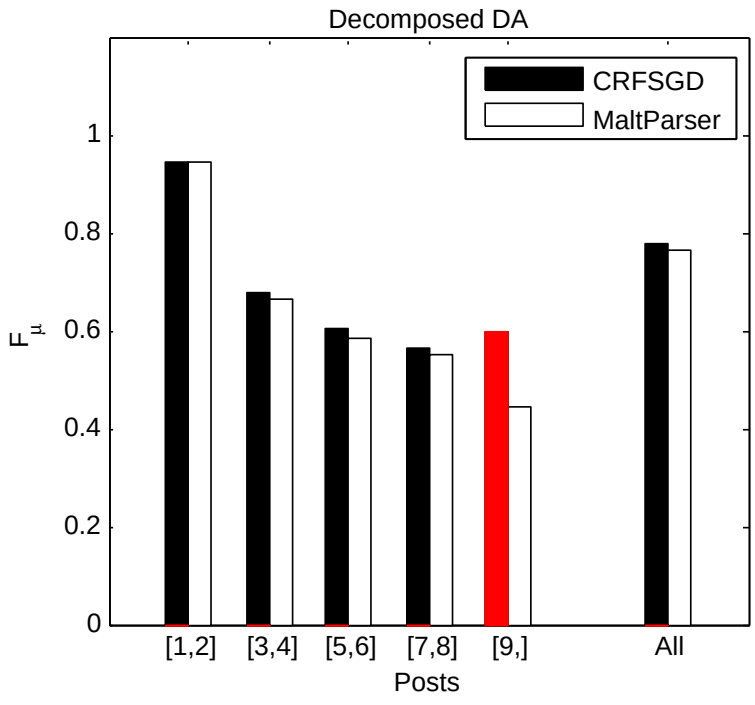




Post Position-based Result Breakdown

- Breakdown of post-level Link and DA F-score based on the decomposition of CRFSGD and MaltParser classifications:





★ the anomaly for CRFSGD comes from the DA component

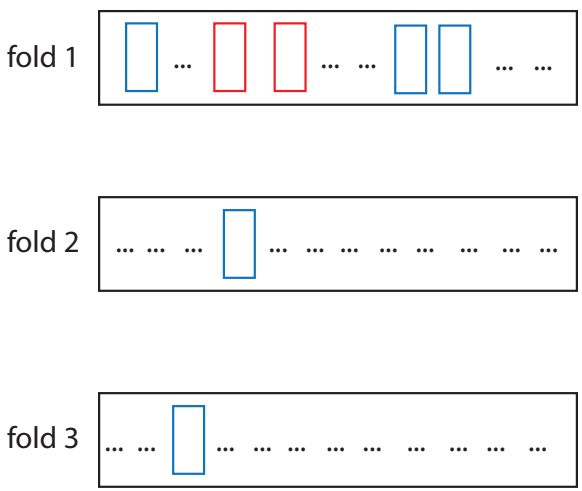
User Profile Feature Analysis

- The user profile feature (UserProf) is the most effective feature for both CRFSGD and MaltParser
- To gain a deeper insight into the behaviour of the feature:
  - use *uscore* to measure the average training–test post ratio per user in cross-validation:

$$uscore_i = \frac{\sum_{j=1}^{n_i} s_{p_{i,j}}}{n_i}$$

User Profile Feature Analysis

- The user profile feature (UserProf) is the most effective feature for both CRFSGD and MaltParser
- To gain a deeper insight into the behaviour of the feature:
  - use *uscore* to measure the average training–test post ratio per user in cross-validation:



train: 2, 3 2 + 2 +2
test: 1 null

Uscore = (2+2+2+4+4)/5

train: 1, 3 4
test: 2 0

train: 1, 2 4
test: 3 0

Uscore = (0+0)/2

Characteristics of online forum data

- Different from plain text documents
  - Complex structures
  - Posts are dynamic
  - Informal language is used
- Different from CQAs and FAQs
  - Broad and shallow vs. specific and in-depth
  - Longer history and more data
  - Multi-purpose
  - Asynchronous

CNET Forums and Sub-forums

Forum	Sub-forum		
Operating Systems	Windows 7	Windows Vista	Windows XP
	Windows 2000/NT	Windows ME	Windows 95/98
	Windows Mobile	Mac OS	Linux
Software	Audio & video	Browsers	CNET Download site
	E-mail, chat, & VoIP	Mac software	Office & productivity
	PC utilities	Photography & design	Spyware, viruses, & security
	Webware	Windows Live	
Hardware	Dell	Desktops	Laptops
	Mac hardware	Networking & wireless	PC hardware
	Peripherals	Storage	
Web Development	Coding & scripting	Web design & hosting	

Table: Data source forums and sub-forums

Thread Characteristic Classification

- Timothy Baldwin, David Martinez, and Richard B. Penman. Automatic thread classification for Linux user forum information access. In *Proceedings of the 12th Australasian Document Computing Symposium (ADCS 2007)*, pages 72–79, Melbourne, Australia, 2007.
- In the context of Linux web user forums
- Focus on classifying threads according to:
  - Task orientation
  - Completeness
  - Solvedness

Reference: Baldwin et al., 2007

Classifying User Forum Participants

- User characteristic classification
  - Clarity
  - Proficiency
  - Positivity
  - Effort

Reference: Lui and Baldwin, 2010

User-level Features in Threaded Discourse

- Describe users based on their posts
- Based on existing techniques
- User-level features for post rating
  - Aggregate: aggregation over features describing individual posts
  - Network-Based: Author Network and Thread Network

Reference: Lui and Baldwin, 2009

An Evaluation of Thread Retrieval in Online Forums

- Treat the task as an information retrieval task
- Findings:
  - thread structure is important in thread ranking
  - selective models outperform inclusive models

Reference: Elsas and Carbonell, 2009

Thread Retrieval Using Thread Structure

- Treat the task as an information retrieval task
- Goals:
  - discover and annotate thread structures, based on interactions between community members
  - improve retrieval performance by exploiting the thread structure

Reference: Seo et al., 2009

Related Work

Related Work

- Build directly on Kim et al. [2010], where the dialogue act set was proposed. The basic methodology was applied to one-to-one live chat data.
- Discourse disentanglement
  - over conversation threads or document threads
  - assume a tree structure, an acyclic graph structure, or a cyclic chain graph structure
- Dialogue act tagging
  - over conversation speech, email, instant messaging, edited documents, or online forums

Related Work

- Joint classification
  - segmentation and dialogue act classification
  - parsing and semantic role labelling (SRL)
  - parsing and named entity recognition (NER)
  - WSD of prepositions and SRL of prepositional phrases
- Research on forums
  - user-level research
  - information retrieval
  - post-level classification
  - initiation-response pair extraction