

(1) Evaluation Result when Considering Links ('Survey'-type Evaluation)

Summary Statistics	
Run ID	GRACE-tt-01
Subtask	Subtask A (survey)
Relevant Level	Level1 (S or A)
Query Method	automatic
Topic Part	TITLE only
Contents/Links Used	contents only
Number of Topics	35
Total number of documents over all topics	
Retrieved:	34996
Relevant:	3893
Rel-ret:	2284

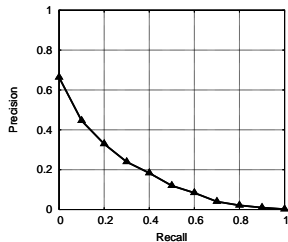
Summary Statistics	
Run ID	GRACE-tt-01
Subtask	Subtask A (survey)
Relevant Level	Level2 (S, A or B)
Query Method	automatic
Topic Part	TITLE only
Contents/Links Used	contents only
Number of Topics	35
Total number of documents over all topics	
Retrieved:	34996
Relevant:	5687
Rel-ret:	3157

Recall Level Precision Averages	
Recall	Precision
0.00	0.6632
0.10	0.4456
0.20	0.3292
0.30	0.2389
0.40	0.1832
0.50	0.1196
0.60	0.0837
0.70	0.0397
0.80	0.0209
0.90	0.0097
1.00	0.0017
Average precision over all relevant documents	
non-interpolated	0.1716

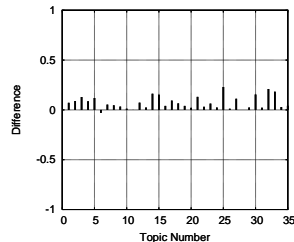
Document Level Averages	
	Precision
At 5 docs	0.4114
At 10 docs	0.3629
At 15 docs	0.3371
At 20 docs	0.3157
At 30 docs	0.2819
At 100 docs	0.2126
At 200 docs	0.1634
At 500 docs	0.1035
At 1000 docs	0.0653
R-Precision (precision after R docs retrieved (where R is the number of relevant documents))	
Exact	0.2062

Recall Level Precision Averages	
Recall	Precision
0.00	0.7922
0.10	0.4811
0.20	0.3576
0.30	0.2642
0.40	0.2159
0.50	0.1460
0.60	0.0976
0.70	0.0476
0.80	0.0246
0.90	0.0066
1.00	0.0021
Average precision over all relevant documents	
non-interpolated	0.1930

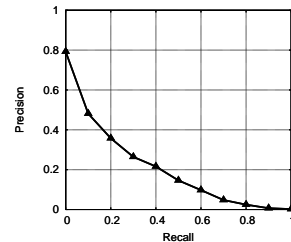
Document Level Averages	
	Precision
At 5 docs	0.4971
At 10 docs	0.4514
At 15 docs	0.4286
At 20 docs	0.3986
At 30 docs	0.3619
At 100 docs	0.2851
At 200 docs	0.2223
At 500 docs	0.1413
At 1000 docs	0.0902
R-Precision (precision after R docs retrieved (where R is the number of relevant documents))	
Exact	0.2491



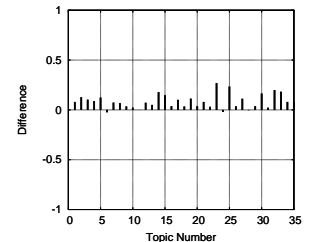
Recall-Precision Curve



Difference from Median in Average Precision per Topic (over 23 'automatic' and 'TITLE only' runs)\*

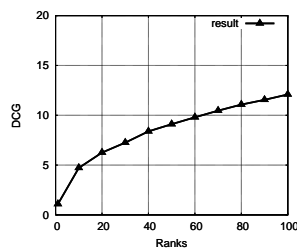


Recall-Precision Curve



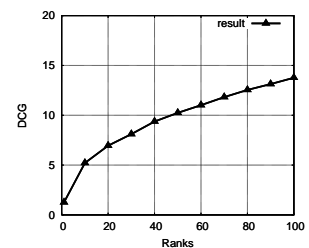
Difference from Median in Average Precision per Topic (over 23 'automatic' and 'TITLE only' runs)\*

Discounted ( $\log_2$ ) Cumulative Gain Averages ( $G_S = 3, G_A = 2, G_B = 1$ )	
	DCG
At 1 doc	1.0857
At 10 docs	4.7150
At 20 docs	6.2519
At 30 docs	7.2502
At 40 docs	8.3831
At 50 docs	9.1038
At 60 docs	9.7938
At 70 docs	10.4621
At 80 docs	11.0719
At 90 docs	11.5486
At 100 docs	12.0872
At 200 docs	15.4686
At 500 docs	20.3703
At 1000 docs	23.3912



Discounted Cumulative Gain Curve ( $G_S = 3, G_A = 2, G_B = 0$ )

Discounted ( $\log_2$ ) Cumulative Gain Averages ( $G_S = 3, G_A = 2, G_B = 0$ )	
	DCG
At 1 doc	1.2571
At 10 docs	5.2086
At 20 docs	6.9474
At 30 docs	8.1051
At 40 docs	9.3705
At 50 docs	10.2429
At 60 docs	11.0119
At 70 docs	11.8217
At 80 docs	12.5414
At 90 docs	13.1294
At 100 docs	13.7638
At 200 docs	17.7729
At 500 docs	23.5349
At 1000 docs	27.1963



Discounted Cumulative Gain Curve ( $G_S = 3, G_A = 2, G_B = 1$ )

\* Topic Numbers: ( 1,2,3,4, 5, 6, 7, 8, 9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31,32,33,34,35 )  
 Topic IDs: ( 1,3,4,6,19,21,22,23,28,29,34,44,45,55,58,61,62,63,65,68,70,71,73,74,76,80,82,84,86,88,91,95,97,98,99 )

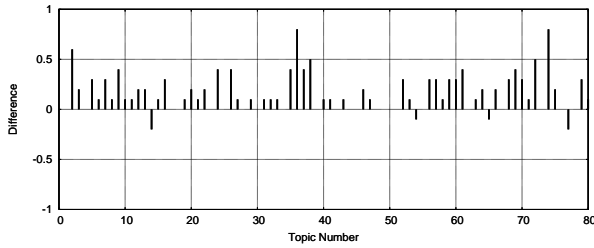
## (2) Evaluation Result when Considering Links ('Target'-type Evaluation)

Summary Statistics	
Run ID	GRACE-tt-01
Subtask	Subtask A (target)
Relevant Level	Level1 (S or A)
Query Method	automatic
Topic Part	TITLE only
Contents/Links Used	contents only
Number of Topics	80

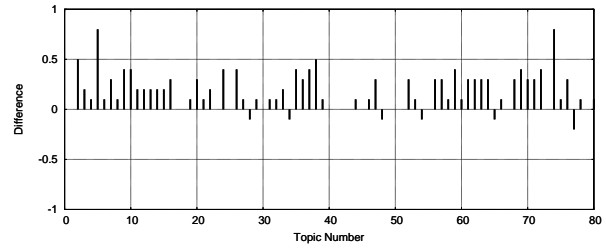
Summary Statistics	
Run ID	GRACE-tt-01
Subtask	Subtask A (target)
Relevant Level	Level2 (S, A or B)
Query Method	automatic
Topic Part	TITLE only
Contents/Links Used	contents only
Number of Topics	80

Document Level Precision Averages, Mean Reciprocal Rank (MRR), and the Number (#) and Percentage (%) of Topics for which No Relevant Document (Level1) were Retrieved				
	Precision	MRR	#	%
At 5 docs	0.4250	0.6190	18	0.2250
At 10 docs	0.3888	0.6316	11	0.1375
At 15 docs	0.3608	0.6354	7	0.0875
At 20 docs	0.3381	0.6368	5	0.0625

Document Level Precision Averages, Mean Reciprocal Rank (MRR), and the Number (#) and Percentage (%) of Topics for which No Relevant Document (Level2) were Retrieved				
	Precision	MRR	#	%
At 5 docs	0.5125	0.7306	11	0.1375
At 10 docs	0.4825	0.7401	5	0.0625
At 15 docs	0.4517	0.7428	2	0.0250
At 20 docs	0.4225	0.7435	1	0.0125

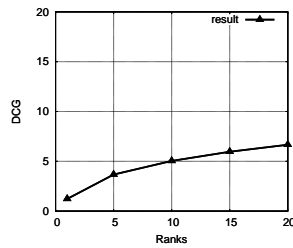


Difference from Median in Precision at Top-ranked 10 Documents per Topic (over 23 'automatic' and 'TITLE only' runs)\*



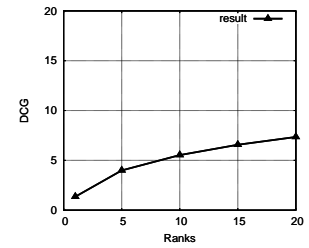
Difference from Median in Precision at Top-ranked 10 Documents per Topic (over 23 'automatic' and 'TITLE only' runs)\*

Discounted ( $\log_2$ ) Cummulative Gain Averages ( $G_S = 3, G_A = 2, G_B = 0$ )	
	DCG
At 1 doc	1.2250
At 5 doc	3.6682
At 10 docs	5.0385
At 15 docs	5.9608
At 20 docs	6.6623



Discounted Cummulative  
Gain Curve  
( $G_S = 3, G_A = 2, G_B = 0$ )

Discounted ( $\log_2$ ) Cummulative Gain Averages ( $G_S = 3, G_A = 2, G_B = 1$ )	
	DCG
At 1 doc	1.3500
At 5 doc	3.9886
At 10 docs	5.5288
At 15 docs	6.5665
At 20 docs	7.3467



Discounted Cummulative  
Gain Curve  
( $G_S = 3, G_A = 2, G_B = 1$ )

\* Topic Numbers: ( 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80 )  
Topic IDs: ( 1, 2, 3, 4, 5, 6, 7, 9, 11, 14, 15, 19, 21, 22, 23, 24, 25, 26, 28, 29, 30, 31, 34, 38, 43, 44, 45, 46, 55, 57, 58, 60, 61, 62, 63, 65, 66, 67, 68, 70, 71, 72, 73, 74, 76, 77, 78, 79, 80 )  
71, 72, 73, 74, 76, 77, 80, 81, 82, 84, 86, 88, 89, 91, 94, 95, 97, 98, 99, 108, 110, 111, 112, 116, 117, 119, 120, 122, 123, 127, 129, 132, 133, 143, 145, 146, 149, 150, 151, 152 )