11-x42 HW3-642

Jiaqiang Ruan

TOTAL POINTS

46 / 50

QUESTION 1

Introduction 0 pts

1.1 Collaboration and Originality 0 / 0
   ✓ - **0 pts** Correct

QUESTION 2

Experiment 1: Baselines 10.9 pts

2.1 Experimental Results 1.9 / 2.4
   ✓ - **0 pts** Correct
   - **0.5 Point adjustment**

   💬 Incorrect representation of win:loss

2.2 Parameters 0.5 / 0.5
   ✓ - **0 pts** Correct

2.3 Discussion 8 / 8
   ✓ - **0 pts** Correct

   💬 Good discussion and observations

QUESTION 3

Experiment 2: The number of feedback
documents 12.5 pts

3.1 Experimental Results 4 / 4
   ✓ - **0 pts** Correct

3.2 Parameters 0.5 / 0.5
   ✓ - **0 pts** Correct

3.3 Discussion 6.5 / 8
   ✓ - **0 pts** Correct
   - **1.5 Point adjustment**

   💬

Overall good discussion though there may be a
small misunderstanding in the last paragraph
under the "Overall" section. The number of
feedback terms remains constant for this
experiment.

QUESTION 4

Experiment 3: The number of feedback
terms 13.3 pts

4.1 Experimental Results 4.8 / 4.8
   ✓ - **0 pts** Correct

4.2 Parameters 0.5 / 0.5
   ✓ - **0 pts** Correct

4.3 Discussion 6.5 / 8
   ✓ - **0 pts** Correct
   - **1.5 Point adjustment**

   💬 Average discussion could have talked about
   more points like cost of adding more terms, etc

QUESTION 5

Experiment 4: Original query vs.
expanded query 13.3 pts

5.1 Experimental Results 4.8 / 4.8
   ✓ - **0 pts** Correct

5.2 Parameters 0.5 / 0.5
   ✓ - **0 pts** Correct

5.3 Discussion 7.5 / 8
   ✓ - **0 pts** Correct
   - **0.5 Point adjustment**

   💬

Good discussion

QUESTION 6

6 Reproducibility 0 / 0

✓ - **0 pts** Correct

Homework 3

1 Collaboration and Originality

1. Did you receive help <u>of any kind</u> from anyone in developing your software for this assignment (Yes or No)? It is not necessary to describe discussions with the instructor or TAs.

   No

2. Did you give help <u>of any kind</u> to anyone in developing their software for this assignment (Yes or No)?

   No

3. Did you examine anyone else's software for this assignment (Yes or No)? It is not necessary to mention software provided by the instructor.

   No.

4. Are you the author of <u>every line</u> of source code submitted for this assignment (Yes or No)? It is not necessary to mention software provided by the instructor.

   Yes

5. Are you the author of <u>every word</u> of your report (Yes or No)?

   Yes

2   Experiment 1: Baselines

2.1   Experimental Results

	Indri			
	BOW		**Query Expansion**	
	Reference System (*given*)	**Your System** (**Exp-1a**)	**Reference System** (**Exp-1b**)	**Your System** (**Exp-1c**)
P@10	0.78	0.75	0.82	0.84
P@20	0.74	0.695	0.765	0.765
P@30	0.6833	0.65	0.7233	0.72
MAP	0.2263	0.2172	0.2546	0.2587
win:loss	-	Loss	Win	Win

1.1 Collaboration and Originality 0 / 0

✓ - **0 pts** Correct

2 Experiment 1: Baselines

2.1 Experimental Results

	Indri			
	BOW		**Query Expansion**	
	Reference System (*given*)	**Your System (Exp-1a)**	**Reference System (Exp-1b)**	**Your System (Exp-1c)**
P@10	0.78	0.75	0.82	0.84
P@20	0.74	0.695	0.765	0.765
P@30	0.6833	0.65	0.7233	0.72
MAP	0.2263	0.2172	0.2546	0.2587
win:loss	-	Loss	Win	Win

2.2 Parameters

The parameter for Indri

Indri:mu=1500

Indri:lambda=0.01

The parameter for fb

fbDocs=10

fbTerms=10

fbMu=0

fbOrigWeight=0.5

2.1 Experimental Results **1.9 / 2.4**

✓ **- 0 pts** Correct

- 0.5 Point adjustment

💬 Incorrect representation of win:loss

2.2 Parameters

The parameter for Indri

Indri:mu=1500

Indri:lambda=0.01

The parameter for fb

fbDocs=10

fbTerms=10

fbMu=0

fbOrigWeight=0.5

2.3 Discussion

Overview: using query expansion improve both reference system and my system. All P@10, P@20, P@30 and MAP metrics improve. Query expansion adds more terms; therefore, it could make the query matches more document. Precision in large range can be improved. For example, P@30 improves a lot. When it comes to the precision in the top range, we would want the expanded query term is accurate and related to the query need, otherwise, it might harm the matching for top results because of bad expanded query. The experiment result show large improvement in P@10 too. So, it means the quality of the document in first retrieval is good and expanded terms suits the query need well.

Examples:

P@10 for Query 738 is 0.9 in reference system and becomes 1.0 in the reference system with query expansion. The original query for 738 is Anthrax hoaxes, and the expansion add words like terrorist, threat. These words are learnt from the first retrieved documents. And "Anthrax hoaxes", the person is probably a terrorist. Adding this information make the matching better. Moreover, the new term makes the query possible to match with other documents related to terrorist but not the person Anthrax hoaxes. This enables the query to match larger range of documents, hence improving the Precision in larger range. The P@30 improves from 0.63 to 0.7 for this query.

The performance for query 793 decrease after query expansion. MAP goes from 0.16 to 0.13. The original query is Bagpipe Band. And the query expansion adds terms like "see", "include" "drum". These words are not so related to the original query, thus harming the performance. The reason is probably that there aren't so many documents strongly related to "Bagpipe Band", therefore when we try to expand query in the first document pool, there are many irrelevant documents. These documents add "bad" terms to the expansion.

2.2 Parameters **0.5 / 0.5**

✓ **- 0 pts** Correct

2.3 Discussion

Overview: using query expansion improve both reference system and my system. All P@10, P@20, P@30 and MAP metrics improve. Query expansion adds more terms; therefore, it could make the query matches more document. Precision in large range can be improved. For example, P@30 improves a lot. When it comes to the precision in the top range, we would want the expanded query term is accurate and related to the query need, otherwise, it might harm the matching for top results because of bad expanded query. The experiment result show large improvement in P@10 too. So, it means the quality of the document in first retrieval is good and expanded terms suits the query need well.

Examples:

 P@10 for Query 738 is 0.9 in reference system and becomes 1.0 in the reference system with query expansion. The original query for 738 is Anthrax hoaxes, and the expansion add words like terrorist, threat. These words are learnt from the first retrieved documents. And "Anthrax hoaxes", the person is probably a terrorist. Adding this information make the matching better. Moreover, the new term makes the query possible to match with other documents related to terrorist but not the person Anthrax hoaxes. This enables the query to match larger range of documents, hence improving the Precision in larger range. The P@30 improves from 0.63 to 0.7 for this query.

 The performance for query 793 decrease after query expansion. MAP goes from 0.16 to 0.13. The original query is Bagpipe Band. And the query expansion adds terms like "see", "include" "drum". These words are not so related to the original query, thus harming the performance. The reason is probably that there aren't so many documents strongly related to "Bagpipe Band", therefore when we try to expand query in the first document pool, there are many irrelevant documents. These documents add "bad" terms to the expansion.

3   Experiment 2: The number of feedback documents

3.1   Experimental Results

	Indri BOW, Reference System (*given*)	Query Expansion, Reference System Initial Results				
		Feedback Documents				
		10 (Exp-2a)	20 (Exp-2b)	30 (Exp-2c)	40 (Exp-2d)	50 (Exp-2e)
P@10	0.78	0.84	0.73	0.75	0.75	0.69
P@20	0.74	0.765	0.68	0.685	0.7	0.64
P@30	0.6833	0.72	0.68	0.68	0.6733	0.6433
MAP	0.2263	0.2587	0.2478	0.2477	0.2451	0.2388
win:loss	-	Win	Win	Win	Win	Win

2.3 Discussion 8 / 8

✓ - **0 pts** Correct

💬 Good discussion and observations

3 Experiment 2: The number of feedback documents

3.1 Experimental Results

	Indri BOW, Reference System (*given*)	Query Expansion, Reference System Initial Results				
		Feedback Documents				
		10 (Exp-2a)	20 (Exp-2b)	30 (Exp-2c)	40 (Exp-2d)	50 (Exp-2e)
P@10	0.78	0.84	0.73	0.75	0.75	0.69
P@20	0.74	0.765	0.68	0.685	0.7	0.64
P@30	0.6833	0.72	0.68	0.68	0.6733	0.6433
MAP	0.2263	0.2587	0.2478	0.2477	0.2451	0.2388
win:loss	-	Win	Win	Win	Win	Win

3.2 Parameters

Indri:mu=1500

Indri:lambda=0.01

fb=Indri

fbDocs=50,40,30,20,10

fbTerms=10

fbMu=0

fbOrigWeight=0.5

3.1 Experimental Results **4 / 4**

✓ **- 0 pts** Correct

3.2   Parameters

Indri:mu=1500

Indri:lambda=0.01

fb=Indri

fbDocs=50,40,30,20,10

fbTerms=10

fbMu=0

fbOrigWeight=0.5

3.3   Discussion

Overall:

All experiments win over the given reference system by MAP. But P@n drops quickly and gets worse than the reference system when feedback document gets larger. It shows that query expansion can improve the matching performance in large range. When it comes to the precision in the top range, we need to tune this parameter carefully.

For feedback document, if it is too small, we cannot get enough expand terms in variety. If it is too large, we might involve irrelevant documents and add "bad" expansion words. Therefore, this parameter should be tuned according to the corpus and possible query. If the corpus is in high quality, we could use a larger feedback document number. If the incoming query is likely to retrieve many related documents, we could have a larger feedback document number.

From our experiments, it seems when fd increases the overall MAP is dropping. More notably, the P@10 drops faster than other metrics. It is because for top results, they are more sensible to the accuracy of the terms used in the query. Adding more terms could mislead the top matchings.

Examples:

When fd increases, non-related and stop words could be added. For query 738, when fd=10, words like terrorist and threat are added. They seem to be relevant. But when fd is 30, "have", "I" are added. They are almost stop words. When fd is 50, "have" has 0.0047 weight, while it only has "0.0034" when fd=30.

3.2 Parameters **0.5 / 0.5**

  ✓ **- 0 pts** Correct

3.3 Discussion

Overall:

All experiments win over the given reference system by MAP. But P@n drops quickly and gets worse than the reference system when feedback document gets larger. It shows that query expansion can improve the matching performance in large range. When it comes to the precision in the top range, we need to tune this parameter carefully.

For feedback document, if it is too small, we cannot get enough expand terms in variety. If it is too large, we might involve irrelevant documents and add "bad" expansion words. Therefore, this parameter should be tuned according to the corpus and possible query. If the corpus is in high quality, we could use a larger feedback document number. If the incoming query is likely to retrieve many related documents, we could have a larger feedback document number.

From our experiments, it seems when fd increases the overall MAP is dropping. More notably, the P@10 drops faster than other metrics. It is because for top results, they are more sensible to the accuracy of the terms used in the query. Adding more terms could mislead the top matchings.

Examples:

When fd increases, non-related and stop words could be added. For query 738, when fd=10, words like terrorist and threat are added. They seem to be relevant. But when fd is 30, "have", "I" are added. They are almost stop words. When fd is 50, "have" has 0.0047 weight, while it only has "0.0034" when fd=30.

4 Experiment 3: The number of feedback terms

4.1 Experimental Results

	Indri BOW, Reference System (*given*)	Query Expansion, Reference System Initial Results					
		Feedback Terms					
		5 (Exp-3a)	10 (Exp-3b)	20 (Exp-3c)	30 (Exp-3d)	40 (Exp-3e)	50 (Exp-3f)
P@10	0.78	0.78	0.84	0.84	0.82	0.82	0.81
P@20	0.74	0.76	0.765	0.735	0.75	0.74	0.745
P@30	0.6833	0.7	0.72	0.73	0.74	0.7333	0.74
MAP	0.2263	0.2479	0.2587	0.2571	0.2579	0.2558	0.2571
win:loss	N/A	Win	Win	Win	Win	Win	win

3.3 Discussion **6.5 / 8**

✓ **- 0 pts** Correct

- 1.5 Point adjustment

💬 Overall good discussion though there may be a small misunderstanding in the last paragraph under the "Overall" section. The number of feedback terms remains constant for this experiment.

ıllı gradescope

4 Experiment 3: The number of feedback terms

4.1 Experimental Results

	Indri BOW, Reference System (*given*)	Query Expansion, Reference System Initial Results					
		Feedback Terms					
		5 (Exp-3a)	10 (Exp-3b)	20 (Exp-3c)	30 (Exp-3d)	40 (Exp-3e)	50 (Exp-3f)
P@10	0.78	0.78	0.84	0.84	0.82	0.82	0.81
P@20	0.74	0.76	0.765	0.735	0.75	0.74	0.745
P@30	0.6833	0.7	0.72	0.73	0.74	0.7333	0.74
MAP	0.2263	0.2479	0.2587	0.2571	0.2579	0.2558	0.2571
win:loss	N/A	Win	Win	Win	Win	Win	win

4.2   Parameters

Indri:mu=1500

Indri:lambda=0.01

fb=Indri

fbDocs=10

fbTerms=50,40,30,20,10,5

fbMu=0

fbOrigWeight=0.5

4.1 Experimental Results **4.8 / 4.8**

✓ **- 0 pts** Correct

4.2   Parameters

Indri:mu=1500

Indri:lambda=0.01

fb=Indri

fbDocs=10

fbTerms=50,40,30,20,10,5

fbMu=0

fbOrigWeight=0.5

4.3   Discussion

Overview: all experiments have higher performance than reference system. MAP, P@10 and P@20 goes up from fbTerms = 5 to fbTerms = 10, then starts decreasing. P@30 keeps increasing when fbTerm goes from 5 to 50. For matching the documents in top range, query terms need to be accurate and related to the query need. Therefore, adding more terms could distract the search engine matching the best document. But for the sake of matching more related documents. It is very useful to have more terms.

Examples:

Use query 738 as example again, the original query term is Anthrax hoaxes. First, "terrosrist", "law" is needed. Then more terms like "threat", "fbi" is added. MAP goes from 0.312 to 0.319. When fbTerms =20, irrelevant words like "who", "have", "false" are added. MAP drops from 0.319 to 0.317.

4.2 Parameters **0.5 / 0.5**

✓ **- 0 pts** Correct

4.3 Discussion

Overview: all experiments have higher performance than reference system. MAP, P@10 and P@20 goes up from fbTerms = 5 to fbTerms = 10, then starts decreasing. P@30 keeps increasing when fbTerm goes from 5 to 50. For matching the documents in top range, query terms need to be accurate and related to the query need. Therefore, adding more terms could distract the search engine matching the best document. But for the sake of matching more related documents. It is very useful to have more terms.

Examples:

Use query 738 as example again, the original query term is Anthrax hoaxes. First, "terrosrist", "law" is needed. Then more terms like "threat", "fbi" is added. MAP goes from 0.312 to 0.319. When fbTerms =20, irrelevant words like "who", "have", "false" are added. MAP drops from 0.319 to 0.317.

5 Experiment 4: Original query vs. expanded query

5.1 Experimental Results

	Indri BOW, Reference System (*given*)	Query Expansion, Reference System Initial Results					
		fbOrigWeight					
		0.0 (Exp-4a)	**0.2 (Exp-4b)**	**0.4 (Exp-4c)**	**0.6 (Exp-4d)**	**0.8 (Exp-4e)**	**1.0 (Exp-4f)**
P@10	0.78	0.75	0.75	0.76	0.75	0.73	0.75
P@20	0.74	0.69	0.685	0.67	0.675	0.69	0.695
P@30	0.6833	0.6667	0.67	0.68	0.68	0.6867	0.65
MAP	0.2263	0.2361	0.2457	0.2485	0.2445	0.2303	0.2172
win:loss	N/A	Win	Win	Win	Win	Win	loss

4.3 Discussion **6.5 / 8**

✓ **- 0 pts** Correct

- 1.5 Point adjustment

💬 Average discussion could have talked about more points like cost of adding more terms, etc

ıllı gradescope

5 Experiment 4: Original query vs. expanded query

5.1 Experimental Results

	Indri BOW, Reference System (*given*)	Query Expansion, Reference System Initial Results					
		fbOrigWeight					
		0.0 (Exp-4a)	0.2 (Exp-4b)	0.4 (Exp-4c)	0.6 (Exp-4d)	0.8 (Exp-4e)	1.0 (Exp-4f)
P@10	0.78	0.75	0.75	0.76	0.75	0.73	0.75
P@20	0.74	0.69	0.685	0.67	0.675	0.69	0.695
P@30	0.6833	0.6667	0.67	0.68	0.68	0.6867	0.65
MAP	0.2263	0.2361	0.2457	0.2485	0.2445	0.2303	0.2172
win:loss	N/A	Win	Win	Win	Win	Win	loss

5.2   Parameters

Indri:mu=1500

Indri:lambda=0.01

fb=Indri

fbDocs=30

fbTerms=10

fbMu=0

fbOrigWeight=1.0, 0.8, 0.6, 0.4, 0.2, 0

5.1 Experimental Results **4.8 / 4.8**

✓ **- 0 pts** Correct

5.2   Parameters

Indri:mu=1500

Indri:lambda=0.01

fb=Indri

fbDocs=30

fbTerms=10

fbMu=0

fbOrigWeight=1.0, 0.8, 0.6, 0.4, 0.2, 0

5.3   Discussion

Overall: The larger fbOrigweight is, the more the query will rely on the original query. In my experiment, if we only rely on original query, the performance will be worse than the reference system. And the best weight is 0.4. Note that, even if we totally use the expanded query, we can use outperform the original query. It is because that the expanded query usually contains the original query terms. If the expanded query is in good quality (which probably means we have corpus in good quality, right numbers for feedback document and feedback terms are picked.), then we can have a smaller fbOriWeight and reply more on the expanded query.

Example:

Most of the query observe a best choice for the fbOriWeight in the middle between 0 and 1.

Query 738: 0.2763(1.0), 0.2822(0.4), 0.2478(1.0)

Query 739: 0.1774(1.0), 0.1798(0.4), 0.1783(1.0)

Therefore, this parameter should be tuned in different system.

5.2 Parameters 0.5 / 0.5

✓ **- 0 pts** Correct

gradescope

5.3 Discussion

Overall: The larger fbOrigweight is, the more the query will rely on the original query. In my experiment, if we only rely on original query, the performance will be worse than the reference system. And the best weight is 0.4. Note that, even if we totally use the expanded query, we can use outperform the original query. It is because that the expanded query usually contains the original query terms. If the expanded query is in good quality (which probably means we have corpus in good quality, right numbers for feedback document and feedback terms are picked.), then we can have a smaller fbOriWeight and reply more on the expanded query.

Example:

Most of the query observe a best choice for the fbOriWeight in the middle between 0 and 1.

Query 738: 0.2763(1.0), 0.2822(0.4), 0.2478(1.0)

Query 739: 0.1774(1.0), 0.1798(0.4), 0.1783(1.0)

Therefore, this parameter should be tuned in different system.

```
------------ Automatic grading for jruan@andrew.cmu.edu ------------

TestName          RawScore          Time (s)    TimePenalty     FinalScore
HW3-Test-0        100.00                  13       0.00           100.00
HW3-Test-1        100.00                  42       0.00           100.00
HW3-Test-2        100.00                  93       0.00           100.00
HW3-Test-3        100.00                  17       0.00           100.00
HW3-Test-4        100.00                  75       0.00           100.00
HW3-Test-5        100.00                 126       0.00           100.00
HW3-Test-6        100.00                  21       0.00           100.00
HW3-Test-7        100.00                  36       0.00           100.00
HW3-Test-8        100.00                  46       0.00           100.00
HW3-Test-9        100.00                  33       0.00           100.00
HW3-Test-10       100.00                  44       0.00           100.00
HW3-Test-11       100.00                  55       0.00           100.00
HW3-Test-12       100.00                  66       0.00           100.00
HW3-Test-13       100.00                  36       0.00           100.00
HW3-Test-14       100.00                  68       0.00           100.00
HW3-Test-15       100.00                  28       0.00           100.00
HW3-Test-16       100.00                  52       0.00           100.00
HW3-Test-17       100.00                  21       0.00           100.00
HW3-Test-18       100.00                  59       0.00           100.00
HW3-Test-19       100.00                  22       0.00           100.00


Summary
FinalGrade:       40.00
TotalRunTime:     953.00
SubmitTime:       Mar22-22:26:07


-- Experiments --
HW3-Exp-1a        21.72
HW3-Exp-1b        25.46
HW3-Exp-1c        25.87
HW3-Exp-2a        25.87
HW3-Exp-2b        24.78
HW3-Exp-2c        24.77
HW3-Exp-2d        24.51
HW3-Exp-2e        23.88
HW3-Exp-3a        24.79
HW3-Exp-3b        25.87
HW3-Exp-3c        25.71
HW3-Exp-3d        25.79
HW3-Exp-3e        25.58
HW3-Exp-3f        25.71
HW3-Exp-4a        23.61
HW3-Exp-4b        24.57
HW3-Exp-4c        24.85
HW3-Exp-4d        24.45
HW3-Exp-4e        23.03
HW3-Exp-4f        21.72
```

5.3 Discussion **7.5 / 8**

✓ **- 0 pts** Correct

- 0.5 Point adjustment

💬 Good discussion

```
------------ Automatic grading for jruan@andrew.cmu.edu ------------

TestName           RawScore          Time (s)    TimePenalty      FinalScore
HW3-Test-0         100.00                 13         0.00           100.00
HW3-Test-1         100.00                 42         0.00           100.00
HW3-Test-2         100.00                 93         0.00           100.00
HW3-Test-3         100.00                 17         0.00           100.00
HW3-Test-4         100.00                 75         0.00           100.00
HW3-Test-5         100.00                126         0.00           100.00
HW3-Test-6         100.00                 21         0.00           100.00
HW3-Test-7         100.00                 36         0.00           100.00
HW3-Test-8         100.00                 46         0.00           100.00
HW3-Test-9         100.00                 33         0.00           100.00
HW3-Test-10        100.00                 44         0.00           100.00
HW3-Test-11        100.00                 55         0.00           100.00
HW3-Test-12        100.00                 66         0.00           100.00
HW3-Test-13        100.00                 36         0.00           100.00
HW3-Test-14        100.00                 68         0.00           100.00
HW3-Test-15        100.00                 28         0.00           100.00
HW3-Test-16        100.00                 52         0.00           100.00
HW3-Test-17        100.00                 21         0.00           100.00
HW3-Test-18        100.00                 59         0.00           100.00
HW3-Test-19        100.00                 22         0.00           100.00


Summary
FinalGrade:        40.00
TotalRunTime:      953.00
SubmitTime:        Mar22-22:26:07


-- Experiments --
HW3-Exp-1a         21.72
HW3-Exp-1b         25.46
HW3-Exp-1c         25.87
HW3-Exp-2a         25.87
HW3-Exp-2b         24.78
HW3-Exp-2c         24.77
HW3-Exp-2d         24.51
HW3-Exp-2e         23.88
HW3-Exp-3a         24.79
HW3-Exp-3b         25.87
HW3-Exp-3c         25.71
HW3-Exp-3d         25.79
HW3-Exp-3e         25.58
HW3-Exp-3f         25.71
HW3-Exp-4a         23.61
HW3-Exp-4b         24.57
HW3-Exp-4c         24.85
HW3-Exp-4d         24.45
HW3-Exp-4e         23.03
HW3-Exp-4f         21.72
```

6 Reproducibility **0 / 0**

✓ **- 0 pts** Correct

```
============> Corrected Automatic Grading Report <====

------------ Automatic grading for jruan@andrew.cmu.edu ------------

TestName            RawScore         Time (s)     TimePenalty     FinalScore
HW3-Test-0          100.00             13           0.00           100.00
HW3-Test-1          100.00             41           0.00           100.00
HW3-Test-2          100.00             92           0.00           100.00
HW3-Test-3          100.00             18           0.00           100.00
HW3-Test-4          100.00             82           0.00           100.00
HW3-Test-5          100.00            128           0.00           100.00
HW3-Test-6          100.00             20           0.00           100.00
HW3-Test-7          100.00             34           0.00           100.00
HW3-Test-8          100.00             43           0.00           100.00
HW3-Test-9          100.00             35           0.00           100.00
HW3-Test-10         100.00             44           0.00           100.00
HW3-Test-11         100.00             52           0.00           100.00
HW3-Test-12         100.00             61           0.00           100.00
HW3-Test-13         100.00             34           0.00           100.00
HW3-Test-14         100.00             66           0.00           100.00
HW3-Test-15         100.00             28           0.00           100.00
HW3-Test-16         100.00             55           0.00           100.00
HW3-Test-17         100.00             20           0.00           100.00
HW3-Test-18         100.00             65           0.00           100.00
HW3-Test-19         100.00             22           0.00           100.00


Summary
FinalGrade:         50.00
TotalRunTime:       953.00
SubmitTime:         Mar22-22:26:07


-- Experiments --
HW3-Exp-1a          21.72
HW3-Exp-1b          25.46
HW3-Exp-1c          25.87
HW3-Exp-2a          25.87
HW3-Exp-2b          24.78
HW3-Exp-2c          24.77
HW3-Exp-2d          24.51
HW3-Exp-2e          23.88
HW3-Exp-3a          24.79
HW3-Exp-3b          25.87
HW3-Exp-3c          25.71
HW3-Exp-3d          25.79
HW3-Exp-3e          25.58
HW3-Exp-3f          25.71
HW3-Exp-4a          23.61
HW3-Exp-4b          24.57
HW3-Exp-4c          24.85
HW3-Exp-4d          24.45
HW3-Exp-4e          23.03
HW3-Exp-4f          21.72
```