Jacqueline Wong, Kevin Nguyen

CS 6301.502

SW Analysis and Comprehension

Team 4

Assignment #2 - Deliverable

**Lucene was used to index the documents and for querying.**

**Additionally, for effectiveness, we listed N/A for when the method changed was not ranked. For calculations, N/A was equivalent to the amount of methods retrieved + 1 (6, 11, 21).**

• Precision and recall @{5,10,20}, and effectiveness for each query of each system

**Rhino**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Query** | **Precision** | | | **Recall** | | | **Effectiveness** | | |
|  | **5** | **10** | **20** | **5** | **10** | **20** | **5** | **10** | **20** |
| **Bug 774083 Query 1** | 0.4 | 0.2 | 0.1 | 1 | 1 | 1 | 2 | 2 | 2 |
| **Bug 774083 Query 2** | 0.4 | 0.2 | 0.1 | 1 | 1 | 1 | 3 | 3 | 3 |
| **Bug 774083 Query 3** | 0.4 | 0.2 | 0.1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Bug 688023 Query 1** | 0 | 0 | 0 | 0 | 0 | 0 | N/A | N/A | N/A |
| **Bug 688023 Query 2** | 0 | 0 | 0 | 0 | 0 | 0 | N/A | N/A | N/A |
| **Bug 688023 Query 3** | 0.2 | 0.1 | 0.1 | 0.25 | 0.25 | 0.5 | 5 | 5 | 5 |
| **Bug 685403 Query 1** | 0 | 0 | 0 | 0 | 0 | 0 | N/A | N/A | N/A |
| **Bug 685403 Query 2** | 0 | 0 | 0 | 0 | 0 | 0 | N/A | N/A | N/A |
| **Bug 685403 Query 3** | 0 | 0 | 0 | 0 | 0 | 0 | N/A | N/A | N/A |
| **Bug 773573 Query 1** | 0.2 | 0.1 | 0.05 | 1 | 1 | 1 | 3 | 3 | 3 |
| **Bug 773573 Query 2** | 0.2 | 0.1 | 0.05 | 1 | 1 | 1 | 4 | 4 | 4 |
| **Bug 773573 Query 3** | 0.2 | 0.1 | 0.05 | 1 | 1 | 1 | 3 | 3 | 3 |
| **Bug 780147 Query 1** | 0 | 0 | 0 | 0 | 0 | 0 | N/A | N/A | N/A |
| **Bug 780147 Query 2** | 0 | 0 | 0 | 0 | 0 | 0 | N/A | N/A | N/A |
| **Bug 780147 Query 3** | 0 | 0 | 0 | 0 | 0 | 0 | N/A | N/A | N/A |

**AspectJ**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Query** | **Precision** | | | **Recall** | | | **Effectiveness** | | |
|  | **5** | **10** | **20** | **5** | **10** | **20** | **5** | **10** | **20** |
| **Bug 482990 Query 1** | 0.2 | 0.1 | 0.05 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Bug 482990 Query 2** | 0.2 | 0.1 | 0.05 | 1 | 1 | 1 | 3 | 3 | 3 |
| **Bug 482990 Query 3** | 0.2 | 0.1 | 0.05 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Bug 470658 Query 1** | 0.2 | 0.1 | 0.05 | 1 | 1 | 1 | 3 | 3 | 3 |
| **Bug 470658 Query 2** | 0 | 0 | 0.05 | 0 | 0 | 1 | N/A | N/A | 15 |
| **Bug 470658 Query 3** | 0.2 | 0.1 | 0.05 | 1 | 1 | 1 | 3 | 3 | 3 |
| **Bug 461323 Query 1** | 0.2 | 0.1 | 0.1 | 0.143 | 0.143 | 0.286 | 5 | 5 | 5 |
| **Bug 461323 Query 2** | 0 | 0 | 0.05 | 0 | 0 | 0.143 | N/A | N/A | 17 |
| **Bug 461323 Query 3** | 0 | 0.1 | 0.1 | 0 | 0.143 | 0.286 | N/A | 6 | 6 |
| **Bug 478003 Query 1** | 0 | 0 | 0 | 0 | 0 | 0 | N/A | N/A | N/A |
| **Bug 478003 Query 2** | 0 | 0 | 0.05 | 0 | 0 | 1 | N/A | N/A | 15 |
| **Bug 478003 Query 3** | 0 | 0 | 0 | 0 | 0 | 0 | N/A | N/A | N/A |
| **Bug 484941 Query 1** | 0 | 0 | 0 | 0 | 0 | 0 | N/A | N/A | N/A |
| **Bug 484941 Query 2** | 0 | 0 | 0 | 0 | 0 | 0 | N/A | N/A | N/A |
| **Bug 484941 Query 3** | 0 | 0 | 0 | 0 | 0 | 0 | N/A | N/A | N/A |

• Average and median precision and recall @{5,10,20}, and effectiveness for each system

**Rhino**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Precision** | | **Recall** | | **Effectiveness** | |
|  | Average | Median | Average | Median | Average | Median |
| **5** | 0.133333 | 0 | 0.416667 | 0 | 4.6 | 6\* |
| **10** | 0.066667 | 0 | 0.416667 | 0 | 7.266667 | 11\* |
| **20** | 0.036667 | 0 | 0.433333 | 0 | 12.6 | 21\* |

\*These values are higher than the amount of retrieved methods. This means the effectiveness of the correct method being retrieved missed most of the time.

**AspectJ**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Precision** | | **Recall** | | **Effectiveness** | |
|  | Average | Median | Average | Median | Average | Median |
| **5** | 0.08 | 0 | 0.342867 | 0 | 4.666667 | 6\* |
| **10** | 0.046667 | 0 | 0.3524 | 0 | 7.333333 | 11\* |
| **20** | 0.04 | 0.05 | 0.514333 | .286 | 11.6 | 15 |

\* These values are higher than the amount of retrieved methods. This means the effectiveness of the correct method being retrieved missed most of the time.

• Average and median precision and recall @{5,10,20}, and effectiveness for **all the systems** (i.e., across systems)

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Precision** | | **Recall** | | **Effectiveness** | |
|  | Average | Median | Average | Median | Average | Median |
| **5** | 0.106667 | 0 | 0.379767 | 0 | 4.633334 | 6\* |
| **10** | 0.056667 | 0 | 0.384534 | 0 | 7.3 | 11\* |
| **20** | 0.038334 | .05 | 0.473833 | .286 | 12.1 | 15\* |

\* These values are higher than the amount of retrieved methods. This means the effectiveness of the correct method being retrieved missed most of the time.

• Comparison of precision, recall, and effectiveness between the three types of queries, defined in step 3.a.

**Rhino**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Precision** | | **Recall** | | **Effectiveness** | |
|  | Average | Median | Average | Median | Average | Median |
| **Query 1** | 0.07 | 0 | 0.4 | 0 | 8.6 | 6\* |
| **Query 2** | 0.07 | 0 | 0.4 | 0 | 9 | 6\* |
| **Query 3** | 0.096667 | 0.1 | 0.466667 | 0.25 | 6.866667 | 5 |

\* These values are higher than the amount of retrieved methods. This means the effectiveness of the correct method being retrieved missed most of the time.

**AspectJ**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Precision** | | **Recall** | | **Effectiveness** | |
|  | Average | Median | Average | Median | Average | Median |
| **Query 1** | 0.073333 | 0.05 | 0.438133 | 0.143 | 6.866667 | 5 |
| **Query 2** | 0.033333 | 0 | 0.342867 | 0 | 9.666667 | 11\* |
| **Query 3** | 0.06 | 0.05 | 0.4286 | 0.143 | 7.066667 | 6 |

\* These values are higher than the amount of retrieved methods. This means the effectiveness of the correct method being retrieved missed most of the time.

**All Systems**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Precision** | | **Recall** | | **Effectiveness** | |
|  | Average | Median | Average | Median | Average | Median |
| **Query 1** | 0.071667 | 0.025 | 0.419067 | 0.0715 | 7.733334 | 5 |
| **Query 2** | 0.051667 | 0 | 0.371434 | 0 | 9.333334 | 6\* |
| **Query 3** | 0.078334 | 0.05 | 0.447634 | 0.25 | 6.966667 | 5 |

\* These values are higher than the amount of retrieved methods. This means the effectiveness of the correct method being retrieved missed most of the time.

**Comparison of precision, recall, and effectiveness between the three types of queries:**

In terms of bug localization, as a general trend Query Types 1 and 3 (the queries with the bug description in them) tended to fare better that Query 2, which only used the bug title as its query. Additionally, in most cases Query types 1 and 3 performed similarly (in precision, recall, and effectiveness) in bug localization. Because of this, we drew the conclusion that bug description had a greater positive effect on query performance when compared with using the bug title in a query. However, it should be noted that some bug queries resulted in full misses and could not find any relevant documents in the top 20 results; this resulted in skewed data when calculating averages in effectiveness.