Test Report for Part2 Assignment CAB201

Name: Phuong Nam Ly

Student number: n10098097

I separate the test into 4 main parts for four levels. Input should be entered in the command line. The Expected Output will be a screenshot from the input file while the Actual Output will be a screenshot of the console output. Comment will explain how I evaluate and justify my statement of completeness and self-assessed CRA.

Test structure

Category		
Input		
Expected Output		
Actual Output		
Comment		

Level 4 Functionalities (15 marks)

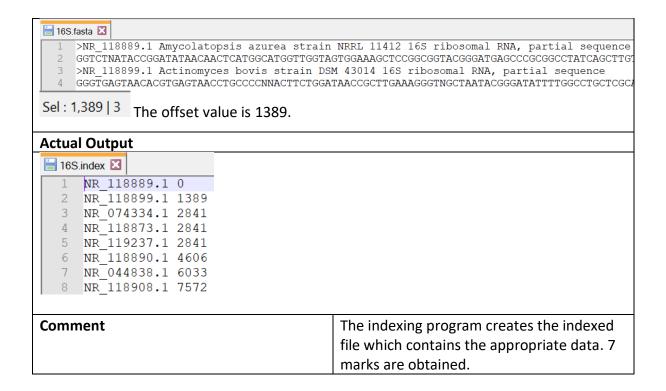
I use the IndexSequence project to create the indexed file, and then put it in the Debug folder of A1.new to use it for the Search project.

Criteria: The program for indexing creates a file as specified by the command line arguments (2 marks)

Criteria: The index file contains appropriate data (5 marks)

Test 4.1: Test the index program

Category	Create a file that displays the correct data	
Input	16S.fasta 16S.index	
Expected Output		
The expected offset for an id sequence can be verified by bolding the texts prior to the id		
sequence and checking the "Sel" value at the right bottom corner.		
Examples:		
For checking the first sequence, NR_118889.1:		
In 16S.fasta ■		
	NRRL 11412 16S ribosomal RNA, partial sequence GTGGAAAGCTCCGGCGGTACGGGATGAGCCCGCGGCCTATCAGCTTG	
Sel: 0 0 The offset value is 0.		
For checking the sequence, >NR_118899.1:		



Criteria: The program for searching creates a file that contains the appropriate data (5 marks)

Test 4.2: Test the searching program (Successful run)

Category	Create a file that displays the correct data
Input	-level4 16S.fasta 16S.index query.txt
	results.txt
Expected Output	
Provided query.txt file	

```
The output file contains all the id sequences
        NR 115365.1
                                     except (NR 999999.9, NR 99999888.9,
        NR 118873.1
    2
                                     NR 312312313123.3) with the corresponding DNA
    3
        NR 999999.9
                                     sequence.
    4
        NR 118941.1
        NR 99999888.9
    6
        NR 037011.1
    7
        NR 118951.1
        NR 119237.1
    9
        NR 312312313123.3
        NR 118718.2
  10
        NR 118735.2
  11
  12
        NR 119037.2
  13
  14
  15
  16
  17
  18
  19
  20
Actual Output
🔚 results.txt 🗵
    NR 118873.1 Archaeoglobus fulgidus DSM 4304 strain VC-16 16S ribosomal RNA, complete sequence
    ATTCTGGTTGATCCTGCCAGAGGCCGCTGCTATCCGGCTGGGACTAAGCCATGCGAGTCAAGGGGCTTGTATCCCTTCGGGGGATGCAAGCACC
    NR 119237.1 Archaeoglobus fulgidus DSM 4304 strain VC-16 16S ribosomal RNA, complete sequence
    ATTCTGGTTGATCCTGCCAGAGGCCGCTGCTATCCGGCTGGGACTAAGCCATGCGAGTCAAGGGGCTTGTATCCCTTCGGGGATGCAAGCACC
    NR_037011.1 Aeromonas salmonicida subsp. achromogenes strain 6263/4/5 16S ribosomal RNA, part
    GAGTTTGATCATGGCTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGAGCGGCAGCGGGAAAGTAGCTTGCTACTTTTGCCGGC
    NR 118951.1 Bacillus azotoformans LMG 9581 strain ATCC 29788 16S ribosomal RNA, partial seque.
    GAGAGTTTNANCCTGGCTCAGGACGACGCTGGCGGCGTGCCTNANACNTNCNAGTCGAGCGGATGATAAAGGAGCTTGCTCCTTTAGATTCA
    NR_118941.1 Streptococcus alactolyticus strain NCDO 1091 16S ribosomal RNA, partial sequence
    GAACGGTGAGTAACGCGTAGGTAACCTGCCTTGTAGCGGGGGATAACTATTGGAAACGATAGCTAATACCGCATAACAGCTTTTGACACATG
    NR 115365.1 Streptomyces albus strain CSSP327 16S ribosomal RNA, partial sequence
    TCACGGAGAGTTTGATCCTGGCTCAGGACGACGCTGGCGGCGTGCTTAACACATGCAAGTCGAACGATGAAGCCCTTCGGGGTGGATTAGTG
 13 NR 118718.2 Treponema bryantii strain RUS-1 16S ribosomal RNA, partial sequence
    15 NR_118735.2 Gottschalkia purinilytica strain ATCC 33906 16S ribosomal RNA, partial sequence
    NR 119037.2 Filifactor villosus strain DSM 1645 16S ribosomal RNA, partial sequence
 18 AATTATATCGGAGNTCTTCGGGCCAAAGAGATAATAAGCTAGTGGCGGACGGGTGCGTAACGTGTGGGTAATCTGCCTTTGTCATAGGAATAA
C:\Users\lyphu\OneDrive - Queensland University of Technology
Error, Id sequence NR_999999.9 not found
Error, Id sequence NR_99999888.9 not found
Error, Id sequence NR_312312313123.3 not found
Program finishes. Press Enter to exit.
                                            The output file contains all 9 id sequences
Comment
                                            with the corresponding DNA lines. 5 marks
                                            are obtained
```

Criteria: Clear error message is provided when the index file does not exist (1 mark)

Test 4.3: Test the index program (Unsuccessful run)

Category	Provide appropriate warning
----------	-----------------------------

Input	RandomName.fasta 16S.index
Expected Output	
Appropriate warning is given through the console.	
Actual Output	
The input file RandomName.fasta does not exist or it is incorrectly formatted Program finishes. Press Enter to exit.	
Comment	The appropriate warning is given. 1 mark is obtained

Criteria: Clear error message is provided when the query file does not exist or is incorrectly formatted (2 marks)

Test 4.4: Test the searching program (Unsuccessful run)

Category	Provide appropriate warning	
Input	-level4 16S.fasta 16S.index RandomName.txt	
	results.txt	
Expected Output		
Appropriate warning is given through the console.		
Actual Output		
C:\Users\lyphu\OneDrive - Queensland University of Technology\Bachelor\CAB201\Assignment\Final		
The input file RandomName.txt does not exist or it is incorrectly formatted		
Program finishes. Press Enter to exit.		
Comment	The appropriate warning is given. 1 mark is	
	obtained	

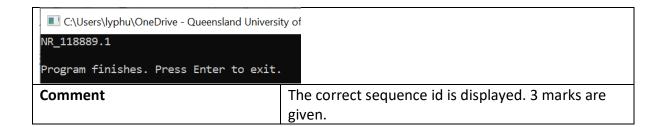
All level 4 functionalities are achieved. 15 marks are obtained.

Level 5 Functionalities (15 marks)

Criteria: The program prints the appropriate data (10 marks)

Test 5.1: Test the searching program with the expected output including the first id sequence (Successful run)

Category	Display the correct data		
Input	-level5 16S.fasta		
	GGTCTNATACCGGATATAACAACTCAT		
Expected Output	Expected Output		
At least NR 11889.1 is displayed.			
☐ 16S.fasta 🔀			
1 >NR_118889.1 Amycolatopsis azure	ea strain NRRL 11412 16S ribosomal RNA, partial sequence		
2 GGTCTNATACCGGATATAACAACTCATGGCATGGTTGGTAGTGGAAAGCTCCGGCGGTACGGGATGAGCCCGCGGCCTATCAGCTTG			
2 NRD 110000 1 Natinomuses havis strain DCM /201/ 165 mihosomal DNN martial seguence			
Actual Output			



Test 5.2: Test the searching program with the expected output including the last id sequence (Successful run)

Category	Display the correct data	
Input	-level5 16S.fasta	
	AATTATATCGGAGNTCTTCGGGCCAAA	
Expected Output		
At least NR_119037.2 is displayed.		
41533 >NR_119037.2 Filifactor villosus strain DSM 1645 16S ribosomal RNA, partial sequence 41534 AATTATATCGGAGNTCTTCGGGCCAAAGAGATAATAAGCTAGTGGCGGACGGGTGCGTAACGTGTGGGTAATCTGCCTTTGTCAT		
Actual Output		
C:\Users\lyphu\OneDrive - Queensland University o		
NR_119037.2		
Program finishes. Press Enter to exit.		
Comment	The correct sequence id is displayed. 3 marks are given.	

Test 5.3: Test the searching program with the expected output including the id sequences on the same metadata line (Successful run)

Category	Display the correct data	
Input	-level5 16S.fasta	
	ATTCTGGTTGATCCTGCCAGAGG	
Expected Output		
At least NR_074334.1, NR_118873.1 and NR_119237.1 are displayed.		
>NR_074334.1 Archaeoglobus fulgidus DSM 4304 16S ribosomal RNA, complete sequence >NR_118873. ATTCTGGTTGATCCTGCCAGAGGCCGCTGCTATCCGGCTGGGACTAAGCCATGCGAGTCAAGGGGCTTGTATCCCTTCGGGGGATGCAAGCACC		
Actual Output		

```
C:\Users\lyphu\OneDrive - Queensland University of
NR_074334.1
NR_118873.1
NR_119237.1
NR_041986.1
NR_044558.1
NR_028242.1
NR_074232.1
NR_074239.1
NR_074522.1
NR_074531.1
NR_074110.1
NR_074177.1
NR 102885.1
NR_102894.1
NR_102902.1
NR_103948.1
NR_113412.1
NR_044726.2
NR_044727.2
NR_044723.2
Program finishes. Press Enter to exit.
                                                 The correct id sequences are displayed. 4
Comment
                                                 marks are given.
```

Criteria: Clear error message is provided sequence does not exist (5 mark)

Test 5.4: Test the searching program (Unsuccessful run)

Category	Provide appropriate warning	
Input	-level5 16S.fasta	
	CTGGTACGGTCAACTTGCTCTAAG	
Expected Output		
Appropriate warning is given through the console.		
Actual Output		
□ C:\Users\lyphu\OneDrive - Queensland University of Technology\Bachelo		
Error, DNA sequence CTGGTACGGTCAACTTGCTCTAAG not found		
Program finishes. Press Enter to exit.		
Comment	The appropriate warning is given. 5 marks is	
	obtained	

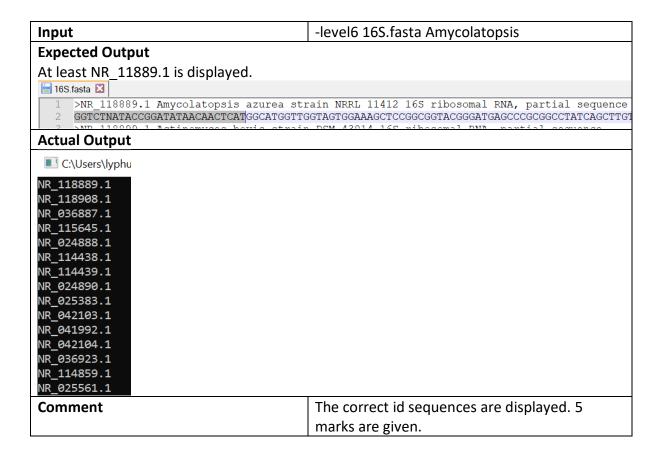
All level 5 functionalities are achieved. 15 marks are obtained.

Level 6 Functionalities (20 marks)

Criteria: The program prints the appropriate data (15 marks)

Test 6.1: Test the searching program with the expected output including the first id sequence (Successful run)

Category	Display the correct data



Test 6.2: Test the searching program with the expected output including the last id sequence (Successful run)

Category	Display the correct data	
Input	-level6 16S.fasta Filifactor	
Expected Output		
At least NR_119037.2 is displayed.		
41533 >NR_119037.2 Filifactor villosus strain DSM 1645 16S ribosomal RNA, partial sequence 41534 AATTATATCGGAGNTCTTCGGGCCAAAGAGATAATAAGCTAGTGGCGGACGGTGCGTAACGTGTGGCTAATCTGCCTTTGTCA		
Actual Output		
C:\Users\lyphu\OneDrive - Queensland University of Tec		
NR_044819.1		
NR_114692.1 NR_041928.1		
NR 074645.1		
NR_119037.2		
Program finishes. Press Enter to exit.		
Comment	The correct id sequences are displayed. 5	
	marks are given.	

Test 6.3: Test the searching program with the expected output including the id sequences on the same metadata line (Successful run)

Category	Display the correct data	
Input	-level6 16S.fasta Archaeoglobus	
Expected Output		
At least NR 074334.1, NR 118873.1 and NR 119237.1 are displayed.		
>NR_074334.1 Archaeoglobus fulgidus DSM 4304 16S ribosomal RNA, complete sequence >NR_118873. ATTCTGGTTGATCCTGCCAGAGGCCGCTGCTATCCGGCTGGGACTAAGCCATGCGAGTCAAGGGGCTTGTATCCCTTCGGGGATGCAAGCACCC		
Actual Output		
■ C:\Users\lyphu\OneDrive - Queensland University		
NR_074334.1 NR_118873.1 NR_119237.1 NR_028831.1 NR_112618.1 NR_028166.1 NR_116822.1 NR_028937.1 NR_118374.1		
NR_074522.1 NR_102885.1 NR_103948.1 NR_125486.1		
Program finishes. Press Enter to exit.		
Comment	The correct id sequences are displayed. 5	
	marks are given.	

Criteria: Clear error message is provided keyword does not exist (5 mark)

Test 6.4: Test the searching program (Unsuccessful run)

Category	Provide appropriate warning	
Input	-level6 16S.fasta RandomKeyword	
Expected Output		
Appropriate warning is given through the console.		
Actual Output		
C:\Users\lyphu\OneDrive - Queensland University o		
Error, keyword RandomKeyword not found		
Program finishes. Press Enter to exit.		
Comment	The appropriate warning is given. 5 marks	
	is obtained.	

All level 6 functionalities are achieved. 20 marks are obtained.

Level 7 Functionalities (10 marks)

Criteria: The program prints the appropriate data (5 marks)

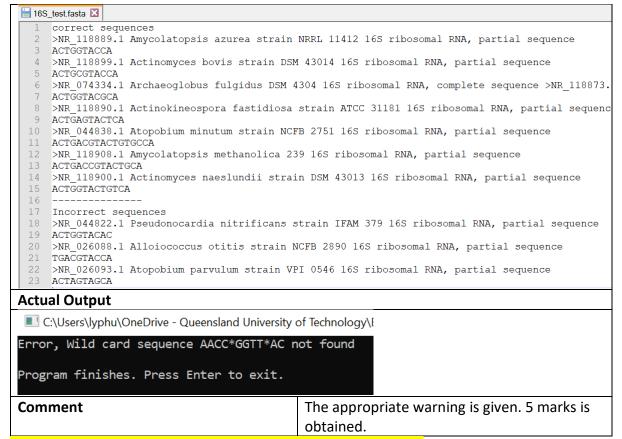
Test 7.1: Test the searching program with a custom fasta file, in which we create correct and incorrect sequences for testing (Successful run)

Category	Display the correct data	
Input	-level7 16S_test.fasta ACTG*GTAC*CA	
Expected Output		
16S_test is included as below. The program should ignore the incorrect sequences and		
output the correct sequences.		
E 16S_test.fasta ■ 16S_test.fasta		
3 ACTGGTACCA 4 >NR 118899.1 Actinomyces bovis strain DSM 5 ACTGCGTACCA 6 >NR_074334.1 Archaeoglobus fulgidus DSM 4 7 ACTGGTACGCA 8 >NR_118890.1 Actinokineospora fastidiosa 9 9 ACTGAGTACTCA 10 >NR 044838.1 Atopobium minutum strain NCF 11 ACTGACGTACTGTGCCA 12 >NR_118908.1 Amycolatopsis methanolica 23 13 ACTGACCGTACTGCA 14 >NR_118900.1 Actinomyces naeslundii strain 15 ACTGGTACTGTCA 16 17 Incorrect sequences 18 >NR_044822.1 Pseudonocardia nitrificans si 19 ACTGGTACAC 20 >NR_026088.1 Alloiococcus otitis strain No	304 16S ribosomal RNA, complete sequence >NR_118873. strain ATCC 31181 16S ribosomal RNA, partial sequence B 2751 16S ribosomal RNA, partial sequence 9 16S ribosomal RNA, partial sequence n DSM 43013 16S ribosomal RNA, partial sequence train IFAM 379 16S ribosomal RNA, partial sequence	
21 TGACGTACCA 22 >NR_026093.1 Atopobium parvulum strain VP 23 ACTAGTAGCA	I 0546 16S ribosomal RNA, partial sequence	
Actual Output		
C:\Users\lyphu\OneDrive - Queensland Univer NR_118889.1 NR_118899.1 NR_074334.1 NR_118873.1 NR_119237.1 NR_118890.1 NR_044838.1 NR_118908.1 NR_118900.1 Program finishes. Press Enter to exit		
Comment	The correct id sequences are displayed. 5	
	marks are given.	

Criteria: Clear error message is provided sequence does not exist (5 mark)

Test 7.2: Test the searching program (Unsuccessful run)

Category	Provide appropriate warning	
Input	-level7 16S_test.fasta AACC*GGTT*AC	
Expected Output		
16S_test is included as below. Appropriate warning is given through the console.		



All level 7 functionalities are achieved. 10 marks are obtained.

In total, 50 marks for L4-6 and 10 marks for L7 = 60 marks should be obtained.