**BLAST Algorithm Steps**:

1. User enters the DNA Sequence that he/she wants to search in the database (The database for this project would be a txt file).

Example Output of the program:

Please enter the DNA Sequence that you would like to match:

User entry: ATGCCCGTCATTCC

1. The program then searches the Database for exact or close matches.
2. The first step for search is to break the user entry into 3 letters:

Example: ATGCCCGTCATTCC

The program breaks this up into 3 letter words: ATG, TGC, GCC, CCC,…………………

1. The program then searches the Database (The txt file) for sequences that match the 3 letter words.

**DNA Sequence Database**:

The 3 Letter Words (User Entry)

ATG

TGC

GCC

CCC

GTCATGCCCGTCATTCC

GGGGATGCCCGGGGG

TTTTATGCCCGTCGAAG

TAATGCCCGTTTTTTTT

GCCATGCCCGTTACCCC

1. After finding the “Hook” by matching the 3 letter words to the database, the algorithm then moves left and right along the DNA sequence of interest.

For example: In the above instance, the first Sequence was of interest since we were able to find the 3 Letter Word: ATG in the sequence.

GTC**ATG**CCCGTCATTCC

1. Once it moves left and right it will keep a score, and try to match the user entry with the DNA sequence of interest from the database.

With every match, the algorithm adds +1 to the overall score, and -1 for any mismatch.

In the above example:

**User Entry:** -- ATGCCCGTCATTCC

**Database (Sequence of Interest**): TC**ATG**CCCGTCATTCC

|  |  |  |
| --- | --- | --- |
| User Entry | Database | Score |
| - | T | -1 |
| - | C | -1 |
| A | A | 1 |
| T | T | 1 |
| G | G | 1 |
| C | C | 1 |
| …….. | ……. | ……………. |

Overall Score: XXXX

1. After the algorithm matches the user entry with the sequence of interest it provides an output to the user that looks something like this.

**User Entry:** -- ATGCCCGTCATTCC

**||| ||||||||||||**

**Database (Sequence of Interest**): TC**ATG**CCCGTCATTCC

(The program will show lines where the letters match, and won’t show any lines where there is no match).

1. It will also provide the user with the score, that it calculated in the step 6 and a percentage value of how much matched.