Project 5 Improvement

Gene Sequence Alignment

Duane Johnson
CS 312 Algorithms / BYU • Section 1 • March 22, 2008

Purpose of Improvement

This improvement adds functionality to the user interface that was lacking in the original project. Specifically, it adds a text area on the right-hand side of the project with a “Show Alignment” button near the top which will call the Extract method and display the results of an extraction / alignment of the selected gene sequences.

Improvement Implementation

User Interface

The user interface has three important elements:

1. The ability to highlight any cell in the DataGridView by double-clicking,
2. A "Show Alignment" button, and
3. A text area where the results are displayed

Implementation

Since the DataGridView is dynamically created at runtime, the double-click event must be programmatically added at runtime as well. For example:

```
public void dataView_CellDoubleClick(
     object sender, DataGridViewCellEventArgs args)
{
    DataGridViewCellStyle selectedStyle = new DataGridViewCellStyle();
    selectedStyle.BackColor = Color.Yellow;

    DataGridViewCellStyle normalStyle = new DataGridViewCellStyle();
    normalStyle.BackColor = Color.White;

    m_dataView[m_selCol, m_selRow].Style = normalStyle;
    m_selCol = args.ColumnIndex;
    m_selRow = args.RowIndex;
    m_dataView[args.ColumnIndex, args.RowIndex].Style = selectedStyle;
    m_dataView[args.ColumnIndex, args.RowIndex].Selected = false;


    Console.WriteLine("double clicked : " +
            args.ColumnIndex.ToString() + ", " +
            args.RowIndex.ToString());
}

public ResultTable(DataGridView dataView, int numberOfSequences)
{
    m_dataView = dataView;
    m_numberOfSequences = numberOfSequences;
```

```
            // Dynamically bind the double click event to our method above
            dataView.CellDoubleClick +=
                     new DataGridViewCellEventHandler(dataView_CellDoubleClick);

              //....

    }
```

In addition, the output of the Extract method must be changed from the Console object to the TextBox object now in the user interface. This is easily accomplished by passing in a TextBox object to the Extract method:

```
        public void Extract(
              GeneSequence seqA,
              GeneSequence seqB,
              System.Windows.Forms.TextBox tbox)
```

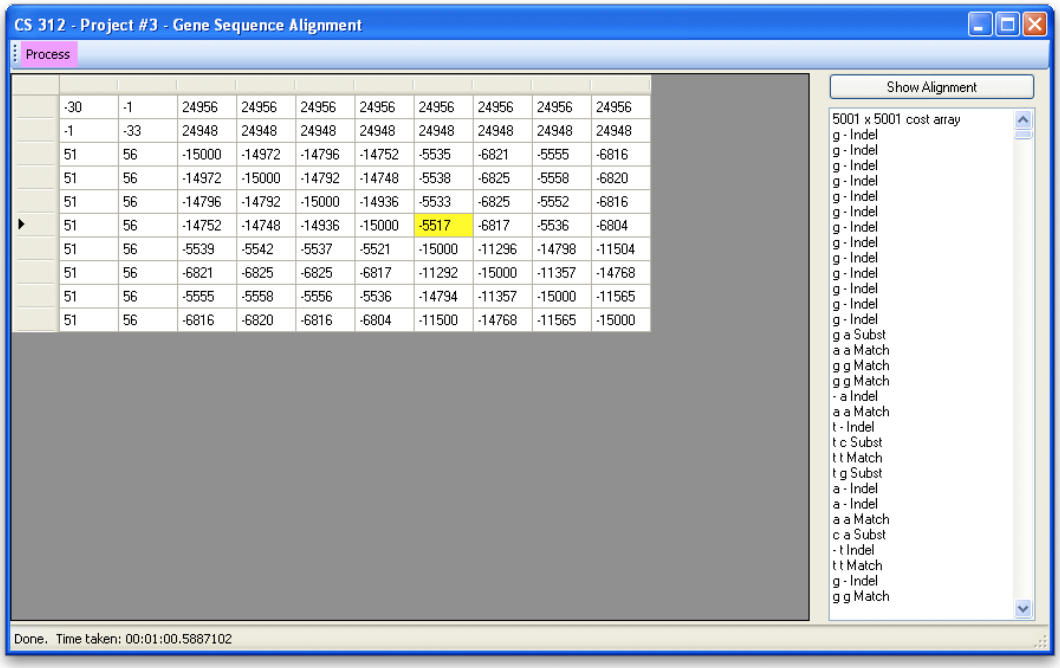
And then, after creating a temporary string, the Extract method assigns the string to the textbox Text property:

```
      String content = "";
      // ...
      content += charA.ToString() + " " + charB.ToString();
      // ...
      tbox.Text = content;
```

Emperical Evidence

SCREENSHOTS

The following screenshots show the results of double-clicking on a cell (highlighted in yellow) and then clicking the "Show Alignment" button:

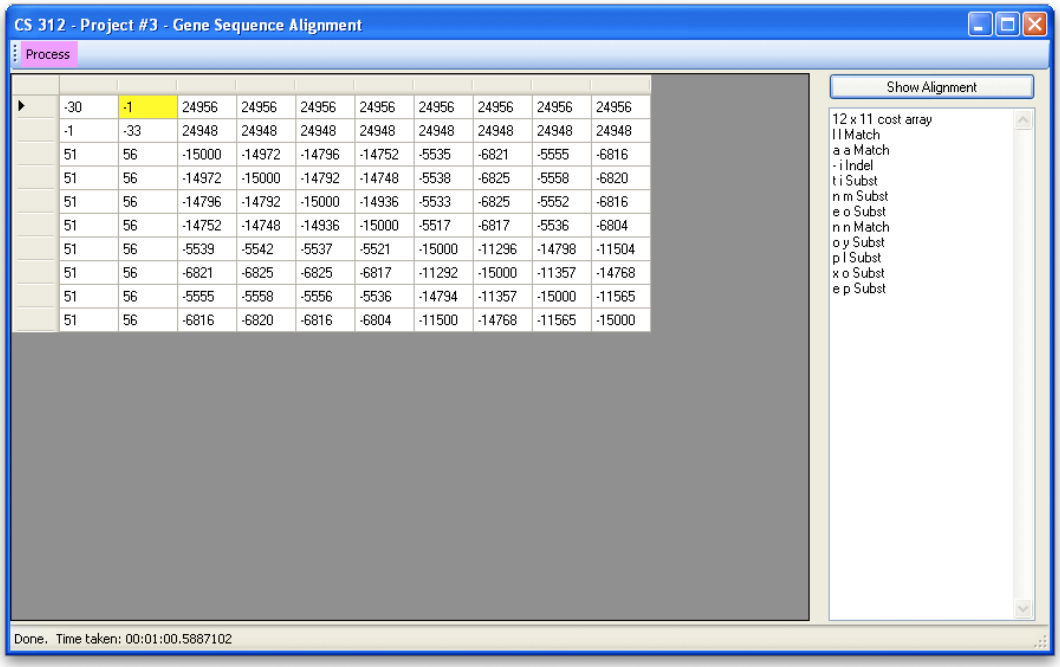


-30	-1	24956	24956	24956	24956	24956	24956	24956	24956
-1	-33	24948	24948	24948	24948	24948	24948	24948	24948
51	56	-15000	-14972	-14796	-14752	-5535	-6821	-5555	-6816
51	56	-14972	-15000	-14792	-14748	-5538	-6825	-5558	-6820
51	56	-14796	-14792	-15000	-14936	-5533	-6825	-5552	-6816
51	56	-14752	-14748	-14936	-15000	-5517	-6817	-5536	-6804
51	56	-5539	-5542	-5537	-5521	-15000	-11296	-14798	-11504
51	56	-6821	-6825	-6825	-6817	-11292	-15000	-11357	-14768
51	56	-5555	-5558	-5556	-5536	-14794	-11357	-15000	-11565
51	56	-6816	-6820	-6816	-6804	-11500	-14768	-11565	-15000

Alignment of Gene Sequences 6 & 7 (row 6, col 7)

Explanation

The first 5000 base-pairs of each sequence are compared and an alignment result is shown in the text box. Wherever a match is found, the word "Match" is used; wherever a substitution is required, the word "Subst" is used; and wherever the sequence required an insertion or deletion, the word "Indel" is used.



-30	-1	24956	24956	24956	24956	24956	24956	24956	24956
-1	-33	24948	24948	24948	24948	24948	24948	24948	24948
51	56	-15000	-14972	-14796	-14752	-5535	-6821	-5555	-6816
51	56	-14972	-15000	-14792	-14748	-5538	-6825	-5558	-6820
51	56	-14796	-14792	-15000	-14936	-5533	-6825	-5552	-6816
51	56	-14752	-14748	-14936	-15000	-5517	-6817	-5536	-6804
51	56	-5539	-5542	-5537	-5521	-15000	-11296	-14798	-11504
51	56	-6821	-6825	-6825	-6817	-11292	-15000	-11357	-14768
51	56	-5555	-5558	-5556	-5536	-14794	-11357	-15000	-11565
51	56	-6816	-6820	-6816	-6804	-11500	-14768	-11565	-15000

Alignment of Test Sequences 1 & 2 (row 1, col 2)

Explanation

Just as in the first example, this screenshot shows the Match / Subst / Indel alignment display in the text box. However, because this is a comparision of test data, the results are not true DNA base pairs, but letters of the alphabet ("exponential" vs. "polynomial").