# Use Case Action Sequence in Detail for Database Conversion

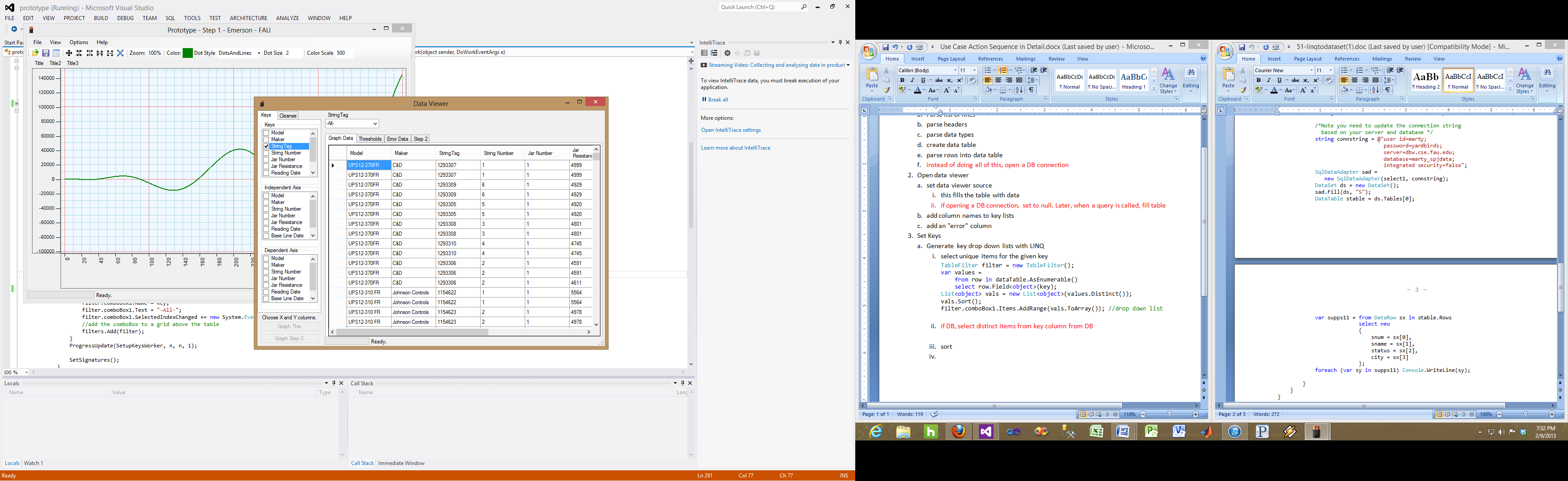
Black indicates stuff already implemented or user actions  
Red indicates stuff to change or add regarding databases.  
Green indicates optional changes.  
Blue indicates email quote  
Purple indicates important concepts  
~~Crossed Out~~ indicates ideas I've rejected/ tasks completed

Note in the documentation that the data source of the graph is the table in the data viewer, which means the table displayed in the data viewer is still relevant.

SQL queries should be constructed with values wrapped in 'quotes' even if value is a number

I'm not sure how to handle batch cleansing at the moment.

~~Future work: I may want to scrap this data viewer window and create something else. Use linq to sql classes (see linqtosql2.doc). Create a rule based system to generate queries~~

1. Open CSV
   1. Parse list of lines
   2. parse headers
   3. parse data types
   4. create data table
   5. parse rows into data table
2. Or, If using a DB:
   1. request server, username, password, database, table  
      
   2. open connection form window, lock dataviewer until connection form closes
   3. ~~show some kind of interface graphic telling if using DB or CSV (refresh button~~) (title bar)
   4. create new states to separate CSV code from DB code
   5. handle DB errors
   6. add refresh button
3. Open data viewer
   1. set data viewer source
      1. this fills the table with data
      2. if opening a DB connection, set to null. Later, when a query is called, fill table
   2. add column names to key lists
      1. if DB, get column names
   3. add an "error" column
      1. if DB, add error column to database
4. Select a key from key list
   1. Generate key drop down lists (TableFIlter objects)   
      
      1. generated when keys list changes, fill immediately
      2. if DB, fill/update only when the drop down is opened, only after refresh button is pressed or empty
      3. CSV: select unique items for the given key
      4. DB: select distinct items from key column from DB  
         
      5. I don't want to call the same query twice if I don't need to. I should store premade TableFilters. TableFilters are created or recalled when a key is chosen; filled or updated when refresh button is pressed. Load from signature table.
      6. **~~Optional:~~** ~~Instead of TableFilters appearing above the table, cause them to appear under a chosen key in the keys list. Need to change the type of the checkedItemsList and change any code referencing checked items.~~
5. Drop Down Filter menu
   1. DB: load from signature table
6. Select a Key Value from filter
   1. Filter the table on each key value
   2. Occurs when a filter drop down changes
   3. If DB, do not filter immediately, instead filter when refresh button is pressed
   4. CSV: Actual data displayed in data viewer uses DataView objects. FilterViewCSV()
   5. DB: select rows matching the selected filters, **sorted** by ind axis FilterViewDB()
   6. ~~I probably don't want to select everything for a single jar here, because it might be too much stuff~~
   7. **~~Optional~~**~~: select top 1000 results. Select next top 1000 results. Select All. Select Top x results.~~
   8. filter step 2 tab and error tab
   9. sort
   10. Remove sort from any DB code (handled by SQL query)
   11. store a table of unique signatures SetSignatures()
   12. DB: store a table of unique signatures, retrieving data by SQL query SetSignaturesDB()
   13. when ind axis is chosen, sort table
   14. when dep axis is chosen, set signature table
   15. DB: remove sort, only set signature table on create/refresh
7. Choose Threshold keys
   1. CSV: SetThresholdTable() 
   2. DB: store a table of unique signatures, retrieving data by SQL query see 5L above. SetThresholdTableDB()
8. Refresh Button
   1. refresh everything
   2. SetDataSource
   3. refresh key lists
   4. Get signature table
   5. fill out filters from signature table
   6. keep track of selected keys and selected key values before refresh, restore them after refresh
   7. ~~Should I allow refresh of certain UI elements with more buttons, or everything at once with one button?~~
   8. ~~Probably not a good idea to refresh everything when only threshold query needs to be sent.~~ situational
9. Cleanse Start
   1. Cleanse All radio buttons seem to be irrelevant
   2. find any rows with something marked in the error column
   3. DB: merge error table into main table before cleansing
      1. rows in the database may or may not already have marks in the error column
      2. ~~there are two ways to handle this: include or ignore pre-marked data~~ include
      3. ~~I could add an option to choose~~
      4. Errors shall be filtered from the main view when loaded FilterViewDB() using the select query
      5. before cleansing, send new query for all data including error data
      6. when error tab is shown, send new query for error data
   4. iterate through each row, mark rows that pass error check as errors
   5. DB: Cleanse data
      1. locally
      2. update database with updated datatable
   6. Repeated Data test
      1. for each row, check if it matches the previous row
      2. ~~This test requires iteration and knowledge of the previous row. Can i make a query that handles this or do I need to iterate through rows?~~
      3. ~~I need to mark rows that are repeated values. Something like: update table set error='repeat' where error=' ' and tag=123 and (previous row's resistance = this row's resistance) Can I do that without iterating through every row?~~**~~if you only want to update the repeated values, I can only think of using an iterator to check if the value exists already in the database~~** ~~I am currently doing this in C# code:~~

~~foreach row A,   
   get A.resistance,   
   if prev == A.resistance,   
      mark A.error with "repeat"~~

~~else~~

~~prev = A.resistance~~

~~This marks any rows that have the same value as the previous row. New values are not marked.~~

~~I imagine I could use transact-sql loops of some sort to do this. I just can't think of any easier way. Some references:~~[~~http://msdn.microsoft.com/en-us/library/ms187953%28v=sql.105%29.aspx~~](http://msdn.microsoft.com/en-us/library/ms187953%28v=sql.105%29.aspx)[~~http://msdn.microsoft.com/en-us/library/ms178642.aspx~~](http://msdn.microsoft.com/en-us/library/ms178642.aspx)

~~Do you think I'm on the right track?~~

* 1. Box/Whisker test
     1. creates views on each signature and runs test on that signature.  
        
     2. ~~DB: create a DB view matching signature and passes test~~
     3. ~~this test is probably going to require some ridiculous sql query or stored procedure~~
     4. ~~create select statements for Q2, Q1, Q3; math; select rows that pass test~~
     5. ~~I need to perform a box/whisker test on specific row sets. The test requires information from three specific rows located at 3n/4, n/2, and n/4 (where n is number of rows), does some math, and then just selects rows above or below thresholds. How do I select a specific row (by row number) from a specific set of rows in a database table?~~

**~~After searching for this, I saw this stackoverflow post that explains how to do it in various RDBMSs ,~~**

[**~~http://stackoverflow.com/questions/16568/how-to-select-the-nth-row-in-a-sql-database-table~~**](http://stackoverflow.com/questions/16568/how-to-select-the-nth-row-in-a-sql-database-table)

1. ~~Cleanse Options remove:~~
   1. ~~Cleanse all won't work with DB~~
   2. ~~Automatic graph won't be relevant if you have to click refresh anyway.~~
2. ~~Tests:~~
   1. ~~Is Using DB queries actually faster than downloading an entire jar reading set and performing tests with LINQ queries?~~ yes
   2. ~~Is iterating through each row (with LINQ) slower or faster than using SQL select statements?~~ slower
3. ~~Check Worker Completed functions for state switches; add states for DB~~
4. ~~Add states for cleanse steps~~
5. Graph, Form1 does not manipulate data
6. ~~Form1 add state for DB connection~~
7. ~~DataViewer add overload for SetDataSource for DB~~
8. Performance:
   1. ~~only refresh parts that need to be refreshed~~
   2. modify graph to only display points that are significantly different
   3. modify graph to show a maximum amount of points
      1. add interface to modify maximum points
   4. look for any point in a use case where the program becomes frozen. Find the associated code and wrap with a background process.
   5. add an ok button to connect form; close on keyboard enter pressed
   6. wrap everything in try catch blocks
   7. limit queries to the DB
      1. grab info from Signature table instead of from database whenever possible. see SetupKeysDB() and filters
   8. put refresh button in a background process
      1. restore choices where appropriate at end of associated background processes
   9. ~~Show progress of queries in a generalized way~~ cant show progress of queries
9. Testing
   1. look for behavior that causes errors
   2. look for behavior that causes crashes
   3. look for behavior that works counter-intuitively
   4. look for unexpected behavior
   5. look for laggy or unresponsive interface
10. **Save user settings**
    1. <http://msdn.microsoft.com/en-us/library/a65txexh.aspx>
11. There is some weird stuff in the StringTag column
    1. String
    2. UPS 1-1
    3. Every jar on the string tag 1114866 shows only one resistance value over 2 years
12. CellNumber goes up to 240
13. ~~The first call of a query to the BatteryReadings table takes so long that it times out, but a call of the same query will be done in less than a second. The database seems to remember results of queries.~~
14. Security
    1. use SQLCommands parameterized to protect against SQL injections

Known Issues

1. ~~when loading table from db, and the table already has error values, should i skip them or add them to the main view. should i add errors to error tab?~~ skip, load error table from database when focused
2. ~~how do i set a value to nothing? set datatable value to '' or null?~~ blank quotes ''
3. graph goes blank if you zoom in too far on nothing
4. graph is lagging if you zoom in too far
5. ~~loading error tab doesn't work?~~
6. ~~table empty after cleanse~~
7. ~~repeats not working~~
8. ~~Thresholds~~
9. ~~Step 2~~
10. ~~block graph button while cleansing~~
11. ~~empty columns after cleansing~~
12. ~~let step2 button be more accessible~~
13. ~~lines drawn between points are sometimes be drawn from top to bottom instead of left to right~~ solved
14. ~~FilterViewDB is called twice?~~
15. Norton thinks this is a virus. Create as a silverlight app?
16. Allow more than one dependent axis
17. graph multiple data sets
18. ~~see if tab to step 2 works~~
    1. ~~enable buttons when step 2 data is loaded~~
    2. ~~or get step2 data when graph step 2 button is pressed~~
19. ~~problem updating in step2 new rows??~~
    1. ~~it seems there is a conflict with "new rows" that seem to be generated from step2 program~~
    2. ~~step2 program deletes duplicate rows~~
    3. ~~winmerge shows no "new rows" only removed duplicates~~
    4. ~~there is a loss of precision with date-times when writing to file~~
    5. ~~can randall's program handle dates with times?~~
20. ~~step2 table grid view doesn’t show RegStat 0s but graph does~~
21. ~~step2 graph~~
22. fix cursor label on datapoints when near a datapoint or on cursor otherwise
    1. and show RegStat value
23. selectable signatures that don’t exist
    1. When opening drop down list, get list from signature table, then filter signature table by selected keys
    2. When another drop down list is opened, the list should be filtered by signatures that only use the already selected keys.
24. ~~third+ filter not showing~~
25. ~~move refresh button to left~~
26. ~~refresh button is marking first ind and dep axis if they were empty~~
27. ~~re-filter datatable if new keys are chosen4~~
28. FAQ:
    1. The filter drop down lists are empty...
       1. Press refresh
    2. How do I un-cleanse?
       1. Uncheck the cleanse options and press Cleanse. It will then merge the error table to the main table.
29. ~~refresh screen on choosing dot style~~
30. CSV:
    1. ~~step 2~~
    2. ~~keys are removed and not returned when checking/unchecking~~
    3. ~~un-cleanse~~
    4. ~~cleanse-all~~
    5. ~~graph step 2 available anytime~~
31. ~~Able to graph without choosing keys~~
32. ~~Readme:~~ 
    1. ~~Note that the DB version does not show every column and every row because we leave that to whatever database software is available to the user.~~
33. Regarding our future work later this semester:
    1. In the context of Baye’s Rule, are temperature, resistance, and voltage independent of one another? In other words,
    2. Does the event of a change in temperature have any effect on the probability of a change in voltage or resistance?
    3. Does the event of a change in voltage…
    4. Does the event of a change in resistance…
    5. If “yes” or “We don’t know” then we should assume resistance, voltage, and temperature are dependent and look for events of change and their effect on others’ probability.
    6. If “No” then we can assume independence between r,t,v and we only need to look for dependence between events occurring along reading date for r, t, or v.
34. “Events” or “Patterns”
    1. These are piecewise or continuous trends in f(x) found within a given block.
    2. A nice way to represent a trend would be with heat maps. See Prince Computer Vision pg11
       1. Place a heat map on a section of the graph. Points are weighted based on their position in the heat map. If the points in the section on the graph add up to a certain degree of probability, then that section of the graph can be labeled with the associated pattern in the heat map.
    3. Growing heat maps:
35. Blocks
    1. These are sections of f(x) specified in ranges of x.
    2. The step2 algorithm may be a good indicator of places to create block ends.