Title: SYSTEMS AND METHOD FOR IMPROVED DATA PROCESSING

Abstract: The present invention enables analysis of calculations contained within cells of spreadsheet applications by determining and displaying a list of direct and indirect predecessors for the calculation in question.
"Systems and Method for Improved Data Processing"

The present invention relates to the field of information processing by electronic means. The invention is particularly concerned with the processing and presentation of information by computer program applications, and in particular that contained in electronic spreadsheets.

Electronic spreadsheets, such as Microsoft® Excel, are known, and offer improvements in creating, editing and using financial models. Electronic spreadsheets permit the user to input and manipulate data, through the use of standard computer input devices, and to view the results of data processing performed by the spreadsheet on the visual display unit. Spreadsheets are essentially made up of a large grid, with each box or cell in the grid having a unique reference number such as A1, G8, Z15 and so on. Each cell can contain one value, which is either an input, or is calculated based on one or more
other values. Modern spreadsheets have extended
this, so that the values can depend on values in
other sheets, and can give rise to graphs or other
reports that do not fit in cells. There is a more
detailed description of what a spreadsheet is in the
existing US Patent 6,138,130 ("System and Method for
Processing Data in an Electronic Spreadsheet in
 Accordance with a Data Type", Adler et al.).

The values in spreadsheets, as stated above, can be
defined in terms of values in other cells (or
elsewhere in modern spreadsheets). These values can
themselves depend on other values and so on. If you
consider the first cell to be at the base of a tree,
and the link to each preceding cell as a branch (and
the links to the preceding cell of the preceding
cells as branches and so on), then this gives rise
to a calculation tree for any particular value. The
term "precedents" is used here to mean the set of
all the cells in this tree except the first cell
(i.e. all the cells that the first cell explicitly
or implicitly depends on). The term "direct
precedents" is used here to mean the cells that a
particular cell explicitly refers to in its
calculation.

It is advantageous in certain circumstances for the
end user to trace the calculation of the value of a
particular cell. This involves displaying all or
part of the calculation tree for that value. Some of
the branches in the calculation tree are more
relevant for determining the value of the cell of
interest than others. The disadvantage of existing spreadsheet programs, or other programs that analyse their output, is that when tracing the calculation of the value of a particular cell, a large number of cells are displayed, many of which are redundant for the purpose of that particular cell calculation. Thus a large amount of redundant information is displayed, rendering the analysis of the cell’s value unnecessarily time-consuming and complex for the end user.

According to an aspect of the present invention, there is provided a computer-implemented method of data processing involving the steps of:

(i) The user selecting a cell to analyse;

(ii) parsing the cell’s calculation rule into separate elements;

(iii) analysing each element and determining its nature;

(iv) determining the cell or cells that are referenced by each element determined in step (ii);

(v) displaying the cells from the previous step, except those which do not directly contribute to the value of the cell chosen by the user, including those which have a blank or zero value.
According to a second aspect of the present
invention there is provided a computer based system,
comprising data storage means, visual display means,
data processing means, including computer software
and computer hardware, and a user interface.

Preferably the data storage means is any suitable
electronic memory unit, such as a computer disc.

Preferably the visual display means is a monitor.

Preferably the data processing means is a computer
program as defined according to the present
invention.

According to a further aspect of the present
invention there is provided a computer programmed to
store given alphanumeric information, and to perform
mathematical calculations and specialised functions
in accordance with user requests, such that only
information related to said request is displayed to
the user, and redundant information is hidden from
the user.

According to a still further aspect of the present
invention there is provided a computer program
product comprising a computer readable medium,
having thereon computer program code means, said
program when loaded performing the method as defined
by the computer program described herein.
For the avoidance of doubt, the program as described herein may be either loaded independently from a suitable spreadsheet program or may perform the operations of the program as part of a suitable spreadsheet program. Said program, if a separate program, then integrates with the spreadsheet program such that it may be accessed directly when running the spreadsheet program by way of, for example, an icon on the toolbar.

The present invention will now be described by way of example only with reference to the following drawing in which:

Fig 1 illustrates a basic spreadsheet containing alphanumeric information.

Fig 2 illustrates an operational model of one embodiment of a data processing method in accordance with the present invention.

Referring now to Fig 1, each cell within the spreadsheet has a value 10; this value may be either: a value 12 which is input by the user or a value which is the result of one or more mathematical or other calculations 14, or it may be a combination thereof. A simple calculated cell may be the result of, for example, a multiplication 16 or an addition 18. Cell G8, 18, is calculated as A3, 14, plus B5, 16. Cell A3, 14, is calculated as A1*A2, and cell B5, 16 is calculated as C1*D6. If the value of a particular cell B5, is zero, 16, then
this cell does not contribute to the value of G8, and thus constitutes redundant data for this calculation tree. Each cell’s calculation rule is parsed for the purpose of determining the calculation tree, that is, broken down into discrete elements. If the element is a reference to a cell or a block of cells, then the program analyses each cell in that block. If the value of the cell is not blank or zero then it is added to the list of direct precedents for the cell of interest.

Referring now to Fig 2, if the element, when parsed is a function, rather than a simple mathematical calculation, then it has zero or more arguments. If the element is the IF function 20, then it has three arguments: a condition, a TRUE part and a FALSE part. The program determines if the condition is true or false 22, and then analyses the cell or cells that the true or false part respectively refers to. Cells with a value that is not blank or zero are added to the list of direct precedents. The cells referred to by the condition and the other argument are not added to the list of direct precedents.

If the element is the HLOOKUP or VLOOKUP function 26, then the program determines the range of cells that are contained in the lookup table 28. The column number 30 and the row number 32 are determined, and the single relevant cell is displayed 34. This cell is then added to the list of
direct precedents, unless its value is blank or zero

36.

If the element is the INDIRECT function 38, then the
program will determine a second cell reference from
a first cell 40 (that reference is the value of the
first cell) and from the cell or block of cells
referred to by the second cell reference, determine
the value 42. Cells with a value that is blank or
zero are not added to the list of direct precedents.

In use, cells contain either a value which has been
input by a user or a value which has been calculated
in some manner. The user chooses a cell of
interest. If the cell contains a calculated value,
the calculation rule for the cell of interest is
parsed into elements. Each element is then analysed
and its nature determined. If the element is a
reference to a particular cell or block of cells,
then each cell with relevant information is added to
a list of direct precedents for the cell of
interest. Any cells with a value that is blank or
has a value of zero is not added to the list of
precedents. The program then moves on to the next
element of the calculation, until such time as all
elements in the calculation have been analysed.

If the element is a function, rather than a simple
mathematical calculation, then each separate
function involves a different series of steps, in
order to determine the list of direct precedents for
the cell of interest.
Such functions may include the IF function, the VLOOKUP function, the HLOOKUP function, and the INDIRECT function.

If the element is the IF function then it has three arguments: a condition, a TRUE part and a FALSE part. For example, the function could be "IF (A1=3, B1, C1)" which means "if the value of A1 is 3 then this function returns the value in B1, otherwise it returns the value in C1". The program determines whether the condition is true or false, and then takes the appropriate argument and analyses the cell or cells that it refers to. Cells with a relevant value are added to the list of direct precedents, and cells with blank or zero values are ignored.

If the element is the HLOOKUP or VLOOKUP function, then it has four arguments: a value to lookup in the first row or column of a table, a reference to the table, the column or row to use and possibly a value saying whether to find an exact match when doing the lookup. For example, the function could be "HLOOKUP ("John", A1:C5, D1, FALSE). This will look for the text "John" in the top row of the table of cells A1:C5. If a match is found, this will give the number of the column to use. The function then evaluates D1 to give the row number to use, and picks the single element out of A1:C5 that is specified by the row and column number it has just worked out. The final argument, FALSE, is to specify that the function should look for an exact match in the top row of the table. If this argument is TRUE,
then the function will find the nearest match (in alphabetical or numerical order) and use that column. VLOOKUP works in the same way except that "row" and "column" are reversed in the above explanation. The present invention determines the row and column that the function will use, and then adds the relevant cell from the table to the list of direct precedents. The other cells in the table, and any cells referred to by the other arguments, are not added.

If the element is the INDIRECT function, then it has two arguments. This function is unusual in that it does not return a value that is specified in one of its arguments. Instead, what it does is look in the cell specified in the first argument and get the value of that cell. It then uses that value as a reference to a second cell or block of cells, and it returns the value in that cell or block of cells. For example, if A1 contains the text "C5" and cell C5 contains the value 100, then INDIRECT (A1) gives the value 100. The second argument (which is optional) specifies what type of address is contained in the cell mentioned in the first argument. The present invention evaluates the first argument and work out what cell it refers to. It then looks in that cell and uses its value to determine what the cell or block of cells is that is really being referred to. It then analyses that block of cells ignoring cells with values of blanks or zeros. Once it has finished analysing the cell or block of cells, or the specialist function and its
arguments, the invention moves on to the next
element. If there is no next element then it
displays the list of direct precedents that it has
gathered.

The invention offers significant advantages in the
ease of navigating complex spreadsheets and avoids
redundant data display. The history of calculations
and equations calculated in a spreadsheet is
accessible with speed and accuracy.

It will be appreciated that improvements and
modifications may be incorporated without departing
from the scope of the invention as defined herein.
Claims

1. A computer-implemented method of data processing involving the steps of:
   (i) a user selecting a cell to analyse;
   (ii) parsing the cell's calculation rule into a plurality of separate elements;
   (iii) analysing each element and determining whether the element is a reference, a
          mathematical calculation or a function;
   (iv) determining a precedent cell or precedent cells that are referenced, directly or
        indirectly, by each element determined in step (ii);
   (v) determining a precedent cell or precedent cells that are a result of an element which
        is a function;
   (vi) displaying the precedent cell or cells from steps (iv) and (v), except those which do
        not directly contribute to the value of the cell chosen by the user, including those
        which have a blank or zero value.

2. A method as claimed in claim 1 wherein the step of displaying the precedent cell or cells further
   comprises organising the precedent cell or cells in a tree structure to enable a hierarchical
   display of precedent cells.

3. A computer based system for implementing the method of claim 1, comprising data storage means,
   visual display means, data processing means,
including computer software and computer hardware, and a user interface.

4. A computer programmed to implement the method of claim 1.

5. A computer program product comprising a computer readable medium, having thereon computer program code means, said program when loaded performing the method as claimed in claim 1.

6. A computer-implemented method of data processing substantially as herein described with reference to Figure 1 and 2.
FIGURE 2

1. Is it an IF function?
   - No: Determine if the condition is TRUE or FALSE
     - TRUE: Analyse the appropriate argument as a formula
     - FALSE: Determine the row number

2. Is it an HLOOKUP or VLOOKUP function?
   - No: Evaluate the final argument
     - HLOOKUP or VLOOKUP: Determine the column number
   - Yes: Find the range that is the lookup table

3. Is it an INDIRECT function?
   - No: Work out what range is really being referred to
     - Yes: Analyse the range really being referred to as if it were a formula