The ‘arrayjob’ package
Management of arrays in (IA)\TeX

Zhuhan Jiang
School of Mathematical and Computer Sciences
University of New England
Armidale
Australia
email: zhuhan@turing.une.edu.au

(Documentation: Denis Girou* (CNRS/IDRIS – France) – Denis.Girou@idris.fr)

Version 1.03
May 3, 2000
Documentation revised May 31, 2000

Abstract

This package provides array data structures in (IA)\TeX, in the meaning of the classical procedural programming
languages like Fortran, Ada or C, and macros to manipulate them. Arrays can be mono or bi-dimensional.
This is useful for applications which require high level programming technics, like algorithmic graphics pro-
grammed in \TeX.

Contents

1 Introduction 2
2 Command reference 2
3 Examples 7
  3.1 Basic examples ................................................. 7
    3.1.1 Plot labels ................................................... 9
    3.1.2 Checkboard drawing ....................................... 10
  3.2 Advanced examples .......................................... 12
    3.2.1 Example with recursion usage ............................. 12
    3.2.2 Structured dynamic diagrams on a grid ..................... 13
    3.2.3 Another example of a structured dynamic diagram on a grid 18

* All errors and misunderstandings are mine.
1 Introduction

One of the big advantages of the (L)\TeX system over common interactive software for text processing is that it offers, also, a programming language, which gives to people who have some knowledge in the algorithmic and programming fields, an exceptional flexibility and power.

Nevertheless, \TeX is a rather specific programming language, based on macros expansion, which implement a lot of unusual constructs but not some ones very familiar in the classical procedural languages, as arrays to store and retrieve arbitrary pieces of data, stored in a structured way. The main reason for which this was not integrated in \TeX is that this is mainly useless in a language focused on text processing (but METAFONT, for instance, has them). Nevertheless, one of the few applications where this is straightforward to use them is to program a mailing system, where nearly the same informations are to be formatted several times, depending of values which can be simply retrieved from an array of data. This was the goal of the ‘formlett’ package [7], written in 1993-1995 by Zhuhan Jiang for dealing with mass letters, invoices and similar tasks of some duplicative nature. It integrate a small but powerful set of macros to manage arrays, which have been extracted to form the present ‘arrayjob’ package. The arrays can be mono or bi-dimensional\(^1\) and they are dynamically allocated (so we have not to declare statically a dimension for them)\(^2\).

Array structures are, at the opposite of text management, often very useful in graphic programming. This is why the \texttt{(AI)Dra\TeX} package from Eitan Gurari [4] (see also [5]) integrate such functionality (but which is bundled inside this package and could not be extracted easily from it), and this is also the case for the \texttt{METAPOST} package [6] (but this one do not use \TeX as programming language)\(^3\). This fact explain why most of our examples in this documentation will concern the area of graphic programming (here using the \texttt{PSTricks} package from Timothy van Zandt [9]).

2 Command reference

\texttt{\newarray} : Define a new array.

Syntax: \texttt{\newarray\langle ArrayName\rangle}

Example: \texttt{\newarray\Values}

\texttt{\delarray} : Delete an array.

Syntax: \texttt{\delarray\langle ArrayName\rangle}

Example: \texttt{\delarray\Values}

Remarks:

\(^{1}\)You can use more than two dimensions, but not in the meaning of the classical programming languages (see page 5).

\(^{2}\)Stephan von Bechtolsheim [1, Volume III, paragraph 20.3, page 136] also demonstrate such macros for array management in the third volume of his huge book, but it was limited to monodimensional arrays.

\(^{3}\)Another required feature, used in conjunction with the managements of arrays, is a generic loop mechanism. \TeX offer the \texttt{\loop} macro and \texttt{L\TeX} the \texttt{\whiledo} macro of the ‘\ifthen’ package (and also the internal \texttt{\for}, \texttt{\while}, \texttt{etc.} macros), but a more high-level structure is to be preferred, as the ones defined too in the \texttt{(AI)Dra\TeX} or \texttt{METAPOST} packages. We will use in our examples the ‘multido’ package [8], also written by Timothy van Zandt, but others are available, like the ‘\repeat’ package written by Victor Eijkhout [2].
1. obviously, elements of a deleted array could not be accessed later,
2. take care that the elements of a deleted array are not themselves deleted.

So, \删除array\Data \newarray\Data \Data(I) can produce a strange behavior. Just avoid to reuse deleted arrays.

\ArrayName: Store or get the content of the array element specified by the indice(s). In the last case, the content is inserted at the current point.

Syntax: \(\ArrayName\)(I)=\{\Content\} or more generally
\(\ArrayName\)(I1,...,In)=\{\Content\} to store a value
\(\ArrayName\)(I) or more generally \(\ArrayName\)(I1,...,In) to retrieve a value

Examples: \Data(6)=\{Paul Adams\}
\Values(19)

\readarray: Store consecutive values in an array, starting from the indice one.

Syntax: \readarray\{(\ArrayName)\}=(\{\Content1\}&...&\{\ContentM\})

Example: \readarray\{Actors\}\{Louise Brooks&Marlene Dietrich&Clark Gable\}

Remarks:
1. the values must be separated by the & character,
2. take care that the trailing spaces are significant, so the previous definition is different from the following one:
\readarray\{Actors\}\{Louise Brooks & Marlene Dietrich & Clark Gable\}
3. you can use \LaTeX macros inside the values.

If you really need to trim the unnecessary left and right spaces, you must apply a special action, like this one:
\checkStrings{\iString}
\esc{\texttt{Strings(\iString)='TrimSpaces{\cachedata}'}}
\makeatletter
% A \TrimSpaces macro adapted from Michael J. Downes <epsmjd@ams.org>
%(posted on c.t.t. June 19, 1998)
% \number'\lz reads past one following space (expanding as it goes)
\long\def\TrimSpaces#1{\expandafter\TrimSpaces@i\number'\^^00#1| |}
% Remove the "0" produced by \number'\~0, and " /" at the end.
\long\def\TrimSpaces@i 0#1 |{\TrimSpaces@ii\empty#1|}
% " /" was removed by \TrimSpaces@i, now remove a trailing "//" or "/ /
\long\def\TrimSpaces@ii #1|#2|{#1}
\makeatother
\multido{\iString=1+1}{4}{% 
\checkStrings{\iString}
\esc{\texttt{Strings(\iString)='TrimSpaces{\cachedata}'}}
}
\Strings(1)="a"
\Strings(2)=" b"
\Strings(3)="c c"
\Strings(4)=" d dd ddd"
\Strings(1)="a"
\Strings(2)=" b"
\Strings(3)="c c"
\Strings(4)=" d dd ddd"
\check : Get the content of the array element specified by the indice(s) and store the result in the macro \cachedata
Syntax: \check{\textit{ArrayName}}(I) or more generally \check{\textit{ArrayName}}(I_1,\ldots,I_n)
Example: \checkActors(2)
\cachedata : Macro where the content is stored after a \check request.
\ifemptydata : True if the last \check request has given an empty result.
\Values(2) = ‘B’
\Actors(3) = ‘Clark Gable’
\Actors(5) not defined.
\Actors(3) = ‘Clark Gable’
\Actors(5) not defined.

% Plain TeX usage
\checkValues(2)\%
\verb+\Values(2)+ = ‘\cachedata’
\checkActors(3)\%
\verb+\Actors(3)+ = ‘\cachedata’
\checkActors(5)\%
\verb+\Actors(5)+ not defined.
\fi
% LaTeX usage
\newcommand{\IsEmptyElement}[2]{\ifthenelse{\boolean{emptydata}}{#1}{#2}}
\checkActors(3)\%
\verb+\Actors(3)+ = \IsEmptyElement{not defined}{‘\cachedata’}
\checkActors(5)\%
\verb+\Actors(5)+ \IsEmptyElement{not defined}{‘\cachedata’}

\ifnormalindex
: See below (Default: \normalindexfalse).
\dataheight
: Counter containing the number of elements in the first dimension, if arrays are bi-dimensional.
Syntax: \dataheight=⟨Number⟩
Remarks:
1. arrays are monodimensional when \dataheight \leq 1,
2. if \normalindexfalse (which is the default value), we have:
\langle ArrayName\rangle(I_1, ..., I_n) = \langle ArrayName\rangle(I_n + (I_{n-1} - 1) \ast \dataheight + \cdots + (I_1 - 1) \ast \dataheight^{n-1})
and if \normalindextrue, we have:
\langle ArrayName\rangle(I_1, ..., I_n) = \langle ArrayName\rangle(I_1 + (I_2 - 1) \ast \dataheight + \cdots + (I_n - 1) \ast \dataheight^{n-1})

\begin{array}{cccccc}
1 & A & B & C & D & E \\
2 & F & G & H & I & J
\end{array}
\Letters(1,2)=‘B’
\Letters(2,1)=‘F’

\begin{array}{cccc}
1 & A & F \\
2 & B & G \\
3 & C & H \\
4 & D & I \\
5 & E & J
\end{array}
\Letters(1,2)=‘F’
\Letters(2,1)=‘B’

% Default is \normalindexfalse
\newarray\Letters
\readarray\Letters{A&B&C&D&E&F&G&H&I&J}
\dataheight=5
\verb+\Letters(1,2)+=‘\Letters(1,2)’
\verb+\Letters(2,1)+=‘\Letters(2,1)’
\normalindextrue
\verb+\Letters(1,2)+=‘\Letters(1,2)’
\verb+\Letters(2,1)+=‘\Letters(2,1)’
\ifexpandarrayelement: Boolean macro to allow or not the element to be evaluated before to be stored in the array (Default: \expandarrayelementfalse).

Syntax: \expandarrayelementtrue or \expandarrayelementfalse
Remark: take care to the possible side effects if you store some macros as values of some array elements without evaluating them, as they can change of content later... (see the following examples).

Syntax:
\newarray\Data \newcount\CounterP \set\CounterP\equal3 \newcounter{CounterL} \set\CounterL\equal3 \def\Town{Madrid} \Data(1)='5' \Data(2)='5' \Data(3)='Roma' \Data(4)='3' \Data(5)='3' \Data(6)='Madrid'
\CounterP=5 \set\CounterL\equal5 \def\Town{Roma}
\multido{iData=1+1}{6}{\Data(iData)='Data(iData)'}

Some other macros exist for monodimensional arrays (and only for them), but with few additive interest:

\array: Store or get the content of the array element specified by the indice. In the last case, the content is inserted at the current point.
Syntax: \array{(ArrayName)}(I)={Content} to store a value
\array{(ArrayName)}(I) to retrieve a value
Examples: \array{Actors}(6)={Joan Crawford} \array{Actors}(3)

\clrarray: Clear the content of the array element specified. A following inquiry on this element will give an empty content.
Syntax: \clrarray{(ArrayName)}(I)
Example: \clrarray{Actors}(2)

\testarray: Get the content of the array element specified by the indice and store the result in the macro \temp@macro⁴.

⁴In L\TeX, this macro should be used inside a package or between the \makeatletter \makeatother macros pair.
3 Examples

We will first show some basic and easy examples, before to look at more advanced ones to solve some complex problems, which require more knowledge of \TeX{} programming technics.

3.1 Basic examples

The immediate thing for which we can use arrays is to store and retrieve informations:

\begin{description}
\item[Actors(1)] Louise Brooks \abor\ 1906–1985
\item[Actors(2)] Marlene Dietrich \abor\ 1902–1992
\item[Actors(3)] Clark Gable \abor\ 1901–1960
\end{description}

But we can also use a general loop macro, as the one provided by the ‘\texttt{multido}’ package, for more powerful usage and a management independent of the number of elements:

\begin{description}
\multido{\iActor=1+1}{\NumberActors}{\item[\Actors(\iActor)] \Dates(\iActor)}
\end{description}

This allow various usage in the formatting of texts. A common usage is for a \textit{mailing} process, when we must compose some similar letters to various people (as we said previously, this was the reason for which these macros were developed in the ‘\texttt{formlett}’ package). Here, the usage of a programming language allow a great flexibility, using some conditionals to format the text differently or even to insert different pieces of text according to the person:

\begin{verbatim}
\newcounter{iActor}
\newcommand{\AccordingSexe}[2]{% \checkSexes{\the\value{iActor}}% \ifthenelse{\equal{\cachedata}{1}}{#1}{#2}}
\end{verbatim}
\whiledo{\value{iActor} < \NumberActors}{% \\
\stepcounter{iActor} \\
\fbox{% \\
\begin{minipage}{0.985\textwidth} \\
   Dear \AccordingSexe{Mr}{Mrs} \Actors(\the\value{iActor}), \\
   I would like to tell you how I admire the great \AccordingSexe{actor}{actress} \\
   you are, etc. \\
\end{minipage}}\[5mm\}}

Dear Mrs Louise Brooks, 
I would like to tell you how I admire the great actress you are, etc.

Dear Mrs Marlene Dietrich, 
I would like to tell you how I admire the great actress you are, etc.

Dear Mr Clark Gable, 
I would like to tell you how I admire the great actor you are, etc.

Nevertheless, people who know a little \TeX as a programming language know that it behaviour is full of pitfalls...
For instance, the following example, which format a table according to the content of entries stored in external arrays, can't work:

\begin{tabular}{|l|c|} \\
\hline \\
multido{iActor=1+1}{\NumberActors}{\Actors(\iActor) & \Dates(\iActor) \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ \ } \\
\hline \\
\end{tabular}

This is because there is an implicit grouping of each entry in a tabular environment and that it do not work with the grouping of the \multido loop. So, this must be program in a different way; storing all the content of the table before really inserting it:
### 3.1.1 Plot labels

A basic usage in graphic programming is to use arrays to retrieve the labels to put on a drawing, as here to label this simple plot:

```latex
\begin{tabular}{|l|c|}
\hline
\textbf{Actors} & \textbf{Dates} \\
\hline
Louise Brooks & 1906–1985 \\
Marlene Dietrich & 1902–1992 \\
Clark Gable & 1901–1960 \\
\hline
\end{tabular}
```

```latex
\begin{pspicture}(-0.5,-0.5)(13,4)
\psaxes[showorigin=false]{->}(13,4)
dataplot[plotstyle=dots,dotstyle=o,dotsize=0.3]{\Data}
\end{pspicture}
```

```latex
\renewcommand{\pshlabel}[1]{\tiny Months(#1)}
\PlotData
```

```latex
\renewcommand{\psvlabel}[1]{\small Levels(#1)}
\PlotData
```
3.1.2 Checkboard drawing

Of course, a bi-dimensional array directly allow to store the state of a checkboard of one of the numberous ones existing, like the reversi game that we will show here (in fact, the only more difficult part to understand in this example concern the definition of the elementary piece, but as this is not related at all to the usage of ‘arrayjob’, this point can be forgotten by most of the readers).
\def\Black{B}
\def\White{W}

\def\Piece#1#2{% 
\psset{unit=0.8,}% 
\pspicture(0,-0.1)(1,0.8)% 
\pscustom[fillstyle=gradient,gradmidpoint=0,gradangle=90, 
gradbegin=#2,gradend=#1]{% 
\psbezier(0,0.2)(0.1,-0.2)(0.9,-0.2)(1,0.2) 
\psline(1,0.2)(1,0.5) 
\psbezier(1,0.5)(0.9,0.9)(0,1,0.9)(0,0.5) 
\psline(0,0.5)(0,0.2)}% 
\psbezier(0,0.5)(0.1,0.1)(0.9,0.1)(1,0.5) 
\endpspicture}%

\def\PieceBlack\Piece{black}{gray}%
\def\PieceWhite\Piece{white}{gray}%

\def\CheckboardHook{}%

\newarray\Checkboard

\def\ShowCheckboard{% 
\dataheight=8 
\pspicture(8,8)% 
\psgrid[subgriddiv=0,gridlabels=0](8,8) 
\CheckboardHook 
\multido{\iColumn=1+1,\iColumnPos=0+1}{8}{% 
\multido{\iRow=1+1,\iRowPos=7+-1}{8}{% 
\checkCheckboard(\iRow,\iColumn)% 
\ifx\cachedata\Black 
\rput(\iColumnPos.5,\iRowPos.5){\PieceBlack} 
\else 
\ifx\cachedata\White 
\rput(\iColumnPos.5,\iRowPos.5){\PieceWhite} 
\fi 
\fi}% 
}% 
\endpspicture}%

\readarray{\Checkboard}{% 
W & &B & & & & &B & 
& W & B & & & & & B & 
W & W & B & & W & & & B & 
W & W & B & & W & & & B & 
B & & W & & W & & & B & 
B & B & W & & W & & & B & 
B & B & B & & W & & & B & 
W & B & B & & B & & & B &}%

\psset{unit=0.6} 
\ShowCheckboard

\Checkboard(2,1)={B} % Black move
\def\CheckboardHook{% 
\psframe[linestyle=none,fillstyle=hlines](0,6)(1,7) 
% White pieces changed by the black move 
\Checkboard(3,1)={B}\Checkboard(4,1)={B} 
\Checkboard(2,2)={B}\Checkboard(3,2)={B} 
\vspace{1cm} 
\ShowCheckboard

\Checkboard(2,1)={B} % Black move
\def\CheckboardHook{% 
\psframe[linestyle=none,fillstyle=hlines](0,6)(1,7) 
% White pieces changed by the black move 
\Checkboard(3,1)={B}\Checkboard(4,1)={B} 
\Checkboard(2,2)={B}\Checkboard(3,2)={B} 
\vspace{1cm} 
\ShowCheckboard
Note also that, if \TeX is obviously not well suited to implement a program to really play at a game like reversi or chess, nevertheless it can be used to solve internally some non trivial algorithmic problems, where the usage of arrays help a lot. For such example (on the coloration of the Truchet’s tiling), see [3].

3.2 Advanced examples

This section will show slightly more difficult to really complex examples, and is mainly for advanced users or for people who want to be able to program complex tasks themselves.

3.2.1 Example with recursion usage

\begin{verbatim}
\makeatletter
% The recursion macro used (from David Carlisle)
\def\Recursion{\ifnum#1>1\relax}
\expandafter\@firstoftwo
\else
\expandafter\@secondoftwo
\fi}
\newcount\IndexRecursion
\IndexRecursion=\z@
\def\PiFrac#1{{% 
\Recursion{#1\relax}
\else
\expandafter\@secondoftwo
\fi}
\newcount\IndexRecursion
\IndexRecursion=\z@
\makeatletter
\newarray\PiValues
\readarray\PiValues{3&7&15&1&292&1&1&1&2&1&3&1&14}
\pi\approx\PiFrac{4}\approx\PiFrac{13}\}
\textbf{\pi as a continued fraction}
\end{verbatim}
\[ \pi \approx 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1}}} \approx 3 + \cfrac{1}{7 + \cfrac{1}{15 + \cfrac{1}{1 + \cfrac{1}{292 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{1 + \cfrac{1}{3 + \cfrac{1}{1 + \cfrac{1}{14}}}}}}}}}}} \]

\text{π as a continued fraction}

### 3.2.2 Structured dynamic diagrams on a grid

A very common case is when we must use some structured data which are to be drawn on a kind of abstract grid.

To draw a diagram like this one (such example illustrate the exchanges of data between processors when using applications on multiprocessors computers in parallel programming):
is easy but rather painful, as we must manage a lot of coordinates. Nevertheless, even if we introduce a little more abstraction level using some minor programming with a loop structure:

we succeed to heavily reduce the number of lines of the code in this specific case, but we gain nothing in genericity. If we have several such diagrams to draw, with each that must be done differently according to the empty and full cells of the array, we must proceed in a completely different way, defining a generic object which can obtain the different values of a data structure previously filled. For such tasks, the ‘arrayjob’ package allow to build very efficient and
powerful tools.
\% Default attributes
\def\ColorBackgroundEmpty{LemonChiffon}
\def\ColorBackgroundFull{Khaki}
\def\AttributeElement#1{\textcolor{red}{#1}}
\def\AttributeLabel#1{\bf #1}

\% Example 1
\newarray\ArrayA
\readarray{\ArrayA}{\%}
A0 & & & &
A0 & A1 & & &
A0 & A1 & A2 & &
A0 & A1 & A2 & A3 &
\ArrayGrid{\ArrayA}{4}{4}

\% Example 2
\newarray\ArrayB
\readarray{\ArrayA}{\%}
A & & & &
& & & &
& & & &
& & &
\readarray{\ArrayB}{\%}
A & & & &
A & & & &
A & & & &
A & & &
\ArrayGrid{\ArrayA}{4}{4}\hfill\ArrayGrid{\ArrayB}{4}{4}

\% Example 3
\newarray\ArrayC
% Data
\readarray{ArrayA}{A0 & A1 & A2 \\
A0 & A1 & A2 \\
A0 & A1 & A2}
\readarray{ArrayB}{A0&A1&A2 & \\
A0&A1&A2 & \\
A0&A1&A2}
\readarray{ArrayC}{A0 & & & \\
& A1 & & \\
& & A2}
% Attributes
\def\ColorBackgroundEmpty{green}
\def\ColorBackgroundFull{red}
\def\AttributeElement#1{\footnotesize\textcolor{white}{#1}}
\def\AttributeLabel#1{\small\bf #1}
\psset{unit=0.7}
\ArrayGrid{ArrayA}{3}{3} \hfill \ArrayGrid{ArrayB}{3}{3} \hfill \ArrayGrid{ArrayC}{3}{3}

% Example 4
% Data
\readarray{ArrayA}{A0\ A1\ A2\ A3 & \\
B0\ B1\ B2\ B3 & \\
C0\ C1\ C2\ C3 & \\
D0\ D1\ D2\ D3}
\readarray{ArrayB}{A0&B0&C0&D0 & \\
A1&B1&C1&D1 & \\
A2&B2&C2&D2 & \\
A3&B3&C3&D3}
% Attributes
\def\AttributeElement#1{\textcolor{red}{#1}}
\def\AttributeLabel#1{\Large\it #1}
\ArrayGrid{ArrayA}{4}{4} \hfill \ArrayGrid{ArrayB}{4}{4}
3.2.3 Another example of a structured dynamic diagram on a grid

The preceding technique is in fact relevant for a lot of problems. Here we give another example about the drawing of linked lists (but do not miss to note that these ones have here no internal existence).

\begin{verbatim}
\usepackage{array}
\usepackage{tikz}
\usetikzlibrary{arrows,calc}
\newlength{\dataheight}
\defaultarrow{linecolor=red,arrowscale=2}
\newcommand{\LinkedListsDraw}[1]{\ifnextchar[\LinkedListsDraw@i{\LinkedListsDraw@i[]}}
\newcommand{\LinkedListsDraw@i}[1]#2#3[{{%\setkeys{psset}{#1} \dataheight=#3 \% \pst@cnta=#3 \multiply\pst@cnta\tw@ \pspicture(0,-\dataheight)(\pst@cnta,0) \multido{iRow=1+1}{#2}{\rput(0,-iRow){% \psframe(1,1) \psline[\dimen=middle,style=\defaultarrow]{->}(0.5,0.5)(2,0.5)} \multido{iColumn=1+1}{#3}{% \ifemptydata \else \pst@cnta=iColumn \advance\pst@cnta@one \pst@cntb=\one \checkLinkedListsTable(iRow,\pst@cnta)% \ifemptydata \else \endframe \endpicture
\end{verbatim}
\ifnum\multidocount=#3 % We test if it is the last one
\else
\pst@cntb=2\@n
\fi
% \pst@cnta=iColumn
\multiply\pst@cnta\tw@
\rput(\the\pst@cnta,-\iRow){% \LinkedListsDraw@ii{\LinkedListsTable(\iRow,\iColumn)}{\the\pst@cntb}}
\fi}}%}
\endpspicture}}%
\def\LinkedListsDraw@ii#1#2{{% put(0,0.1){% \psset{unit=0.8} \psgrid[subgriddiv=0,gridlabels=0](2,1) % Element \rput(0.5,0.5){#1} % Label for this element \ifnum#2=\@ne \psline(1,0)(2,1) % Mark of the end of the list \else \psline[style=LinkedListsArrowStyle]{->}(1.5,0.5)(2.5,0.5) % Link to the next element \fi}}}
\makeatother
\newarray\LinkedListsTable
\readarray{\LinkedListsTable}{11&12&13&14&% 21&&&&% 31&32&33&&}
\LinkedListsDraw{3}{4}
\vspace{1cm}
\newpsstyle{LinkedListsArrowStyle}{linecolor=green,doubleline=true}
\readarray{\LinkedListsTable}{a&&&&&% A&B&C&D&E&% 0&1&2&3&4&% +&\times\\times\times\times\times}
\ LinkedListsDraw{4}{5}
3.2.4 Management of heaps and linked lists

Here we will demonstrate a really more complex usage, but very useful. In fact, if the ‘arrayjob’ package primarily allow to manage only arrays, it allow a lot more with some efforts. Above it, we can implement the common internal data structures frequently used in programming, like heaps, simple or double linked lists, trees, associative arrays, etc. Here we will show how to define macros to build and manage heaps and simple linked lists.

\makeatletter
\def\PstDebug{\z@} % For debugging, set \def\PstDebug{1}
% Code for heaps management
% -------------------------
\def\HeapMaxDepth{100} % No more than 100 elements in the heap
\newarray\Heap
% Add an element at top of the heap
\def\HeapPush#1{% 
\multido{\iElem=1+1}{\HeapMaxDepth}{% \checkHeap(\iElem)% \ensuremath{\text{\ifemptydata}}% \ifnum\PstDebug=\@ne\typeout{Push Heap(\iElem)=#1}\fi% Debug
% Get (in the \cachedata macro) and delete the element at top of the heap
\def\HeapPop{%
\Multido{\iElem=\HeapMaxDepth+-1}{\HeapMaxDepth}{%\checkHeap{\iElem}%\ifemptydata\else
  \ifnum\PstDebug=1\typeout{Delete Heap(\iElem)=\cachedata}\fi\% Debug
  \Heap{\iElem}={}%
\multidostop%\fi}%
% Print the current state of the heap
\def\HeapPrint{%
\checkHeap{\@ne}%\ifemptydata\typeout{The heap is empty!}\% Empty heap\else
\multido{\iElem=\HeapMaxDepth+-1}{\HeapMaxDepth}{%\checkHeap{\iElem}%\ifemptydata\else
  \typeout{Heap(\iElem)=\cachedata}%
\fi}%\fi\typeout{}}%
% Draw the current state of the heap
\def\HeapDraw{%
% To compute the size of the picture
\Multido{\iSize=0+1}{\HeapMaxDepth}{%\checkHeap{\the\multidocount}%\ifemptydata\multidostop%\fi}%
\ifnum\iSize=0\typeout{The heap is empty!}\% Empty heap\else
\pspicture(1,\iSize)
\psgrid[subgriddiv=0,gridlabels=0](1,\iSize)
\multido{\iPos=0+1}{\HeapMaxDepth}{%\checkHeap{\the\multidocount}%
\HeapPush{Germany}\HeapPrint\HeapDraw
\hfill
\HeapPush{France}\HeapPrint\HeapDraw
\hfill
\HeapPush{Italy}\HeapPrint\HeapDraw
\hfill
\HeapPush{Spain}\HeapPrint\HeapDraw
\hfill
\HeapPop\typeout{Popped element: 'cachedata'}\HeapPrint\HeapDraw
\hfill
\HeapPop\typeout{Popped element: 'cachedata'}\HeapPrint\HeapDraw
\hfill
\HeapPop\typeout{Popped element: 'cachedata'}\HeapPrint\HeapDraw
\hfill
\HeapPop\typeout{Popped element: 'cachedata'}\HeapPrint\HeapDraw

\def\PstDebug{1}
\For debugging, set \def\PstDebug{1}
% Code for linked lists management
% -----------------------------------
\def \LinkedListMaxDepth {100} \text{ No more than 100 elements in the list}
\newarray \LinkedList
\dataheight=2 \text{ Two cells by element: the content and the pointer}

% Add an element at the end of the list
% In fact, we define it here as a simple heap, but the difference is that
% we will be able to get and to delete any element in it
\def \LinkedListAdd #1 {% 
  \edef \@tempa {\@ne} \text{ Pointer initialization}
  \multido {\iElem = 2+1} {\LinkedListMaxDepth} {% 
    \checkLinkedList (\@tempa,2) \text{ We got the content of this element}
    \ifemptydata % No pointed element, so we will insert the new one here
      \LinkedList (\iElem,1) = {#1} \text{ We got the content of this element}
    \else % We update the pointer for the last preceding element
      \expandarrayelementtrue \text{ We must evaluate the content of the counter}
      \LinkedList (\@tempa,2) = {\iElem} \text{ We must evaluate the content of the counter}
      \expandarrayelementfalse \text{ We must evaluate the content of the counter}
      \edef \@tempa {\cachedata} \text{ We update the pointer for the last preceding element}
    \fi
  }%multidostop
  \ifnum \PstDebug = \@ne
    \typeout {Add LinkedList (\@tempa) -> \iElem} \text{ Debug}
    \typeout { LinkedList (\iElem) = #1} \text{ Debug}
  \fi
}%LinkedListAdd

% Get (in the \cachedata macro) the next element in the list
\def \LinkedListGetNext #1 {% 
  \checkLinkedList (1,2) \text{ We got the pointer for the first element}
  \edef \@tempa {\@ne} \text{ Pointer initialization}
  \ifx \cachedata \empty % No pointed element, so we will insert the new one here
    \LinkedList (\iElem,1) = {#1} \text{ We got the content of this element}
  \else % We update the pointer for the last preceding element
    \edef \cachedata {#1} \text{ We update the pointer for the last preceding element}
    \Multido {\iElem = 2+1} {\LinkedListMaxDepth} {% 
      \edef \@tempb {\@tempa} \text{ We update the pointer for the last preceding element}
      \edef \@tempa {\cachedata} \text{ We update the pointer for the last preceding element}
      \checkLinkedList (\@tempa,2) \text{ We got the pointed element}
    }%multidostop
    \edef \@tempc {#1} \text{ We update the pointer for the last preceding element}
    \Multido {\iElem = 2+1} {\LinkedListMaxDepth} {% 
      \edef \@tempb {\@tempa} \text{ We update the pointer for the last preceding element}
      \edef \@tempa {\cachedata} \text{ We update the pointer for the last preceding element}
      \checkLinkedList (\@tempa,2) \text{ We got the pointed element}
    }%multidostop
    \edef \@tempc {#1} \text{ We update the pointer for the last preceding element}
  \fi
}%LinkedListGetNext
\ifemptydata
  \typeout{Element ‘#1’ has no successor!\string-j}\%
\else
  \checkLinkedList(\cachedata,1)\%
  \fi
\multidostop
\else
  \checkLinkedList(\@tempa,2)\% We got the pointer to the next element
  \ifemptydata
     \typeout{Element ‘#1’ not found!\string-j}\%
     \multidostop\% No pointed element: end of the list
  \fi
  \fi}
\fi}
\%
\def\LinkedListDelete{\@ne}{% 
\checkLinkedList(1,2)\% We got the pointer for the first element
  \edef\@tempa{\@ne}\% We got the pointer to the next element
  \if\cachedata\empty\else\edef\@tempc{#1}\%
    \multido{}{\LinkedListMaxDepth}{% 
      \edef\@tempb{\@tempa} \edef\@tempa{\cachedata}\%
      \checkLinkedList(\cachedata,1)\% We got the pointed element
      \if\cachedata\@tempc\%
        \checkLinkedList(\@tempa,2)\% This is the element to delete: we will update the pointer of
        \expandarrayelementtrue \LinkedList(\@tempb,2)=\cachedata\%
        \expandarrayelementfalse \ifnum\PstDebug=\@ne
          \typeout{Delete LinkedList(\@tempb)=#1} \%
          \typeout{space space space space space space LinkedList(\@tempb)->\cachedata} \%
        \fi
      \fi \fi}
\multidostop
\else
  \checkLinkedList(\@tempa,2)\%
  \multidostop
  \else
    \checkLinkedList(\@tempa,2)\% We got the pointer to the next element
    \ifemptydata
      \typeout{Element ‘#1’ not found and so not deleted!\string-j}\%
      \multidostop\% No pointed element: end of the list
    \fi
  \fi
\fi}
\def\LinkedListPrint{%
\checkLinkedList(1,2)\% We got the pointer for the first element
\ifx\cachedata\empty\%
Empty list
\typeout{The list is empty!}\%
\else\% We got the pointed element
\multido{}{\LinkedListMaxDepth}{% We got the pointed element
\checkLinkedList(\cachedata,1)\%
\typeout{LinkedList=\cachedata}\%
\checkLinkedList(\@tempa,2)\% We got the pointer to the next element
\ifemptydata\multidostop\%
No pointed element: end of the list
\fi\}
\fi\%
\typeout{}}
%
\def\LinkedListDraw{%
\psset{subgriddiv=0,gridlabels=0}
\checkLinkedList(1,2)\% We got the pointer for the first element
\ifx\cachedata\empty\%
Empty list
\typeout{The list is empty!}\%
\else\%
To compute the size of the picture
\Multido{\iSize=3+3}{\LinkedListMaxDepth}{% We got the pointer to the next element
\checkLinkedList(\cachedata,2)\% We got the pointed element
\ifemptydata\multidostop\%
\fi\%
\checkLinkedList(1,2)\%
}\pspicture(\iSize,1.5)\%
\multido{\iPos=0+3}{\LinkedListMaxDepth}{% We got the pointed element
\edef\@tempa{\cachedata}\%
\checkLinkedList(\cachedata,1)\%
\rput(\iPos,0){% We got the pointed element
\psgrid(2,1)\%
\rput(0.5,0.5){\cachedata}\%
\ifnum\iPos=\z@\else\rput(-1.5,0.5){\psline[linecolor=red,arrowscale=2]{->[}1.5,0)\% We got the pointer to the next element
\fi\%
\checkLinkedList(\@tempa,2)\%
}\pspicture(\iSize,1.5)\%
\multido{\iPos=0+3}{\LinkedListMaxDepth}{% We got the pointed element
\edef\@tempa{\cachedata}\%
\checkLinkedList(\cachedata,1)\%
\rput(\iPos,0){% We got the pointed element
\psgrid(2,1)\%
\rput(0.5,0.5){\cachedata}\%
\ifnum\iPos=\z@\else\rput(-1.5,0.5){\psline[linecolor=red,arrowscale=2]{->[}1.5,0)\% We got the pointer to the next element
\fi\%
}
\ifemptydata
  \rput{(\iPos,0)}{\psline[linecolor=red](1,0)(2,1)} % End of the list
\endpspicture
\fi}
\fi}}
\makeatother
\psset{unit=1.5}
\LinkedListPrint\LinkedListDraw
\LinkedListAdd{Germany}\LinkedListPrint\LinkedListDraw
\LinkedListAdd{France}\LinkedListPrint\LinkedListDraw
\LinkedListAdd{Italy}\LinkedListPrint\LinkedListDraw
\LinkedListAdd{Spain}\LinkedListPrint\LinkedListDraw
\LinkedListGetNext{France}\typeout{Next element after 'France' is: '{\cachedata}'}%
\LinkedListGetNext{Germany}\typeout{Next element after 'Germany is: '{\cachedata}'}%
\LinkedListGetNext{Spain}
\LinkedListGetNext{Unknown}
\LinkedListDelete{Unknown}
\LinkedListDelete{France}\LinkedListPrint\LinkedListDraw
\LinkedListDelete{Germany}\LinkedListPrint\LinkedListDraw
\LinkedListDelete{Spain}\LinkedListPrint\LinkedListDraw
\LinkedListDelete{Italy}\LinkedListPrint\LinkedListDraw
We can write an extended version of the inclusion of an element, where we do not still store each element at the end of the list, but insert it at a special position. For instance, we can decide to use the list to sort the elements, inserting a new one at its sorted position. In the next example, we will manage integer numbers in this way (of course, we can do the same thing with arbitrary strings, but it would be more difficult to program in \TeX, mainly if we want to be able to use accented letters).

\begin{verbatim}
\makeatletter
\def\LinkedListSortedAdd#1{% 
  \edef\@tempa{\@ne} 
  \edef\@tempb{\@ne} 
  \edef\@tempc{\z@} 
  \multido{\iElem=2+1}{\LinkedListMaxDepth}{% 
    \checkLinkedList(\@tempa,2)% We got the pointer of this element 
    \checkLinkedList(\@tempb,\@tempa)% 
    \edef\@tempb{\@tempa} 
    %
    %\edef\@tempc{\z@}% \edef\@tempc{\@tempc+1}%
  } 
  \checkLinkedList(\@tempc,\@tempb)%
}\end{verbatim}
\ifemptydata
  \LinkedList{\iElem,1}={#1}\%
  \% No pointed element, so we will insert the new one here
  \LinkedList{\@tempb,2}={\iElem}\%
  \expandarrayelementtrue
  \LinkedList{\@tempb,2}={\iElem}\%
  \ifnum\PstDebug=\@ne
    \typeout{Add LinkedList(\@tempb)->\iElem} \ Debug
    \typeout{\space\space\space\space LinkedList(\iElem)=#1} \ Debug
  \fi
  \% We update the pointer for the last element
  \ifnum\@tempc=\z@\else
    \LinkedList{\iElem,2}={\@tempc}\%
    \ifnum\PstDebug=\@ne
      \typeout{Update LinkedList(\iElem)->\@tempc} \ Debug
    \fi
  \fi
  \% We update the pointer of the new element to the position
  \% of the next greater one, if it exist
  \else\edef\@tempa{\cachedata}\%
  \checkLinkedList{\cachedata,1}\%
  \ifnum\@tempc<\cachedata\edef\@tempc{\@tempa}\%
    \typeout{New element less than this one} \ Debug
  \else\edef\@tempb{\@tempa}\%
    \typeout{New element greater or equal than this one} \ Debug
  \fi\fi\fi
\makeatother
\LinkedListPrint
\LinkedListSortedAdd{732}\LinkedListPrint\LinkedListDraw
\LinkedListSortedAdd{487}\LinkedListPrint\LinkedListDraw
\LinkedListSortedAdd{718}\LinkedListPrint\LinkedListDraw
\LinkedListSortedAdd{962}\LinkedListPrint\LinkedListDraw
\LinkedListDelete{487}\LinkedListPrint\LinkedListDraw
\LinkedListDelete{732}\LinkedListPrint\LinkedListDraw
3.2.5 Associative arrays

To finish, we will give a complete solution to a classical problem: to count the number of occurrences of the various letters in a sentence. For this, we will first build some macros to deal with associative arrays, as they have been popularized by scripting languages like AWK and Perl.

\makeatletter
\% Internally, we use two "standard" arrays to define one associative array
\newarray\AssociativeArray@Names
\newarray\AssociativeArray@Values
\newcount\AssociativeArrayNbValues
\AssociativeArrayNbValues=z\$
\newif\ifAssociativeArray@ElementFound
% To store one element
\def\AssociativeArray(#1)=#2{%
\expandarrayelementtrue
\AssociativeArray@ElementFoundfalse
\edef\@tempa{#1}%
\Multido\{iValue=\One+%\One\}{\AssociativeArrayNbValues}{%
\checkAssociativeArray@Names(iValue)\%
\ifx\@tempa\cachedata
% This element already exist: we replace it current value
\checkAssociativeArray@Values(iValue)% Debug
\propertypageIn #1, replace ‘cachedata’\space by ‘#2’% Debug
\AssociativeArray@Values(iValue)=#2%
\AssociativeArray@ElementFoundtrue
\multidostop
\fi}
\ifAssociativeArray@ElementFound
\else
% New element
\advance\AssociativeArrayNbValues+1
\AssociativeArray@Names(\AssociativeArrayNbValues)={#1}%
\AssociativeArray@Values(\AssociativeArrayNbValues)={#2}%
\fi}
%
% To get one element
\def\checkAssociativeArray(#1){%
\edef\@tempa{#1}%
\edef\@tempb{999999}%
\Multido\{iValue=\One+%\One\}{\AssociativeArrayNbValues}{% 
\checkAssociativeArray@Names(iValue)\%
\ifx\@tempa\cachedata
% We have found it by name
\edef\@tempb{iValue}%
\multidostop
\fi}
% We have now to get its value
\checkAssociativeArray@Values(\@tempb)\%
%
% Simple macro to print all the associative array
\def\printAssociativeArray{%
\Multido\{iValue=\One+%\One\}{\AssociativeArrayNbValues}{% 
\checkAssociativeArray@Names(iValue)\%
\if\assocarray@i:BS\texttt{.FirstNames(\cachedata)}='\AssociativeArray@Values(iValue)'\space\%
\fi}
30
\makeatother
\let\FirstNames\AssociativeArray
\let\checkFirstNames\checkAssociativeArray
\let\printFirstNames\printAssociativeArray
\FirstNames(Crawford)={Joan}
\FirstNames(Tierney)={Gene}
\FirstNames(Lake)={Veronika}
\printFirstNames
\checkFirstNames(Lake)
\verb+\FirstNames(Lake)+='\cachedata'
\checkFirstNames(Monroe)
\ifemptydata\verb+\FirstNames(Monroe) undefined!+'\else\verb+\FirstNames(Monroe)+='\cachedata'\fi

1: \FirstNames(Crawford)=Joan 2: \FirstNames(Tierney)=Gene 3: \FirstNames(Lake)=Veronika
\FirstNames(Lake)=Veronika \FirstNames(Monroe) undefined!

Now, we can define the macros to read a sentence, to cut it letter by letter, to count the occurrences of each of them, and finally to draw a summary plot of the results.

\makeatletter

\% A \PerChar style macro adapted from Juergen Schlegelmilch
\% (<schlegel@Informatik.Uni-Rostock.de> - posted on c.t.t. January 27, 1998)
\def\PerChar#1#2#3\@nil{% #1#2%
  \edef\@tempa{#3}%
  \ifx\@tempa\empty
    \else
    \DoPerChar#1#3\@nil
  \fi
}
\def\DoPerChar#1#2#3\@nil{% #1#2% 
  \edef\@tempa{#3}%
  \ifx\@tempa#3%
    \else\@tempa\empty
  \else
    \DoPerChar#1#3\@nil
  \fi
}
\% The action to do for each character: we increase the counter
\% for this character by one
\def\ActionPerChar{\% \#1=character
\checkAssociativeArray(#1)\%
\ifemptydata
\@tempcnta=\@ne
\else
\@tempcnta=\cachedata
\fi
\AssociativeArray(#1)={\the\@tempcnta}}

\% To store in an associative array the number of occurrences by characters
\% (spaces are not counted)
def\StatSentence{\% \#1=sentence
\AssociativeArrayNbValues=\z@ \PerChar{\ActionPerChar}{#1}}

\% To draw a plot of the number of occurrences of the characters of a sentence
\def\DrawOccurrences{\% \#1=sentence
\StatSentence{#1}
\pst@cnta=\AssociativeArrayNbValues
\advance\pst@cnta\@ne
\% To know the maximum of the numbers
\pst@cntb=\z@
\Multido{iValue=\@ne+\@ne}{\AssociativeArrayNbValues}{\checkAssociativeArray@Values(iValue)\%
\ifnum\cachedata>\pst@cntb
\pst@cntb=\cachedata
\fi}
\advance\pst@cntb\@ne
\% The drawing itself
\pspicture(-0.5,-0.5)(\pst@cnta,\pst@cntb)
\psaxes[axesstyle=frame,labels=y](\pst@cnta,\pst@cntb)
\multido{iValue=\@ne+\@ne}{\AssociativeArrayNbValues}{\checkAssociativeArray@Values(iValue)\%
\draw@OneOccurrence(iValue)}
\endpspicture}

\% To draw the value for one occurrence
\def\draw@OneOccurrence{\% \#1=Letter
\checkAssociativeArray@Values(#1)\%
pdplot(#1,\cachedata)
rput[B]{\iValue,-0.5}{\checkAssociativeArray@Names(#1)}}
We will now write another version, using another array to store indexes, to allow us to sort the letters found in the sentence read.

\makeatletter
\% To get in #2 the ASCII code of the #1 character
\def\Character@AsciiCode#1#2{\chardef#2='#1\relax}
\% A new "standard" array to store the sorted indexes
\newarray\AssociativeArray@Indexes
\% A simple macro to print the array of indexes (for debugging)
\def\PrintIndexes{\multido{iIndexes=\one+\one}{\AssociativeArrayNbValues}{\checkAssociativeArray@Indexes(iIndexes)\BS\texttt{indexes(iIndexes)}='\cachedata'\space}}
\% We redefine the macro to store an element in the associative array,
\% to allow to sort it elements, using an additional array of indexes
\def\AssociativeArray(#1)=#2{%
\expandarrayelementtrue
\edef\@tempa{#1}\edef\@tempb{#2}\
\Multido{\iValue=\@ne+\@one}{\AssociativeArrayNbValues}{% 
  \checkAssociativeArray@Names{\iValue}\ifx\@tempa\cachedata\%
  \checkAssociativeArray@Values{\iValue}\typeout{In #1, replace ‘\cachedata’ space by ‘#2’}\%
  \AssociativeArray@Values{\iValue}={\@tempb}\% \AssociativeArray@ElementFoundtrue\multidostop\fi}
\ifAssociativeArray@ElementFound\else\else\fi
% New element: we must insert it at it sorted position
\@tempcnta=\@one\expandafter\Character@AsciiCode\@tempa\@tempx
\Multido{\iValue=\@one+\@one}{\AssociativeArrayNbValues}{% 
  \checkAssociativeArray@Names{\iValue}\ifnum\@tempx<\@tempy\%
  \checkAssociativeArray@Indexes{\iValue}\@tempcntb=\cachedata\% 
  \advance\@tempcntb\@one\% \AssociativeArray@Indexes{\iValue}={\the\@tempcntb}\% \else\fi\seven\#1\fi}
% We change the way we draw each number of occurrences
\def\Draw@OneOccurrence{\% 
  \checkAssociativeArray@Indexes{\#1}\@ptcnta=\cachedata\% 
  \checkAssociativeArray@Values{\#1}\psdot{\the\@ptcnta}{\cachedata}\% 
  \rput{[B]}{\the\@ptcnta,-0.5}{\AssociativeArray@Names{\#1}}\% 
\makeatother
\DrawOccurrences{We are Wednesday}
\PrintIndexes

\indexes(1)='1' \indexes(2)='4' \indexes(3)='2' \indexes(4)='6' \indexes(5)='3' \indexes(6)='5' \indexes(7)='7' \indexes(8)='8'

\DrawOccurrences{Tuesday 8 February}

Now, we can easily redefine the internal macro which plot each occurrence, to obtain various kinds of graphics:

\makeatletter
\def\DrawOneOccurrence#1{% \checkAssociativeArray@Indexes(#1)% \pst@cnta=\cachedata \checkAssociativeArray@Values(#1) \psline[linestyle=dotted](\the\pst@cnta,0)(\the\pst@cnta,\cachedata) \psdot(\the\pst@cnta,\cachedata) \rput[B](\the\pst@cnta,-0.5){\AssociativeArray@Names(#1)}}
I do not believe that it can be correct

I don’t know a better solution than this one
4 Thanks

I would like to thank Rolf Niepraschk <niepraschk@ptb.de> to have carefully read a preliminary version of this document and to have sent me several good advices to clarify and improve some points.

References

[1] Stephan von Bechtolsheim, \TeX\ in Practice, four volumes, Springer-Verlag, New York, 1993 (see also CTAN: macros/tip/arraymac.tip).


[4] (AI)Dra\TeX, by Eitan M. Gurari, Ohio State University, Columbus, USA, CTAN:graphics/dratex (see also http://www.cis.ohio-state.edu/~gurari/systems.html).


