Visual presentation of multidimensional dataflow programs

Joey Paquet John Plaice
School of Computer Science and Engineering
The University of New South Wales
Sydney 2052 Australia
{jpaquet,plaice}@cse.unsw.edu.au
1998

Abstract
In order to make multidimensional dataflows easier to understand, two visual formalisms are introduced, through a number of examples. The multidimensional dataflow diagram considers dimensions to be values, which can be passed along flow lines. In addition, flows known to vary in a number of dimensions are appropriately labeled. The multidimensional evaluation tree presents the demand-driven process of eduction in a multiple-dimensional space.

1 Introduction
This paper presents two new visual formalisms for presenting the multidimensional aspects of Lucid programs. The multidimensional dataflow diagrams or networks allow one to present the structure of a program. The multidimensional evaluation tree allows one to understand the demand-driven evaluation process.

These formalisms are quite intuitive, so they are essentially presented through a number of example Lucid programs, beginning from simple, one-dimensional examples, towards more interesting multi-dimensional ones. We begin by giving a description of the problem solved, then give the Lucid approach to the resolution of the problem, which is generally far from the conventional imperative approach.

In fact, the rest of the paper will be an introduction to Lucid, using these formalisms as support.

Lucid has two primitive operators, called intensional operators. First, for intensional navigation (@) through a multidimensional context space; second for querying the current context of evaluation of the program (#).

These two operators can be used to create user-defined functions. The following examples use commonly used functions for the generation of streams of values. Lucid programs using these functions can be thought of as multidimensional dataflow networks, which are simply dataflow networks carrying multidimensional objects. The dimensionality of the objects carried is annotated in the graphs as tags on the edges. We give the dataflow network representation along with the Lucid program for each of the examples.

2 The natural numbers
This first example is really simple. However, it captures all the essential aspects of intensional programming. The problem is to extract a value from the stream representing the natural numbers, beginning from the ubiquitous number 42:
Let us arbitrarily pick the fourth value of the stream, which is assigned tag number three. Let also the whole stream vary in the \( d \) dimension. The program doing this is simply the following, which is represented in a dataflow graph in Figure 1:

\[
N \oplus .d \, 3 \\
\text{where} \\
\text{dimension } d; \\
N = 42 \text{ fby.d } N+1; \\
\text{end;}
\]

![Dataflow graph for the natural numbers problem](image1)

Figure 1: Dataflow graph for the natural numbers problem

![Totally exploded dataflow network for the natural numbers problem](image2)

Figure 2: Totally exploded dataflow network for the natural numbers problem.

\((42, 43, 44, 45, \ldots)\)
With not much intuition, one can readily expect the program to return the value 45. However, as computers do not rely on intuition for the evaluation of programs, we will give more details of the evaluation method.

To see how the program is evaluated, we rewrite it in terms of the basic @ and # operators, which is represented in Figure 2:

```plaintext
N @.d 3
where
dimension d;
N = if #.d <= 0
   then 42
   else (N+1) @.d (#.d-1)
fi;
end;
```

Figure 3 shows how evaluation takes place by generating successive demands for the appropriate
values of $N$, until the final computation can be effected. The tree notation should be easily understood. A demand is made at the beginning of one of the long vertical arrows, and its result is found at the head of the arrow. Changes of context are of the form $d : 1$.

3 The Hamming Problem

The Hamming problem is a simple problem that consists in generating the stream of all numbers of the form $2^i3^j5^k$ in increasing order and without repetition. The statement of the problem is simple indeed, but its programmation in an imperative language is far from being so simple. However, this problem finds a trivial solution when thought of as a dataflow problem. The graph corresponding to the dataflow solution is given in Figure 4.

![Figure 4: Dataflow network of the Hamming program.](image)

$H$
where

\[ H = 1 \text{ fby merge(merge(2*H,3*H),5*H)}; \]
merge($x,y$) = if ($xx\leq yy$) then $xx$ else $yy$
where $xx = x$ upon ($xx\leq yy$);
$yy = y$ upon ($yy\leq xx$);
end;
end

The previous examples defined and manipulated only one-dimensional intensions. The most interesting feature of Lucid is its ability to naturally define and manipulate multidimensional intensions. The following examples use multiple dimensions. Dataflow networks are normally used for the expression of one-dimensional systems. We here generalize dataflow networks by permitting the edges to carry multidimensional tokens. The edges in the dataflow graphs are tagged with the dimension names of the tokens they carry.

4 Matrix Transposition

Scientific programming is mostly about the mathematical manipulation of matrices. A common operation on matrices is the transposition. An imperative program to do matrix transposition
ifthenelse

merge(X,Y)

upon

leq

ifthenelse

upon

leq

Figure 5: Dataflow network of the \textit{merge} function.

typically copies all the matrix elements through an iterative process. In Lucid the transpose is simply done by renaming the dimensions in which varies the matrix.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{Dataflow network for the \textit{transpose} program.}
\end{figure}

\texttt{transpose.d0,d1(M)} = \texttt{realign.tmp,d0(realign.d0,d1(realign.d1,tmp(A)))}

where
dimension tmp;
\texttt{realign.a,b(X) = X @.a #.b;}
end;

The \texttt{tmp} dimension used here is local to the \texttt{where} clause and is used as a temporary dimension to transpose the matrix. Also, the \texttt{realign} function is here defined locally to the expression defining the \textit{transpose} function. Here it would not be possible to use the \texttt{realign} function in other definitions in the outermost \texttt{where} clause.

\section{5 Matrix Multiplication}

Matrix multiplication is one of the most basic and common problems in scientific computation. To multiply two $n \times n$ matrices $A$ and $B$ we have to multiply, pointwise, the rows of $A$ and the
columns of \( B \) and add together the values produced. More precisely, the \((i,j)\)-th element of the product is the sum of the \(n\) values

\[ A_{i,k} \times B_{k,j}, \quad k \in 1..n. \]

The required program is:

\[
\text{mm} \cdot x,y(M1,M2,n) = \text{sum} \cdot z(\text{product} \cdot x,y,z(M1,M2),n)
\]

where

\[
\text{dimension } z; \\
\text{product} \cdot d1,d2,d3(M,N) = \text{realign} \cdot d2,d3(M) \ast \text{realign} \cdot d1,d3(N); \\
\text{realign} \cdot a,b(X) = X @.a #.b; \\
\text{sum} \cdot d(X,n) = Y @.t \log n
\]

where

\[
\text{dimension } t; \\
Y = X fby.t (\text{firstOfPair} \cdot d(Y) + \text{secondOfPair} \cdot d(Y)); \\
\text{firstOfPair} \cdot a(Z) = Z @.a (#.a*2); \\
\text{secondOfPair} \cdot a(Z) = Z @.a (#.a*2+1); \\
\text{end}; \\
\text{end};
\]

In this program, matrix \( A \) is turned so that its variation in dimension \( y \) is instead changed into variation in dimension \( z \). Similarly, \( B \) is turned so that its variation in dimension \( x \) is changed into variation in dimension \( z \). As a result,

\[
\text{product} \cdot x,y,z(A,B)_{(x,y,z)} = A_{(x,y,z)} \times B_{(x,y,z)}
\]
\[ A_{(x,z,z)} \times B_{(z,y,z)} = A_{(x,z)} \times B_{(z,y)} \]

The last equality holds since both \( A \) and \( B \) are supposed to be constant in dimension \( k \).

The 3-dimensional \texttt{product} stream is collapsed into two dimensions by running a sum in the \( z \) dimension, exactly as in the problem statement.

Figure 9: Dataflow network for the \texttt{product} function.

Figure 10: Dataflow networks for \texttt{firstOfPair.d(Y)} and \texttt{secondOfPair.d(Y)}

6 Conclusion

We feel that the visual formalisms used in this presentation allow a novice user to readily understand what is happening in a Lucid program, hence help make intensional programming more accessible.
Figure 11: Dataflow network for the \textit{sum} function.

\begin{align*}
x & : i \\
y & : j \\
z & : 0 \\
t & : 0 \\
Y(0..2) & + (Y(0..2+1)) \\
t & : 1 \\
z & : 0 \\
t & : 0 \\
(Y(0..2) & + (Y(0..2+1))) \\
y & : 0 \\
x & : B \rightarrow B_{0j} \\
A_{0i}B_{0j} \\
z & : 1 \\
y & : 0 \\
x & : B \rightarrow B_{1j} \\
A_{1i}B_{1j} \\
\sum_{k=0..1} A_{ik}B_{kj} \\
z & : 1 \\
t & : 0 \\
z & : 2 \\
y & : 0 \\
x & : B \rightarrow B_{2j} \\
A_{2i}B_{2j} \\
z & : 3 \\
y & : 0 \\
x & : B \rightarrow B_{3j} \\
A_{3i}B_{3j} \\
\sum_{k=2..3} A_{ik}B_{kj} \\
\sum_{k=0..3} A_{ik}B_{kj}
\end{align*}

Figure 12: Evaluation tree of the Matrix Multiplication program