Calculating the Warshall/Floyd Path Algorithm

by

Roland C. Backhouse

92/09
This is a series of notes of the Computing Science Section of the Department of Mathematics and Computing Science Eindhoven University of Technology. Since many of these notes are preliminary versions or may be published elsewhere, they have a limited distribution only and are not for review. Copies of these notes are available from the author.

Copies can be ordered from:
Mrs. F. van Neerven
Eindhoven University of Technology
Department of Mathematics and Computing Science
P.O. Box 513
5600 MB EINDHOVEN
The Netherlands
ISSN 0926-4515

All rights reserved
editors: prof.dr.M.Rem
prof.dr.K.M.van Hee.
CALCULATING THE WARSHALL/FLOYD PATH ALGORITHM

Roland C. Backhouse
Department of Mathematics and Computing Science,
Eindhoven University of Technology,
P.O. Box 513,
5600 MB Eindhoven,
The Netherlands.

May 18, 1992

Abstract

A calculational derivation is given of an all-pairs path algorithm two instances of which are Warshall’s reachability algorithm and Floyd’s shortest-path algorithm. The derivation provides an elementary example of the importance of the so-called star-decomposition rule.
1 Algebraic Framework

This paper presents a calculational derivation of an all-pairs path algorithm, two well-known instances of which are Warshall's (reachability) algorithm and Floyd's shortest-path algorithm. The calculations presented here are essentially the same as those in [1, 2]. The presentation has been brought up-to-date in that explicit rather than implicit use is made of invariant properties. Moreover notational refinements enhance the clarity of the derivation.

Like [1, 2] the framework for the current derivation is regular algebra. The axioms of regular algebra — the algebra of regular languages — are now widely known and publicised. (See e.g. [6, 4, 3].) The fact that the elementary operators involved in several path-finding algorithms obey the axioms of regular algebra is also widely known and this knowledge will be assumed.

The specific details of the framework are that \((S, +, \cdot, *, 0, 1)\) is a regular algebra. That is, \(S\) is a set on which are defined two binary operators \(+\) and \(\cdot\) and one unary operator \(*\) (written as a postfix of its argument). Addition \((+)\) is associative, commutative and idempotent. Multiplication \((\cdot)\) distributes over addition and is associative but is not necessarily commutative. The basic properties of \(*\) that we use here are, for all \(a, b \in S:\)

\(1\)
\[ a^* = 1 + a \cdot a^* \]

\(2\)
\[ a \cdot (b \cdot a)^* = (a \cdot b)^* \cdot a \]

\(3\)
\[ (a + b)^* = (a^* \cdot b)^* \cdot a^* \]

\(4\)
\[ 1^* = 1 \]

Rule \((2)\) will be referred to as the “leapfrog rule” whilst rule \((3)\) will be called the “star-decomposition rule.” The main contribution made by Backhouse and Carré [2] was to show that these four rules are at the heart of several elimination techniques for solving shortest-path and other path-finding problems. This paper provides the most elementary instance of this thesis. Backhouse and van Gasteren [3] have recently shown how a class of algorithms that includes Dijkstra’s shortest-path algorithm [7] can be derived from these four rules. That derivation is longer than the one here since the
underlying assumptions are more complicated (and the algorithm is more efficient). The current paper can thus be viewed as an elementary introduction to the principal ideas.

An important theorem that we exploit is that if \((S, +, \cdot, *, 0, 1)\) is a regular algebra then so too is \((\mathcal{M}_N(S), +, \cdot, *, 0, 1)\) where \(\mathcal{M}_N(S)\) is the set of \(N \times N\) matrices whose elements are drawn from \(S\). In proving this theorem appropriate definitions must be given of matrix addition, multiplication and star, and of the null and identity matrices \(0\) and \(1\). For the first two the usual definitions of matrix addition and multiplication are taken; for the last two: \(0\) is the \(N \times N\) matrix all of whose entries are 0 and the identity matrix is a 0-1 matrix that is everywhere 0 but for its diagonal elements (which are all 1). It is not so easy to explain the definition of \(A^*\), for matrix \(A\), in a few words. Appropriate definitions are given in [1, 4]. The former also includes a proof of the theorem. A proof has also recently been published by Kozen [8]. (As remarked by Kozen the proof of the theorem is an elementary exercise, although it does not appear explicitly in any of the standard references.)

With these preliminaries we can now proceed to the task at hand. Given is a matrix \(A \in \mathcal{M}_N(S)\) and required is to derive an algorithm to compute \(A^+ = A \cdot A^*\). By an "algorithm" we mean an (imperative) program in which in no assignment or test is the star operator applied to a matrix. (Application of star to matrix elements is however allowed.) The matrix \(A^+\) is called the transitive closure of \(A\).

The heuristic underlying the algorithm is to use the star-decomposition rule to reduce the given matrix \(A\) to the null matrix by successively nullifying columns of the matrix. In order to express this formally we need to introduce some additional notation. Specifically, for all integers \(k, 0 \leq k < N\) we define \(\bullet k\) to be a \(N \times 1\) matrix (i.e. a column vector) that is everywhere 0 but for its \(k\)th entry which is 1. (We index rows and columns beginning at index 0.) The notation \(\bullet \circ\) is used for the transpose of \(\bullet k\), that is a \(1 \times N\) matrix (a row vector) whose only non-null entry is the \(k\)th. We also define \([k]\) to be the (matrix) product \(\bullet k \cdot k \bullet\) and \([\geq k]\) to be the sum of \([i]\) over all \(i\) in the range \(k \leq i < N\).

The specific properties that we assume of these expressions are as follows. First, by definition,

\[
(5) \quad [k] = \bullet k \cdot k \bullet
\]
Second, by range splitting on the summation defining $\geq k$, for all $k < N$,

$$[\geq k] = [k] + [\geq (k + 1)]$$

Third, $[\geq N]$ is an empty summation. Thus

$$[\geq N] = 0$$

Fourth, $[\geq 0]$ is the identity matrix:

$$[\geq 0] = 1$$

Finally, because it is a criterion for deciding when we have an algorithm we remark that, for all $N \times N$ matrices $X$, $i \cdot X \cdot j$ is the $(i, j)$th element of $X$.

## 2 First Steps

We begin our calculation by noting that $[\geq k] \cdot X$, for $N \times N$ matrix $X$, is a copy of $X$ but for its first $k$ columns which are all null. Let us consider the expression $([\geq k] \cdot X)^*$ for $k < N$. We have:

$$([\geq k] \cdot X)^* = [k] \cdot X + [\geq (k + 1)] \cdot X^*$$

$$= [k] \cdot X^* + ([\geq (k + 1)] \cdot X^* - ([k] \cdot X^*)^*)^*$$

This little calculation is interesting because the pattern $([\geq i] \cdot Y)^*$ recurs in the first and last lines. In the first line $i$ is just $k$ and $Y$ is $X$. In the last line $i$ is $k + 1$ and $Y$ is $X \cdot ([k] \cdot X)^*$. It invites us to seek a particularly simple recurrence. Let us introduce the function $M$ defined by

$$M(X, k) = X \cdot ([\geq k] \cdot X)^*$$

Then, for $k < N$ we have:

$$M(X, k) = X \cdot ([k] \cdot X)^* \cdot ([\geq (k + 1)] \cdot X \cdot ([k] \cdot X)^*)^*$$

$$= M(X \cdot ([k] \cdot X)^*, k + 1)$$
Noting also that
\[
\begin{align*}
M(X,0) &= \{ \text{(8), definition } \} \\
&= X \cdot (1 \cdot X)^* \\
&\quad \{ 1 \text{ is the unit of multiplication, definition } \} \\
&= X^+ \\
\end{align*}
\]
and
\[
\begin{align*}
M(X,N) &= \{ \text{definition, (7) } \} \\
&= X \cdot (0 \cdot X)^* \\
&\quad \{ 0 \text{ is zero of multiplication, } 0^* = \{(1)\} \} \\
&= X \\
\end{align*}
\]
we have established the correctness of the following algorithm to compute \( A^+ \):
\[
\begin{align*}
\{ A^+ &= M(A,0) \} \\
X, k &:= A, 0 \\
; \quad \{ \text{Invariant: } A^+ = M(X,k) \} \\
d&k \neq N \rightarrow X, k := X \cdot ([k] \cdot X)^*, k + 1 \\
\od \\
\{ A^+ &= M(X,N) \} \\
\{ A^+ &= X \} \\
\end{align*}
\]

3 The Algorithm

There is one more step to be taken before we have an algorithm in which \( \cdot \) is applied only to elements and not to matrices. We take the expression \( X \cdot ([k] \cdot X)^* \) and rewrite it using (5).
\[
\begin{align*}
X \cdot ([k] \cdot X)^* &= \{ (1) \} \\
&= X \cdot 1 + X \cdot [k] \cdot X \cdot ([k] \cdot X)^* \\
&\quad \{ X \cdot 1 = X, (5) \} \\
&= X + X \cdot \bullet k \cdot k \cdot X \cdot (\bullet k \cdot k \cdot X)^* \\
\end{align*}
\]
\[ X + X \cdot k \cdot (k \cdot X \cdot k)^* \cdot k \cdot X \]

As remarked earlier, \( k \cdot X \cdot k \) is the \((k, k)\)th element of \( X \) and since it is the argument of the only application of \( \ast \) we have obtained our algorithm:

\[
\begin{align*}
\{ A^+ &= M(A, 0) \} \\
X, k &:= A, 0 \\
\{ \text{Invariant: } A^+ &= M(X, k) \} \\
do &\ k \neq N \rightarrow \\
X, k &:= X + X \cdot k \cdot (k \cdot X \cdot k)^* \cdot k \cdot X, k + 1 \\
o d \\
\{ A^+ &= M(X, N) \} \\
\{ A^+ &= X \}
\end{align*}
\]

4 Implementation Freedom

The algorithm we have obtained is not quite Warshall’s algorithm or Floyd’s algorithm (even after suitable interpretation of the operators). The reason is that at element level the assignment in the body of the loop is a \textit{simultaneous} assignment to all matrix elements. Spelling this out in detail, the matrix assignment

\[ X := X + X \cdot k \cdot (k \cdot X \cdot k)^* \cdot k \cdot X \]

is directly implemented as the simultaneous assignment

\[
\text{simultaneously_for } i := 0 \text{ to } N - 1 \text{ and } j := 0 \text{ to } N - 1 \text{ do}
\begin{align*}
i \cdot X \cdot j &:= \\
i \cdot X \cdot j + i \cdot X \cdot k \cdot (k \cdot X \cdot k)^* \cdot k \cdot X \cdot j
\end{align*}
\]

(Writing \( i \cdot X \cdot j \) conventionally as \( x_{ij} \) this takes on the more familiar appearance:

\[
\text{simultaneously_for } i := 0 \text{ to } N - 1 \text{ and } j := 0 \text{ to } N - 1 \text{ do}
\begin{align*}
x_{ij} &:= x_{ij} + x_{ik} \cdot (x_{kk})^* \cdot x_{kj}
\end{align*}
\]
But, of course, the problem of the simultaneous assignment remains.)

Exploitation of the, as yet unused, idempotency of addition and star, however, gives unlimited freedom in the order in which the matrix elements are assigned. They may be assigned sequentially as in Warshall’s and Floyd’s algorithms, or completely in parallel!

To explain why this is so consider the function

\[(10) \quad X \mapsto X + X \cdot \bullet k \cdot (k \cdot X \cdot \bullet k)^* \cdot k \cdot \bullet X\]

which, as we know, is equal to the function

\[(11) \quad X \mapsto X \cdot ([k] \cdot X)^*\]

Let this function be called \(f\). The body of the loop is then the assignment

\[X, k := f.X, k + 1\]

Forget the matrix structure of \(X\) and just regard \(X\) as the name of the set of variables \(\{i, j : 0 \leq i, j < N : i \cdot X \cdot \bullet j\}\). Then the fact that the elements of the set may be assigned in an arbitrary order rests on three key properties of the function:

(a) \(f\) is idempotent,

(b) \(X \leq f.X\),

(c) \(f\) is monotone non-decreasing in each of its arguments.

The verification of the idempotency of \(f\) proceeds as follows:

\[f.(f.X)\]

\[= \quad \{ (11) \}\]

\[X \cdot ([k] \cdot X)^* \cdot (\{a* = (1+a)^*\})\]

\[= \quad \{ (1) \}\]

\[X \cdot ([k] \cdot X)^* \cdot (1 + [k] \cdot X \cdot ([k] \cdot X)^*)^*\]

\[= \quad \{ a** = a* and a* = a* \cdot a* \}\]

\[X \cdot ([k] \cdot X)^*\]

\[= \quad \{ (11) \}\]

\[f.X\]
Property (b) is immediate from (10), and (c) follows from the monotonicity of addition, multiplication and star.

The claim that the combination of these three properties permits the conversion of the simultaneous assignment to a parallel assignment may or may not be well known. (It is not well known among colleagues to whom I have spoken.) Its proof is remarkably simple and is given in the appendix.

This concludes the derivation of the Warshall/Floyd algorithm. Note that the total calculation (including the discussion of implementation freedom) takes roughly twenty elementary steps which is about what it should be for such a compact algorithm.

References


Appendix

Let $X$ be a finite set of variables and suppose the type of each of the variables is a set ordered by the (partial) relation $\leq$. Suppose $f$ is an endofunction on the domain of $X$ satisfying the properties:

(a) $f$ is idempotent,

(b) $\forall (X :: X \leq f.X)$,

(c) $f$ is monotone non-decreasing in each of its arguments.

Then the assignment

$$X := f.X$$

is equivalent to the assignment

parfor $x \in X$ do $x := f.x$.

The proof is by induction on the size of $X$. The basis is of course trivial. For the induction step the following lemma suffices.

**Lemma 12** Let $\oplus$ and $\otimes$ be binary operators such that

(a) $x \oplus y = (x \oplus y) \oplus (x \otimes y)$

$\otimes y = (x \oplus y) \otimes (x \otimes y)$

(i.e. the function $(x, y \mapsto x \oplus y, x \otimes y)$ is idempotent.)

(b) $x \leq x \oplus y$

$y \leq x \otimes y$

(c) Both $\oplus$ and $\otimes$ are monotone in both their arguments.

Then the simultaneous assignment

$$x, y := x \oplus y, x \otimes y$$

can be implemented by the sequential assignment

$x := x \oplus y$

$; y := x \otimes y$
or by the sequential assignment
\[
y := x \odot y
\]
; \quad x := x \oplus y

**Proof** We have to show that
\[\tag{13} (x \oplus y) \odot y = x \odot y\]
and
\[\tag{14} x \oplus (x \odot y) = x \oplus y\]
The first is proved as follows:

\[
x \odot y
\leq \quad \{ \text{(b) and (c)} \}
\]
\[
(x \odot y) \odot y
\]
\[
\leq \quad \{ \text{(b) and (c)} \}
\]
\[
(x \oplus y) \odot (x \odot y)
\]

Hence, by (a),
\[
x \odot y = (x \oplus y) \odot y
\]

The second identity is proved similarly.
\[\square\]

**Acknowledgements**

Thanks go to Wim Feijen, Joop van den Eijnde and Lambert Meertens for their critical comments and suggestions for improvement.

Preparation of the report was expedited by the use of the proof editor developed by Paul Chisholm [5].
### In this series appeared:

<table>
<thead>
<tr>
<th>Volume</th>
<th>Author(s)</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>89/1</td>
<td>E.Zs.Lepoeter-Molnar</td>
<td>Reconstruction of a 3-D surface from its normal vectors.</td>
</tr>
<tr>
<td>89/2</td>
<td>R.H. Mak</td>
<td>A systolic design for dynamic programming.</td>
</tr>
<tr>
<td></td>
<td>P.Struik</td>
<td></td>
</tr>
<tr>
<td>89/3</td>
<td>H.M.M. Ten Eikelder C. Hemerik</td>
<td>Some category theoretical properties related to a model for a polymorphic lambda-calculus.</td>
</tr>
<tr>
<td>89/4</td>
<td>J.Zwiers W.P. de Roever</td>
<td>Compositionality and modularity in process specification and design: A trace-state based approach.</td>
</tr>
<tr>
<td>89/5</td>
<td>Wei Chen T.Verhoeff J.T.Udding</td>
<td>Networks of Communicating Processes and their (De-)Composition.</td>
</tr>
<tr>
<td>89/6</td>
<td>T.Verhoeff</td>
<td>Characterizations of Delay-Insensitive Communication Protocols.</td>
</tr>
<tr>
<td>89/7</td>
<td>P.Struik</td>
<td>A systematic design of a parallel program for Dirichlet convolution.</td>
</tr>
<tr>
<td>89/9</td>
<td>K.M. van Hee P.M.P. Rambags</td>
<td>Discrete event systems: Dynamic versus static topology.</td>
</tr>
<tr>
<td>89/10</td>
<td>S.Ramesh</td>
<td>A new efficient implementation of CSP with output guards.</td>
</tr>
<tr>
<td>89/11</td>
<td>S.Ramesh</td>
<td>Algebraic specification and implementation of infinite processes.</td>
</tr>
<tr>
<td>89/12</td>
<td>A.T.M.Aerts K.M. van Hee</td>
<td>A concise formal framework for data modeling.</td>
</tr>
<tr>
<td>89/14</td>
<td>H.C.Haesen</td>
<td>ELDA, data manipulatie taal.</td>
</tr>
<tr>
<td>89/15</td>
<td>J.S.C.P. van der Woude</td>
<td>Optimal segmentations.</td>
</tr>
<tr>
<td>89/16</td>
<td>A.T.M.Aerts K.M. van Hee</td>
<td>Towards a framework for comparing data models.</td>
</tr>
<tr>
<td>89/17</td>
<td>M.J. van Diepen K.M. van Hee</td>
<td>A formal semantics for Z and the link between Z and the relational algebra.</td>
</tr>
<tr>
<td>Page</td>
<td>Authors/Contributors</td>
<td>Title/Description</td>
</tr>
<tr>
<td>-------</td>
<td>----------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>90/2</td>
<td>K.M. van Hee P.M.P. Rambags</td>
<td>Dynamic process creation in high-level Petri nets, pp. 19.</td>
</tr>
<tr>
<td>90/4</td>
<td>A. Peeters</td>
<td>Decomposition of delay-insensitive circuits, p. 25.</td>
</tr>
<tr>
<td>90/5</td>
<td>J.A. Brzozowski J.C. Ebergen</td>
<td>On the delay-sensitivity of gate networks, p. 23.</td>
</tr>
<tr>
<td>90/8</td>
<td>M.B. Josephs</td>
<td>Receptive Process Theory, p. 16.</td>
</tr>
<tr>
<td>90/9</td>
<td>A.T.M. Aerts P.M.E. De Bra K.M. van Hee</td>
<td>Combining the functional and the relational model, p. 15.</td>
</tr>
<tr>
<td>90/10</td>
<td>M.J. van Diepen K.M. van Hee</td>
<td>A formal semantics for Z and the link between Z and the relational algebra, p. 30. (Revised version of CSNotes 89/17).</td>
</tr>
<tr>
<td>90/11</td>
<td>P. America F.S. de Boer</td>
<td>A proof system for process creation, p. 84.</td>
</tr>
<tr>
<td>90/12</td>
<td>P.America F.S. de Boer</td>
<td>A proof theory for a sequential version of POOL, p. 110.</td>
</tr>
<tr>
<td>90/14</td>
<td>F.S. de Boer</td>
<td>A proof system for the language POOL, p. 70.</td>
</tr>
<tr>
<td>90/15</td>
<td>F.S. de Boer</td>
<td>Compositionality in the temporal logic of concurrent systems, p. 17.</td>
</tr>
<tr>
<td>90/16</td>
<td>F.S. de Boer C. Palamidessi</td>
<td>A fully abstract model for concurrent logic languages, p. 23.</td>
</tr>
<tr>
<td>90/17</td>
<td>F.S. de Boer C. Palamidessi</td>
<td>On the asynchronous nature of communication in logic languages: a fully abstract model based on sequences, p. 29.</td>
</tr>
<tr>
<td>Volume/Year</td>
<td>Authors</td>
<td>Title</td>
</tr>
<tr>
<td>------------</td>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>91/03</td>
<td>J. P. Katoen, L. A. M. Schoenmakers</td>
<td>Parallel Programs for the Recognition of $P$-invariant Segments, p. 16.</td>
</tr>
<tr>
<td>91/05</td>
<td>D. de Reus</td>
<td>An Implementation Model for GOOD, p. 18.</td>
</tr>
<tr>
<td>91/06</td>
<td>K. M. van Hee</td>
<td>SPECIFICATIEMETHODEN, een overzicht, p. 20.</td>
</tr>
<tr>
<td>91/07</td>
<td>E. Poll</td>
<td>CPO-models for second order lambda calculus with recursive types and subtyping, p. 49.</td>
</tr>
<tr>
<td>91/12</td>
<td>E. van der Sluis</td>
<td>Relational Catamorphism, p. 31.</td>
</tr>
<tr>
<td>91/14</td>
<td>P. Lemmens</td>
<td>The PDB Hypermedia Package. Why and how it was built, p. 63.</td>
</tr>
</tbody>
</table>

An example of proving attribute grammars correct: the representation of arithmetical expressions by DAGs, p. 25.

Transforming Functional Database Schemes to Relational Representations, p. 21.

Transformational Query Solving, p. 35.

Some categorical properties for a model for second order lambda calculus with subtyping, p. 21.


Assertional Data Reification Proofs: Survey and Perspective, p. 18.


Z and high level Petri nets, p. 16.

Formal semantics for BRM with examples, p. 25.

A compositional proof system for real-time systems based on explicit clock temporal logic: soundness and completeness, p. 52.

The GOOD based hypertext reference model, p. 12.

Embedding as a tool for language comparison: On the CSP hierarchy, p. 17.

A compositional proof system for dynamic process creation, p. 24.

Correctness of Acceptor Schemes for Regular Languages, p. 31.

An Algebra for Process Creation, p. 29.
Some algorithms to decide the equivalence of recursive types, p. 26.


The modelling and analysis of queueing systems with QNM-ExSpect, p. 23.

Specifying fault tolerant programs in deontic logic, p. 15.

Asynchronous communication in process algebra, p. 20.

A note on compositional refinement, p. 27.

A compositional semantics for fault tolerant real-time systems, p. 18.

Real space process algebra, p. 42.

Program derivation in acyclic graphs and related problems, p. 90.

Conservative fixpoint functions on a graph, p. 25.

Discrete time process algebra, p.45.

The fine-structure of lambda calculus, p. 110.

On stepwise explicit substitution, p. 30.


Composition and decomposition in a CPN model, p. 55.

Demonic operators and monotype factors, p. 29.