MASTER COPY

PATH ALGEBRAS : A MULTISET-THEORETIC APPROACH

Thesis Submitted for the Degree of

Doctor of Philosophy

at the University of Southampton

July 1976

by Ahnont Wongseelashote

Department of Electronics

The University

Southampton.

ABSTRACT

FACULTY OF ENGINEERING AND APPLIED SCIENCE

DEPARTMENT OF ELECTRONICS

Doctor of Philosophy

PATH ALGEBRAS : A MULTISET-THEORETIC APPROACH

by Ahnont Wongseelashote

This thesis develops an algebraic theory for path problems such as that of finding the shortest or more generally, the k shortest paths in a network, enumerating elementary paths in a graph. It differs from most earlier work in that the algebraic structure appended to a graph or a network of a path problem is not axiomatically given as a starting point of the theory, but is derived from a novel concept called a 'path space'. This concept is shown to provide a coherent framework for the analysis of path problems, and the development of algebraic methods for solving them. A number of solution methods are derived, which are analogous to the classical techniques of solving linear algebraic equations, and the applicability of these methods to different classes of path problems is examined in detail.

The thesis also presents in particular an algebra which is appropriate for the formulation and solution of k-shortest-paths problems. This algebra is a generalization of Giffler's Schedule Algebra for computing all the numerical labels of paths in a network. It is shown formally that these labels can be calculated by using direct methods of linear algebra and an algorithm similar to the long-division procedure of ordinary arithmetic. Such a method is then modified to yield an algorithm for finding k shortest elementary paths in a network.

TO HAIDEH

PATH ALGEBRAS

A Multiset-Theoretic Approach

CONTENTS

ACKNOWLEDGEMENTS

The author is especially grateful to his supervisor, Dr. Bernard A. Carré for his help and encouragement and also for his kind understanding of the author's numerous personal problems during the preparation of this thesis. He also wishes to acknowledge the benefit he derived from reading the draft of his supervisor's book before its publication.

The author is indebted to Dr. E. Keith Lloyd of the Department of Mathematics for very fruitful discussions of this work, and for his most helpful suggestions and criticisms of the author's draft papers.

If this thesis is worthy of attention by mathematicians, it is the fruit of those who taught the author mathematics, especially Dr. Haya Freedman and the late Dr. Leslie W. Ahrendt. Otherwise, only the author is responsible.

Mrs. B. Betts and Mrs. J. Grace very kindly undertook the typing at short notice, and the author wishes to thank them for their helpful and competent assistance.

Finally, the author wishes to acknowledge the financial support which he received from the Bank of Thailand throughout this work.

INTRODUCTION

This thesis is concerned with the mathematical (abstract) study of path problems. The term 'path problem' has, in recent years, been widely used to describe a certain class of mathematical problems, many of which have real world applications. These problems had previously been studied separately by a great number of authors in different branches of engineering, Operational Research and Computing Science. Numerous procedures for solving them had also been proposed separately. It was not until the use of algebraic methods for these problems became widespread that mathematicians began their search for a unified theory which would be useful for their solution.

Any theory which would adequately describe a path problem must be able to incorporate the two different mathematical aspects of the problem: one algebraic, the other structural. The algebraic aspect of the problem is usually described by a semiring (section 0.2) which is often called a path algebra, while the structural aspect is described by a graph (section 0.4). The roles played by these two mathematical constructs in the abstract study of path problems will be seen in Chapter 1 where we present a retrospective study of the existing theory of path problems. An important point which emerges from this study is that the path algebra of a path problem can be naturally derived from a more basic concept, which we describe as a 'path space'. Chapters 2 and 3 therefore develop this concept. Then, in chapter 4 we show how the interaction between the two mathematical aspects of a path problem can be fruitfully analysed with the help of a path space. The usefulness of this approach is further demonstrated in Chapter 5 where we analyse and extend various methods of solving path problems.

The final chapter treats a specific path algebra in detail. This path algebra is shown to be especially useful for solving the k shortest path problems, particularly in the case where only elementary paths are required. A novel algorithm for obtaining such paths is also given.

In order to make this thesis self-contained and to prevent possible confusion of terminology, a preliminary chapter on background mathematics is also included.

CHAPTER 0

BACKGROUND MATHEMATICS

0.1 Sets, Relations and Functions

As usual in mathematics, we do not formally define what a set is, but rather think of it intuitively as a collection of distinct objects which can be distinguished (at least theoretically) from those objects which do not belong to the set. Moreover, throughout this thesis, whenever the word "set" is used, it means a subset of a given set. We shall use the notation 2^X to denote the set of all subsets of a given set X, and $2^{(X)}$ to denote the set of all finite subsets of X, i.e. those subsets which contain only a finite number of objects obtained from the given set X. We shall also write

(i) $x \in A$ for "x is an object or element of a set A", and $x \notin A$ for its negation.

(ii) $A \subseteq B$ for "A is a subset of a set B", and $A \nsubseteq B$ for its negation.

(iii) $A \setminus B$ for "the set of all objects which belong to A but not to B".

The notation $\{x|P(x)\}$ will always be used to denote the set of all x such that the proposition $P(x)$ is valid. For example, the set $\{x_i \mid i \in I\}$ denotes the set of all elements indexed by some set I. Now let $\{A_i|i \in I\}$ be a collection (set) of sets indexed by some set I. Then the set $\bigcup_{i \in I} A_i$ of elements which belong to at least one set A_i is called the union of the sets A_i , and the set $\bigcap_{i \in I} A_i$ of elements which belong to every set A_i is called the intersection of the sets A_i. For $I = \{1,2\}$, we also write $A_1 \cup A_2$ and $A_1 \cap A_2$ for the union and intersection of A_1 and A_2 respectively.

By a Cartesian product $A \times B$ of two given sets A, B, we shall mean, as usual, the set made up of all ordered pairs (x, y) with $x \in A$ and $y \in B$; similarly for the Cartesian product $A \times B \times C$.

By a relation on a set X , we mean a subset of the Cartesian product $X \times X$. When A is a relation on X , we usually write

$$x \, A \, y \quad \text{for } (x, y) \in A.$$

We then say that A is reflexive iff† $x A x$ for all $x \in X$, transitive iff whenever $x A y$ and $y A z$, $x A z$ also; symmetric iff $x A y$ always implies $y A x$, and anti-symmetric iff from $x A y$ and $y A x$ always follows $x = y$. Among all types of relations the following two classes are very useful to us. The first is the class of equivalence relations. A relation $\sim$ is said to define an equivalence relation on a set X iff it is reflexive, symmetric and transitive. The most important property of such a relation is that it partitions the set X into disjoint equivalence subsets of X, i.e. those subsets of which the elements are in equivalence relation to one another. Also conversely, every disjoint partition of X always defines an equivalence relation on X. The second useful class of relations is formed by orderings. A relation $\leq$ is said to define an ordering on a set X iff it is

---

† "iff" denotes "if and only if"

reflexive, anti-symmetric and transitive. An ordering is said to be <u>total</u> on X iff for any $x, y \in X$, we either have $x \leq y$ or $y \leq x$. A set X considered with an ordering $\leq$, usually written as $(X, \leq)$, is called an <u>ordered set</u>. Several classes of ordered sets will be considered below.

Let $(X, \leq)$ be an ordered set. A subset A of X is said to contain a <u>least</u> (<u>greatest</u>) element x', also written as min A (max A), iff $x' \in A$ and $x' \leq x$ $(x \leq x')$ for all $x \in A$. If every non-empty subset of X has this property, then $(X, \leq)$ is called a <u>well ordered</u> (<u>dually well ordered</u>) set.

Let A be a subset of X. An element $u \in X$ is called an <u>upper</u> (<u>lower</u>) <u>bound</u> of A in X iff $x \leq u$ $(u \leq x)$ for all $x \in A$. The set A is then said to be <u>bounded above</u> (<u>below</u>). Moreover, u is called the <u>least upper</u> (<u>greatest lower</u>) <u>bound</u> of A in X if u is also the least (greatest) element among the upper (lower) bounds of A in X. With these definitions, we can now define the following interesting classes of ordered sets.

An ordered set $(X, \leq)$ is called a <u>join</u> (<u>meet</u>) <u>semilattice</u> iff every finite subset A of X has a least upper bound sup A <u>(greatest lower bound inf A),</u> a <u>complete</u> <u>join</u> (<u>meet</u>) <u>semilattice</u> iff every non-empty subset of X has a least upper (greatest lower) bound, and a <u>conditionally complete join (meet) semilattice</u> iff every non-empty subset of X which is bounded above (below) has a least upper (greatest lower) bound. If $(X, \leq)$ is both a (complete, conditionally complete) join and meet semilattice, then it is simply called a <u>(complete, conditionally complete) lattice.</u>

Now for any two sets A,B, a <u>function</u> f from A to B, written $f: A \to B$, is defined as a subset of the Cartesian product $A \times B$ such that $(x, y) \in f$ for each $x \in A$ and a single $y \in B$. Usually, y is denoted by $f(x)$ and referred to as the <u>image</u> of x under f. The set of all the images under f is usually called the <u>range</u> of f and denoted by $f(A)$. The <u>domain</u> of f is just the set A. A function $f: A \to B$ is said to be a <u>surjection</u> iff $f(A) = B$, an <u>injection</u> iff from $f(x)=f(x')$ always follows $x=x'$, and a <u>bijection</u> iff it is both an injection and a surjection. When a set A is said to be in <u>one-to-one correspondence</u> with another set B, we mean there is a bijection from A to B.

Let $f: A \to B$ and $g: B \to C$ be two functions. Then the function $gf: A \to C$ defined by $gf(x) = g(f(x))$ for all $x \in A$ is called the <u>composition</u> of f by g (in that order). More generally, let $g : B' \to C$ where $B' \subseteq B$, we can define the composition gf as above whenever $f(A) \subseteq B'$. Now a function $f: A \to B$ is said to <u>extend</u> or be an <u>extension</u> of $g: A' \to B'$ iff $A' \subseteq A$, $B' \subseteq B$ and $f(x) = g(x)$ for all $x \in A'$.

By a <u>binary operation</u> o on a set X, we mean a function $o : X \times X \to X$. Each image $o(x, y)$ is more familiarly written as $x \circ y$. A binary operation o is said to be <u>commutative</u> iff $x \circ y = y \circ x$ for any x, y; <u>associative</u> iff $(x \circ y) \circ z = x \circ (y \circ z)$ for any x, y, z. An element e of a set X is said to be an <u>identity</u> for a binary operation o on X iff $e \circ x = x = x \circ e$ for all $x \in X$. Note that whenever such an element exists, it is unique. Now for each $x \in X$, we can define the <u>k^{th} power</u> of x with respect to a binary operation o on X to be an element x^k of X which is obtained recursively as follows.

$$x^k = x \circ x^{k-1} \quad \text{for any} \quad k \in \{2, 3, \ldots\} \ , \ x^1 = x \ .$$

It is also convenient to define $x^0 = e$, the identity for $\circ$, whenever e exists. Finally, a non-empty subset A of a set X is said to be closed with respect to a binary operation $\circ$ on X iff $x \circ y \in A$ whenever $x, y \in A$.

0.2 Monoids and Semirings

By a monoid $(X,\circ)$ we mean a non-empty set X equipped with an associative binary operation $\circ$ for which there is an element $e \in X$ acting as the identity. For any two subsets A, B of a monoid $(X, \circ)$, we define a new set

$$AB = \{x \circ y \mid x \in A, y \in B\} \ ,$$

called the complex product of A and B induced by $\circ$. With respect to this complex product operation, the set 2^X can be easily seen to form a monoid with $\{e\}$ as the identity.

A monoid $(X, \circ)$ is said to be commutative iff $\circ$ is commutative, cancellative iff $x \circ z = y \circ z$ or $z \circ x = z \circ y$ for any $x, y, z \in X$ always implies $x = y$, and a group iff each element of X has an inverse, i.e. for every $x \in X$, there is $y \in X$ such that $x \circ y = e = y \circ x$. Note that whenever an inverse element exists, it is unique by the associativity of $\circ$. A monoid $(X, \circ)$ is said to be locally finite (Eilenberg (1974))iff each $x \in X$ admits only a finite number of factorizations $x = x_1 \circ x_2 \circ \ldots \circ x_n$ with $x_i \neq e$ for all $i \in \{1, 2, \ldots, n\}$. It follows that if $x \circ y = e$, then $x = e = y$, and hence a non-trivial group cannot be locally finite.

We shall also be interested in monoids (X, o) which are also ordered by some relation $\leqslant$ such that the following condition is satisfied:

For any $x, y \in X$ such that $x \leqslant y$, we have $x \circ u \leqslant y \circ u$ and $u \circ x \leqslant u \circ y$ for all $u \in X$.

Such a monoid will be called an ordered monoid, and denoted by $(X, \leqslant, o)$. Furthermore, whenever the above condition is satisfied, we shall say that the binary operation o is compatible with the ordering $\leqslant$ on X. An ordered monoid $(X, \leqslant, o)$ is said to be Archimedean iff from $x > e$ and $y > e$, we can always find a positive integer n such that $x^n > y$. Here $x > y$ denotes, as usual, $x \geqslant y$ (or $y \leqslant x$) and $x \neq y$.

Let X, Y be two sets equipped with binary operations o_X and o_Y respectively. Then a function $f: X \to Y$ is called a homomorphism iff

$$f(x \; o_X \; x') = f(x) \; o_Y \; f(x') \qquad \text{for all } x, x' \in X .$$

If in addition, both X and Y possess identity elements, say e_X and e_Y respectively and (X, o_X) is a monoid, then f is called a monoid homomorphism whenever $f(e_X) = e_Y$ is also satisfied. Note that $(f(X), o_Y)$ then becomes a monoid also.

A semiring $(X, +, o)$ is a set X on which two binary operations $+$ and o, called addition and multiplication respectively, are defined such that

(i) $(X, +)$ is a commutative monoid with $\theta \in X$ as identity for $+$,

(ii) (X, o) is a monoid with $e \in X$ as identity for o, and

(iii) Multiplication is distributive over addition, i.e. $x \circ (y+z) = x \circ y + x \circ z$ and $(x+y) \circ z = x \circ z + y \circ z$ for any $x, y, z, \in X$.

We note here that the expressions on the right-hand sides of the above two equalities are not ambiguous if we assume the convention that multiplication is performed before addition. Indeed, this convention will be implicitly assumed for all the semirings discussed throughout this thesis.

Now the identity e for multiplication will be called the unit and the identity θ for addition will be called the zero of the semiring. Note that the latter definition is inspired by the fact that

$$x \circ \theta = \theta = \theta \circ x \quad \text{for all } x \text{ in any semiring.}$$

This is because $x = x \circ e = x \circ (e+\theta) = x \circ e + x \circ \theta = x + x \circ \theta$ for all x implies that $x \circ \theta = \theta$, and similarly for $\theta = \theta \circ x$. Note also that whenever $e = \theta$ in a semiring $(X, +, \circ)$, we necessarily have $X = \{\theta\}$ (because $x = x \circ e = x \circ \theta = \theta$).

Let $A \subseteq X$. Then A is said to be a subsemiring of a semiring $(X, +, \circ)$ iff $(A, +, \circ)$ is itself a semiring. Consequently, for a given semiring $(X, +, \circ)$, we infer that a non-empty subset A of X is a subsemiring iff A is closed with respect to addition and multiplication, and that $\theta, e \in A$ also.

In the present study, it is often convenient to express certain properties of a semiring $(X, +, \circ)$ in terms of the relation $\prec$ defined on X by

(0.1) $\quad x \prec y \quad \text{iff} \quad x + y = y \quad \text{for any } x, y \in X$.

It is easily seen that $\prec$ is anti-symmetric because addition is commutative, and that $\prec$ is transitive because addition is associative. But $\prec$ is not generally reflexive. However, it is always so whenever $e + e = e$ holds. For these reasons, we shall, for convenience, refer to $\prec$ as the pseudo-ordering of the semiring $(x, +, o)$. The following properties of $\prec$ will be found especially useful for the present study. Their proofs are straightforward and hence omitted.

(0.2) $\theta \prec x$ for all $x \in X$.

(0.3) $x \prec y$ implies $x \circ u \prec y \circ u$ and $u \circ x \prec u \circ y$ for all $u \in X$. Moreover, if $u \prec w$, then $x \circ u \prec y \circ w$ and $u \circ x \prec w \circ y$ always.

(0.4) $x \prec y$ and $u \prec w$ imply $x + u \prec y + w$ for any $x, y, u, w \in X$.
Moreover, if $e + e = e$ holds, then $x + u \prec y + u$ for all $u \in X$.

(0.5) $x \prec z$ and $y \prec z$ imply $x + y \prec z$ for any $x,y,z \in X$. More generally, $x_i \prec z$ for all $i \in \{1,2,\ldots,k\}$ implies that

$$x_1 + x_2 + \ldots + x_k \prec z.$$

A semiring $(X, +, o)$ is said to be a ring iff $(x, +)$ is also a group, commutative iff o is commutative, idempotent iff $e + e = e$, and a Q-semiring (Yoeli (1961)) iff $x + e = e$ for all $x \in X$. Note that every Q-semiring is idempotent and every idempotent ring $(X, +, o)$ is trivial, i.e. $X = \{\theta\}$. The latter follows because for all $x \in X$,

$$x = x + \theta = x + x + (-x) = x + (-x) = \theta.$$

A commutative ring $(X, +, o)$ in which $e \neq 0$ and $x \circ y = \theta$ always implies either $x = \theta$ or $y = \theta$ is called an <u>integral domain</u>. A commutative ring $(X, +, o)$ in which the set of non-zero elements forms a group with respect to multiplication is called a <u>field.</u>

Now let $(X, +_X, o_X)$ be a semiring and Y be any set equipped with two binary operations $+_Y$ and o_Y, and with identities θ_y for $+_Y$ and e_Y for o_Y. Then a function $f\colon X \to Y$ is called a <u>semiring homomorphism</u> iff f is a monoid homomorphism with respect to $(X, +_X)$ and (X, o_X). Note that $(f(X), +_Y, o_Y)$ then becomes a semiring. If in addition, f is a bijection, then f is called a <u>semiring isomorphism</u>, and X, Y are also said to be <u>isomorphic</u> as semirings.

In the present study, we shall also be interested in the concept of a complete semiring (Eilenberg (1974)) which can be defined as follows.

Let (X, o) be a monoid with $e \in X$ as the identity for o and consider a <u>formal sum</u> $\sum_{i \in I} x_i$ for an arbitrary indexing set I to be a well defined element of X which satisfies (0.6) to (0.8) below.

(0.6) If $I = \{ i \}$, then $\sum_{i \in I} x_i = \{x_i\}$

(0.7) If $I = \bigcup_{j \in J} I_j$ is a disjoint partition of I, then

$$\sum_{i \in I} x_i = \sum_{j \in J} \left(\sum_{i \in I_j} x_i \right).$$

(0.8) $$z \circ \left(\sum_{i\varepsilon I} x_i \right) = \sum_{i\varepsilon I} (z \circ x_i) \ , \quad \text{and}$$

$$\left(\sum_{i\varepsilon I} x_i \right) \circ z = \sum_{i\varepsilon I} (x_i \circ z) \ .$$

The set X is then said to form a complete semiring. Note that a complete semiring is also a semiring if one defines $x_1 + x_2$ as $\sum_{i\varepsilon I} x_i$ with $I = \{1, 2\}$ and θ as $\sum_{i\varepsilon I} x_i$ with $I = \phi$. As an example of a complete semiring, consider the set $N_\infty = N \cup \{\infty\}$ which is obtained from the usual semiring $(N, +, \cdot)$ of non-negative integers by augmenting it with the element ∞ and extending addition and multiplication by the following rules

(i) For any $n \varepsilon N$, $n + \infty = \infty + n = \infty + \infty = \infty$,

(ii) For any $n \varepsilon N$ such that $n \neq 0$, $n \infty = \infty n$, and

(iii) $\infty \infty = \infty$, $0 \infty = 0 = \infty 0$.

A formal sum $\sum_{i\varepsilon I} x_i$ can then be defined in N_∞ as follows:

If $x_i = 0$ for all but a finite number of $i \varepsilon I$, then $\sum_{i\varepsilon I} x_i$ is the addition of all the non-zero x_i . Otherwise, $\sum_{i\varepsilon I} x_i = \infty$.

Therefore, N_∞ is a complete semiring. Moreover, the set N_∞ also forms a semiring with respect to the extended addition and multiplication as defined above.

0.3 Matrices

By an $(m \times n)$ matrix A over a set X, we mean a function $A : \{1, 2, \ldots, m\} \times \{1, 2, \ldots, n\} \to X$. Each image $A(i,j)$ is usually written as A_{ij}, also called the (i,j)-entry or (i, j)-element of the matrix A. An $(m \times n)$ matrix A can also be visualized as an array of $m \times n$ elements, namely

$$\begin{bmatrix} A_{11}, & A_{12}, & \ldots, & A_{1n} \\ A_{21}, & A_{22}, & \ldots, & A_{2n} \\ A_{m1}, & A_{m2}, & \ldots, & A_{mn} \end{bmatrix}$$

The entries $A_{i1}, A_{i2}, \ldots, A_{in}$ form the i^{th} row of A and the entries $A_{1j}, A_{2j}, \ldots, A_{mj}$ form the j^{th} column of A. The matrix A above has m rows and n columns : it is an $(m \times n)$ matrix.

If $(X, +, \circ)$ is a semiring, we can define addition $A + B$ for any two $(m \times n)$ matrices over X by

(0.9) $$(A + B)_{ij} = A_{ij} + B_{ij} \qquad \text{for all } i,j;$$

and multiplication $A \circ B$ for two conformable matrices, say A is $m \times r$ and B is $r \times n$, by

(0.10) $$(A \circ B)_{ij} = \sum_{k=1}^{r} (A_{ik} \circ B_{kj}) \text{ for all } i, j.$$

Therefore, whenever $(X, +, \circ)$ is a semiring, we can always make the set $\mathcal{M}_n(X)$ of all $(n \times n)$ matrices over X into a semiring by defining additions and multiplication of any two $(n \times n)$ matrices as above.

Moreover, the unit I and zero Θ of this semiring is given by

$$I_{ij} = \begin{cases} e & \text{if } i = j \\ \theta, & \text{otherwise} \end{cases} \quad \text{and} \quad \Theta_{ij} = \theta \quad \text{for all } i, j.$$

Here e and θ are respectively the unit and zero of the semiring $(X, +, o)$.

If X is actually a complete semiring, then we can also make $\mathcal{M}_n(X)$ into a complete semiring by defining multiplication as before but addition is now replaced by the following definition of a formal sum $\sum_{k \in I} A_k$.

$$(0.11) \qquad \left(\sum_{k \in I} A_k\right)_{ij} = \sum_{k \in I} (A_k)_{ij} \qquad \text{for all } i, j$$

Now if X is ordered by some relation $\leqslant$, we can also define

$$(0.12) \qquad A \leqslant B \text{ iff } A_{ij} \leqslant B_{ij} \text{ for all } i, j,$$

where A and B are any two $(m \times n)$ matrices

Therefore, whenever $(X, \leqslant)$ is an ordered set, we can also make $\mathcal{M}_n(X)$ into an ordered set with respect to the ordering defined by (0.12). Moreover, if addition and multiplication of the semiring $(X, +, o)$ is compatible with the ordering $\leqslant$, then so is addition and multiplication of the semiring $(\mathcal{M}_n(X), +, o)$ with the ordering defined by (0.12).

0.4 Graphs

By a graph G, we mean an ordered pair (W,V), where W is a finite set of elements called nodes, and V is a set of ordered pairs of nodes called arcs. For convenience, we shall always assume that W has n nodes which are designated as $x_1, x_2, \ldots, x_n$. In such a graph G, we define a path p of order k which begins at a node x_{i_0} and ends at a node x_{i_k} to be a sequence

$$(x_{i_0}, x_{i_1})\,(x_{i_1}, x_{i_2}) \ldots (x_{i_{k-1}}, x_{i_k})$$

of k consecutive arcs. $x_{i_1}, x_{i_2}, \ldots, x_{i_{k-1}}$ will be referred to as intermediate nodes of p. A path q is said to be a subpath of p iff q is a path which can be traversed when traversing p. The path p is said to be closed iff $x_{i_0} = x_{i_k}$, and elementary iff $x_{i_r} \neq x_{i_s}$ whenever $r \neq s$ (except of course, for closed paths where $x_{i_0} = x_{i_k}$ must be satisfied). For convenience, we shall always assume that the set P of all paths in the graph G also contains the null path θ_i for each node x_i which can be defined as a closed path of order zero with x_i as its beginning and end.

Now for any two given paths p and q in which the end of p is just the beginning of q , we may combine the two paths into one single path pq by concatenating the sequence of arcs of p with those of q. Such an operation will be referred to as path concatenation and will be denoted simply by juxtaposition. Since the concatenation of two paths is not always defined, it is not a binary operation as defined in section 0.1. However, it is convenient to regard path concatenation as a kind of partial binary operation defined on the set P of all paths in G, and to consider some of its properties which are similar to those of binary operations. Thus for instance, path concatenation is associative, and $p\theta_j = p = \theta_i p$ for all paths p which begin at x_i and end at x_j.

The following observation will be useful in our subsequent study. Suppose p is a non-elementary path, then in traversing the path p, we must come across at least one elementary closed path, say ω_1. Accordingly, we may factorize p as follows.

$p = p_1 \omega_1 q_1$, where p_1 and q_1 are subpaths of p, one of which may be null but not both; if p_1 is not null, then p_1 has the same beginning as p and the same end as ω_1; if q_1 is not null, then q_1 has the same beginning as ω_1 and the same end as p.

Now since the beginning and end of ω_1 are the same, $p_1 q_1$ is also a path with the same beginning and end as those of p. In other words, $p_1 q_1$ is just the path obtained from p by deleting ω_1 from p. If $p_1 q_1$ is again non-elementary, we may again factorize $p_1 q_1$ as above and obtain $p_1 q_1 = p_2 \omega_2 q_2$ for some elementary closed path ω_2. We can again factorize $p_2 q_2$ if it is also non-elementary and so on until we finally obtain an elementary path $p_s q_s$ for some $s \geq 1$. For convenience, we shall refer to $p_s q_s$ as a <u>contraction</u> of p and the above process for obtaining $p_s q_s$ as the <u>contraction process</u>.

By a <u>graph G over a set L</u>, we mean a triple $G=(W,V,v)$, where (W,V) is a graph and $v : V \to L$ is a bijection. If in addition, L is a subset of a monoid (X, o) or a semiring $(X, +, o)$, G will be respectively referred to as a <u>graph over a monoid or a semiring</u>. For a graph G over a monoid (X, o), the bijection $v : V \to L$ can be extended to the function $v : P \to X$ as follows.

$$(0.13) \qquad v(p) = \begin{cases} e & \text{if } p = \theta_i \text{ for all } i \\ v(x_{i_0}, x_{i_1}) o v(x_{i_1}, x_{i_2}) o \ldots o\, v(x_{i_{k-1}}, x_{i_k}) & \text{,otherwise.} \end{cases}$$

Note that, for economy of notation, we have here used the same notation v for both functions and that (0.13) is well defined because o is associative, and that $v(pq) = v(p)ov(q)$ for any two paths p and q for which their path concatenation is defined. Thus in view of our previous remark, v can be considered as a kind of partial homomorphism from the set P of all paths in G to the set X. Let us note also that whenever G is a graph over a semiring, $v(p)$ will always be defined in terms of the multiplication of the semiring.

Now let us note the one-to-one correspondence between graphs and matrices over a set L. It is well known that for any given $(n \times n)$ matrix A, one can always define a graph $G = (W, V, v)$ over the set L of all the (i, j)-entries of A by taking

W to be the set of all the columns $A_1, A_2, \ldots, A_n$ of A,

V to be the Cartesian product $W \times W$, and

$v: V \rightarrow L$ to be given by $v(A_i, A_j) = A_{ij}$ for all i,j.

However, if A is a matrix over a semiring $(X, +, o)$, it is convenient to redefine V to be the set of ordered pairs (A_i, A_j) such that $A_{ij} \neq \theta$, where θ is the zero of the semiring. The resulting graph will be denoted by $G(A)$ and will be simply referred to as the <u>graph of the matrix A</u> over the semiring $(X, +, o)$.

Conversely, it is also well known that for any graph G over a set L, one can always define a matrix $A(G) = A$ over $L \cup \{\theta\}$, where $\theta \notin L$ is some special symbol, by

$$(0.14) \qquad A_{ij} = \begin{cases} v(x_i, x_j) & \text{if } (x_i, x_j) \varepsilon V \\ \theta & \text{, otherwise} \end{cases}$$

If G is a graph over a semiring, θ is usually chosen to be the zero of the semiring. For convenience, we shall call $A(G)$ the <u>arc-value matrix</u> of the graph G. We note that as a consequence of choosing a fixed numbering for the nodes of G, the arc-value matrix $A(G)$ is unique. If we apply a permutation to the numbering of nodes of G, then the new arc-value matrix A' is equivalent to the existing arc-value matrix A of G in the sense of matrix equivalence, i.e. $A' = Q \circ A \circ Q^T$, where Q is the permutation matrix obtained by applying the same permutation sequence to the rows of the unit matrix I, and Q^T is defined by $(Q^T)_{ij} = Q_{ji}$ for all i, j. Finally, we note that $A(G(A)) = A$ and $G(A(G)) = G$ always hold.

From the above discussion, we see that any graph G is completely described by its arc-value matrix $A(G) = A$. Moreover, the power of this matrix describes all the paths in Q completely. More specifically, let $P_{ij}^{(k)}$ denote the set of all paths from x_i to x_j which have order exactly k, then it can be shown that A^k, the k^{th} power of the arc-value matrix A, is given by

$$(0.15) \qquad (A^k)_{ij} = \sum_{p \varepsilon P_{ij}^{(k)}} v(p) \quad .$$

A neat way of proving (0.15) is to introduce the function $\sigma : 2^{(P)} \to X$, where $2^{(P)}$ denotes the set of all finite subsets of P, as follows.

$$(0.16) \qquad \sigma(Q) = \sum_{p\varepsilon Q} v(p) \quad , \quad \sigma(\phi) = \theta$$

This function has the following two properties which are easily verified.

$$(0.17) \qquad \sigma(Q_1 \cup Q_2 \cup \ldots \cup Q_k) = \sigma(Q_1) + \sigma(Q_2) + \ldots + \sigma(Q_k) \text{ whenever}$$

$$Q_i \cap Q_j = \phi \quad \text{for} \quad i \neq j .$$

$$(0.18) \qquad \sigma(Q_{ir_1} Q_{r_1 r_2} \ldots Q_{r_k j}) = \sigma(Q_{ir_1}) \circ \sigma(Q_{r_1 r_2}) \circ \ldots \circ \sigma(Q_{r_k j}) ,$$

where each Q_{rs} is a finite subset of paths from x_r to x_s and

$$Q_{rs} \, Q_{st} = \{pq \mid p\varepsilon \, Q_{rs} \text{ and } q \, \varepsilon \, Q_{st}\} .$$

The proof of (0.15) now follows from these two properties. For in view of (0.17) and (0.18), it suffices to show that

$$(0.19) \qquad P_{ij}^{(k)} = \bigcup_{r_1, r_2, \ldots, r_{k-1}} (P_{ir_1}^{(1)} \, P_{r_1 r_2}^{(1)} \ldots P_{r_{k-1} j}^{(1)}) ,$$

and that the terms in the union form a disjoint partition of $P_{ij}^{(k)}$. This can be shown by considering the equivalence relation $\sim$ defined on $P_{ij}^{(k)}$ as follows:

$p \sim q$ iff p and q have the same intermediate nodes.

The set $P_{ij}^{(k)}$ is then partitioned into its equivalence subsets by this relation. Now a glance at the general term on the right-hand side of (0.19) will confirm that it is in fact one such equivalence subset. The proof of (0.15) is therefore completed.

Finally, we note that from (0.15), it follows that $A^n = \Theta$ whenever the graph G has no non-null closed paths, because $P_{ij}^{(n)} = \phi$ for all i, j in such a graph.

CHAPTER 1

PATH PROBLEMS IN RETROSPECT

The first abstract (mathematical) study of path problems that appeared in print was by Moisil (1960). This work of Moisil was inspired by an earlier work of A.G. Lunts (also known as A.G. Lunc) on the application of Boolean algebra to the analysis of relay-contact electrical circuits. Moisil showed that the theorem obtained by Lunts (1950) for matrices over a Boolean algebra in fact holds for matrices over a less restrictive algebraic structure, namely

THEOREM 1.1 Let $(X, +, \circ)$ be a commutative semiring (see section 0.2) which also satisfies

(1.1) For any $x \in X$, $x + x \circ y = x$ for all $y \in X$.

Then for any matrix $A \in \mathcal{M}_n(X)$ such that $A_{ii} = e$ for all i,

(1.2) $I \prec A \prec A^2 \prec \ldots \prec A^{n-1} = A^n = \ldots ,$

where $\prec$ denotes the pseudo-ordering of the semiring $(\mathcal{M}_n(X), +, \circ)$, see section 0.2.

In fact, the assumption of commutativity in the above theorem is superfluous. This point is evident from the following result discovered by Yoeli (1961).

THEOREM 1.2 Let $(X, +, \circ)$ be a Q-semiring (see section 0.2) Then for any matrix $A \in \mathcal{M}_n(X)$ such that $A_{ii} = e$ for all i, we have

$$I \prec A \prec A^2 \prec \dots \prec A^{n-1} = A^n = \dots$$

That the above theorem is equivalent to theorem 1.1 without the assumption of commutativity can be easily seen to be a consequence of the fact that (1.1) says no more than $y \prec e$ for all $y \in X$.

Theorem 1.1 was used by Moisil (1960) to solve special cases of the following problems.

PROBLEM 1.1 Shortest Path

Let G be a graph over the additive group $(R^+, +)$ of non-negative real numbers (see section 0.4). For any two nodes x_i, x_j in G, determine

$$\min \{v(p) | p \in P_{ij}\} ,$$

where P_{ij} is the set of all paths in G which begin at x_i and end at x_j.

PROBLEM 1.2 Maximal Capacity Path

Let G be a graph over $(R^+ \cup \{\infty\}, \wedge)$ where the binary operation $\wedge$ is defined on $R^+ \cup \{\infty\}$ by $a \wedge b = \min \{a,b\}$ For any two nodes x_i, x_j in G, determine

$$\max \{v(p) | p \in P_{ij}\} .$$

PROBLEM 1.3 Most Reliable Path

Let G be a graph over the multiplicative monoid $(\{x | 0 \leq x \leq 1\}, .)$ For any two nodes x_i, x_j in G, determine

$$\max \{v(p) | p \in P_{ij}\} .$$

We note here that the relevance of theorem 1.2 (and hence theorem 1.1) to some of the above problems was also noted by Yoeli (1961). Now in order to solve these three problems by means of theorem 1.2, let us first interpret theorem 1.2 in terms of the graph G(A) of the matrix A (see section 0.4) as follows.

If G(A) is a graph over a Q-semiring such that $v(x_i,x_i) = e$ for all nodes x_i in G(A), then

(1.3) $\sigma(P_{ij}^{(0)}) \prec \sigma(P_{ij}^{(1)}) \prec \sigma(P_{ij}^{(2)}) \prec \ldots \prec \sigma(P_{ij}^{(n-1)}) = \sigma(P_{ij}^{(n)}) = \ldots,$

where σ and $P_{ij}^{(k)}$ are as defined in section 0.4 above.

In fact, the assumption that $v(x_i,x_i) = e$ enables us to conclude that $\sigma(P_{ij}^{[s]}) = \sigma(P_{ij}^{(s)})$ for all $s \in N$, where

$$P_{ij}^{[s]} = \bigcup_{k=0}^{s} P_{ij}^{(k)} \quad \text{for all } s \in N .$$

To see this, let

$$A^{[s]} = I+A+A^2+\ldots+A^s \quad \text{for all } s \in N.$$

Then from (0.15) to (0.17), it follows that the matrix $A^{[s]}$ is given by

(1.4) $$(A^{[s]})_{ij} = \sigma(P_{ij}^{[s]}) \quad \text{for all } s \in N .$$

But the assumption $A_{ii} = e$ for all i is equivalent to $I \prec A$, and hence $A^{[s]} = A^s$ for all $s \in N$. Consequently, the above claim is justified and (1.3) can now be rewritten as

(1.5) $\sigma(P_{ij}^{[0]}) \prec \sigma(P_{ij}^{[1]}) \prec \sigma(P_{ij}^{[2]}) \prec \ldots \prec \sigma(P_{ij}^{[n-1]}) = \sigma(P_{ij}^{[n]}) = \ldots$

This result essentially enabled Moisil (1960) to identify problems 1.1 to 1.3 above in the case where the graph G is such that $v(x_i, x_i) = e$ for all nodes x_i in G with the more general problem of determining the (i, j)- entry of the matrix A^{n-1}, where A is the arc-value matrix of the graph G.

To see how reasonable this identification is, let us consider problem 1.1. It is clear that in this case, the additive group $(R^+, +)$ can be embedded in the Q-semiring $(R^+ \cup \{\infty\}, \wedge, +)$ where $\wedge$ is as defined in problem 1.2 above. Now if $v(x_i, x_i) = 0$ for all i is satisfied, theorem 1.2 yields

$$\min \left\{v(p) \mid p \in P_{ij}{}^{[s]}\right\} = \min \left\{v(p) \mid p \in P_{ij}{}^{[n-1]}\right\} \text{ for all } s \geq n-1$$

as interpreted from (1.5) above. Since $P_{ij} = \bigcup_{k=0}^{\infty} P_{ij}{}^{(k)}$,

$$\min \left\{v(p) \mid p \in P_{ij}\right\} = \min \left\{v(p) \mid p \in P_{ij}{}^{[n-1]}\right\}$$

Therefore, problem 1.1 is equivalent to the determination of the (i, j)- entry of the matrix $A^{[n-1]} = A^{n-1}$. Similarly, one can easily show that problems 1.2 and 1.3 are also equivalent to the determination of the (i,j)- entry of the matrix $A^{[n-1]} = A^{n-1}$, where A is the arc-value matrix of the corresponding graph.

In fact, the restriction that $v(x_i, x_i) = e$ for all i can be dropped if one identifies the above three problems with the problem of determining the (i, j)- entry of the matrix $A^{[n-1]}$ rather than A^{n-1}. This is because theorem 1.2 can be easily seen to be equivalent to the following

THEOREM 1.3 Let $(X, +, \circ)$ be a Q-semiring. Then for any matrix $A \in \mathcal{M}_n(X)$, we have

$$(1.6) \qquad I \prec A^{[1]} \prec A^{[2]} \prec \ldots \prec A^{[n-1]} = A^{[n]} = \ldots$$

It is somewhat surprising to note that under the same hypothesis as that of theorem 1.3, Pair (1967, p. 278) later obtained a weaker conclusion than (1.6) namely

$$I + A + A^2 + .. + A^s = I + A + A^2 + ...+ A^n \text{ for all } s \geq n ,$$

while Peteanu (1970, p.167) erroneously concluded that

$$A + A^2 + ... + A^n = A + A^2 + ... + A^{n-1}$$

To see that $A + A^2 + ... + A^n \neq A + A^2 + ... + A^{n-1}$ for some $(n \times n)$ matrix A over a Q-semiring, consider the (2×2) matrix $A = \begin{bmatrix} \theta & e \\ e & \theta \end{bmatrix}$ over the two-element Q-semiring $X = \{\theta, e\}$. For clearly,

$$A + A^2 = \begin{bmatrix} e & e \\ e & e \end{bmatrix} \neq A$$

However, from theorem 1.3, we always have $A + A^2 + ... + A^{n+1} = A + A^2 + ... + A^n$, a result which was also proved by Benzaken (1968).

From the above discussion, we may now tentatively define a path problem as follows.

DEFINITION 1.1 Let G be a graph over a Q-semiring and A be its arc-value matrix. Then by a <u>path problem</u>, we mean the determination of one or more entries of the matrix $A^{[n-1]}$.

From this definition, a path problem can therefore be solved by computing the matrix $A^{[n-1]}$. Now as a consequence of theorem 1.3 and the fact that a Q-semiring is necessarily an idempotent semiring, one can compute the matrix $A^{[n-1]}$ by recursively squaring the matrix $I + A$ until one obtains the matrix $(I + A)^k$ where k is the least positive integer not less than $n-1$. As a matter of fact, this method was widely in use until the following algorithm was discovered.

ALGORITHM 1.1

Step 1. Set $B^{\{0\}} = A$

Step 2. Compute $B^{\{k\}}$ recursively for $k = 1$ until n by

$$B_{ij}^{\{k\}} = B_{ij}^{\{k-1\}} + B_{ik}^{\{k-1\}} \circ B_{kj}^{\{k-1\}} \text{ , where } B_{ij}^{\{k\}} = (B^{\{k\}})_{ij}.$$

The above algorithm was first shown by Roy (1959) to compute the matrix A^{n-1} over the two-element Q-semiring $X = \{\theta, e\}$ whenever $A_{ii} = e$ for all i is also satisfied. But from the result of Warshall (1962), this algorithm actually computes the matrix $A + A^2 + \ldots + A^n$ without the assumption that $A_{ii} = e$ for all i. That Warshall's result can be extended to matrices over the semiring $(R^+ \cup \{\infty\}, \wedge, +)$ was first realized by Floyd (1962). On the other hand, Tomescu (1968) subsequently generalizes the result of Roy (1959) to matrices over a commutative semiring satisfying (1.1), i.e. the algebraic structure of Moisil (1960). However, as was noted by Benzaken (1968) and proved by Murchland (1965) as well as Robert and Ferland (1968), algorithm 1.1 is valid for computing the matrix $A + A^2 + \ldots + A^n$ over a Q-semiring. It will be seen later that this algorithm is in fact valid for computing the matrix $A + A^2 + \ldots + A^n$ over any semiring provided that A satisfies a certain condition. Moreover, it is somewhat interesting to note that algorithm 1.1 is in fact a particular form of a more general result obtained by McNaughton and Yamada (1960) in automata theory.

If we now replace step 1 of algorithm 1.1 by

Step 1' Set $B^{\{0\}} = I + A$,

then the resulting algorithm was shown by Pair (1967) to compute the matrix $A^{[n]}$ over any Q-semiring. But in view of theorem 1.3, this modified algorithm must in fact yield $A^{[n-1]}$. This modified algorithm is of little interest to us because to obtain $A^{[n-1]}$, we can simply use algorithm 1.1 to compute $A + A^2 + \ldots + A^n$ and then add the unit matrix I to it which would then yield the required result. Moreover, algorithm 1.1 is of interest in its own right because in several practical problems, we often require the matrix $A + A^2 + \ldots + A^n$ rather than the matrix $A^{[n-1]}$. As a matter of fact, we could simultaneously consider the determination of one or more entries of the matrix $A + A^2 + \ldots + A^n$ and the matrix $A^{[n-1]}$ in our definition of a path problem . However, for simplicity of exposition, we shall omit this consideration throughout this chapter.

Let us now consider the following variant of problem 1.1

PROBLEM 1.4 Longest Path

Let G be a graph over the additive group $(R^+, +)$ of non-negative real numbers. For any two nodes x_i, x_j in G, determine

$$\max \{v(p) | p \in P_{ij}\} \text{ , if it exists.}$$

We note that $\max\{v(p) | p \in P_{ij}\}$ may not exist if the graph G over $(R^+, +)$ contains a closed path c such that $v(c) > 0$. So let us assume that the graph G satisfies the following condition.

(1.7) $v(c) = 0$ for every closed path c in G.

With this assumption, it can be verified that

$$\max\{v(p) \mid p \in P_{ij}^{[s]}\} = \max\{v(p) \mid p \in P_{ij}^{[n-1]}\} \text{ for all } s \geq n-1$$

Now since $P_{ij} = \bigcup_{k=0}^{\infty} P_{ij}^{(k)}$, it then follows that

$$\max\{v(p) \mid p \in P_{ij}\} = \max \{v(p) \mid p \in P_{ij}^{[n-1]}\}$$

Therefore, we can also identify problem 1.4 in the case where condition (1.7) is satisfied with the problem of determining the (i,j) -entry of the matrix $A^{[n-1]}$ over the semiring $(R^+, \vee, +)$ where the binary operation $\vee$ is defined on R^+ by $a \vee b = \max\{a, b\}$. However, even this modified case of problem 1.4 does not fit into our previous definition of a path problem because $(R^+, \vee, +)$ is not a Q-semiring. Therefore, definition 1.1 has to be modified if one also wants to consider problem 1.4 as a path problem. It is precisely for this reason that Peteanu (1970) obtained a generalization of theorem 1.2 which can be stated as follows.

THEOREM 1.4. Let $(X, +, o)$ be an idempotent semiring (see section 0.2). Then for any matrix $A \in \mathcal{M}_n(X)$ such that $A_{ii} = e = (A^{n-1})_{ii}$ for all i, we have

$$I \prec A \prec A^2 \prec \ldots \prec A^{n-1} = A^n = \ldots$$

Let us now interpret this theorem in terms of the graph $G(A)$. Theorem 1.4 effectively says that if $G(A)$ is such that

(1.8) $\sigma(P_{ii}^{(k)}) = e$ for all i, and for all $k \in \{0,1,\ldots,n-1\}$,

then (1.5) holds. This interpretation suggests that one may replace condition (1.8) by the following.

(1.9) $v(\omega) \prec e$ for every elementary closed path ω in G.

To see why this is so, let us first establish the following useful result, where for convenience, a graph G satisfying (1.9) will be said to be absorptive.

LEMMA 1.1 For any non-elementary path $p \in P_{ij}$ of an absorptive graph G over a semiring $(X, +, o)$, there is always an elementary path $\bar{p} \in P_{ij}$ such that $v(p) \prec v(\bar{p})$.

PROOF This result can be obtained by employing the contraction process discussed in section 0.4 above. For using this process, the non-elementary path p can be factorized as follows

$$p = p_1 \, \omega_1 \, q_1, \quad p_1 \, q_1 = p_2 \, \omega_2 \, q_2, \ldots, \quad p_{s-1} q_{s-1} = p_s \, \omega_s \, q_s ,$$

where $p_s \, q_s$ is a contraction of p.

Consequently, $v(p) \prec v(p_1 q_1) \prec v(p_2 q_2) \prec \ldots \prec v(p_s q_s) = v(\bar{p})$, where $\bar{p} = p_s q_s$ is the required elementary path. Note that we have here used property (0.3) of the pseudo-ordering $\prec$ of the semiring. ∇

With the aid of lemma 1.1, we now show that (1.8) is implied by (1.9). For let $p \in P_{ii}^{(k)} \setminus \theta_i$. If p is elementary, then $v(p) \prec e$ by assumption. If p is non-elementary, then by lemma 1.1, there exists an elementary path $\bar{p} \in P_{ii}$ such that $v(p) \prec v(\bar{p})$. But $v(\bar{p}) \prec e$ because $\bar{p}$ is an elementary closed path and hence $v(p) \prec e$. Therefore, $v(p) \prec e$ for all $p \in P_{ii}^{(k)} \setminus \theta_i$, and hence by property (0.5) of $\prec$, we have $\sigma(P_{ii}^{(k)} \setminus \theta_i) \prec e$. But then

$$\sigma(P_{ii}^{(k)}) = \sigma(P_{ii}^{(k)} \setminus \theta_i) + \sigma(\theta_i)$$

$$= \sigma(P_{ii}^{(k)} \setminus \theta_i) + e \text{ , since } \sigma(\theta_i) = v(\theta_i) = e$$

$$= e \quad \text{as required.}$$

On the other hand, (1.9) is implied by (1.8) because it follows from the idempotency of addition that $v(\omega) \prec v(\omega)$, and hence $v(\omega) \prec v(\omega) + \sigma(P_{ii}^{(k)} \setminus \omega) = \sigma(P_{ii}^{(k)}) = e$, where k is chosen so that $\omega \in P_{ii}^{(k)}$.

Therefore, the hypothesis of theorem 1.4 can be replaced by (1.9). In fact, using (1.9), one can obtain the following result which coincides with theorem 1.4 whenever the semiring is also idempotent. For convenience, let us call a matrix A <u>absorptive</u> iff its graph $G(A)$ is absorptive.

THEOREM 1.5 Let $(X, +, \circ)$ be any semiring. Then for any absorptive matrix $A \in \mathcal{M}_n(X)$, we have

$$A^{[n-1]} = A^{[n]}$$

The validity of this theorem will be proved below. But first, let us obtain the following,

LEMMA 1.2 Let G be an absorptive graph over any semiring $(X, +, o)$. Then the following condition holds.

(1.10) For any finite subset B such that $E_{ij} \subseteq B \subseteq P_{ij}$, we have $\sigma(B) = \sigma(E_{ij})$, where E_{ij} is the set of all elementary paths in P_{ij} of G.

PROOF First, we show that $v(p) \prec \sigma(E_{ij})$ for all $p \in B \setminus E_{ij}$. By lemma 1.1, we have $v(p) \prec v(\bar{p})$ for some $\bar{p} \in E_{ij}$ whenever $p \in B \setminus E_{ij}$. Therefore, it follows from properties (0.2) and (0.4) of $\prec$ that $v(p) \prec v(\bar{p}) + \sigma(E_{ij} \setminus \bar{p}) = \sigma(E_{ij})$ as required.

Consequently, it follows from property (0.5) of $\prec$ that $\sigma(B \setminus E_{ij}) \prec \sigma(E_{ij})$ for all i,j, and hence

$$\sigma(B) = \sigma(B \setminus E_{ij}) + \sigma(E_{ij}) = \sigma(E_{ij}) \text{ for all } i, j. \qquad \nabla$$

We note here that Brucker (1974, p.34) erroneously concluded that condition (1.10) implies $A^{[n]} = A^{[n-1]}$, where A is the arc-value matrix of G. The error in his argument lies in the fact that every elementary path in a graph with n nodes has at most n arcs (and not $(n-1)$ arcs as claimed by Brucker (1974)). However, it is true that every elementary <u>open</u> path in a graph with n nodes has at most $(n-1)$ arcs, and hence $E_{ij} \subseteq P_{ij}^{[n-1]}$ for $i \neq j$. Therefore, $\sigma(P_{ij}^{[n-1]}) = \sigma(E_{ij})$ for $i \neq j$ follows from lemma 1.2. That $\sigma(P_{ii}^{[n-1]}) = \sigma(E_{ii})$ also holds cannot be deduced from lemma 1.2 because in general $E_{ii} \not\subseteq P_{ii}^{[n-1]}$, but this can be deduced from the fact that since G is absorptive,

$$\sigma(P_{ii}^{[n-1]}) = e = \sigma(E_{ii}).$$

Since lemma 1.2 can always be used to obtain $\sigma(P_{ij}^{[n]}) = \sigma(E_{ij})$ for all i, j, the validity of theorem 1.5 is thereby established. Moreover, we have just shown that

(1.11) $A^* = A^{[n-1]}$, where $(A^*)_{ij} = \sigma(E_{ij})$ for all i, j.

We note here that (1.11) and theorem 1.5 in its present form are due to a result originally obtained by Carré (1971) for matrices over a semiring which also satisfies additional assumptions. However, his proof of this result can easily be rewritten without the use of additional assumptions. This point was also noted by Shier (1973) and essentially by Gondran (1975). It is also of interest to note that Iri (1962) had earlier shown that (1.11) is always true whenever lemma 1.1 and the idempotency of addition are assumed valid.

Although of no interest to the study of path problems in general, the question concerning the validity of the converse of theorem 1.5 is a sound mathematical question. It turns out that this question has a negative answer since we have found a counter-example to this converse even when the idempotency of addition is assumed. This example will be given later in this chapter where it is more appropriate (see page 46).

Let us note in passing that in stating condition (1.10), Brucker (1974) failed to stipulate the finiteness of B. Such a stipulation is necessary because when B is an infinite set, $\sum_{p \varepsilon B} v(p)$ may not be a well-defined element of the semiring.

The above discussion suggests that we may now modify our previous definition of a path problem as follows.

DEFINITION 1.2 Let G be an absorptive graph over a semiring (X, +, o) and A be its arc-value matrix. Then by a <u>path problem</u> we mean the determination of one or more entries of the matrix $A^{[n-1]}$.

Since any graph G over a Q-semiring is obviously absorptive, it follows that this definition is more general than definition 1.1 above. Moreover, this definition allows us to include the following as a path problem. For convenience, a graph G which has no non-null closed paths will be said to be <u>acyclic</u>. Clearly, an acyclic graph is also absorptive.

PROBLEM 1.5 Parts Requirement (Vazonyi (1954))

Let G be an acyclic graph over the multiplicative monoid $(N, \cdot)$ of non-negative integers. For any two nodes x_i, x_j in G, determine

$$\sum_{p \varepsilon P_{ij}} v(p) ,$$

which is just the ordinary sum of a finite number of v(p) for each $p \in P_{ij}$. Note that P_{ij} is finite because G is assumed to be acyclic.

From definition 1.2, we see that a path problem is again solved by computing the matrix $A^{[n-1]}$, where A is an absorptive matrix over any semiring. For the case where the semiring is also idempotent, it is evident that the previous method for computing $A^{[n-1]}$ by recursively squaring the matrix $I + A$ remains valid.

However, better methods are available in this case.

It was effectively demonstrated by Carré (1971) that in solving a path problem as given by definition 1.2, an analogy with the classical methods of solving ordinary linear equations can be fruitfully exploited, since the matrix $A^{[n-1]}$ can be viewed as a solution of the matrix equation $Y = A \circ Y + I$. To this end, he developed several methods which are analogous to both the elimination and iterative techniques of linear algebra (see e.g. Fox (1964)), for solving the matrix equation $Y = A \circ Y + B$. Some of these methods were also seen by him to correspond to already well known algorithms for solving problem 1.1. Moreover, algorithm 1.1 was noted by him to correspond to his Jordan method for solving the matrix equation $Y = A \circ Y + A$. (This was also shown subsequently in Backhouse and Carré (1975)). Since the idempotency of addition played a major role in the work of Carré (1971) as well as Backhouse and Carré (1975), it is therefore of interest to find out the extent to which these variants of linear algebraic methods can be applied without the assumption of idempotency of addition. An attempt along this line was made by Gondran (1975) where he claimed in particular that the Gauss and Jordan methods as developed by Carré (1971) remain valid without the idempotency assumption. He in fact gave a proof of the Jordan method in substantiating his claim. Unfortunately, his proof cannot be taken as valid for a reason to be given in section 5.2 below. Thus the task remains for us to justify his claim. This justification would yield in particular the validity of algorithm 1.1 for computing $A + A^2 + \ldots + A^n$, where A is an absorptive matrix over a semiring, a result which was obtained by Roy (1975). It is interesting to note here that Brucker (1974) also showed that algorithm 1.1 with step 1 replaced by step 1' above is in fact valid for computing the matrix $A^{[n-1]}$ whenever the graph $G(A)$ satisfies (1.10) above.

Let us now consider the following natural generalization of problem 1.1.

PROBLEM 1.6 k Shortest Paths

Let G be a graph over the additive group $(R^+, +)$ of non-negative real numbers and k a positive integer. For any two nodes x_i, x_j in G, determine

$$k\text{-}\min\{v(p) \mid p \in P_{ij}\} ,$$

which is just the set consisting of the first, the second, ..., t^{th} smallest elements of the set $\{v(p) \mid p \in P_{ij}\}$, where t is the largest positive integer such that $t \leq k$.

Note that when $k = 1$, problem 1.6 coincides with problem 1.1. The method used in solving previous problems in the setting of a graph over a semiring were also extended to this problem by several authors which include Pair (1967), Giffler (1968), Derniame and Pair (1971), Minieka and Shier (1973), Shier (1974), Gondran (1975) and Roy (1975). However, unlike previous problems, a relevant semiring for solving this problem, which we shall from now on refer to as a <u>k shortest path algebra</u>, is not immediately apparent. In consequence, several analogous proposals for a k shortest path algebra were made by these authors. The following k shortest path algebra $(\mathcal{V}_{k\text{-min}}, \oplus, \odot)$ is inspired by their work.

Let $\mathcal{V}$ denote the set of all well-ordered subsets of R^+ including ϕ, and define the function k-min : $\mathcal{V} \to \mathcal{V}$ by

$$(1.12) \qquad \text{k-min}(A) = \begin{cases} \phi & \text{if } A = \phi \\ \{a_1, a_2, \ldots, a_t\}, & \text{otherwise} \end{cases}$$

Here $a_1 < a_2 < \ldots < a_t$ are t successively smallest elements of A and t is the largest index such that $t \leqslant k$.

Now set $\mathcal{V}_{k\text{-min}} = \{A \in \mathcal{V} \mid \text{k-min}(A) = A\}$ and define two binary operations $\oplus$ and $\odot$ on $\mathcal{V}_{k\text{-min}}$ by

$$A \oplus B = \text{k-min}\ (A \cup B), \text{ and}$$

$$A \odot B = \text{k-min}\ (AB), \text{ where } AB = \{a+b \mid a\varepsilon A,\ b\varepsilon B\}$$

It can be verified that $(\mathcal{V}_{k\text{-min}}, \oplus, \odot)$ forms an idempotent semiring. Now since $\text{k-min}\{x\} = \{x\}$ holds for all $x \in R^+$, one can identify R^+ with the subset of $\mathcal{V}_{k\text{-min}}$ which consists of only singleton subsets of R^+. This identification then allows us to view problem 1.6 above in terms of a graph over $(\mathcal{V}_{k\text{-min}}, \oplus, \otimes)$. Also it can be verified (cf. Shier (1974)) that

$$(1.13) \qquad \text{k-min}\left\{v(p) \mid p \in P_{ij}^{[s]}\right\} = \text{k-min}\left\{v(p) \mid p \in P_{ij}^{[nk-1]}\right\} \text{ for } s \geq nk-1.$$

Since $P_{ij} = \bigcup_{k=0}^{\infty} P_{ij}^{(k)}$, it follows that

$$\text{k-min}\{v(p) \mid p \in P_{ij}\} = \text{k-min}\left\{v(p) \mid p \in P_{ij}^{[nk-1]}\right\} .$$

Consequently, problem 1.6 is equivalent to the determination of $\text{k-min}\left\{v(p) \mid p \in P_{ij}^{[nk-1]}\right\}$ which is just the (i, j)-entry of the matrix $A^{[nk-1]} = I \oplus A \oplus \dots \oplus A^{nk-1}$, where A is the arc-value matrix of the graph G over $(\mathcal{V}_{k\text{-min}}, \oplus, \otimes)$.

Similarly, we may consider the following generalization of problem 1.4.

PROBLEM 1.7 k Longest Paths

Let G be a graph over the additive group $(R^{+}, +)$ of non-negative real numbers. For any two nodes x_i, x_j in G, determine

$$\text{k-max}\{v(p) \mid p \in P_{ij}\} \text{ , if it exists.}$$

Note that k-max(A) for any subset A of the set $\mathcal{V}'$ of all dually well ordered subsets of R^{+} and the empty set ϕ can be defined dually from (1.12), and that $\text{k-max}\{v(p) \mid p \in P_{ij}\}$

may not exist for the same reason as in problem 1.4. Therefore, condition (1.7) must be assumed to guarantee its existence. Moreover, with this assumption, one can verify similarly that

$$\text{k-max}\left\{v(p) \mid p \in P_{ij}^{[s]}\right\} = \text{k-max}\left\{v(p) \mid p \in P_{ij}^{[nk-1]}\right\} \text{ for all } s \geq nk-1,$$

and hence

$$\text{k-max}\{v(p) \mid p \in P_{ij}\} = \text{k-max}\left\{v(p) \mid p \in P_{ij}^{[nk-1]}\right\} .$$

Therefore, one can also identify problem 1.7 with the determination of $\text{k-max}\left\{v(p) \mid p \in P_{ij}^{[nk-1]}\right\}$ or the (i, j)-entry of the matrix $A^{[nk-1]}$, where A is the arc-value matrix of the graph G over the <u>k longest path algebra</u> $(\mathcal{V}'_{\text{k-max}}, \oplus, \odot)$ which is constructed from the function $\text{k-max} : \mathcal{V}' \to \mathcal{V}'$ in a dual manner from the derivation of the k shortest path algebra above.

We may now rephrase our definition of a path problem so as to include problems 1.6 and 1.7 as follows, where for convenience, a matrix A over a semiring is said to be <u>n_0-stable</u> iff there exists a non-negative integer n_0 such that $A^{[n_0+1]} = A^{[n_0]}$. Clearly, A is n_0-stable iff $A^{[s]} = A^{[n_0]}$ for all $s \geq n_0$.

DEFINITION 1.3 Let G be a graph over a semiring $(X, +, o)$ and A its arc-value matrix which is also n_0-stable. Then by a <u>path problem</u> we mean the determination of one or more entries of the matrix $A^{[n_0]}$.

By this definition, a path problem is solved by computing the matrix $A^{[n_0]}$. Again, for the case where the semiring is also idempotent, this matrix can be obtained by resursively squaring the matrix $I+A$ until we obtain $(I + A)^k$ where k is the least positive integer not less than n_0. However, since A is n_0-stable, $A^{[n_0]}$ obviously satisfies the matrix equation $Y = A \circ Y + I$ and hence, it is more fruitful to follow Carré (1971) and also more recently Carré (1976) by considering methods of solving the matrix equation $Y = A \circ Y + B$ in this case (see section 5.3 below).

Now observe that definition 1.3 is more general than definition 1.2 because by theorem 1.5, an absorptive matrix A is always $(n-1)$-stable. At this point, it interesting to ask which other matrices are n_0-stable for some positive integer n_0. For matrices over commutative semirings, Gondran (1975) found an answer in Theorem 1.6 below. For convenience, a graph G over a semiring $(X, +, \circ)$ will be said to be <u>q-regular</u> iff it satisfies

(1.14) There exists a positive integer q such that

$$v(\omega)^q \prec e + v(\omega) + v(\omega)^2 + \ldots + v(\omega)^{q-1}$$

for every elementary closed path ω in G, where $v(\omega)^q$ denotes the q-th power of $v(\omega)$.

Also for convenience, a matrix $A \in \mathcal{M}_n(X)$ will, be said to be <u>q-regular</u> iff $G(A)$ is q-regular.

THEOREM 1.6. Any matrix A over a commutative semiring $(X, +, \circ)$ is n_0-stable if it is q-regular for some positive integer q. Moreover, if G(A) has t elementary non-null closed paths, then

$$n_0 = nt(q-1) + (n-1) .$$

In fact, Gondran (1975) did not obtain an explicit value for n_0 but noted that it corresponds to the maximum order of paths in G(A) which do not traverse any elementary non-null closed paths more than $(q-1)$ times. We shall prove theorem 1.6 via the following

LEMMA 1.3 Let G be a <u>q-regular</u> graph over a <u>commutative</u> semiring $(X, +, \circ)$, $\Omega^{(q)}$ the set of all paths in G which do not traverse any elementary non-null closed path in G more that $(q-1)$ times, and Q be any set of paths in G which also contains all the subpaths of any path in Q . Then for any $p \in Q$ but $p \notin \Omega^{(q)}$, there exists $H \subseteq Q \cap \Omega^{(q)}$ such that $v(p) \prec \sigma(H)$.

PROOF Let $p \in Q$ but $p \notin \Omega^{(q)}$. Then by assumption, p must traverse at lease one elementary non-null closed path more than $(q-1)$ times. Let us suppose that p traverses exactly k elementary non-null closed paths for more than $(q-1)$ times each, say s_1 times for ω_1, s_2 for ω_2 and so on. Now by the commutativity of $\circ$, we may use the contraction process to obtain

$$v(p) = v(\omega_1)^{s_1} \circ v(\omega_2)^{s_2} \circ \dots \circ v(\omega_k)^{s_k} \circ v(p_s q_s) ,$$

where $p_s q_s$ is taken without loss of generality to be an elementary open path of Q. Let us write

$$y_i = e + v(\omega_i) + v(\omega_i)^2 + \ldots + v(\omega_i)^{q-1}$$

for all $i \in \{1, 2, \ldots, k\}$. Then we claim that

$$v(\omega_i)^{s_i} \prec y_i \quad \text{for all} \quad i \in \{1, 2, \ldots, k\} .$$

To justify this claim it suffices to show that condition (1.14) above always implies

$$v(\omega)^m \prec e + v(\omega) + v(\omega)^2 + \ldots + v(\omega)^{q-1}$$

for all $m \geq q$. This result can be easily shown by mathematical induction on t, the detail of which will be omitted here since a more general result will be proved later in lemma 1.4 below.

Therefore, granting that this claim is justified, it follows from property (0.3) of $\prec$ that

$$v(p) \prec y_1 \circ y_2 \circ \ldots \circ y_k \circ v(p_s q_s) .$$

But by definition, $y_i = \sigma(H_i)$ for all $i \in \{1, 2, \ldots, k\}$, where

$$H_i = \left\{\theta_{j_i}, \omega_i, \omega_i^2, \ldots, \omega_i^{q-1}\right\} \text{ for all } i \in \{1, 2, \ldots, k\} ,$$

and hence

$$v(p) \prec \sigma(H_1) \circ \sigma(H_2) \circ \ldots \circ \sigma(H_k) \circ \sigma(p_s q_s) = \sigma(H) ,$$

where $H = H_1 H_2 \ldots H_k \cdot \{p_s q_s\}$.

We now claim that $H \subseteq Q \cap \Omega^{(q)}$.

For let $x \in H$, then $x = \omega_1^{t_1} \omega_2^{t_2} \ldots \omega_k^{t_k} p_s q_s$ where $t_i \leq q-1$ for all $i \in \{1, 2, \ldots, k\}$ can be considered as a path in

Q which traverses all the elementary closed paths $\omega_1, \omega_2, \ldots, \omega_k$ and the open path $p_s q_s$ in any manner which defines a path. Note that there must exist at least one possibility since p itself was assumed to traverse $\omega_1, \omega_2, \ldots, \omega_k$ and $p_s q_s$ in the first place. Now x is also in $\Omega^{(q)}$ because it does not traverse any elementary closed path ω_i more than $q-1$ times. Hence $x \in Q \cap \Omega^{(q)}$, and therefore $H \subseteq Q \cap \Omega^{(q)}$ as claimed, and the lemma is proved. ∇

The proof of theorem 1.6 can be seen to follow from lemma 1.1 by observing that

$$\text{(1.15)} \qquad \Omega_{ij}^{(q)} \subseteq P_{ij}^{[s]} \subseteq P_{ij} \quad \text{for all} \quad s \geq n_0 = nt(q-1) + (n-1),$$

where $\Omega_{ij}^{(q)} = P_{ij} \cap \Omega^{(q)}$ and t denotes the number of all elementary non-null closed paths in G. This observation follows because the maximum order of a path which does not traverse any elementary non-null closed path more than $(q-1)$ times corresponds to that of the path which traverses exactly $t(q-1)$ elementary non-null closed paths plus one open path which amounts to $nt(q-1) + (n-1) = n_0$.

Now let $p \in P_{ij}^{[s]}$ for any $s \geq n_0$, but $p \notin \Omega_{ij}^{(q)}$. Then $p \notin \Omega^{(q)}$ because otherwise,

$$p \in P_{ij}^{[s]} \cap \Omega^{(q)} = P_{ij}^{[s]} \cap P_{ij} \cap \Omega^{(q)} = P_{ij}^{[s]} \cap \Omega_{ij}^{(q)} = \Omega_{ij}^{(q)},$$

a contradiction.

Therefore, by lemma 3.1, we have $v(p) \prec \sigma(H)$ for some $H \subseteq \Omega_{ij}^{(q)}$, and hence

$$v(p) \prec \sigma(H) + \sigma(\Omega_{ij}^{(q)} \setminus H) = \sigma(\Omega_{ij}^{(q)})$$

for all $p \in P_{ij}^{[s]} \setminus \Omega_{ij}^{(q)}$ follows from properties (0.2) and (0.4) of $\prec$.

Consequently, it follows from property (0.5) of $\prec$ that

$$\sigma(P_{ij}^{[s]} \setminus \Omega_{ij}^{(q)}) \prec \sigma(\Omega_{ij}^{(q)}), \text{ and hence}$$

$\sigma(P_{ij}^{[s]}) = \sigma(P_{ij}^{[s]} \setminus \Omega_{ij}^{(q)}) + \sigma(\Omega_{ij}^{(q)}) = \sigma(\Omega_{ij}^{(q)})$ which proves theorem 1.6 above.

In problem 1.6, the k shortest path algebra $(\mathcal{V}_{k\text{-min}}, \oplus, \odot)$ can be seen to be a commutative semiring and condition (1.14) is easily seen to be satisfied by $q = k$. Hence theorem 1.6 enables us to infer that the arc-value matrix A in problem 1.6 is n_0-stable where $n_0 = nt(k-1) + (n-1)$. However, from (1.13), we see that A is in fact $(nk-1)$-stable. It is therefore of interest to note that theorem 1.7 below, which is due essentially to Roy (1975), yields (1.13) directly when applied to problem 1.6. For convenience, a graph G over a semiring $(X, +, \circ)$ will be said to be <u>q-absorptive</u> iff it satisfies

$$(1.16) \quad v(\omega_1) \circ v(\omega_2) \circ \ldots \circ v(\omega_q) \prec e + v(\omega_1) + \ldots + v(\omega_1) \circ v(\omega_2) \circ \ldots \circ v(\omega_{q-1})$$

for every q-tuple $(\omega_1, \omega_2, \ldots, \omega_q)$ of elementary closed paths in G.

Also for convenience, a matrix $A \in \mathcal{M}_n(X)$ will be said to be <u>q-absorptive</u> iff $G(A)$ is q-absorptive

THEOREM 1.7 Any matrix A over a commutative semiring is $(nq-1)$-stable if it is q-absorptive.

We note here that the original result obtained by Roy (1975) was that A is nq-stable. That A is in fact $(nq-1)$-stable can be proved via the following

LEMMA 1.4 Let G be a q-absorptive graph over a commutative semiring $(X, +, o)$, $\Omega^{[q]}$ be the set of all paths in G which do not traverse more than $(q-1)$ elementary non-null closed path in G, and Q be any set of paths in G which also contains all the subpaths of any path in Q. Then for any $p \in Q$ but $p \notin \Omega^{[q]}$, there exists a subset $H \subseteq Q \cap \Omega^{[q]}$ such that $v(p) \prec \sigma(H)$.

PROOF First, let us show that condition (1.16) implies that

$$v(\omega_1) \circ v(\omega_2) \circ \ldots \circ v(\omega_s) \prec e + v(\omega_1) + \ldots +$$

$$+ v(\omega_1) \circ v(\omega_2) \circ \ldots \circ v(\omega_{q-1})$$

for every s-tuple $(\omega_1, \omega_2, \ldots, \omega_s)$ and for all $s \geq q$.

We show this by mathematical induction on s. For $s = q$, the result is obviously true. So we may assume its validity for $q \leq s < t$ as our induction hypothesis and show that it is also valid for $s = t$.

By this induction hypothesis, we have

$$v(\omega_2) \circ v(\omega_3) \circ \ldots \circ v(\omega_t) \prec e + v(\omega_2) + \ldots + v(\omega_2) \circ v(\omega_3) \circ \ldots \circ v(\omega_q),$$

and hence it follows from properties (0.2) and (0.4) of $\prec$ that

$$v(\omega_1)\circ v(\omega_2)\circ\ldots\circ v(\omega_t) \prec v(\omega_1)+v(\omega_1)\circ v(\omega_2)+\ldots+v(\omega_1)\circ v(\omega_2)\circ\ldots\circ v(\omega_q)$$

$$\prec e+v(\omega_1)+\ldots+v(\omega_1)\circ v(\omega_2)\circ\ldots\circ v(\omega_q)$$

$$= e+v(\omega_1)+\ldots+ v(\omega_1) \ \ldots \ v(\omega_{q-1}) \ .$$

which yields the required result.

Now let $p \in Q$ but $p \notin \Omega^{[q]}$. Then by definition, p must traverse at least q elementary non-null closed paths. Thus by the contraction process, we may write

$$p = p_1\omega_1 q_1, \ p_1q_1 = p_2\omega_2q_2, \ldots, p_{s-1}q_{s-1} = p_sq_s, \quad \text{where}$$

$s \geq q$ and p_sq_s is a contraction of p. Since $\circ$ is commutative, we then have $v(p) = v(\omega_1)\circ v(\omega_2)\circ\ldots\circ v(\omega_s)\circ v(p_sq_s)$. It then follows from the above result and property (0.3) of $\prec$ that
$v(p) \prec (e+v(\omega_1)+\ldots+v(\omega_1)\circ v(\omega_2)\circ\ldots\circ v(\omega_{q-1}))\circ v(p_sq_s)$

If p_sq_s is not a closed path, then $v(p) \prec \sigma(H_1)$ where

$$H_1 = \{\theta_i, \ \omega_1, \ \omega_1 \ \omega_2, \ldots, \omega_1 \ \omega_2\ldots \ \omega_{q-1}\}\circ\{p_sq_s\}$$

$$= \{p_sq_s, \omega_1p_sq_s, \ \omega_1\omega_2p_sq_s, \ldots, \omega_1\omega_2\ldots\omega_{q-1}p_sq_s\}$$

If otherwise, then $p_sq_s = \omega$ is an elementary non-null closed path, and hence

$$v(p) \prec v(\omega)+v(\omega)\circ v(\omega_1)+\ldots+ v(\omega)\circ v(\omega_1)\circ\ldots\circ v(\omega_{q-1})$$

It then follows from properties (0.2) and (0.4) of $\prec$ that

$$v(p) \prec e + v(\omega) + v(\omega)\circ v(\omega_1) + \ldots + v(\omega)\circ v(\omega_1)\circ\ldots\circ v(\omega_{q-1})$$
$$= e + v(\omega) + \ldots + v(\omega)\circ v(\omega_1)\circ\ldots\circ v(\omega_{q-2})$$
$$= \sigma(H_2) \ ,$$

where $H_2 = \{\theta_i, \omega, \omega\omega_1, \ldots, \omega\omega_1 \ldots \omega_{q-2}\}$.

Since H_1 and H_2 can both be shown to be a subset of $Q \cap \Omega^{[q]}$ by an argument similar to that used at the end of the proof of lemma 1.3 above, it follows that in both cases, $v(p) \prec \sigma(H)$ for some $H \subseteq Q \cap \Omega^{[q]}$ as required. ∇

The proof of theorem 1.7 now follows from lemma 1.4 by observing that

$$(1.17) \qquad \Omega_{ij}^{[q]} \subseteq P_{ij}^{[s]} \subseteq P_{ij} \quad \text{for all} \quad s \geq nq-1,$$

where $\Omega_{ij}^{[q]} = P_{ij} \cap \Omega^{[q]}$. This observation follows because the maximum order of a path which does not traverse more than $(q-1)$ elementary non-null closed paths corresponds to that of the path which traverses exactly $(q-1)$ elementary non-null closed paths plus one elementary open path which amounts to $n(q-1)+(n-1) = nq-1$. The rest of the proof follows from an argument similar to that used at the end of the proof of theorem 1.6, and hence its detail will be omitted.

Let us now consider the following

PROBLEM 1.8 Simple Paths

Let Σ be a finite set of letters, also known as an alphabet. By a word over Σ , we mean a finite sequence of letters written one after another in a definite order. A word will be said to be simple iff all its letters are distinct. A word without any letters will be called the empty word and is denoted by λ. The operation which combines two words into one is known as concatenation, and is denoted by juxtaposition. It is well known that the set Σ^* of all words over Σ (including λ) forms a monoid with respect to concatenation, also known as the free monoid generated by Σ. With these preliminaries, we can now state our present problem as follows.

Let G be a graph over an alphabet Σ such that each arc in G is assigned a distinct letter of Σ. We may then consider G as a graph over the monoid Σ^*. For any two nodes x_i, x_j in G, determine

$$\text{sim}\ \{v(p) \mid p \in P_{ij}\}\ ,$$

which denotes the set of all simple words $v(p)$ for each $p \in P_{ij}$.

This problem can be formulated and solved as a path problem in the sense of definition 1.3 as follows

Let $\mathcal{V} = 2^{\Sigma^*}$ and $\mathcal{V}_{sim} = \{A \in \mathcal{V} \mid \text{sim}(A) = A\}$, where sim(A) denotes the set of all simple words in A. Define two binary operations $\oplus$ and $\odot$ on $\mathcal{V}_{sim}$ by

$$A \oplus B = \text{sim}\ (A \cup B),\ \text{ and}$$

$$A \odot B = \text{sim}\ (AB),\ \text{where } AB = \{ab \mid a \in A,\ b \in B\}$$

It can then be verified that $(\mathcal{V}_{sim}, \oplus, \odot)$, called the simple path algebra, forms an idempotent semiring.

Now since $\text{sim}\{x\} = \{x\}$ for all $x \in \Sigma$ implies that one can identify the letters of Σ with the singleton subsets of $\mathcal{V}_{sim}$, the graph G in problem 1.8 can be considered as a graph over $(\mathcal{V}_{sim}, \oplus, \odot)$. Moreover, if m is the total number of arcs in G, it can be verified that

$$(1.18) \qquad \text{sim}\,\{v(p)\,|\,p \in P_{ij}^{[s]}\} = \text{sim}\,\{v(p)\,|\,p \in P_{ij}^{[m]}\} \text{ for all } s \geq m.$$

Since $P_{ij} = \bigcup_{k=0}^{\infty} P_{ij}^{(k)}$, it follows that

$$\text{sim}\{v(p)\,|\,p \in P_{ij}\} = \text{sim}\,\{v(p)\,|\,p \in P_{ij}^{[m]}\} .$$

Therefore, problem 1.8 is equivalent to the determination of $\text{sim}\{v(p)\,|\,p \in P_{ij}^{[m]}\}$ which is just the (i, j)-entry of the matrix $A^{[m]}$, where A is the arc-value matrix of the graph G over $(\mathcal{V}_{sim}, \oplus, \otimes)$.

Now we note that although the graph G in problem 1.8 satisfies (1.14) and (1.16) when considered as a graph over the simple path algebra, we cannot use theorems 1.6 or 1.7 to deduce the validity of (1.18) because the simple path algebra is not necessarily a commutative semiring. This example therefore gives us the motivation for obtaining other results analogous to theorems 1.6 or 1.7 but without the commutativity assumption. These results will be given later (see theorems 4.6 and 4.10 below).

Problem 1.8 also provides us with a counter example to the validity of the converse of theorem 1.5, which we discussed earlier (see page 30). For it suffices to take a graph G with $m = n-1$ arcs which also contains an elementary closed path say ω. Then clearly,

$$v(\omega) \oplus \{\lambda\} = \text{sim}\{v(\omega),\lambda\} = \{v(\omega),\lambda\} \neq \{\lambda\} ,$$

which shows that (1.9) does not hold here. On the other hand, (1.18) can be expressed as

$$A^{[s]} = A^{[n-1]} \qquad \text{for all } s \geq n-1 .$$

Let us now make an observation which will open the way to an alternative approach to the abstract study of path problems. This observation concerns the semirings in each of the above problems. We have seen that while the relevant semirings for solving problems 1.1 to 1.5 were self-evident, those for solving problems 1.6 to 1.8 were not. Nevertheless, the methods employed to construct these latter semirings, whether it be the k shortest or the k longest or the simple path algebras, are all carried out via a certain function, say r, which is defined on a certain set $\mathcal{V}$ of subsets of the monoid (X, o) which is also closed under union and complex product and contains {e}, ϕ . By means of r, one can then define

(1.19) $$\mathcal{V}_r = \{A \in \mathcal{V} \mid r(A) = A\} ,$$

and two binary operations $\oplus$ and $\otimes$ on $\mathcal{V}_r$ by

(1.20) $$A \oplus B = r(A \cup B), \text{ and}$$

(1.21) $$A \otimes B = r(AB) , \quad \text{where } AB = \{a \circ b \mid a \in A, b \in B\} .$$

The properties that the function r must possess are obviously those which will make $(\mathcal{V}_r, \oplus, \otimes)$ become a semiring. For this purpose, the following properties can be seen to suffice.

(1.22) $$r(\phi) = \phi$$

(1.23) $$r(A \cup B) = r(r(A) \cup B)$$

(1.24) $$r(AB) = r(r(A)B) = r(A\, r(B))$$

In fact, one can also apply the above method to construct the semiring for solving problems 1.1 by taking $r = \min$ and $\mathcal{V}$ as the set of all well ordered subsets of R^+ and ϕ. The semiring $(\mathcal{V}_{\min}, \oplus, \odot)$ so obtained can be easily seen to be isormorphic to the semiring $(R^+ \cup \{\infty\}, \wedge, +)$ which we obtained earlier for problem 1.1. The same remark applies also to problems 1.2 and 1.4, where r can again be taken as min or max, whichever is appropriate. The only exception is problem 1.5. We shall return to this point later. Meanwhile, let us note that all these problems except problem 1.5 can also be equivalently expressed in terms of the function r as the determination of

(1.25) $$r\{v(p) | p \in P_{ij}\} \text{ , if it exists.}$$

Note that in general, $r\{v(p)|p \in P_{ij}\}$ may not exist because the set $\{v(p)|p \in P_{ij}\}$ may not belong to $\mathcal{V}$, the domain of the function r. Examples of this situation have already been seen in problems 1.4 and 1.7 above. Moreover, it can be observed that condition (1.7) which was assumed in order to guarantee the existence of (1.25) in these two problems in fact imposes a restriction on the graph under consideration. Any graph which does not satisfy condition (1.7) is therefore in some sense not compatible with the domain of the function r and hence not compatible with the algebraic structure $(\mathcal{V}_r, \oplus, \odot)$ over which the graph is to be considered in these two problems. It turns out that this question of compatibility can be fruitfully analysed in its full generality if $\mathcal{V}$ has the following two properties.

(1.26) $\mathcal{V}$ contains all the finite subsets of the monoid $(X, \circ)$ including the empty set ϕ.

(1.27) If $A \varepsilon \mathcal{V}$, and $B \subseteq A$, then $B \varepsilon \mathcal{V}$ also.

The details of this analysis will be given later but in a slightly more general framework (see section 4.2 below). It suffices to note here that all the appropriate $\mathcal{V}$-sets of the above problems except problem 1.5 do possess properties (1.26) and (1.27) above.

In summary, all the above problems except problem 1.5 can all be described as a path problem in accordance with the following definition

DEFINITION 1.4 Let G be a graph over a monoid $(X, \circ)$, $\mathcal{V}$ a set of all the subsets of X which has properties (1.26), (1.27) and is also closed with respect to union and complex product, and r a function defined on $\mathcal{V}$ which satisfies (1.22) to (1.24). Then by a path problem we mean the determination of $r\{v(p) | p \varepsilon P_{ij}\}$ for one or more pairs (i, j), provided, of course, that they exist.

Let us now give another example of a path problem in accordance with the above definition, namely

PROBLEM 1.9 Elementary Paths

Let Σ^* be the set of all words (inlcuding λ) over an alphabet Σ (see problem 1.8 above). A word x of Σ^* is said to be an abbreviation of another word y of Σ^* iff x can be obtained from y by removing at least one (and possibly all) of the letters of y . (Note that every word with at least one letter has the abbreviation λ). For instance, the word "mary" is an abbreviation of the word "elementary". Now for any set A of words of Σ^* , let $b(A)$ be the set of all words which are obtained from A by deleting all those words which also have all their abbreviations in A. With these preliminaries, we can now state our present problem as follows.

Let G be a graph over an alphabet Σ such that each arc in G is assigned a distinct letter of Σ . We may then consider G as a graph over the monoid Σ^*.

For any two nodes x_i, x_j in G, determine

$$b\,\{v(p)\,|\,p \in P_{ij}\} \, .$$

We note that the above problem can also be considered as a path problem in accordance with definition 1.3 or, better still, definition 1.2, because the graph G can be easily verified to be absorptive when considered over the elementary path algegra $(\mathcal{V}_b, \oplus, \odot)$ which can be defined via (1.20) and (1.21) with $r = b$ and $\mathcal{V} = 2^{\Sigma^*}$. We note also that $(\mathcal{V}_b, \oplus, \odot)$ coincides with the free distributive pseudo-lattice of Benzaken (1968) which was also used by him to enumerate elementary paths (see also Murchland (1965)) in a graph.

In fact, any problem which satisfies definition 1.4 and has the property that

(1.28) $$r\left\{ v(p)\,|\,p \in P_{ij}^{[n_0+1]}\right\} = r\left\{v(p)\,|\,p \in P_{ij}^{[n_0]}\right\} \text{ for some } n_0 \in N$$

can easily be shown to satisfy definition 1.3 by considering the graph G of the corresponding problem to be over the semiring $(\mathcal{V}_r, \oplus, \ominus)$. The converse is also true in the case where the semiring $(X, +, \circ)$ is idempotent. For one can then define $\mathcal{V} = 2^X$ and r as follows.

$$r(A) = \begin{cases} \{\sup A\} & \text{if } \sup A \in X \text{ and } \sup A \neq \theta \\ \phi & , \text{ otherwise} \end{cases}$$

That $\mathcal{V}$ so defined has properties (1.26) and (1.27) is obvious and that r has properties (1.22) to (1.24) can be shown in a manner similar to the proof of theorem 3.1 to be given later.

Now since the arc-value matrix A of G is n_0-stable, it follows from (0.16) and (1.4) that

$$(1.29) \qquad \sum_{p \in P_{ij}^{[n_0+1]}} v(p) = \sum_{p \in P_{ij}^{[n_0]}} v(p)$$

But condition (1.29) is equivalent to (1.28) because $\sup A = \sum_{x \in A} x$ whenever $A \neq \phi$ is a finite set, and therefore, the converse is verified.

Let us now return to consider the difficulty which prevents problem 1.5 from being a path problem in accordance with definition 1.4. This difficulty in fact arises from the non-idempotency of addition in N. However, this shortcoming can be eliminated by using a concept of multisets, to be introduced in the next chapter. It is precisely the aim of this thesis to show exactly how this can be done and also to demonstrate its usefulness for solving all the above problems and many others (see chapter 5 below).

CHAPTER 2

MULTISETS

2.1 Complete Lattice of Multisets

A natural way of generalizing the intuitive notion of a set is to remove the restriction that all its elements are distinct. We then have the corresponding notion of a collection which may contain an identical element repeated a finite number of times. Such a collection is usually called a multiset (Knuth (1969)) whenever the number of repetitions is finite. However, here we shall take the term "multiset" to mean a collection of elements in which infinite repetitions of an element are allowed†, and call a multiset non-singular whenever none of its elements are repeated an infinite number of times. Thus formally,

DEFINITION 2.1 Let N_∞ be the complete semiring of non-negative integers as defined in section 0.2. A multiset A with elements from a given set X is a function $A:X \to N_\infty$. Each image $A(x)$ will be called the multiplicity of x in A, which is just the number of times x occurs in A. A multiset A is said to be empty, written $A = \phi$, iff $A(x) = 0$ for all $x \in X$; non-singular iff $A(x) \neq \infty$ for all $x \in X$.

From the above definition of a multiset, it follows that any set can simply be regarded as a multiset in which the multiplicity of each element is at most unity. The connection of sets with multisets is even more fundamental in that to each multiset A, there

---

† This corresponds to an $\mathcal{N}$-subset in Eilenberg(1974), where $\mathcal{N} = N_\infty$

is always a unique set $d(A) = \{x | A(x) \neq 0\}$, called the support of the multiset A. Note that $d(A)$ is just the set of all distinct elements of the multiset A. This concept of a support enables us to distinguish the following types of multisets which are especially useful to our present study.

DEFINITION 2.2.

(i) A multiset A is said to be quasi-finite iff its support is a finite set, i.e. $d(A)$ contains only a finite number of elements. A is then said to be finite iff it is also non-singular.

(ii) A multiset A is said to be quasi-countable iff its support is a countable set, i.e $d(A)$ is either a finite set or it can be put in one-to-one correspondence with the set of positive integers. A is then said to be countable iff it is also non-singular.

(iii) Let $(X, \leq)$ be an ordered set. A multiset A with elements from X is said to be well ordered (dually well ordered) iff its support is a well ordered (dually well ordered) set.

When there is no chance of confusion, we shall use capital letters to denote multisets and lower case letters to denote elements. By virtue of the fact that multisets are mere generalizations of sets, it is natural to make extensive use of set-theoretic notation whenever confusion is not possible. Thus for instance, we shall write $x \in A$ to indicate that x is an element of A, i.e. $A(x) \neq 0$. However, the notation $\{x | P(x)\}$ will be reserved exclusively for sets. For finite multisets it is often convenient to exhibit their elements as a list enclosing symbols between braces, but note that for multisets which contain only a single element, say $\{x\}$, we find it convenient to omit its braces whenever its meaning is clear from the context. Finally, a quasi-finite or quasi-countable multiset may also be written in extenso such as $\{1, \ldots, 1, 2, 3\}$ or $\{1, 2, 3, \ldots\}$, provided that the suppressed elements are obvious.

Now let N_∞^X denote the set of all multisets with elements taken from a set X. Then N_∞^X is easily seen to be ordered by the following relation to be called <u>multiset inclusion</u> and denoted by $\subseteq$ in analogy with set inclusion.

(2.1) $A \subseteq B$ iff $A(x) \leq B(x)$ for all $x \in X$,

where $\leq$ denotes the extension of the ordering "less than or equal to" to N_∞ by defining $\infty \leq \infty$ and $n < \infty$ for all $n \in N$.

The multiset A is then said to be a <u>submultiset</u> of B.

THEOREM 2.1 $(N_\infty^X, \subseteq)$ is a complete lattice

PROOF From (2.1), it follows easily that the least upper bound $\bigcup_{i\in I} A_i$ of an arbitrary collection $\{A_i | i \in I\}$ of multisets indexed by some set I is given by

(2.2) $$\left(\bigcup_{i\in I} A_i\right)(x) = \sup\{A_i(x) | i\in I\} = \begin{cases} \max\{A_i(x) | i\in I\}, & \text{if it exists} \\ \infty, & \text{otherwise} \end{cases}$$

Similarly, its greatest lower bound $\bigcap_{i\in I} A_i$ is given by

(2.3) $\left(\bigcap_{i\varepsilon I} A_i\right)(x) = \inf\{A_i(x) | i\varepsilon I\} = \min\{A_i(x) | i\ \varepsilon I\}$.

∇

Note that our use of the notations $\bigcup$ and $\bigcap$ in the proof of the above theorem is justified by the fact that when each A_i is a set, $\bigcup_{i\varepsilon I} A_i$ and $\bigcap_{i\varepsilon I} A_i$ coincide respectively with the set-theoretic union and intersection as defined in section 0.1.

2.2 Multisums, Multiproducts and Closures

Several interesting operations can be defined on multisets. But first let us note that since multisets are formally defined as functions, two multisets are considered equal iff they are equal as functions, i.e.

(2.4) $A = B$ iff $A(x) = B(x)$ for all $x \varepsilon X$.

DEFINITION 2.3 For any two multisets A, B of N_∞^X, the multiset $A \uplus B$ which is defined by

(2.5) $(A \uplus B)(x) = A(x) + B(x)$ for all $x \varepsilon X$

is called the <u>multisum</u> of A and B. Here $+$ denotes the extended addition defined on N_∞, see section 0.2.

Various properties of the multisum operation can be immediately derived from the corresponding properties of the extended addition on N_∞. Thus for instance, the multisum operation is commutative and associative because the extended addition has these properties; also $A \uplus \phi = A$ for all $A \varepsilon N_\infty^X$ because $A(x) + 0 = A(x)$ always.

More generally, one can define a multisum $\biguplus_{i\varepsilon I} A_i$ for an arbitrary collection $\{A_i | i \varepsilon I\}$ of multisets indexed by some set I as follows.

(2.6) $$(\biguplus_{i\varepsilon I} A_i)(x) = \sum_{i\varepsilon I} A_i(x) \quad \text{for all} \quad x \varepsilon X.$$

Note that the right-hand side is meaningful because N_∞ is a complete semiring (see section 0.2), and recall that $\sum_{i\varepsilon I} A_i(x)$ denotes the sum of all non-zero $A_i(x)$ if there are only finitely many such $A_i(x)$, otherwise $\sum_{i\varepsilon I} A_i(x) = \infty$.

We shall also write $\biguplus_{i\varepsilon I} A_i$ as $\biguplus_{i=1}^{\infty} A_i$ if $I = \{1, 2, \ldots\}$.

In fact, by virtue of theorem 2.1, we could also define $\biguplus_{i\varepsilon I} A_i$ as the least upperbound of all the multisums $\biguplus_{j\varepsilon J} A_j$, where J ranges over all the finite subsets of I, i.e.

(2.7) $$\biguplus_{i\varepsilon I} A_i = \bigcup_{J\varepsilon 2(I)} (\biguplus_{j\varepsilon J} A_j)$$

This definition is easily seen to be equivalent to (2.6).

Another multiset operation of interest is that induced by a binary operation on the set X. More precisely,

DEFINITION 2.4 Let o be a binary operation defined on a set X. For any two multisets A, B of N_∞^X, the multiset $A \circ B$ which is defined by

(2.8) $$(A \circ B)(x) = \sum_{x=y\circ z} A(y)B(z)$$

is called the <u>multiproduct</u> of A and B. Here juxtaposition of $A(y)$ with $B(z)$ denotes the extended multiplication defined on

N_∞ , see section 0.2.

Note that (2.8) is meaningful because N_∞^X is a complete semiring and that the right-hand side of (2.8) denotes the sum of all non-zero $A(y)B(z)$ which satisfies the equality $x = y \circ z$, provided that there are finitely many such $A(y)B(z)$, otherwise it is just ∞ . We note also that in view of (2.6) above, the multiproduct $A \circ B$ may be better understood if one writes

$$(2.9) \qquad A \circ B = \biguplus_{(x,y) \in A \times B} \{x \circ y\} ,$$

which is easily seen as a generalization of the complex product induced by the binary operation $\circ$, see section 0.2. Furthermore, our use of the same notation for multiproduct as for the binary operation which induces it does not lead to confusion since multisets are denoted by capital letters here.

Now as one might expect from the way we define the multiproduct operation, various properties of this operation do not depend only on the extended addition and multiplication on N_∞ but also on the corresponding properties of the binary operation which induces it. Thus for instance, it is easily seen that the multiproduct operation is commutative and associative if the binary operation which induces it also has these properties; also $\{e\}$ is the identity for the multiproduct operation whenever e is the identity for the binary operation. But there is one property of the multiproduct operation which is independent of the binary operation which induces it, namely, the multiproduct operation is always distributive over the multisum operation. Actually, a more general distributive law holds which can be seen in the following theorem, where all the above properties of the multisum and multiproduct operations are conveniently summarized.

THEOREM 2.2. Let (X, o) be a monoid. Then $(N_\infty^X, \uplus, o)$ is a semiring with unit $\{e\}$ and zero ϕ. Moreover, N_∞^X is a complete semiring if $\biguplus_{i\varepsilon I} A_i$ as defined by (2.6) is taken as a formal sum in N_∞^X.

PROOF Since N_∞ is a semiring (see section 0.2), it follows from the above discussion that $(N_\infty^X, \uplus, o)$ is a semiring also. Now the multisum $\biguplus_{i\varepsilon I} A_i$ as defined by (2.6) can be seen to possess properties (0.6) to (0.8) of a formal sum in a complete semiring because N_∞ is itself a complete semiring, and hence N_∞^X is a complete semiring as claimed. We shall not verify properties (0.6) to (0.7) for $\biguplus_{i\varepsilon I} A_i$ here since their validity is easily seen from the corresponding properties of the formal sum in N_∞. Property (0.8) can be verified as follows. For all $x \varepsilon X$, we have

$$\left(B o \biguplus_{i\varepsilon I} A_i\right)(x) = \sum_{x=yoz} \left[B(y)\left(\biguplus_{i\varepsilon I} A_i\right)(z)\right]$$

$$= \sum_{x=yoz} \left[B(y) \sum_{i\varepsilon I} A_i(z)\right]$$

$$= \sum_{x=yoz} \sum_{i\varepsilon I} B(y)A_i(z) \quad \text{by (0.8) for } N_\infty$$

$$= \sum_{i\varepsilon I} \sum_{x=yoz} B(y)A_i(z) \quad \text{by (0.7) for } N_\infty$$

$$= \sum_{i\varepsilon I} (BoA_i)(x)$$

$$= \left(\biguplus_{i\varepsilon I} (BoA_i)\right)(x)$$

Therefore, by (2.4), $B \circ \biguplus_{i\in I} A_i = \biguplus_{i\in I} (B \circ A_i)$.

Similarly, we can show that $\left(\biguplus_{i\in I} A_i\right) \circ B = \biguplus_{i\in I} (A_i \circ B)$. ∇

In this thesis, we shall have occasion to consider sequences of multisets. When we do, the following result will be useful. But first note that a sequence $A_1, A_2, \ldots.$ of multisets will be simply written as (A_i) and is said to be non-decreasing iff $A_i \subseteq A_{i+1}$, and non-increasing iff $A_{i+1} \subseteq A_i$ for all $i \in \{1,2,\ldots\}$.

THEOREM 2.3 (i) For any two non-decreasing sequences (A_i) and (B_i) of multisets of N_∞^X , we have

$$\bigcup_{i=1}^{\infty} (A_i \uplus B_i) = \left(\bigcup_{i=1}^{\infty} A_i\right) \uplus \left(\bigcup_{i=1}^{\infty} B_i\right)$$

(ii) For any two non-increasing sequences (A_i) and (B_i) of multisets of N_∞^X , we have

$$\bigcap_{i=1}^{\infty} (A_i \uplus B_i) = \left(\bigcap_{i=1}^{\infty} A_i\right) \uplus \left(\bigcap_{i=1}^{\infty} B_i\right)$$

PROOF For all $x \in X$, we have

$$\left(\bigcup_{i=1}^{\infty} (A_i \uplus B_i)\right)(x) = \sup_i \{A_i(x) + B_i(x)\}$$

$$= \sup_m \{\sup_n \{A_m(x)+B_n(x)\}\},$$ since both $(A_m(x))$ and $(B_n(x))$ are non-decreasing

$$= \sup_m \{A_m(x) + \sup_n\{B_n(x)\}\}$$

$$= \sup_m \{A_m(x)\} + \sup_n\{B_n(x)\}$$

$$= \left(\bigcup_{i=1}^{\infty} A_i\right)(x) + \left(\bigcup_{i=1}^{\infty} B_i\right)(x)$$

Therefore, (i) follows from (2.4), and (ii) can be proved in a similar manner. ∇

THEOREM 2.4 (i) If (A_i) is a non-decreasing sequence of multisets of N_∞^X, then for all $B \in N_\infty^X$, we have

$$B \circ \bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} (B \circ A_i) \quad \text{and} \quad \left(\bigcup_{i=1}^{\infty} A_i\right) \circ B = \bigcup_{i=1}^{\infty} (A_i \circ B)$$

(ii) If (A_i) is a non-increasing sequence of multisets of N_∞^X, then for all $B \in N_\infty^X$, we have

$$B \circ \bigcap_{i=1}^{\infty} A_i = \bigcap_{i=1}^{\infty} (B \circ A_i) \quad \text{and} \quad \left(\bigcap_{i=1}^{\infty} A_i\right) \circ B = \bigcap_{i=1}^{\infty} (A_i \circ B)$$

PROOF (i) If $B(y)\sup\{A_i(z)\} \neq 0$ for only finitely many pairs (y, z) such that $x = y \circ z$, then $\sum_{x=y\circ z} B(y) \sup\{A_i(z)\}$ is a finite sum, and hence by an argument similar to the proof of (i) in the above theorem, we have

$$\sum_{x=y\circ z} B(y) \sup_i\{A_i(z)\} = \sup_i \sum_{x=y\circ z} B(y)A_i(z)$$

But then, $\left(B \circ \bigcup_{i=1}^{\infty} A_i\right)(x) = \left(\bigcup_{i=1}^{\infty} (B \circ A_i)\right)(x)$ as required.

So we may suppose that $B(y) \sup_i \{A_i(z)\} \neq 0$ for infinitely many pairs (y,z) such that $x = y \circ z$, which means that $\left(B \circ \bigcup_{i=1}^{\infty} A_i\right)(x) = \infty$.

Now suppose that $\left(\bigcup_{i=1}^{\infty} (B \circ A_i)\right)(x) \neq \infty$. But then (2.2) implies that $\max\{B \circ A_i(x)\} \neq \infty$ exists, say $(B \circ A_{i_0})(x)$. Since $(B \circ A_{i_0})(x) \neq \infty$, it follows that $\sum_{x=yoz} B(y)\, A_{i_0}(z)$ is a finite sum, and that $B(y)\, A_{i_0}(z) \neq \infty$ for any (y, z) such that $x = y \circ z$.

Consequently, $A_{i_0}(z) = \max_i\{A_i(z)\}$ also, and hence

$$\left(B \circ \bigcup_{i=1}^{\infty} A_i\right)(x) = \sum_{x=yoz} B(y) \sup_i\{A_i(z)\}$$

$$= \sum_{x=yoz} B(y)\, A_{i_0}(z)$$

$$= (B \circ A_{i_0})(x)$$

which implies that $\left(B \circ \bigcup_{i=1}^{\infty} A_i\right)(x) \neq \infty$, a contradiction.

(ii) Just as in (i) above, we may suppose that $B(y) \inf_i\{A_i(z)\} \neq 0$ for infinitely many pairs (y, z) such that $x = y \circ z$, i.e. $\left(B \circ \bigcap_{i=1}^{\infty} A_i\right)(x) = \infty$. Now since

$$B(y) \inf_i\{A_i(z)\} \leq B(y)\, A_i(z) \quad \text{for all } i,$$

it follows that for all i, we have $B(y)\, A_i(z) \neq 0$ for infinitely many such pairs (y, z), i.e. $(B \circ A_i)(x) = \infty$. Consequently,

$$\left(\bigcap_{i=1}^{\infty} (B \circ A_i) \right)(x) = \inf_i \{B \circ A_i(x)\} = \infty \text{ as required.} \qquad \nabla$$

A useful fact which we shall need later in this section is that both the multisum and multiproduct operations are compatible with the multiset inclusion, as defined by (2.1). This is expressed by the following

THEOREM 2.5 For any $A, B \in N_{\infty}^{X}$ such that $A \subseteq B$, we always have

$$A \uplus C \subseteq B \uplus C \ , \quad A \circ C \subseteq B \circ C \quad \text{and} \quad C \circ A \subseteq C \circ B$$

for all $C \in N_{\infty}^{X}$.

PROOF Trivial ∇

DEFINITION 2.5 For any $A \in N_{\infty}^{X}$, where $(X, \circ)$ is a monoid, the multisets

$$A^* = \biguplus_{k=0}^{\infty} A^k \quad \text{and} \quad A^+ = \biguplus_{k=1}^{\infty} A^k$$

are respectively called the closure and weak closure of A.

From the name closure, one might expect the usual properties of closure to hold, namely

(i) $A \subseteq A^*$

(ii) $A^* \subseteq B^*$ whenever $A \subseteq B$

(iii) $(A^*)^* = A^*$.

In fact, (i) and (ii) are valid but (iii) is not always true. For instance, $\phi^* = \{e\}$, but $(\phi^*)^* = \{e\}^* = \{e, e, \ldots\}$, and therefore $(\phi^*)^* \neq \phi^*$.

Therefore, A^* is not the closure of A in the conventional sense. Nevertheless, we call A^* the closure of A for want of a more appropriate name. The same remarks also apply to the weak closure. As an example where $(A^+)^+ \neq A^+$, consider the multiset $A = \{1\}$ of N_∞^N where N is the additive monoid of non-negative integers.

The following identities for closures are useful but obvious.

$$(2.10) \qquad A^* = A \circ A^* \uplus \{e\} = A^* \circ A \uplus \{e\} \quad \text{for all } A \in N_\infty^X .$$

THEOREM 2.6 For any two given multisets $A, B \in N_\infty^X$, $A^* \circ B$ and $B \circ A^*$ are respectively the least solution of

$$Y = A \circ Y \uplus B \quad \text{and} \quad Y = Y \circ A \uplus B$$

with respect to multiset inclusion.

PROOF It is easily seen from (2.10) that $A^* \circ B$ is a solution of $Y = A \circ Y \uplus B$. On the other hand, $Y = A \circ Y \uplus B$ always implies that for any $k \in N$,

$$Y = A^{k+1} \circ Y \uplus A^{[k]} \circ B \text{ , where } A^{[k]} = \{e\} \uplus A \uplus \dots \uplus A^k .$$

Therefore, $A^{[k]} \circ B \subseteq Y$ for all $k \in N$,

and hence $\bigcup_{k=0}^{\infty} (A^{[k]} \circ B) \subseteq Y$. But from (2.7) and (i) of theorem 2.4, it follows that $A^* \circ B = \bigcup_{k=0}^{\infty} (A^{[k]} \circ B)$, and hence the least solution of $Y = A \circ Y \uplus B$. Similarly, $B \circ A^*$ can be seen to be the least solution of $Y = Y \circ A \uplus B$. ∇

The above theorem is very useful for establishing identities involving closures. As an illustration of this usefulness, let us show how to derive the following identities which we shall need later.

THEOREM 2.7 For any $A, B \in N_\infty^X$, we have

$$(A \uplus B)^* = A^* \circ (B \circ A^*)^* = (A^* \circ B)^* \circ A^*$$

PROOF From theorem 2.6, $(A \uplus B)^*$ is the least solution of $Y = (A \uplus B) \circ Y \uplus \{e\}$. On the other hand, we shall show that $A^* \circ (B \circ A^*)^*$ is also the least solution and hence establish $(A \uplus B)^* = A^* \circ (B \circ A^*)^*$. That $A^* \circ (B \circ A^*)^*$ is a solution can be seen as follows.

$$(A \uplus B) \circ A^* \circ (B \circ A^*)^* \uplus \{e\}$$

$$= A \circ A^* \circ (B \circ A^*)^* \uplus B \circ A^* \circ (B \circ A^*)^* \uplus \{e\}$$

$$= A \circ A^* \circ (B \circ A^*)^* \uplus (B \circ A^*)^*, \text{ by } (2.10)$$

$$= (A \circ A^* \uplus \{e\}) \circ (B \circ A^*)^*$$

$$= A^* \circ (B \circ A^*)^*, \text{ by } (2.10)$$

To see that it is also the least solution, let us rewrite the above equation as $Y = A \circ Y \uplus B \circ Y \uplus \{e\}$. It then follows from theorem 2.6 that

$A^* \circ (B \circ Y \uplus \{e\}) \subseteq Y$, and hence

$B \circ A^* \circ B \circ Y \uplus B \circ A^* \subseteq B \circ Y$, by theorem 2.5

Again, using theorem 2.6, we obtain

$$(B \circ A^*)^* \circ B \circ A^* \subseteq B \circ Y .$$

But then

$$\begin{aligned} A \circ Y \uplus (B \circ A^*)^* &= A \circ Y \uplus (B \circ A^*)^* B \circ A^* \uplus \{e\} \\ &\subseteq A \circ Y \uplus B \circ Y \uplus \{e\} \text{ , by theorem 2.5} \\ &= Y \end{aligned}$$

and hence $A^* \circ (B \circ A^*)^* \subseteq Y$ as required.

Similarly, we can show that $(A \uplus B)^* = (A^* \circ B)^* \circ A^*$. ∇

2.3 Hereditary Semirings and Their Closed Multisets.

DEFINITION 2.6 A non-empty subset $\mathcal{V}$ of N_∞^X is said to be hereditary iff whenever $A \in \mathcal{V}$ and $B \subseteq A$, then $B \in \mathcal{V}$.

A typical example of a hereditary subset of N_∞^X is N_∞^X itself, but so are the following subsets of N_∞^X .

THEOREM 2.8 The following are hereditary subsets of N_∞^X

(i) The set $\mathcal{N}_X$ of all non-singular multisets of N_∞^X

(ii) The set $\mathcal{P}_X$ of all quasi-finite multisets of N_∞^X and ϕ.

(iii) The set $\mathcal{Q}_X$ of all quasi-countable multisets of N_∞^X and ϕ.

(iv) The set $\mathcal{W}_X$ $(\mathcal{W}_X')$ of all well ordered (dually well ordered) multisets of N_∞^X and ϕ, where X is an ordered set.

PROOF (i) Let $A \in \mathcal{N}_X$ and $B \subseteq A$. Then $B(x) \leq A(x)$ implies that $B(x) \neq \infty$ whenever $A(x) \neq \infty$. Hence $B \in \mathcal{N}_X$

(ii) Let $A \in \mathcal{P}_X$ and $B \subseteq A$. Then $B(x) \leq A(x)$ implies that $B(x) = 0$ whenever $A(x) = 0$. Consequently, $d(B) \subseteq d(A)$, and hence $d(B)$ is a finite set, i.e. $B \in \mathcal{P}_X$ as required.

Both (iii) and (iv) can also be proved in the same fashion as (ii) above. ∇

THEOREM 2.9 The intersection of an arbitrary collection of hereditary subsets of N_∞^X is again a hereditary subset of N_∞^X .

PROOF Let $\mathcal{V} = \bigcap \{\mathcal{U} \mid \mathcal{U}$ is a hereditary subset of $N_\infty^X\}$. Suppose that $A \in \mathcal{V}$ and $B \subseteq A$. Then by definition of intersection, $A \in \mathcal{U}$ for every $\mathcal{U}$. Since each $\mathcal{U}$ is a hereditary subset, we have $B \in \mathcal{U}$ for every $\mathcal{U}$ also. Hence $B \in \mathcal{V}$ as required. ∇

The above theorem enables us to construct more hereditary subsets of N_∞^X from those already given in theorem 2.8 above. Some such hereditary subsets of N_∞^X which are of special interest later are

(i) The set $\mathcal{F}_X = \mathcal{P}_X \cap \mathcal{N}_X$ of all finite multisets of N_∞^X and ϕ.

(ii) The set $\mathcal{U}_X = \mathcal{C}_X \cap \mathcal{N}_X \cap \mathcal{W}_X$ (or $\mathcal{U}'_X = \mathcal{C}_X \cap \mathcal{N}_X \cap \mathcal{W}'_X$)

of all countable and well ordered (or dually well ordered) multisets of N_∞^X and ϕ , where X is an ordered set.

DEFINITION 2.7 Let (X, o) be a monoid. A hereditary subset $\mathscr{V}$ of N_∞^X is called a hereditary semiring iff $\mathscr{V}$ contains all the finite multisets of N_∞^X including ϕ and is closed with respect to the multisum and multiproduct operation.

Note that by the above definition, a hereditary semiring is in fact a subsemiring of N_∞^X with respect to the multisum and multiproduct operations. Moreover, every hereditary semiring is also a lattice with respect to multiset inclusion, because by the hereditary property of $\mathscr{V}$, $A \cap B \subseteq A \cup B \subseteq A \uplus B$ yields $A \cap B$, $A \cup B \in \mathscr{V}$ whenever $A \uplus B \in \mathscr{V}$.

THEOREM 2.10 The following are hereditary semirings

(i) The set N_∞^X of all multisets with elements in a monoid (X, o)

(ii) The set $\mathscr{P}_X$ of all quasi-finite multisets of N_∞^X and ϕ, where (X, o) is a monoid.

(iii) The set $\mathscr{C}_X$ of all quasi-countable multisets of N_∞^X and ϕ, where (X, o) is a monoid.

(iv) The set $\mathscr{W}_X$ ($\mathscr{W}_X'$) of all well ordered (dually well ordered) multisets of N_∞^X and ϕ, where $(X, \leq, o)$ is a totally ordered monoid.

(v) The set $\mathscr{N}_X$ of all non-singular multisets of N_∞^X, where (X, o) is a locally finite monoid.

(vi) The set $\mathscr{F}_X$ of all finite multisets of N_∞^X and ϕ, where (X, o) is a monoid.

(viii) The set $\mathcal{U}_X(\mathcal{U}'_X)$ of all countable and well ordered (dually well ordered) multisets of N_∞^X and ϕ, where $(X, \leq, o)$ is a totally ordered group.

PROOF (i) is trivial.

(ii) Since $\mathcal{P}_X$ obviously contains all the finite multisets of N_∞^X, it remains to show that it is closed with respect to the multisum and multiproduct operations. To do this let us note the following two properties of supports, namely

(2.11) $$d(A \uplus B) = d(A) \cup d(B) \quad \text{and} \quad d(A \circ B) = d(A)\, d(B),$$

where juxtaposition denotes complex product.

Since (2.11) is easy to verify, we shall omit its proof here. Now the required result follows from (2.11) because the union and complex product of two finite sets are themselves finite.

(iii) follows from an argument similar to (ii) by using the fact that the union and complex product of two countable sets are themselves countable.

(iv) also follows from an argument similar to (ii) if we can show that the union and complex product of two well ordered (dually well ordered) sets are themselves well ordered (dually well ordered). So let A, B be two well ordered sets. If Y is any subset of $A \cup B$, then we can write $Y = Y_A \cup Y_B$, where Y_A and Y_B can be chosen so that $Y_A \subseteq A$ and $Y_B \subseteq B$. But then

$$\min Y = \min \{\min Y_A, \min Y_B\}$$

exists because X is a totally ordered set. Therefore $A \cup B$ is a well ordered set.

Now let $Y \subsetneq AB$. Then by the definition of complex product $Y = Y_A Y_B$ for some $Y_A \subseteq A$, $Y_B \subseteq B$. Since the binary operation is also compatible with the ordering on X,

$$\min Y = \min Y_A \circ \min Y_B$$

exists, and hence AB is also a well ordered set. The case for dually well ordered sets can be demonstrated in a dually fashion.

(v) Since $\mathcal{N}_X$ obviously contains all the finite multisets, it remains to show that it is also closed with respect to the multisum and multiproduct operations. $\mathcal{N}_X$ is closed with respect to multisum because $A(x) + B(x) \neq \infty$ whenever $A(x) \neq \infty$, $B(x) \neq \infty$. Now recall that $(X, \circ)$ is a locally finite monoid means that for each fixed $x \in X$, there is only a finite number of factorizations $x = y \circ z$ with $y \neq e$, $z \neq e$. Consequently, $A(y)\,B(z) \neq 0$ for only a finite number of pairs (y, z) such that $x = y \circ z$, and hence $\sum_{x=y\circ z} A(y)\,B(z)$ is a finite sum. Since $A(y)\,B(z) \neq \infty$ because $A(y) \neq \infty$ and $B(z) \neq \infty$, we must then have

$$(A \circ B)(x) = \sum_{x=y\circ z} A(y)\,B(z) \neq \infty \text{ , as required.}$$

(vi) Since $\mathcal{F}_X = \mathcal{P}_X \cap \mathcal{N}_X$ and $\mathcal{P}_X$ and $\mathcal{N}_X$ were both seen to be closed with respect to the multisum operation in (ii) and (v), so is $\mathcal{F}_X$. Now while $\mathcal{P}_X$ was seen in (ii) to be closed with respect to the multiproduct operation, $\mathcal{N}_X$ may not be so without additional assumptions. However $\mathcal{P}_X \cap \mathcal{N}_X$ is so because for any $A, B \in \mathcal{P}_X$, $A(x) \neq 0$ and $B(x) \neq 0$ for only finitely many $x \in X$ implies that there can be only finitely many pairs (y, z) with $A(y) \neq 0$ and $B(z) \neq 0$ such that $x = y \circ z$. Consequently, $A(y)\, B(z) \neq 0$ for only finitely many such pairs, and hence $\sum_{x=y\circ z} A(y)\, B(z)$ is a finite sum. Since $A(y)\, B(z) \neq \infty$ whenever $A(y) \neq \infty$, $B(z) \neq \infty$, it then follows that

$$(A \circ B)(x) = \sum_{x=y\circ z} A(y)\, B(z) \neq \infty \text{ whenever } A, B \in \mathcal{N}_X \text{ also.}$$

(vii) By the above arguments, it is obvious that $\mathcal{U}_X = \mathcal{Q}_X \cap \mathcal{N}_X \cap \mathcal{W}_X$ contains all the finite multisets of N_∞^X and is also closed with respect to the multisum operation. It therefore remains to show that $\mathcal{U}_X$ is also closed with respect to the multiproduct operation. Again, while $\mathcal{Q}_X$ and $\mathcal{W}_X$ were both seen to be closed with respect to the multiproduct operation in (iii) and (iv), $\mathcal{N}_X$ may not be so. However, we shall show that $\mathcal{Q}_X \cap \mathcal{N}_X \cap \mathcal{W}_X$ is necessarily so. We shall establish this claim by examining the following two cases, where $A, B \in \mathcal{U}_X = \mathcal{Q}_X \cap \mathcal{N}_X \cap \mathcal{W}_X$

(a) Suppose both $A, B \in \mathcal{W}'_X$. Then $d(A)$ and $d(B)$ are both dually well ordered sets. But $d(A)$ and $d(B)$ are both countable and well ordered by assumptions. Hence $d(A)$ and $d(B)$ must be finite sets, i.e. $A, B \in \mathcal{P}_X$. But $A, B \in \mathcal{N}_X$ also, and therefore $A, B \in \mathcal{F}_X = \mathcal{P}_X \cap \mathcal{N}_X$. Consequently, $A \circ B \in \mathcal{F}_X$ follows from (vi) above. Since $\mathcal{F}_X \subseteq \mathcal{N}_X$, we have $A \circ B \in \mathcal{N}_X$ as required.

(b) Suppose $A \notin \mathcal{W}'_X$, i.e. $d(A)$ is not a dually well ordered set. Then we claim that for each fixed $x \in X$, $A(y)B(z) \neq 0$ for only finitely many pairs (y, z) such that $x = y \circ z$. For suppose otherwise, then $A, B \in \mathcal{C}_X$ implies that $A(y)\, B(z) \neq 0$ for only countably many such pairs (y,z), say (y_k, z_k) for all $k \in \{1,2,\dots\}$. Now since $d(A)$ is not a dually well ordered set, the sequence (y_k) must contain a strictly increasing subsequence, say

$$y_{k_1} < y_{k_2} < \dots < y_{k_i} < \dots$$

Since $(X, \circ)$ is a totally ordered group, it follows that

$$x \circ y_{k_1}^{-1} > x \circ y_{k_2}^{-1} > \dots > x \circ y_{k_i}^{-1} > \dots,$$

i.e. $z_{k_1} > z_{k_2} > \dots > z_{k_i} > \dots$ because $x = y_k \circ z_k$.

This means that $d(B)$ can not be well ordered, i.e. $B \notin \mathcal{W}_X$, a contradiction. The rest follows from the end argument of (v) or (vi) above and the case for $\mathcal{U}'_X = \mathcal{C}_X \cap \mathcal{N}_X \cap \mathcal{W}'_X$ can also be shown dually. ∇

Let us note in passing that the intersection of an arbitrary collection of hereditary semirings is again a hereditary semiring. This fact implies that $\mathcal{F}_X$ is the least hereditary semiring with respect to set inclusion.

DEFINITION 2.8 Let $\mathcal{V}$ be a subset of N_∞^X, where $(X, \circ)$ is a monoid. A multiset $A \in \mathcal{V}$ is said to be <u>closed</u> in $\mathcal{V}$ iff $A^* \in \mathcal{V}$, where A^* denotes the closure of A.

From this definition, we can establish the following useful results.

LEMMA 2.1 A multiset A of a hereditary semiring $\mathscr{V}$ is closed in $\mathscr{V}$ iff $A^+ \in \mathscr{V}$.

PROOF If $A^+ \in \mathscr{V}$, then $A^* = \{e\} \uplus A^+ \in \mathscr{V}$. If $A^* \in \mathscr{V}$, then $A^+ \subseteq A^*$ implies that $A^+ \in \mathscr{V}$, by the hereditary property of $\mathscr{V}$. ∇

LEMMA 2.2 A multiset A of a hereditary subset $\mathscr{V}$ of N_∞^X is closed in $\mathscr{V}$ iff for any $B \subseteq A$, $B^* \in \mathscr{V}$ also.

PROOF Sufficiency is obvious, while necessity follows directly from the fact that $B^* \subseteq A^*$ whenever $B \subseteq A$, and the hereditary property of $\mathscr{V}$. ∇

THEOREM 2.11

(i) Every multiset is closed in N_∞^X.

(ii) ϕ is the only multiset closed in $\mathscr{F}_X$.

(iii) Every multiset is closed in $\mathscr{Q}_X$.

PROOF (i) is trivial.

(ii) Since $\phi^* = \{e\}$, it follows that ϕ is closed in $\mathscr{F}_X$. Now suppose that $A \neq \phi$ is closed in $\mathscr{F}_X$. Then by assumption, A contains at least one element, say x, and by lemma 2.2 above, $\{x\}^* \in \mathscr{F}_X$ because $\{x\} \subseteq A$. But $\{x\}^* = \{e, x, x^2, \ldots\} \notin \mathscr{N}_X$ if $\{x\}^* \in \mathscr{P}_X$ and hence

$\{x\}^* \notin \mathscr{F}_X = \mathscr{P}_X \cap \mathscr{N}_X$, a contradiction.

(iii) Let us first note that

$$d\left(\biguplus_{i \in I} A_i\right) = \bigcup_{i \in I} d(A_i) . \tag{2.12}$$

For let $x \in d\left(\biguplus_{i\in I} A_i\right)$, then $\sum_{i\in I} A_i(x) \neq 0$ implies $A_i(x) \neq 0$ for at least one $i \in I$, i.e. $x \in d(A_i)$ for at least one $i \in I$, and hence $d\left(\biguplus_{i\in I} A_i\right) \subseteq \bigcup_{i\in I} d(A_i)$.
But $\bigcup_{i\in I} d(A_i) \subseteq d\left(\biguplus_{i\in I} A_i\right)$ follows from the fact that $d(A_i) \subseteq d\left(\biguplus_{i\in I} A_i\right)$ for all $i \in I$. Therefore (2.12) is verified. Thus $d(A^*) = \bigcup_{k=0}^{\infty} d(A^k) = \bigcup_{k=0}^{\infty} d(A)^k$, by (2.11).
Now if $A \in \mathcal{C}_X$ i.e. $d(A)$ is countable, then so is $d(A)^k$ for all $k \in N$, because the complex product of a finite number of countable sets is itself countable. Since a countable union of countable sets is also a countable set, $d(A^*)$ is therefore a countable set, and hence $A^* \in \mathcal{C}_X$ as required. ∇

We note here that the above theorem characterizes completely the nature of closed multisets of N_∞^X, $\mathcal{F}_X$ and $\mathcal{C}_X$. For other hereditary subsets, the nature of their closed multisets are much more difficult to characterize. However, for $\mathcal{N}_X$, $\mathcal{W}_X$ $(\mathcal{W}'_X)$ and $\mathcal{N}_X \cap \mathcal{W}_X$ $(\mathcal{N}_X \cap \mathcal{W}'_X)$, we have the following

THEOREM 2.12 A necessary condition for a multiset A to be closed in $\mathcal{N}_X$ is that $x \neq e$ for every $x \in A$. This condition is also sufficient if (X, o) is a locally finite monoid.

PROOF Suppose that $x = e$ for some $x \in A$. Then clearly, $\{x\}^* = \{e, x, x^2, \ldots\} = \{e, e, \ldots\} \notin \mathcal{N}_X$. But by lemma 2.2, $\{x\}^* \in \mathcal{N}_X$ because $\{x\} \subseteq A$, a contradiction.

Now if (X, o) is a locally finite monoid, then each $x \in X$ admits only a finite number of factorizations $x = x_1 \circ x_2 \circ \ldots \circ x_n$ with $x_i \neq e$ for all $i \in \{1, 2, \ldots, n\}$. So the largest index for such a factorization exists, say n_0. Now by definition of A^k, $x \in A^k$ iff $x = x_1 \circ x_2 \circ \ldots \circ x_k$ for some $x_1, x_2, \ldots, x_k \in A$. Consequently, $A^k(x) = 0$ for all $k > n_0$, and therefore $A^*(x) = \biguplus_{k=0}^{n_0} A^k(x) \neq \infty$, if $A(x) \neq \infty$, which proves the sufficiency. ∇

THEOREM 2.13 If a multiset A is closed in $\mathcal{W}_X$ ($\mathcal{W}'_X$), where $(X, \leqslant, o)$ is a totally ordered monoid such that $x^2 = x$ always implies $x = e$, then we have $x \geqslant e$ ($x \leqslant e$) for every $x \in A$. On the other hand, for a multiset A to be closed in $\mathcal{W}_X$ ($\mathcal{W}'_X$), where $(X, \leqslant, o)$ is an Archimedean totally ordered monoid, it is sufficient that $x \geqslant e$ ($x \leqslant e$) for every $x \in A$.

PROOF Suppose that A is a closed multiset of $\mathcal{W}_X$ and that $x < e$ for some $x \in A$. Then $x^2 \leqslant x < e$. But by assumption, $x^2 = x$ implies $x = e$, and hence $x^2 < x$. Similarly, $x^4 < x^2$ and so on. Therefore, $\{x\}^* = \{e, x, x^2, x^4, \ldots\} \notin \mathcal{W}_X$. But by lemma 2.2, $\{x\} \subsetneq A$ implies $\{x\}^* \in \mathcal{W}_X$, a contradiction. Therefore, $x \geqslant e$ for all $x \in A$ which proves the first part of the theorem.

To prove the second part, let us first verify the following two special cases.

(i) If $x = e$ for every $x \in A$ of $\mathcal{W}_X$, then $A^* \in \mathcal{W}_X$

(ii) If $x > e$ for every $x \in A$ of $\mathcal{W}_X$, then $A^* \in \mathcal{W}_X$.

Since (i) is obvious, we prove only (ii). In view of (2.11) and (2.12), this is equivalent to showing that $\bigcup_{k=0}^{\infty} d(A)^k$ is a well ordered set.

Now suppose otherwise and let B be a non-empty subset of $\bigcup_{k=0}^{\infty} d(A)^k$. Then B must contain a sequence (b_k) of elements in B satisfying $b_1 > b_2 > \ldots > b_k \ldots$ Observe that $b_i \geq e$ for all $i \in \{1, 2, \ldots\}$. So if $b_1 = e$, then

$$b_1 = b_2 = \ldots , \text{ a contradiction.}$$

Therefore, we may suppose that $b_1 \neq e$, i.e. $b_1 > e$ and let $b_0 = \min d(A)$. Since $b_0 > e$ by assumption, it then follows from the Archimedean property that $b_0^n > b_1$ for some positive integer n. Consequently, $b_0^n > b_1 > b_2 > \ldots > b_k > \ldots$.

Now suppose that $b_k \in d(A)^s$ for some $s > n$. Then $b_k = x \circ y$ for some $x \in d(A)^n$, $y \in d(A)^{s-n}$. Since $y \geq e$, $b_k = x \circ y \geq x \geq \min d(A)^n = b_0^n$, a contradiction. So for all $k \in \{1, 2, \ldots\}$, $b_k \in \bigcup_{k=0}^{n} d(A)^k$. Since $\bigcup_{k=0}^{n} d(A)^k$ is a well ordered set whenever $d(A)$ is, it follows that the above sequence (b_k) must terminate, i.e.

$$b_k = b_{k+1} = \ldots, \text{ for some } k,$$

which then yields a contradiction.

Now for the general case, let us write $A = B \uplus C$, where B and C are such that $x = e$ for every $x \in B$ and $x > e$ for every $x \in C$. By theorem 2.7, we know that $A^* = (B \uplus C)^* = B^* \circ (C \circ B^*)^*$. Since B has the property of (i) and $(C \circ B^*)$ has the property of (ii), it follows from the above argument that $B^* \varepsilon \mathcal{W}_X$, $(C \circ B^*)^* \varepsilon \mathcal{W}_X$, and hence $A^* \varepsilon \mathcal{W}_X$ as required.

The case for $\mathcal{W}'_X$ can be established in a dual fashion. ∇

COROLLARY 2.1 In the case where $(X, \leqslant, \circ)$ is a totally ordered cancellative monoid which is conditionally complete with respect to $\leqslant$, the condition that $x \geqslant e (x \leqslant e)$ for every $x \varepsilon A$ is both necessary and sufficient for the multiset A to be closed in $\mathcal{W}_X$ $(\mathcal{W}'_X)$.

PROOF By cancellativity, it follows that $x = e$ whenever $x^2 = x$, and hence the necessity of the condition follows from the above theorem. Now its sufficiency also follows from the above theorem if $(X, \leqslant, \circ)$ can be shown to be Archimedean. Suppose not, i.e. $x > e$, $y > e$, but $x^n \leqslant y$ for all $n \varepsilon N$. Then the set $\{x^n | n \varepsilon N\}$ must have a least upper bound, say x_0, because $(X, \leqslant)$ is assumed to be conditionally complete. Therefore,

$$x \circ x_0 = x \circ \sup\{x^n | n \varepsilon N\} = \sup \{x^{n+1} | n \varepsilon N\} \leqslant x_0 .$$

But by cancellativitity, $x > e$ implies $x \circ x_0 > x_0$, a contradiction. ∇

THEOREM 2.14 For a multiset A to be closed in $\mathcal{N}_X \cap \mathcal{W}_X$ $(\mathcal{N}_X \cap \mathcal{W}'_X)$, where $(X, \leqslant, \circ)$ is an Archimedean totally ordered cancellative monoid, it is both necessary and sufficient that $x > e$ $(x < e)$ for every $x \varepsilon A$.

PROOF By cancellativity, it follows that $x = e$ whenever $x^2 = x$, and hence by theorem 2.13, $x \geq e$ for every $x \in A$ is a necessary condition for A to be closed in $\mathcal{W}_X$. But for A to be closed in $\mathcal{N}_X$ also, it follows from theorem 2.12 that $x \neq e$ for every $x \in A$ is necessary. Therefore, the necessity condition is verified. Now by theorem 2.13 above, $x > e$ for every $x \in A$ is a sufficient condition for A to be closed in $\mathcal{W}_X$. We now show that it is also sufficient for A to be closed in $\mathcal{N}_X$ also, i.e. $A^*(x) \neq \infty$ for all $x \in A$.

First, we claim that $\bigcap_{k=0}^{\infty} (A^k \circ A^*) = \phi$. For suppose otherwise, and let $T = \bigcap_{k=0}^{\infty} (A^k \circ A^*)$. Since $(A^k \circ A^*)$ is a non-increasing sequence of multisets, it follows from (ii) of theorem 2.4 that

$$A \circ T = A \circ \bigcap_{k=0}^{\infty} (A^k \circ A^*) = \bigcap_{k=0}^{\infty} (A^{k+1} \circ A^*) = T$$

Let $t_1 \in T$, then $t_1 = x \circ t_2$ for some $x \in A$ and $t_2 \in T$. By cancellativity, $x > e$ implies $t_1 = x \circ t_2 > t_2$. Similarly, we can show that $t_2 > t_3$ for some $t_3 \in T$ and so on. Therefore, if $T \neq \phi$, we can always obtain a strictly decreasing sequence $t_1 > t_2 > \ldots.$ of elements in T which then implies that T is not a well ordered multiset. But $T \subseteq A^*$ and A^* is well ordered by theorem 2.13 above, it must follow that T is also well ordered, a contradiction. Therefore our claim is justified.

We next claim that $\bigcap_{k=0}^{\infty} (A^k \circ A^*) = \phi$ implies that $A^k(x) \neq 0$ for only finitely many k. For suppose otherwise and let (k_i) be a subsequence of (k) such that $A^{k_i}(x) \neq 0$ for all $i \in N$. Since $A^{k_i} \subseteq A^{k_i} \circ A^*$ for all $i \in N$, it follows that $A^{k_i} \circ A^* (x) \neq 0$ for all $i \in N$ also. But this implies that $\bigcap_{i=0}^{\infty} A^{k_i} \circ A^* \neq \phi$, a contradiction. Hence for all $x \in X$, $A^k(x) \neq 0$ for only finitely many k as claimed, and therefore, $A^*(x) = \sum_{k=0}^{\infty} A^k(x) \neq \infty$ as required.

The case for $\mathcal{N}_X \cap \mathcal{W}'_X$ can also be proved dually. ∇

CHAPTER 3

P-SPACES AND THEIR PATH ALGEBRAS

3.1 P-Spaces

DEFINITION 3.1 A quadruple $(X, o, \mathscr{V}, r)$ is called a path space or p-space for short iff it has the following properties

(i) (X, o) is a monoid

(ii) $\mathscr{V}$ is a hereditary semiring of N_∞^X

(iii) r is a function on $\mathscr{V}$ which satisfies (3.1) to (3.3) below.

(3.1) $$r(\phi) = \phi$$

(3.2) $$r(A \uplus B) = r(r(A) \uplus B)$$

(3.3) $$r(A \circ B) = r(r(A) \circ B) = r(A \circ r(B))$$

For convenience, such a function r will always be referred to as a reduction function†. As examples of p-spaces, we offer the following.

---

† This name was inspired by the reduction function studied in Wongseelashote (1976). However, its abstract formulation was not conceived without the influence of the concept of an extraction function introduced by Roy (1975).

EXAMPLE 3.1 (X, o, N_∞^X, d), where (X, o) is a monoid and $d(A)$ is the support of A (see section 2.1).

EXAMPLE 3.2 $(X, o, \mathcal{W}_X, \min)$, where $(X, \leq, o)$ is a totally ordered monoid and min is defined on $\mathcal{W}_X$ as follows.

$$\min(A) = \begin{cases} \phi \text{ if } A = \phi \\ \{\min d(A)\} \text{ , otherwise} \end{cases}$$

Note that $(X, o, \mathcal{W}'_X, \max)$ can be dually defined.

EXAMPLE 3.3 $(X, o, \mathcal{W}_X, k\text{-min})$, where $(X, \leq, o)$ is a totally ordered monoid and k-min is defined on $\mathcal{W}_X$ as follows.

$$k\text{-min}(A) = \begin{cases} \phi \text{ if } A = \phi \\ \{a_1, a_2, \ldots, a_t\} \text{ , otherwise} \end{cases}$$

Here $a_1 < a_2 < \ldots < a_t$ are t successively smallest elements of A, and t is the largest index such that $t \leq k$. Note that $(X, o, \mathcal{W}'_X, k\text{-max})$ can also be defined in a dual fashion and that for $k = 1$, this example coincides with example 3.2 above.

EXAMPLE 3.4 $(X, o, \mathcal{W}_X, r_u)$, where $(X, \leq, o)$ is a totally ordered monoid and r_u is define on $\mathcal{W}_X$ for some given $u \in X$ by

$$r_u(A) = \begin{cases} \phi \text{ if } A = \phi \text{ or } u < x \text{ for every } x \in A \\ \{a_1, a_2, \ldots, a_j\} \text{ , otherwise.} \end{cases}$$

Here $a_1 < a_2 < \ldots < a_j$ are j successively smallest elements of A and j is the largest index such that $a_j \leq u$. Note that one can also define $(X, o, \mathcal{W}'_X, r'_u)$ dually.

EXAMPLE 3.5 $(R, +, \mathcal{W}_R, \Delta\text{-min})$, where $(R, +)$ is the additive group of real numbers and Δ-min is defined on $\mathcal{W}_R$ for some $\Delta \geq 0$ as follows.

$$\Delta\text{-min}(A) = \begin{cases} \phi & \text{if } A = \phi \\ \{a_1, a_2, \ldots, a_k\}, & \text{otherwise} \end{cases}$$

Here $a_1 < a_2 < \ldots < a_k$ are k successively smallest elements of A and k is the largest index such that $a_k \leq a_1 + \Delta$. Note that $(R, +, \mathcal{W}'_R, \Delta\text{-max})$ can also be defined in a dual fashion and that for $\Delta = 0$, this example also coincides with example 3.2 in the case where $(X, o) = (R, +)$.

EXAMPLE 3.6 $(\Sigma^*, \cdot, N_\infty^{\Sigma^*}, \text{sim})$, where $(\Sigma^*, \cdot)$ is the free monoid generated by an alphabet Σ and sim is defined on $N_\infty^{\Sigma^*}$ as follows.

$$\text{sim}(A) = \begin{cases} \phi & \text{if } A = \phi \\ \{x \in A \mid x \text{ is a simple word}\}, & \text{otherwise} \end{cases}$$

For the meaning of a simple word in Σ^*, see problem 1.8 above.

EXAMPLE 3.7 $(\Sigma^*, \cdot, N_\infty^{\Sigma^*}, b)$, where $(\Sigma^*, \cdot)$ is the free monoid generated by an alphabet Σ and b is defined on $N_\infty^{\Sigma^*}$ as follows

$$b(A) = \begin{cases} \phi & \text{if } A = \phi \\ \{x \in A \mid x \text{ is not an abbreviation of any } y \in A\}, & \text{otherwise.} \end{cases}$$

For the meaning of abbreviation, see problem 1.9 above.

EXAMPLE 3.8 Let $(N, +)$ be the additive monoid of non-negative integers and $\hat{N}$ a set of arbitrary objects disjoint from N which can also be put in one-to-one correspondence with N, i.e. for each $n \in N$, there is a unique $\hat{n} \in \hat{N}$ and vice versa. $\{n, \hat{n}\}$ will be called a twin pair. Now define a binary operation o on $X = N \cup \hat{N}$ by the following rules:

(i) $m \mathbin{o} n = \hat{m} \mathbin{o} \hat{n} = m + n$ for any $m, n \in N$

(ii) $\hat{m} \mathbin{o} n = m \mathbin{o} \hat{n} = \widehat{m \mathbin{o} n} = \widehat{m + n}$ for any $m, n \in N$

It can be verified that (X, o) so defined is a commutative monoid with 0 as the identity for o (cf. theorem 6.1 below). Moreover, this monoid is also locally finite because $(N, +)$ has this property.

Let $r = \mathcal{N}_X \to \mathcal{N}_X$ be defined by $r(A) = \phi$ if $A = \phi$, else $r(A)$ is the multiset obtained from A by deleting all its twin pairs.

The quadruple $(X, o, \mathcal{N}_X, r)$ then forms a p-space (cf. theorem 6.3 below).

The concept of p-spaces as defined here turns out to be closely related to that of a semiring. In fact, given one, the other can always be obtained. In the remainder of this section, we shall show how to construct a p-space from a given semiring. The converse construction is more important to the solution of path problems, and will be treated in the next section where it is more appropriate.

THEOREM 3.1 Let $(X, +, o)$ be a semiring. Then the quadruple $(X, o, \mathscr{F}_X, s)$, where s is defined by (3.4) below, forms a p-space.

$$(3.4) \qquad s(A) = \begin{cases} \left\{ \sum_{i=1}^{n} a_i \right\} & \text{if } A = \{a_1, a_2, \ldots, a_n\} \text{ and } \sum_{i=1}^{n} a_i \neq \theta \\ \phi \ , & \text{otherwise.} \end{cases}$$

PROOF From (3.4), it follows easily that property (3.1) is satisfied by s. In fact,

$$(3.5) \qquad s\{\theta, \theta, \ldots, \theta\} = s\{\theta\} = \phi = s(\phi) \text{ always.}$$

That property (3.2) is also satisfied by s can be seen as follows.

Let $A, B \in \mathscr{F}_X$. If $A = \phi$, then $s(A) = \phi$, and hence $s(A \uplus B) = s(B) = s(s(A) \uplus B)$ as required. So suppose that $A \neq \phi$, say $A = \{a_1, a_2, \ldots, a_n\}$. If $B = \phi$, then $s(A \uplus B) = s(A)$. But $s(s(A)) = s(A)$, because if $\sum_{i=1}^{n} a_i = \theta$, then $s(s(A)) = s(\phi) = \phi = s(A)$, while if $\sum_{i=1}^{n} a_i \neq \theta$, then $s(s(A)) = s\left(\sum_{i=1}^{n} a_i\right) = \sum_{i=1}^{n} a_i = s(A)$.

Consequently, $s(s(A) \uplus B) = s(s(A)) = s(A) = s(A \uplus B)$ as required. So we may now suppose that $B \neq \phi$ also, say $B = \{b_1, b_2, \ldots, b_m\}$. Then there are two cases to consider

(i) Either $\sum_{i=1}^{n} a_i = \theta$ or $\sum_{j=1}^{m} b_j = \theta$.

If $\sum_{i=1}^{n} a_i = \theta$, then $s(A) = \phi$, and hence

$s(s(A) \uplus B) = s(B)$.

Now we claim that $s(A \uplus B) = s(B)$ also.

For if $\sum_{j=1}^{m} b_j = \theta$, then $s(A \uplus B) = \phi = s(B)$ because $\sum_{i=1}^{n} a_i + \sum_{j=1}^{m} b_j = \theta$, while if $\sum_{j=1}^{m} b_j \neq \theta$, then

$s(A \uplus B) = \sum_{i=1}^{n} a_i + \sum_{j=1}^{m} b_j = \sum_{j=1}^{m} b_j = s(B)$, and hence our claim is justified.

Therefore $s(A \uplus B) = s(B) = s(s(A) \uplus B)$ as required. So we may now suppose $\sum_{i=1}^{n} a_i \neq \theta$ but $\sum_{j=1}^{m} b_j = \theta$. Then $s(B) = \phi$ and $\sum_{i=1}^{n} a_i + \sum_{j=1}^{m} b_j \neq \theta$. Consequently,

$$s(s(A) \uplus B) = s\left(\sum_{i=1}^{n} a_i \uplus B \right) = \sum_{i=1}^{n} a_i + \sum_{j=1}^{m} b_j = s(A \uplus B) \text{ as required.}$$

(ii) Both $\sum_{i=1}^{n} a_i \neq \theta$ and $\sum_{j=1}^{m} b_j \neq \theta$.

If $\sum_{i=1}^{n} a_i + \sum_{j=1}^{m} b_j = \theta$, then

$$s(s(A) \uplus B) = s(\sum_{i=1}^{n} a_i \uplus B) = \phi = s(A \uplus B) .$$

So suppose that $\sum_{i=1}^{n} a_i + \sum_{j=1}^{m} b_j \neq \theta$, then

$$s(s(A) \uplus B) = s(\sum_{i=1}^{n} a_i \uplus B) = \sum_{i=1}^{n} a_i + \sum_{j=1}^{m} b_j = s(A \uplus B).$$

Therefore, s satisfies property (3.2) above. It remains to show that s also satisfies (3.3) above.

Let $A, B \in \mathscr{F}_X$. If $A = \phi$, then

$$s(A \circ B) = \phi = s(s(A) \circ B) \text{ because } s(A) = \phi \text{ also.}$$

If $B = \phi$, then $s(A \circ B) = \phi = s(s(A) \circ B)$ always.

So let us suppose that $A = \{a_1, a_2, \ldots, a_n\}$ and $B=\{b_1,b_2,\ldots,b_m\}$. Then there are two cases to consider

(i) $\left(\sum_{i=1}^{n} a_i\right) \circ \left(\sum_{j=1}^{m} b_j\right) = \theta$

In this case, $s(A \circ B) = \phi$ by (3.4) because

$$\sum_{i=1}^{n} \sum_{j=1}^{m} a_i \circ b_j = \left(\sum_{i=1}^{n} a_i\right) \circ \left(\sum_{j=1}^{m} b_j\right) = \theta$$

Now we claim that $s(s(A) \circ B) = \phi$ also, and hence the required result. If $\sum_{i=1}^{n} a_i = \theta$, then $s(A) = \phi$ by (3.4) implies that

$s(s(A) \circ B) = s(\phi) = \phi$ as claimed.

So suppose $\sum_{i=1}^{n} a_i \neq \theta$. But then

$$s(s(A) \circ B) = s\left(\{\sum_{i=1}^{n} a_i\} \circ B\right) = \phi \text{ by (3.4) because}$$

$$\sum_{j=1}^{m}\left(\sum_{i=1}^{n} a_i\right) \circ b_j = \left(\sum_{i=1}^{n} a_i\right) \circ \left(\sum_{j=1}^{m} b_j\right) = \theta \text{ , and hence our}$$

claim is justified.

(ii) $\left(\sum_{i=1}^{n} a_i\right) \circ \left(\sum_{j=1}^{m} b_j\right) \neq \theta$.

In this case, it must be that both

$$\sum_{i=1}^{n} a_i \neq \theta \text{ and } \sum_{j=1}^{m} b_j \neq \theta \text{ .}$$

But then $s(s(A) \circ B) = s\left(\left\{\sum_{i=1}^{n} a_i\right\} \circ B\right)$

$$= \sum_{j=1}^{m}\left(\sum_{i=1}^{n} a_i\right) \circ b_j$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{m} a_i \circ b_j$$

$$= \left(\sum_{i=1}^{n} a_j\right) \circ \left(\sum_{j=1}^{m} b_j\right)$$

$$= s(A \circ B) \text{ as required}$$

Similarly, one can show that $s(A \circ s(B)) = s(A \circ B)$. ∇

For convenience, the p-space $(X, \circ, \mathscr{F}_X, s)$ as defined in the above theorem will be said to be induced by the semiring $(X, +, \circ)$. We emphasize here that each semiring may induce more than one p-space, since it may be possible to choose a larger domain for the function s above or to invent other reduction functions from a given semiring. For instance, if X is also a complete semiring, we can define the p-space $(X, \circ, N_\infty^X, s')$ from X by defining $s' : N_\infty^X \to N_\infty^X$ as follows.

$$(3.6) \qquad s'(A) = \begin{cases} \left\{\sum_{i\varepsilon I} a_i\right\} & \text{if } A = \biguplus_{i\varepsilon I} \{a_i\} \text{ and } \sum_{i\varepsilon I} a_i \neq \theta \\ \phi\ , & \text{otherwise.} \end{cases}$$

Note that s' so defined is an extension of the function s in (3.4) above, and the verification that s' is a reduction function can also be carried out in exactly the same manner as in theorem 3.1.

3.2 Path Algebras of P-Spaces.

Given a p-space $(X, \circ, \mathscr{V}, r)$, a multiset A of $\mathscr{V}$ is said to be reduced iff $A = r(A)$. From (3.1), it follows that ϕ is reduced, and therefore the set

$$\mathscr{V}_r = \{A \in \mathscr{V} \mid A = r(A)\}$$

is non-empty. Moreover, for any $A \in \mathscr{V}$, $r(A) \in \mathscr{V}_r$ because

$$\begin{aligned} r(r(A)) &= r(r(A) \uplus \phi) \\ &= r(A \uplus \phi) \quad , \text{by } (3.2) \\ &= r(A) \end{aligned}$$

Now for any $A, B \in \mathcal{V}_r$, we can define two binary operations $\oplus$ and $\odot$ on $\mathcal{V}_r$ as follows.

$$(3.7) \qquad A \oplus B = r(A \uplus B)$$

$$(3.8) \qquad A \odot B = r(A \circ B)$$

The triple $(\mathcal{V}_r, \oplus, \odot)$ will be called the <u>path algebra</u> of the p-space $(X, \circ, \mathcal{V}, r)$.

THEOREM 3.2 The path algebra of any p-space $(X, \circ, \mathcal{V}, r)$ forms a semiring with unit $r(e)$ and zero ϕ . Moreover, this semiring is idempotent whenever $r(e \uplus e) = r(e)$ holds and commutative whenever $\circ$ is commutative.

PROOF Since for any $A \in \mathcal{V}$, $r(A)$ was seen above to belong to $\mathcal{V}_r$, it follows that $\mathcal{V}_r$ is closed with respect to $\oplus$ and $\odot$ as defined by (3.7) and (3.8). Now the binary operations $\oplus$ and $\odot$ can be seen to possess all the properties of a semiring as follows.

(i) $A \oplus \phi = r(A \uplus \phi) = r(A) = A$ for all $A \in \mathcal{V}_r$

(ii) $A \oplus B = r(A \uplus B) = r(B \uplus A) = B \oplus A$ for any $A,B \in \mathcal{V}_r$

(iii) $A \oplus (B \oplus C) = r(A \uplus r(B \uplus C))$

$= r(A \uplus B \uplus C)$, by (3.2)

$= r(r(A \uplus B) \uplus C)$, by (3.2)

$= (A \oplus B) \oplus C$ for any $A,B, C \in \mathcal{V}_r$

(iv) $r(e) \circledcirc A = r(r(e) \circ A)$

$= r(e \circ A)$, by (3.3)

$= r(A)$

$= A \qquad$ for all $A \in \mathcal{V}_r$

(v) $A \circledcirc (B \circledcirc C) = r(A \circ r(B \circ C))$

$= r(A \circ B \circ C)$, by (3.3)

$= r(r(A \circ B) \circ C)$, by (3.3)

$= (A \circledcirc B) \circledcirc C \qquad$ for any $A, B, C \in \mathcal{V}_r$

(vi) $A \circledcirc (B \oplus C) = r(A \circ r(B \uplus C))$

$= r(A \circ (B \uplus C))$, by (3.2)

$= r(A \circ B \uplus A \circ C)$

$= r(r(A \circ B) \uplus r(A \circ C))$, by (3.2)

$= A \circledcirc B \oplus A \circledcirc C \quad$ for any $A, B, C \in \mathcal{V}_r$

Similarly, $(A \oplus B) \circledcirc C = A \circledcirc C \oplus B \circledcirc C$ for any $A, B, C \in \mathcal{V}_r$

Moreover, if $r(e \uplus e) = r(e)$ holds, we have

(vii) $A \oplus A = r(A \uplus A)$

$= r(A \circ r(e \uplus e))$ by (3.3)

$= r(A \circ r(e))$

$= r(A \circ e)$ by (3.3)

$= r(A)$

$= A \qquad$ for all $A \in \mathcal{V}_r$.

and if $\circ$ is commutative, then

(viii) $A \circledcirc B = r(A \circ B) = r(B \circ A) = B \circledcirc A$ for any $A, B \in \mathcal{V}_r$.

∇

An equivalent way of obtaining the path algebra from any given p-space is by making use of the following relation $\sim$ defined on $\mathcal{V}$ as follows.

(3.9) $\qquad A \sim B$ iff $r(A) = r(B)$

It is easy to see that $\sim$ is an equivalence relation, and hence the set $\mathcal{V}$ is partitioned into its equivalence subsets by this relation. Let $[A]$ denote the equivalence subset containing A and $\mathcal{V}/\sim$ denote the set of all these equivalence subsets. Then for any $[A], [B] \in \mathcal{V}/\sim$, we can define two binary operations $\oplus_r$ and $\odot_r$ on $\mathcal{V}/\sim$ as follows.

(3.10) $$[A] \oplus_r [B] = [A \uplus B]$$

(3.11) $$[A] \odot_r [B] = [A \circ B]$$

Now since it is possible that for $A \neq C$, $[A] = [C]$, one may question whether these operations are well-defined, i.e. independent of equivalence subset representatives. Indeed, they are because of the way we define a reduction function. For suppose $[A] = [C]$ and $[B] = [D]$, i.e. $r(A) = r(C)$ and $r(B) = r(D)$. Then it follows from (3.2) that

$$\begin{aligned} r(A \uplus B) &= r(r(A) \uplus B) \\ &= r(r(A) \uplus r(B)) \\ &= r(r(C) \uplus r(D)) \\ &= r(C \uplus r(D)) \\ &= r(C \uplus D) \end{aligned}$$

Similarly, we can show that $r(A \circ B) = r(C \circ D)$ by (3.3). Therefore, $[A \uplus B] = [C \uplus D]$ and $[A \circ B] = [C \circ D]$ which justify our claim.

THEOREM 3.3 $(\mathcal{V}/\sim, \oplus_r, \odot_r)$ is a semiring with unit $[e]$ and zero $[\phi]$, and is isomorphic to the path algebra.

PROOF In view of theorem 3.2 it suffices to show that there is a semiring isomorphism from the path algebra to $(\mathcal{V}/\sim, \oplus_r, \odot)$. In fact, the required isomorphism is given by the function $f: \mathcal{V}_r \to \mathcal{V}/\sim$ which is defined by $f(A) = [A]$ for all $A \in \mathcal{V}_r$. For we have

$$\begin{aligned} f(A \oplus B) &= [A \oplus B] \\ &= [r(A \uplus B)] \text{ , by (3.7)} \\ &= [A \uplus B] \text{ , since } r(r(A \uplus B)) = r(A \uplus B) \\ &= [A] \oplus_r [B] \\ &= f(A) \oplus_r f(B) \text{ ,} \end{aligned}$$

and likewise, $f(A \odot B) = f(A) \odot_r f(B)$,

Since $f(\phi) = [\phi]$ and $f(r(e)) = [r(e)] = [e]$ also, it remains to show that f is a bijection.

Suppose $f(A) = f(B)$. Then by definition, $[A] = [B]$ or $r(A) = r(B)$. But then $A = B$ because $A, B \in \mathcal{V}_r$. Therefore, f is an injection.

Now let $[A] \in \mathcal{V}/\sim$ be given. Then

$$f(r(A)) = [r(A)] = [A], \text{ since } r(r(A)) = r(A).$$

Therefore, f is also a surjection as required. ∇

It is interesting to note that in fact $\mathcal{V}_r = r(\mathcal{V})$, the range of the reduction function r. For if $A \in \mathcal{V}_r$, then $A = r(A)$ implies that $A \in r(\mathcal{V})$, while on the other hand,

$A \in r(\mathscr{V})$ implies that $A = r(B)$ for some $B \in \mathscr{V}$. But then $A \in \mathscr{V}_r$ because $r(B) \in \mathscr{V}_r$. The above claim is therefore justified.

In view of this observation, theorem 3.3 in fact says that $\mathscr{V}/\sim$ is isomorphic to $r(\mathscr{V})$ as semirings. This conclusion resembles the usual homomorphism theorem in algebra except that the reduction function r may not be a homomorphism.

We note in passing that an elegant proof of theorem 3.2 can in fact be obtained by considering the cannonical surjection $\pi : \mathscr{V} \to \mathscr{V}/\sim$ which is defined by $\pi(A) = [A]$ for any $A \in \mathscr{V}$. Now from (3.10) and (3.11), we conclude that π is a semiring homomorphism from $\mathscr{V}$ to $\mathscr{V}/\sim$, and hence $(\mathscr{V}/\sim, \oplus_r, \odot_r)$ is a semiring. But $(\mathscr{V}/\sim, \oplus_r, \odot_r)$ was seen to be isomorphic to $(\mathscr{V}_r, \oplus, \odot)$ in theorem 3.3 above, and hence $(\mathscr{V}_r, \oplus, \odot)$ is a semiring as required.

Let us now consider the path algebra of the p-space $(X, \circ, \mathscr{F}_X, s)$ induced by a given semiring $(X, +, \circ)$ in theorem 3.1 above. The relationship between this path algebra and the semiring $(X, +, \circ)$ is expressed in the following.

THEOREM 3.4 The path algebra of the p-space $(X, \circ, \mathscr{F}_X, s)$ induced by a given semiring $(X, +, \circ)$ is isomorphic to $(X, +, \circ)$ as semirings.

PROOF Let us define a function $f: X \to \mathscr{V}_s$, where $\mathscr{V} = \mathscr{F}_X$, by

$$(3.12) \qquad f(x) = \begin{cases} \{x\} & \text{if } x \neq \theta \\ \phi & \text{, otherwise.} \end{cases}$$

Then f can be seen to be a semiring homomorphism from (i) and (ii) below.

(i) If $x = \theta$, then by (3.12), $f(x) = \phi$, and hence

$$f(x) \oplus f(y) = \phi \oplus f(y)$$
$$= f(y)$$
$$= f(x + y), \text{ since } x = \theta .$$

Similarly, if $y = \theta$, then $f(x) \oplus f(y) = f(x + y)$ as required. So we may suppose that $x \neq \theta$ and $y \neq \theta$. But then

$$f(x) \oplus f(y) = s(f(x) \uplus f(y))$$
$$= s(x \uplus y)$$
$$= \begin{cases} \{x+y\} & \text{if } x + y \neq \theta \\ \phi, & \text{otherwise} \end{cases}$$
$$= f(x+y) \text{ as required.}$$

(ii) If $x = \theta$, then $f(x) = \phi$ by (3.12), and hence

$$f(x) \odot f(y) = \phi \odot f(y)$$
$$= \phi$$
$$= f(x)$$
$$= f(x \circ y), \text{ since } x = \theta = x \circ y.$$

Similarly, if $y = \theta$, then $f(x) \odot f(y) = f(x \circ y)$ as required. So we may suppose that $x \neq \theta$, $y \neq \theta$. But then

$$f(x) \odot f(y) = s(f(x) \circ f(y))$$
$$= s(x \circ y)$$
$$= \begin{cases} \{x \circ y\} & \text{if } x \circ y \neq \theta \\ \phi, & \text{otherwise} \end{cases}$$
$$= f(x \circ y) \text{ as required.}$$

Now we show that f is a bijection.

Suppose that $f(x) = f(y)$, then $\{x\} = \{y\}$ if $x, y \neq \theta$, and hence $x = y$.

If $x = \theta$, then $f(x) = \phi$. But then $f(y) = \phi$ also, which implies that $y = \theta = x$.

Similarly, if $y = \theta$, then $x = \theta = y$.

Therefore f is an injection.

To see that f is also a surjection, let $A \in \mathscr{V}_s$,

Then

$$A = s(A)$$

$$= \begin{cases} \left\{\sum_{i=1}^{n} a_i\right\} & \text{if } A = \{a_1, a_2, \ldots, a_n\} \text{ and } \sum_{i=1}^{n} a_i \neq \theta \\ \phi & \text{, otherwise} \end{cases}$$

$$= f\left(\sum_{i=1}^{n} a_i\right) \quad \text{as required.} \qquad \nabla$$

3.3 A Variety of P-Spaces

In future study, we shall find it useful to distinguish several types of p-spaces in accordance with their monoids, hereditary semirings or reduction functions. Therefore, let us make the following

DEFINITION 3.2 A p-space $(X, o, \mathscr{V}, r)$ is said to be commutative iff (X, o) is a commutative monoid.

DEFINITION 3.3 A p-space $(X, o, \mathscr{V}, r)$ is said to be finite iff $\mathscr{V} = \mathscr{F}_X$.

DEFINITION 3.4 A p-space $(X, o, \mathscr{V}, r)$ is said to be

(i) <u>intensive</u> iff $r(A) \subsetneq A$ for every $A \varepsilon \mathscr{V}$,

(ii) <u>idempotent</u> iff $r(e \uplus e) = r(e)$ holds,

(iii) <u>q-stationary</u> iff whenever $A^* \varepsilon \mathscr{V}$, we have $r(A^*) = r(e \uplus A \uplus \ldots \uplus A^q)$ for some $q \varepsilon N$, and

(iv) <u>complete</u> iff whenever $r(A_i \uplus B) = r(B)$ for every $A_i \varepsilon \mathscr{F}_X$ such that $\biguplus_{i \varepsilon I} A_i = A \varepsilon \mathscr{V}$, we have $r(A \uplus B) = r(B)$.

The following theorems are consequences of the above definitions.

THEOREM 3.5 Every finite p-space is 0-stationary.

PROOF Let $(X, o, \mathscr{V}, r)$ be a finite p-space. Then it follows from (ii) of theorem 2.11 that $A^* \varepsilon \mathscr{V} = \mathscr{F}_X$ iff $A = \phi$, and hence

$$r(A^*) = r(\phi^*) = r(e) \text{ as required.} \qquad \nabla$$

THEOREM 3.6 Every finite p-space is complete.

PROOF Let $(X, o, \mathscr{V}, r)$ be a finite p-space and suppose that $r(A_i \uplus B) = r(B)$ for every $A_i \varepsilon \mathscr{F}_X$ such that $\biguplus_{i \varepsilon I} A_i = A \varepsilon \mathscr{V}$. Now since $\mathscr{V} = \mathscr{F}_X$, it follows that $r(A_i \uplus B) = r(B)$ for only finitely many A_i, i.e. $A_i \oplus B = r(B)$ for only finitely many A_i, say $A_1, A_2, \ldots, A_k$. But then $A_1 \oplus A_2 \oplus \ldots \oplus A_k \oplus B = r(B)$, i.e. $r(A_1 \uplus A_2 \uplus \ldots \uplus A_k \uplus B) = r(B)$, i.e. $r(A \uplus B) = r(B)$ as required. ∇

THEOREM 3.7 A p-space $(X, o, \mathcal{V}, r)$ is idempotent iff $r(A \uplus A) = r(A)$ for all $A \in \mathcal{V}$.

PROOF Sufficiency is obvious whereas necessity follows from (vii) in the proof of theorem 3.2. ∇

THEOREM 3.8 A p-space $(X, o, \mathcal{V}, r)$ is idempotent iff $r(A \uplus B) = r(A \cup B)$ for any $A, B \in \mathcal{V}$.

PROOF If $r(A \uplus B) = r(A \cup B)$ for any $A, B \in \mathcal{V}$, then in particular, $r(A \uplus A) = r(A \cup A) = r(A)$, and hence the p-space is idempotent by theorem 3.7 above.

Now suppose that the p-space is idempotent. For ease of exposition, let $Y = A \cap B$ and $Z = A \cup B$.

Then clearly, $Y \subseteq Z$.

If $Y = Z$, then $A = B$, and hence

$$r(A \uplus B) = r(A \uplus A) = r(A) = r(A \cup A) = r(A \cup B)$$

as required.

If $Y \neq Z$, then we can define the multiset $Z \ominus Y$ by

$$(Z \ominus Y)(x) = \begin{cases} \infty & \text{if } Z(x) = \infty \\ Z(x) - Y(x), & \text{otherwise} \end{cases}$$

From this definition, it is easily verified that

$$Z = Y \uplus (Z \ominus Y) \qquad \text{always, and hence}$$

$$
\begin{aligned}
r(A \uplus B) &= r(Y \uplus Z) \\
&= r(Y \uplus Y \uplus (Z \boxminus Y)) \\
&= r(r(Y \uplus Y) \uplus (Z \boxminus Y)) \text{ , by (3.2)} \\
&= r(r(Y) \uplus (Z \boxminus Y)) \\
&= r(Y \uplus (Z \boxminus Y)) \text{ , by (3.2)} \\
&= r(Z) \\
&= r(A \cup B) \quad \text{as required.}
\end{aligned}
$$
∇

COROLLARY 3.1 A p-space $(X, o, \mathcal{V}, r)$ is idempotent iff the following condition holds.

(3.13) For any $A, B \in \mathcal{V}$ such that $A \subseteq B$, we have $r(A) \prec r(B)$, where $\prec$ denotes the pseudo-ordering of the path algebra.

PROOF Suppose (3.13) holds, then in particular $A \subseteq A$ implies $r(A) \prec r(A)$, i.e. $r(A \uplus A) = r(A)$, and hence the p-space is idempotent.

Now suppose that the p-space is idempotent and let $A \subseteq B$, i.e. $A \cup B = B$. Then it follows from theorem 3.8 that

$$r(A \uplus B) = r(A \cup B) = r(B) \text{ , i.e. } r(A) \prec r(B)$$

as required. ∇

Let us now make a special note concerning the significance of theorem 3.8 above from the view point of lattice theory. From theorem 3.2 above, we know that the path algebra $(\mathcal{V}_r, \oplus, \odot)$ of any idempotent p-space is an idempotent semiring, and hence the pseudo-ordering $\prec$ of $(\mathcal{V}_r, \oplus, \odot)$ is in fact an ordering. Moreover, $(\mathcal{V}_r, \prec)$ is also a join-semilattice (see section 0.1 above). In fact, the least upper bound of any finite subset $\{A_1, A_2, \ldots, A_k\}$ of $\mathcal{V}_r$ is given by $A_1 \oplus A_2 \oplus \ldots \oplus A_k$. Therefore the addition $\oplus$ of the idempotent semiring $(\mathcal{V}_r, \oplus, \odot)$ coincides with what is usually known in lattice theory as the join operation, and is usually denoted by $\vee$ (see e.g. Birkhoff (1967)). Therefore in future, we shall use the notation $\vee$ in place of $\oplus$ to emphasize the fact that the p-space under consideration is idempotent. Now theorem 3.8 essentially says that the reduction function r is a join-morphism (Birkhoff (1967))from the semi-lattice $(\mathcal{V}, \subseteq)$ to the semi-lattice $(\mathcal{V}_r, \prec)$, i.e. $r(A \cup B) = r(A) \vee r(B)$. This follows because

$$r(A \cup B) = r(A \uplus B) = r(r(A) \uplus r(B)) = r(A) \vee r(B).$$

In fact, for complete idempotent p-space, its reduction function can be seen to be a complete join-morphism (Birkhoff (1967)) as follows.

THEOREM 3.9 Let $(X, o, \mathcal{V}, r)$ be an idempotent p-space. Then the following are all equivalent

(i) $r(\bigcup_{i\in I} A_i) = \bigvee_{i\in I} r(A_i)$ whenever $\bigcup_{i\in I} A_i \in \mathcal{V}$

(ii) $r(\biguplus_{i\in I} A_i) = \bigvee_{i\in I} r(A_i)$ whenever $\biguplus_{i\in I} A_i \in \mathcal{V}$

(iii) The p-space is complete

PROOF We show that (i) implies (ii), (ii) implies (iii) and (iii) implies (i).

That (i) implies (ii) can be seen from (2.7) as follows.

$$r\left(\biguplus_{i\in I} A_i\right) = r\left[\bigcup_{J\in 2^{(I)}} \left(\biguplus_{j\in J} A_j\right)\right]$$

$$= \bigvee_{J\in 2^{(I)}} r\left(\biguplus_{j\in J} A_j\right)$$

$$= \bigvee_{J\in 2^{(I)}} \left(\bigvee_{j\in J} r(A_j)\right)$$

$$= \bigvee_{i\in I} r(A_i) \quad \text{as claimed.}$$

That (ii) implies (iii) can be seen as follows. Let $r(A_i \uplus B) = r(B)$ for every $A_i \in \mathcal{R}_X$ such that $\biguplus_{i\in I} A_i = A \in \mathcal{V}$. Then by (ii), we have

$$r(A \uplus B) = r\left(\biguplus_{i\in I} A_i \uplus B\right) = \bigvee_{i\in I} r(A_i) \vee r(B).$$

But by theorem 3.8 and our assumption, we have

$$r(A_i) \vee r(B) = r(A_i \uplus B) = r(B) \quad \text{for every} \quad i \in I,$$

and hence $r(A_i) \prec r(B)$ for every $i \in I$.

Consequently, $\bigvee_{i\in I} r(A_i) \prec r(B)$, i.e.

$$\bigvee_{i\in I} r(A_i) \vee r(B) = r(B), \text{ and hence}$$

$$r(A \uplus B) = r(B), \text{ which proves (iii).}$$

That (iii) implies (i) can be seen as follows.

Let $Z = \bigcup_{i\in I} A_i$. But then $A_i \subseteq Z$ for all $i \in I$, and hence by corollary 3.1, $r(A_i) \prec r(Z)$ for all $i \in I$, which says that

$r(Z)$ is an upper bound for $\{r(A_i) | i \varepsilon I\}$. We now show that $r(Z)$ is in fact the least upper bound. So let $Y \varepsilon \mathcal{V}_r$ be such that

$$r(A_i) \prec Y \text{ for all } i \varepsilon I.$$

Now for each $x \varepsilon Z = \bigcup_{i \varepsilon I} A_i$, $x \varepsilon A_i$ for some $i \varepsilon I$, and hence by corollary 3.1 above, $r(x) \prec r(A_i)$. for some $i \varepsilon I$. Since $\prec$ is transitive, it follows that $r(x) \prec Y$. Thus we have just shown that for every $x \varepsilon Z$,

$$r(x \uplus Y) = r(Y).$$

Since $Z = \biguplus_{j \varepsilon J} \{x_j\}$, it then follows from the completeness assumption that $r(Z \uplus Y) = r(Y) = Y$ or equivalently, $r(Z) \prec Y$. Consequently, $r(Z) = \bigvee_{i \varepsilon I} r(A_i)$ as required.

∇

Let us now examine which of the properties in definition 3.4 above are possessed by the p-spaces in the above examples. A straight-forward verification will show that the p-spaces of all the above examples are complete and intensive. In fact, we are unable to find an example of a p-space which is not complete. As an example of a p-space which is not intensive, we mention the p-space $(N, \cdot, \mathcal{F}_N, s)$ induced by the usual semiring $(N, +, \cdot)$ of non-negative integers (see theorem 3.1 above). For if we let $A = \{1, 2, 3\}$, then $s(A) = \{1 + 2 + 3\} = \{6\} \not\subset A$. The p-spaces in all but example 3.8 can also be easily seen to be idempotent.

However, in examples 3.3 to 3.5, if one replaces $<$ everywhere by $\leqslant$, then the resulting p-spaces will not be idempotent. Similarly, in examples 3.6 and 3.7, if one replaces the requirement for sim(A) and b(A) to be sets by multisets, then the resulting p-spaces will not be idempotent either. The p-spaces in examples 3.6 and 3.7 are also q-stationary. In fact, it is easily seen that in example 3.6, $q = |\Sigma|$, where $|\Sigma|$ denotes the number of elements in Σ, and in example 3.7, $q = 0$. If we assume that the totally ordered monoid $(X, \leqslant, \circ)$ in examples 3.2 and 3.3 also has the property that $x^2 = x$ implies $x = e$ for all $x \in X$, then the p-spaces in these example are also q-stationary. In fact $q = 0$ for example 3.2 and $q = 2^{(k-2)}$ for example 3.3. For by theorem 2.13, we know that $A^* \in \mathcal{W}_X$ implies that $x \geqslant e$ for every $x \in A$, and hence

$$\min(A^*) = \{e\} = \min\{e\},$$

and $$k\text{-}\min(A^*) = k\text{-}\min\left\{e \uplus A \uplus A^2 \uplus \dots \uplus A^{2^{(k-2)}}\right\}.$$

The latter can be shown as follows. If $x = e$ for every $x \in A$, then $A^* = \{e, e, \dots\}$ and hence

$$k\text{-}\min(A^*) = \{e\} = k\text{-}\min\left\{e \uplus A \uplus A^2 \uplus \dots \uplus A^{2^{(k-2)}}\right\}.$$

So we may suppose that $x > e$ for some $x = a$, say. But then $e < a < a^2 < a^4 < \dots < a^{2^{(k-2)}}$ are contained in $e \uplus A \uplus \dots \uplus A^{2^{(k-2)}}$ and since $x \in A^s$ for any $s > 2^{(k-2)}$ always implies that $x \geqslant y$ for some $y \in A^{2^{(k-2)}}$, the required result follows. We note that if we replace $<$ by $\leqslant$ everywhere in example 3.3 or if we assume that $x^{j+1} = x^j$ implies $x = e$ for all $x \in X$ and all $j \in N$, then it can be similarly shown that $q = k-1$ in these cases.

CHAPTER 4

COMPATIBILITY AND STABILITY

4.1 Networks and their Graphs

The formulation and solution of path problems, using the concept of multisets of chapter 2, which we shall discuss in the next chapter, actually permits us to describe a path problem in the setting of a more general framework than that of a graph over a monoid, i.e. multiple arcs can now be taken into consideration. More precisely, let us make the following

DEFINITION 4.1 A network $\mathcal{N}$ over a finite set L is an ordered pair (W,U) where W is a finite set, and U is a multiset with elements from the cartesian product $W \times L \times W$. The elements of L are called labels, elements of W are nodes, and elements of U are arcs. More precisely, each triple (x, a, x') of the multiset U will be called an arc beginning at node x, ending at node x', and carrying the label a .

We note that by this definition, a graph G over a set L as defined in section 0.4 is just a network in which no two arcs with the same beginning and end are allowed, whatever labels they may carry. Thus a graph over a set L is just a particular instance of a network. In view of this connection, we can also generalize other concepts which are previously defined for graphs in section 0.4 to networks as follows, where for convenience, the nodes of the network $\mathcal{N}$ are designated as $x_1, x_2, \ldots, x_n$.

DEFINITION 4.2 Let $\mathcal{N}$ be a network. Then a path p of $\mathcal{N}$ is a finite ordered sequence

$$(x_{i_0}, a_1, x_{i_1})(x_{i_1}, a_2, x_{i_2}) \dots (x_{i_{k-1}}, a_k, x_{i_k})$$

of consecutive arcs in $\mathcal{N}$. The integer $k > 0$ is called its order, x_{i_0} its beginning, x_{i_k} its end, and x_{i_1} to $x_{i_{k-1}}$ its intermediate nodes. The path p is said to be closed iff $x_{i_0} = x_{i_k}$, elementary iff $x_{i_r} \neq x_{i_s}$ whenever $r \neq s$ (except, of course, for closed paths, where $x_{i_0} = x_{i_k}$ always).

In the present study, we shall only be interested in networks over a finite subset of a monoid, which we shall simply refer to as networks over a monoid. This is to say that the labels of the network under consideration are elements of a monoid (X, o), say. Consequently, for each path

$$p = (x_{i_0}, a_1, x_{i_1})(x_{i_1}, a_2, x_{i_2}) \dots (x_{i_{k-1}}, a_k, x_{i_k})$$

in $\mathcal{N}$, the element $a_1 \circ a_2 \circ \dots \circ a_k$ is well defined (because o is associative) and will be called the label of the path p in $\mathcal{N}$. For convenience, we shall also introduce the concept of a null path θ_i for each node x_i in $\mathcal{N}$ which can be defined as a closed path of order zero which begins and ends at x_i, and has label e, the identity for o. For convenience, each node x_i has exactly one null path by definition.

A network can generally be represented by a diagram such as figure 4.1 below, where the nodes are represented by letters $x_1, x_2, \dots, x_n$ and an arc such as (x_i, a, x_j) by $x_i \xrightarrow{a} x_j$.

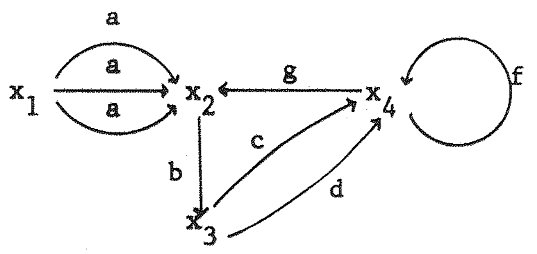

Figure 4.1 A network $\mathcal{N}$

It would appear simpler to draw figure 4.1 as figure 4.2 below, which contains essentially the same informations as figure 4.1. But figure 4.2 can be seen to be a graph over the set $\bar{L} \subseteq N_\infty^X$ (see section 2.1). This suggests that a network $\mathcal{N}$ over a finite set $L \subseteq X$ is in fact equivalent to a graph over $\bar{L} \subseteq N_\infty^X$. Indeed, this is so. For let $\mathcal{N} = (W, U)$ be a given network over a finite set $L \subseteq X$, we can define a graph $G(\mathcal{N}) = (W, V, v)$ over $\bar{L} \subseteq N_\infty^X$ by taking W to be the set of nodes in $\mathcal{N}$, V to be a subset of $W \times W$ such that $(x_i, x_j) \in V$ iff $(x_i, a, x_j) \in U$ for at least one $a \in L$, and $v : V \to \bar{L}$ to be defined by $v(x_i, x_j)$ = the multiset of all labels of arcs in U which begins at node x_i and ends at node x_j. And conversely, we can define a network $\mathcal{N}(G) = (W,U)$ from a graph $G = (W, V, v)$ over a subset of N_∞^X by taking W to be the set of nodes in G, and U to be a multiset such that $(x_i, a, x_j) \in U$ iff $a \in v(x_i, x_j)$.

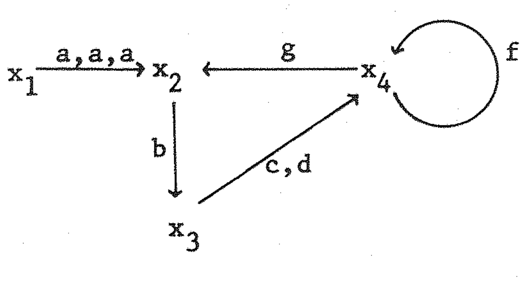

Figure 4.2 The graph $G(\mathcal{N})$ of the network $\mathcal{N}$ in figure 4.1.

From the above discussion, we see that given a network $\mathcal{N}$, we can define the <u>label matrix</u> M of $\mathcal{N}$ to be the arc-value matrix of the graph $G(\mathcal{N}) = (W, V, v)$, i.e.

$$(4.1) \qquad M_{ij} = \begin{cases} v(x_i, x_j) & \text{if } (x_i, x_j) \in V \\ \phi & \text{, otherwise} \end{cases}$$

In this way, all the previous results concerning the arc-value matrix of a graph over a semiring can be translated into results concerning the label matrix of a network over a monoid. For instance, the matrix M^k, the kth power of M is given by (0.15) as follows.

$$(4.2) \qquad (M^k)_{ij} = \biguplus_{p \in P_{ij}^{(k)}} v(p) \quad ,$$

where $v(p)$ and $P_{ij}^{(k)}$ are defined with respect to the graph $G(\mathcal{N})$ over the complete semiring N_∞^X.

Note that in this instance, $v(p)$ as given by (0.13) is the multiset consisting of all the labels of all the paths in $\mathcal{N}$ which traverse the same nodes as the path p in $G(\mathcal{N})$.

Here it is convenient to introduce the function $v : 2^P \to N_\infty^X$, where P denotes the set of all paths in $G(\mathcal{N})$. This function is defined by

$$(4.3) \qquad v(Q) = \biguplus_{p \in Q} v(p) \qquad \text{for all } Q \in 2^P.$$

We note here that strictly speaking, a different notation should be employed in place of v on the left-hand side of (4.3) above. However, our use of the same v does not cause any confusion because the argument of the function $v : 2^P \to N_\infty^X$ are sets which are denoted by capital letters, whereas paths will always be denoted by lower case letters. We note also that the function $v : 2^P \to N_\infty^X$ is in fact an extension of the function $\sigma : 2^{(P)} \to N_\infty^X$ as defined by (0.16) for the graph $G(\mathcal{N})$ over the complete semiring N_∞^X. Accordingly, properties (0.17) and (0.18) of σ can also be obtained for v, namely

(4.4) $$v(Q_1 \cup Q_2 \cup \dots \cup Q_k) = v(Q_1) \uplus v(Q_2) \uplus \dots \uplus v(Q_k) \quad \text{whenever} \quad Q_i \cap Q_j = \phi \ \text{ for } \ i \neq j.$$

(4.5) $$v(Q_{ir_1} Q_{r_1 r_2} \dots Q_{r_k j}) = v(Q_{ir_1}) \circ v(Q_{r_1 r_2}) \circ \dots \circ v(Q_{r_k j}),$$

where each Q_{rs} is a subset of paths from x_r to x_s and

$$Q_{rs} Q_{st} = \left\{ pq \mid p \in Q_{rs}, \ q \in Q_{st} \right\}.$$

In fact, a stronger property than (4.4) is possessed by $v : 2^P \to N_\infty^X$ as stated in theorem 4.1 below. But first, let us establish the following:

LEMMA 4.1 Let $Q_1, Q_2 \in 2^P$ be such that $Q_1 \subseteq Q_2$.
Then $v(Q_1) \subseteq v(Q_2)$.

PROOF Since $Q_1 \subseteq Q_2$, it follows that $Q_2 = Q_1 \cup (Q_2 \setminus Q_1)$. Therefore, it follows from theorem 2.5 and (4.4) above that

$$v(Q_1) \subseteq v(Q_1) \uplus v(Q_2 \setminus Q_1) = v(Q_2). \qquad \nabla$$

THEOREM 4.1 If (Q_i) is a sequence of subsets of P such that $Q_i \cap Q_j = \phi$ whenever $i \neq j$, then

$$v(\bigcup_{i=1}^{\infty} Q_i) = \biguplus_{i=1}^{\infty} v(Q_i) \quad .$$

PROOF Let $Q = \bigcup_{i=1}^{\infty} Q_i$ and $S_k = Q_1 \cup Q_2 \cup \ldots \cup Q_k$. Then in view of (2.7) and (4.4) above, it suffices to show that $v(Q) = \bigcup_{k=1}^{\infty} v(S_k)$.

By lemma 4.1 above, $S_k \subseteq Q$ implies that $v(S_k) \subseteq v(Q)$ for all k, i.e. $v(Q)$ is an upper bound for $\{v(S_k) | k \in \{1,2,\ldots\}\}$. We claim that $v(Q)$ is in fact the least upper bound. In order to justify this claim, let us define a function $\tilde{v} : N_{\infty}^X \rightarrow 2^P$ by $\tilde{v}(A) = \{p \in P | v(p) \subseteq A\}$. This function has the following three properties.

(i) $A \subseteq B$ implies $\tilde{v}(A) \subseteq \tilde{v}(B)$.

This follows because any $p \in \tilde{v}(A)$ implies $v(p) \subseteq A \subseteq B$, and hence $p \in \tilde{v}(B)$ also.

(ii) $Q \subseteq \tilde{v}(v(Q))$ for any $Q \in 2^P$.

For let $p \in Q$, then $v(p) \subseteq \biguplus_{p \in Q} v(p) = v(Q)$ implies $p \in \tilde{v}(v(Q))$ as required.

(iii) $v(\tilde{v}(A)) \subseteq A$ for any $A \in N_{\infty}^X$.

This follows because

$$v(\tilde{v}(A)) = \biguplus_{p \in \tilde{v}(A)} v(p) = \biguplus_{v(p) \subseteq A} v(p) \subseteq A \; .$$

Now let Y be any other upper bound for $\{v(S_k) | k \in \{1,2,\ldots\}\}$. Then $v(S_k) \subseteq Y$ for all k, and hence from (i) above, it follows that $\tilde{v}(v(S_k)) \subseteq \tilde{v}(Y)$ for all k. But then it follows from (ii) that $S_k \subseteq \tilde{v}(v(S_k)) \subseteq \tilde{v}(Y)$, and hence $Q \subseteq \tilde{v}(Y)$. Therefore, by lemma 4.1

and (iii) above, we have $v(Q) \subseteq v(\tilde{v}(Y)) \subseteq Y$ as claimed. ∇

Theorem 4.1 is useful in establishing the following theorem, where all the essential properties concerning the label matrix M that we shall require later are conveniently summarized.

THEOREM 4.2 Let M be the label matrix of a network $\mathcal{N}$ over a monoid $(X,\circ)$. Then we have the following

(i) M^k is given by $(M^k)_{ij} = v(P_{ij}^{(k)})$ for all i,j, where $P_{ij}^{(k)}$ is defined with respect to the graph $G(\mathcal{N})$ over the complete semiring N_∞^X

(ii) $M^{[s]} = \biguplus_{k=0}^{s} M^k$ is given by $(M^{[s]})_{ij} = v(P_{ij}^{[s]})$ for all i,j,

where $P_{ij}^{[s]} = \bigcup_{k=0}^{s} P_{ij}^{(k)}$ for all $s \in N$.

(iii) $M^* = \biguplus_{k=0}^{\infty} M^k$ and $M^+ = \biguplus_{k=1}^{\infty} M^k$ are respectively given by

$$(M^*)_{ij} = v(P_{ij}) \text{ and } (M^+)_{ij} = v(P_{ij} \setminus P_{ij}^{(0)}) \text{ for all } i,j,$$

where $P_{ij} = \bigcup_{k=0}^{\infty} P_{ij}^{(k)}$.

PROOF (i) is just (4.2) above, (ii) follows directly from (i) and (4.4), and (iii) follows directly from (i) and Theorem 4.1 above. ∇

THEOREM 4.3 Let $\mathcal{N}$ be a network over a monoid $(X,\circ)$ with n nodes. For any $k \in \{0, 1, 2,\ldots,n\}$, let $Q_{ij}^{\{k\}}$ be the set of all paths of $P_{ij} \setminus P_{ij}^{(0)}$ in the graph $G(\mathcal{N})$ that do not use any node x_r such that $r > k$ as an intermediate node. Then for all i,j, we have

(i) $v(P_{ij}^{(0)}) = v(Q_{ij}^{\{0\}})$,

(ii) $v(P_{ij} \setminus P_{ij}^{(0)}) = v(Q_{ij}^{\{n\}})$, and

(iii) $v(Q_{ij}^{\{k\}}) = v(Q_{ij}^{\{k-1\}}) \uplus v(Q_{ik}^{\{k-1\}}) \circ v(Q_{kk}^{\{k-1\}})^* \circ v(Q_{kj}^{\{k-1\}})$

for all $k \in \{1, 2,\ldots,n\}$.

PROOF (i) Observe that each path in $Q_{ij}^{\{0\}}$ is not allowed to have intermediate nodes, and hence its order is at most unity; in fact, it is exactly unity because $Q_{ij}^{\{0\}}$ does not contain null paths. On the other hand, each path of order unity can never have intermediate nodes. Therefore, $P_{ij}^{(1)} = Q_{ij}^{\{0\}}$ and hence $v(P_{ij}^{(1)}) = v(Q_{ij}^{\{0\}})$.

(ii) Since the network has only n nodes, each path in $P_{ij} \setminus P_{ij}^{(0)}$ cannot use more than n nodes as intermediate nodes, and hence $P_{ij} \setminus P_{ij}^{(0)} \subseteq Q_{ij}^{\{n\}}$.

But $Q_{ij}^{\{n\}} \subseteq P_{ij} \setminus P_{ij}^{(0)}$ always. Therefore,

$$P_{ij} \setminus P_{ij}^{(0)} = Q_{ij}^{\{n\}} \text{ and hence } v(P_{ij} \setminus P_{ij}^{(0)}) = v(Q_{ij}^{\{n\}}).$$

(iii) Let $p \in Q_{ij}^{\{k\}}$. Then by definition, p has no intermediate nodes x_r such that $r > k$. But then either $p \in Q_{ij}^{\{k-1\}}$ which means that p does not also have x_k as intermediate node, or p has x_k as intermediate node, in which case, p may use x_k as an intermediate node more than once, i.e. the path p can be factorized into paths in $Q_{ij}^{\{k-1\}}$ as

$$p = ab \quad \text{or} \quad p = a\, c_1\, c_2 \ldots c_t b \quad ,$$

where $a \in Q_{ik}^{\{k-1\}}$, $b \in Q_{kj}^{\{k-1\}}$ and $c_1, c_2, \ldots, c_t \in Q_{kk}^{\{k-1\}}$. Conversely, such a path p is always a path in $Q_{ij}^{\{k\}}$. Therefore

$$Q_{ij}^{\{k\}} = Q_{ij}^{\{k-1\}} \cup Q_{ik}^{\{k-1\}} \left(\bigcup_{t=0}^{\infty} (Q_{kk}^{\{k-1\}})^t \right) Q_{kj}^{\{k-1\}}$$

and hence the required result follows from (4.5) and theorem 4.1 above. ∇

We note here the above theorem is a generalization of a result obtained by McNaughton and Yamada (1960) in automata theory.

4.2 Compatibility of Networks with P-Spaces

In the definition of a path problem to be given in the next chapter, we shall find it necessary to impose the condition that $rv(P_{ij}) = r(v(P_{ij}))$ is a well-defined multiset for a given network $\mathcal{N}$ over a monoid (X,o) and a given p-space $(X, o, \mathcal{V}, r)$. That $rv(P_{ij})$ may not be well defined can best be seen from the following example.

Let $\mathcal{N}$ be a network over the additive group $(R,+)$ of real numbers and suppose that there exists at least one closed path in $\mathcal{N}$ whose label a is negative. Then we claim that $\min(v(P_{ij}))$ is not defined, i.e. $v(P_{ij}) \notin \mathcal{W}_R$, where $\mathcal{W}_R$ denotes the set of all well ordered multisets of N_∞^R and ϕ. For we know from corollary 2.1 that $v(P_{ij}) \in \mathcal{W}_R$ iff $x \geqslant 0$ for every $x \in v(P_{ij})$. But this is not possible since $a \in v(P_{ij})$ and $a < 0$ was assumed. Therefore, the claim is justified.

From the above example, we see that the question concerning the definability of $rv(P_{ij})$ for a given network and a given p-space $(X, o, \mathcal{V}, r)$ is closely related to the nature of the labels of closed paths in the network. Thus any network $\mathcal{N}$ in which the labels of closed paths are of the wrong nature is in some sense not compatible with the given p-space $(X, o, \mathcal{V}, r)$. It turns out that this troublesome nature of the labels of closed paths in a network can be formally described in terms of closed multisets of the hereditary semiring $\mathcal{V}$ of the p-space and the question of compatibility can be translated into the question of whether or not the multiset $v(P_{ii} \setminus \theta_i)$ is closed in $\mathcal{V}$. To prove this result, let us first define formally what we mean by compatibility.

DEFINITION 4.3 Let $\mathcal{N}$ be a given network over a monoid (X,o) and $(X, o, \mathcal{V}, r)$ be a given p-space. Then the network $\mathcal{N}$ is said to be <u>compatible</u> with the p-space iff $v(P) \in \mathcal{V}$, where P is the set of all paths in $G(\mathcal{N})$.

Since our previous understanding of compatibility from the above example is that $v(P_{ij}) \varepsilon \mathcal{V}$ so that $rv(P_{ij})$ is defined for all i,j, definition 4.3 might seem somewhat surprising. However, this definition is in fact in agreement with our previous understanding as the following lemma shows.

LEMMA 4.2 Let $\mathcal{N}$ be a given network over a monoid (X,o) and $(X, o, \mathcal{V}, r)$ be a given p-space. Then $\mathcal{N}$ is compatible with the p-space iff $v(P_{ij}) \varepsilon \mathcal{V}$ for all i,j, where P_{ij} is the set of all paths in $G(\mathcal{N})$ which begin at x_i and end at x_j.

PROOF Suppose first that $\mathcal{N}$ is compatible with $(X, o, \mathcal{V}, r)$ i.e. $v(P) \varepsilon \mathcal{V}$, where P is the set of all paths in $G(\mathcal{N})$. But by lemma 4.1, $P_{ij} \subseteq P$ for all i,j, implies that $v(P_{ij}) \subseteq v(P)$ for all i,j,. Therefore, it follows from the hereditary property of $\mathcal{V}$ that $v(P_{ij}) \varepsilon \mathcal{V}$ for all i,j .

Now suppose that $v(P_{ij}) \varepsilon \mathcal{V}$ for all i,j,. But then by (4.4), we have

$$v(P) = v(\bigcup_{i,j} P_{ij}) = \biguplus_{i,j} v(P_{ij}) \varepsilon \mathcal{V} \quad \text{as required.} \qquad \nabla$$

THEOREM 4.4 Let $\mathcal{N}$ be a given network over a monoid (X,o) and $(X,o,\mathcal{V},r)$ be a given p-space. Then $\mathcal{N}$ is compatible with the p-space iff $v(P_{ii} \setminus \theta_i)$ is a closed multiset of $\mathcal{V}$ for all i , where P_{ii} denotes the set of all closed paths in $G(\mathcal{N})$ which begin and end at x_i and θ_i is the null path for x_i in $G(\mathcal{N})$.

PROOF Suppose first that $\mathcal{N}$ is compatible with the p-space, i.e. $v(P) \varepsilon \mathcal{V}$. Since $v(P_{ii} \setminus \theta_i) \subseteq v(P)$ for all i, we have

$v(P_{ii} \setminus \theta_i) \in \mathscr{V}$ for all i. Now let $B \subseteq v(P_{ii} \setminus \theta_i)$. Then $B = v(C)$ for some $C \subseteq P_{ii} \setminus \theta_i$. Since $P_{ii} \setminus \theta_i$ contains all the non-null closed paths which begin and end at node x_i, it follows that $\bigcup_{k=1}^{\infty} C^k \subseteq P_{ii} \setminus \theta_i$, and hence by theorem 4.1 and lemma 4.1 above,

$$B^+ = \biguplus_{k=1}^{\infty} v(C)^k = v(\bigcup_{k=1}^{\infty} C^k) \subseteq v(P_{ii} \setminus \theta_i) \quad .$$

Therefore, it follows from the hereditary property of $\mathscr{V}$ that $B^+ \in \mathscr{V}$. Consequently, it follows from lemma 2.1 and lemma 2.2 above that $v(P_{ii} \setminus \theta_i)$ is closed in $\mathscr{V}$.

Now suppose that $v(P_{ii} \setminus \theta_i)$ is a closed multiset of $\mathscr{V}$ for all i . Then in view of lemma 4.2 above, it suffices to show that $v(P_{ij}) \in \mathscr{V}$ for all i,j. We first show that $v(Q_{ij}^{\{k\}}) \in \mathscr{V}$ for all $k \in \{0, 1, 2,\ldots,n\}$, where $Q_{ij}^{\{k\}}$ is as defined in theorem 4.3 above.

Since $\mathscr{F}_X \subseteq \mathscr{V}$ by the definition of a hereditary semiring it follows that $v(Q_{ij}^{\{0\}}) \in \mathscr{V}$ for all $i,j,$. In particular, $v(Q_{11}^{\{0\}}) \in \mathscr{V}$. But $v(Q_{11}^{\{0\}}) \subseteq v(P_{11} \setminus \theta_1)$, and hence by lemma 2.2 above, $v(Q_{11}^{\{0\}})^* \in \mathscr{V}$ also. It then follows from (iii) of theorem 4.3 that $v(Q_{ij}^{\{1\}}) \in \mathscr{V}$ for all i,j. The same argument can then be repeated to show that $v(Q_{ij}^{\{2\}}) \in \mathscr{V}$ for all i,j, and so on for $v(Q_{ij}^{\{3\}}),\ldots,v(Q_{ij}^{\{n\}})$.

Therefore, from (ii) of theorem 4.3, we conclude that $v(P_{ij} \setminus P_{ij}^{(0)}) \in \mathscr{V}$. But then by (4.4),

$$v(P_{ij}) = v(P_{ij} \setminus P_{ij}^{(0)}) \uplus v(P_{ij}^{(0)}) \in \mathscr{V} \text{ also.} \qquad \nabla$$

COROLLARY 4.1 (i) Every acyclic network over a monoid is compatible with any p-space. Here a network over a monoid is said to be <u>acyclic</u> iff it has no non-null closed paths.

(ii) Any network over a monoid (X,o) is compatible with the p-spaces (X,o,N_∞^X,r) and $(X,o,\mathcal{G}_X,r)$.

(iii) A network $\mathcal{N}$ over a monoid (X,o) is compatible with the p-space $(X,o,\mathcal{F}_X,r)$ iff $\mathcal{N}$ is acyclic.

PROOF (i) follows from the above theorem because for all i, $v(P_{ii} \setminus \theta_i) = \phi$ is always a closed multiset of the hereditary semiring $\mathcal{V}$ of any p-space $(X,o,\mathcal{V},r)$.

(ii) follows from the above theorem because by (i) and (iii) of theorem 2.11, $v(P_{ii} \setminus \theta_i)$ is always closed in N_∞^X and $\mathcal{G}_X$ for all i.

(iii) follows from the above theorem because by (ii) of theorem 2.11, $v(P_{ii} \setminus \theta_i)$ is closed in $\mathcal{F}_X$ iff $v(P_{ii} \setminus \theta_i) = \phi$ for all i.

COROLLARY 4.2 A network $\mathcal{N}$ over a locally finite monoid (X,o) is compatible with a given p-space $(X,o,\mathcal{N}_X,r)$ iff the label a of any elementary closed path in $\mathcal{N}$ is such that $a \neq e$.

PROOF Suppose first that $\mathcal{N}$ is compatible with $(X,o,\mathcal{N}_X,r)$. Then it follows from the above theorem that $v(P_{ii} \setminus \theta_i)$ is closed in $\mathcal{N}_X$ for all i. But then by theorem 2.12, we know that $x \neq e$ for every $x \in v(P_{ii} \setminus \theta_i)$ for all i, and hence the label a of any elementary closed path in $\mathcal{N}$ satisfies $a \neq e$.

Now assume that the label a of any elementary closed path in a network $\mathcal{N}$ over a locally finite monoid (X,o) is such that $a \neq e$. Then we claim that $x \neq e$ for every $x \in v(P_{ii} \setminus \theta_i)$ for all i. For let $x \in v(P_{ii} \setminus \theta_i)$, then $x \in v(p)$ for some $p \in P_{ii} \setminus \theta_i$. If p is an elementary closed path in $G(\mathcal{N})$, then x is the label of an elementary closed path in $\mathcal{N}$, and hence $x \neq e$ by assumption. So we may suppose that p is non-elementary. But then we may write $p = p_1\omega_1 q_1$ for some elementary closed pathe ω_1 in $G(\mathcal{N})$ as in the contraction process explained in section 0.4 above. Consequently, $v(p) = v(p_1) \circ v(q_1)$. Hence $x = y_1 \circ a_1 \circ z_1$ for some $y_1 \in v(p_1)$, $a_1 \in v(\omega_1)$, $z_1 \in v(q_1)$. But then $x \neq e$ because $a_1 \neq e$ by assumption. For suppose otherwise, then $y_1 \circ a_1 \circ z_1 = e$ would imply that $y_1 = a_1 = z_1 = e$ because $(X,\circ)$ is a locally finite monoid (see section 0.2), a contradiction. Therefore, $x \neq e$ for every $x \in v(P_{ii} \setminus \theta_i)$ as claimed. But from theorem 2.12, this means that $v(P_{ii} \setminus \theta_i)$ is closed in $\mathcal{N}_X$, and hence $\mathcal{N}$ is compatible with the p-space $(X,\circ,\mathcal{N}_X,r)$ as required. ∇

COROLLARY 4.3 Let $\mathcal{N}$ be a network over a totally ordered monoid $(X,\leq,\circ)$ which also satisfies the condition that $x^2 = x$ always implies $x = e$, then for $\mathcal{N}$ to be compatible with the p-space $(X,\circ,\mathcal{W}_X,r)$ (or$(X,\circ,\mathcal{W}'_X,r)$), it is necessary that the label a of any elementary closed path in $\mathcal{N}$ satisfies $a \geq e$ (or $a \leq e$). On the other hand, this condition is sufficient for $\mathcal{N}$ to be compatible with $(X,\circ,\mathcal{W}_X,r)$ (or $(X,\circ,\mathcal{W}'_X,r)$) in the case where $(X,\circ)$ is an Archimedean totally ordered monoid.

PROOF Suppose first that $\mathcal{N}$ is compatible with $(X,\circ,\mathcal{W}_X,r)$. Then it follows from the above theorem that $v(P_{ii} \setminus \theta_i)$ is closed in $\mathcal{W}_X$ for all i. But then by theorem 2.13 above, we have $x \geq e$ for every $x \in v(P_{ii} \setminus \theta_i)$ for all i, and hence the label a of any elementary closed path in $\mathcal{N}$ must satisfy $a \geq e$ as required.

Now suppose that the label a of any elementary closed path in a network $\mathcal{N}$ over an Archimedean totally ordered monoid $(X,\circ)$ satisfies $a \geq e$. Then we claim that $x \geq e$ for every $x \in v(P_{ii} \setminus \theta_i)$ for all i. For let $x \in v(P_{ii} \setminus \theta_i)$, then $x \in v(p)$ for some $p \in P_{ii} \setminus \theta_i$. If p is an elementary closed path in $G(\mathcal{N})$, then x is the label of an elementary closed path in $\mathcal{N}$, and hence $x \geq e$ by assumption. So we may assume that p is non-elementary. But then by the contraction process explained in section 0.2, we may factorize p as follows

$$p = p_1\omega_1 q_1 \quad , \quad p_1q_1 = p_2\omega_2 q_2 , \ldots, p_{s-1}q_{s-1} = p_s\omega_s q_s$$ where p_sq_s is a contraction of p .

Consequently, $v(p) = v(p_1) \circ v(\omega_1) \circ v(q_1)$ implies that $x = y_1 \circ a_1 \circ z_1$ for some $y_1 \in v(p_1)$, $a_1 \in v(\omega_1)$ and $z_1 \in v(q_1)$. Since $a_1 \geqslant e$ by assumption, it follows that $x \geqslant y_1 \circ z_1$. But $y_1 \circ z_1 \in v(p_1) \circ v(q_1) = v(p_1q_1) = v(p_2) \circ v(\omega_2) \circ v(q_2)$ implies that

$$y_1 \circ z_1 = y_2 \circ a_2 \circ z_2$$

for some $y_2 \in v(p_2)$, $a_2 \in v(\omega_2)$ and $z_2 \in v(q_2)$

Again, since $a_2 \geqslant e$ by assumption, it follows that $y_1 \circ z_1 \geqslant y_2 \circ z_2$ and hence $x \geqslant y_2 \circ z_2$.

By repeating the above argument for $y_2 \circ z_2$ and so on, we obtain $x \geqslant y_2 \circ z_2 \geqslant y_3 \circ z_3, \ldots, y_{s-1} \circ z_{s-1} \geqslant y_s \circ z_s$.

But $y_s \circ z_s \in v(p_s) \circ v(q_s) = v(p_sq_s)$, and p_sq_s is necessarily an elementary closed path in $G(\mathcal{N})$, it follows that $y_s \circ z_s \geqslant e$, and hence $x \geqslant e$ as required. Consequently, $v(P_{ii} \setminus \theta_i)$ is closed in $\mathcal{W}_X$ by theorem 2.13, and therefore $\mathcal{N}$ is compatible with $(X,\circ,\mathcal{W}_X,r)$.

The case for the p-space $(X,\circ,\mathcal{W}'_X,r)$ can be proved in a dually fashion. ∇

COROLLARY 4.4 Let $\mathcal{N}$ be a network over a totally ordered cancellative monoid $(X,\leqslant,\circ)$ which is also conditionally complete. Then $\mathcal{N}$ is compatible with $(X,\circ,\mathcal{W}_X,r)$ (or $(X,\circ,\mathcal{W}'_X,r)$) iff the label a of any elementary closed path in $\mathcal{N}$ satisfies $a \geqslant e$ (or $a \leqslant e$).

PROOF This is a special case of corollary 4.3 above because any cancellative monoid has the property that $x^2 = x$ implies $x = e$, and any totally ordered cancellative monoid which is also conditionally complete was seen in the proof of corollary 2.1 above to be Archimedean. ∇

COROLLARY 4.5 Any network $\mathcal{N}$ over an Archimedean totally ordered group $(X, \leq, o)$ is compatible with the p-space $(X, o, \mathcal{U}_X, r)$ (or $(X, o, \mathcal{U}'_X, r)$) where $\mathcal{U}_X = \mathcal{C}_X \cap \mathcal{N}_X \cap \mathcal{W}_X$ (or $\mathcal{U}'_X = \mathcal{C}_X \cap \mathcal{N}_X \cap \mathcal{W}'_X$) iff the label a of any elementary closed path in $\mathcal{N}$ satisfies $a > e$ (or $a < e$).

PROOF Suppose first that $\mathcal{N}$ is compatible with the p-space $(X, o, \mathcal{U}_X, r)$. Then it follows from the above theorem that $v(P_{ii} \setminus \theta_i)$ is closed in $\mathcal{U}_X$ for all i. Now by (ii) of corollary 4.1, $v(P_{ii} \setminus \theta_i)$ is always closed in $\mathcal{C}_X$, and by theorem 2.14 , $v(P_{ii} \setminus \theta_i)$ is closed in $\mathcal{N}_X \cap \mathcal{W}_X$ implies $x > e$ for every $x \in v(P_{ii} \setminus \theta_i)$. Therefore, in particular, $a > e$ for every label a of any elementary closed path in $\mathcal{N}$.

Now suppose that the label a of any elementary closed path in $\mathcal{N}$ satisfies $a > e$. Then by an argument similar to corollary 4.3 above, we can show by using, in addition, the cancellativity assumption that $x > e$ for every $x \in v(P_{ii} \setminus \theta_i)$. But by theorem 2.14, this means that $v(P_{ii} \setminus \theta_i)$ is closed in $\mathcal{N}_X \cap \mathcal{W}_X$ and hence in $\mathcal{C}_X \cap \mathcal{N}_X \cap \mathcal{W}_X$ also (see (ii) of corollary 4.1 above). Therefore $\mathcal{N}$ is compatible with the p-space $(X, o, \mathcal{U}_X, r)$ as required.

The case for $\mathcal{U}'_X$ can be shown dually. ∇

4.3 Stability of P-Spaces with respect to Networks

In many concrete instances of path problems to be defined in the next chapter, the problem of determining $rv(P_{ij})$ for a given network over a monoid (X, o) which is also compatible with a given p-space $(X, o, \mathcal{V}, r)$ is actually equivalent to the computation of $rv(P_{ij}^{[n_0]})$ for some positive integer n_0. For instance, consider again the p-space $(R, +, \mathcal{W}_R, \min)$, where $(R, +)$ is the additive group of real numbers.

As seen in the beginning of the previous section (also corollary 4.4), any network $\mathcal{N}$ which is compatible with this p-space must satisfy the condition that the label a of any elementary closed path in $\mathcal{N}$ is such that $a > 0$. Therefore, the label $b \in v(p)$ for any $p \in P_{ij}$ must satisfy $b \geq a$ for some

$a \in v(\bar{p})$ where $\bar{p}$ is a contraction of p in $G(\mathcal{N})$. Consequently, for any $b \in v(P_{ij})$, $b \geq a$ for some $a \in v(E_{ij})$ where E_{ij} is the set of all elementary closed paths in $G(\mathcal{N})$. Therefore, $b \geq a_0 = \min v(E_{ij})$ for all $b \in v(P_{ij})$.

Now since $a_0 \in v(P_{ij})$ also, it follows that

$$a_0 = \min v(P_{ij}) \quad ,$$

i.e. $$\min v(P_{ij}) = \min v(E_{ij}) \quad .$$

Now an elementary <u>open</u> path in a network with n nodes cannot have more than (n-1) arcs, and hence $v(E_{ij}) \subseteq v(P_{ij}^{[n-1]})$ for $i \neq j$. By an argument similar to that used above, we then have

$$\min v(P_{ij}^{[n-1]}) = \min v(E_{ij}) \quad \text{for } i \neq j$$

Now for $i = j$, $P_{ij}^{[n-1]}$ contains the null path θ_i, and since $v(\theta_i) = \{0\}$, we have

$$\min v(P_{ij}^{[n-1]}) = \{0\} = \min v(E_{ij}) \quad .$$

Therefore, $\min v(P_{ij}) = \min v(P_{ij}^{[n-1]}) = \min v(E_{ij})$ for all i,j as required.

Thus in this section, we shall concern ourselves with the problem of finding sufficient conditions for the above situation in general.

Its relevance to the solution of path problems will be shown in the next chapter. Now for convenience, let us introduce the following.

DEFINITION 4.4 Let $\mathcal{N}$ be a network over a monoid (X,o). Then a p-space $(X,o,\mathcal{V},r)$ is said to be <u>n_0-stable</u> with respect to $\mathcal{N}$ iff for some positive integer n_0,

$$\text{(4.6)} \qquad rv(P_{ij}^{[n_0+1]}) = rv(P_{ij}^{[n_0]}) \qquad \text{for all } i,j.$$

Moreover, the p-space is said to be <u>completely n_0-stable</u> with respect to $\mathcal{N}$ iff $\mathcal{N}$ is compatible with the p-space and also

$$\text{(4.7)} \qquad rv(P_{ij}) = rv(P_{ij}^{[n_0]}) \qquad \text{for all } i,j.$$

Now from a given matrix $A \in \mathcal{M}_n(\mathcal{V})$, let us define a matrix $A_r \in \mathcal{M}_n(\mathcal{V}_r)$ by

$$\text{(4.8)} \qquad (A_r)_{ij} = r(A_{ij}) \qquad \text{for all } i,j.$$

Using this definition, we can now obtain

LEMMA 4.3 Let $\mathcal{N}$ be a network over a monoid (X,o) and M its label matrix (as defined by (4.1) above). Then a p-space $(X,o,\mathcal{V},r)$ is n_0-stable with respect to $\mathcal{N}$ iff $M_r^{[n_0]} = M_r^{[n_0+1]}$, where $M_r^{[k]} = I_r \oplus M_r \oplus M_r^{\,2} \oplus \ldots \oplus M_r^{\,k}$, and completely n_0-stable with respect to its compatible network $\mathcal{N}$ iff $M_r^* = M_r^{[n_0]}$, where $M_r^* = (M^*)_r$.

PROOF In view of definition 4.4, it suffices to show that $(M_r^{[k]})_{ij} = rv(P_{ij}^{[k]})$ for all i,j, and all $k \in N$, and that

$(M_r^*)_{ij} = rv(P_{ij})$ for all i,j. The latter follows easily from (4.8) and (iii) of theorem 4.2. One way of proving the former is to consider the matrix M_r as the arc-value matrix of the graph $G(M_r)$ over the path algebra $(\mathcal{V}_r, \oplus, \odot)$ and note that all the paths in $G(M_r)$ are exactly those in $G(\mathcal{N})$. Hence the result follows if we can show that $(M_r^{[k]})_{ij} = rv(P_{ij}^{[k]})$, where $P_{ij}^{[k]}$ is defined with respect to $G(M_r)$. But the right-hand side is just another way of writing $\bigoplus_{p \in P_{ij}^{[k]}} v(p)$,

and hence the required result follows from (1.4) above. Alternatively, one can use (ii) of theorem 4.3 and the following consequence of (4.8):

$$(A \oplus B)_r = A_r \oplus B_r \quad \text{and} \quad (A \odot B)_r = A_r \odot B_r .$$ ∇

LEMMA 4.4 Let $\mathcal{N}$ be a network over a monoid (X,o). A p-space $(X, o, \mathcal{V}, r)$ is n_0-stable with respect to $\mathcal{N}$ iff

$$(4.9) \qquad rv(P_{ij}^{[s]}) = rv(P_{ij}^{[n_0]}) \qquad \text{for all} \quad s \geq n_0$$

Hence, if the p-space is complete and $\mathcal{N}$ is compatible with the p-space, then this p-space is also completely n_0-stable with respect to $\mathcal{N}$.

PROOF From lemma 4.3 above, we see that (4.9) is equivalent to $M_r^{[s]} = M_r^{[n_0]}$ for all $s \geq n_0$. This we can prove by mathematical induction as follows. For $s = n_0$, the result is trivially so. So let us suppose that the result is true for all s such that $n_0 \leq s < t$ and show that it is also true for $s = t$ as follows.

$$M_r^{[t]} = I_r \oplus M_r \otimes M_r^{[t-1]}$$

$$= I_r \oplus M_r \otimes M_r^{[n_0]} , \quad \text{by induction hypothesis}$$

$$= M_r^{[n_0+1]}$$

$$= M_r^{[n_0]}$$

Hence $M_r^{[s]} = M_r^{[n_0]}$ for all $s \geq n_0$ as required.

For the second part of the lemma, let us first use mathematical induction to show that

$$(4.10) \quad r(v(P_{ij}^{(k)}) \uplus v(P_{ij}^{[n_0]})) = rv(P_{ij}^{[n_0]}) \quad \text{for all } k \geq n_0 + 1$$

For $k = n_0 + 1$, this is just (4.6) above. So let us assume the result to hold for all k such that $n_0 + 1 \leq k < t$ and show that it also holds for $k = t$ as follows.

From the first part of the lemma, we know that

$$rv(P_{ij}^{[t-1]}) = rv(P_{ij}^{[n_0]}) \quad \text{for all } t \geq n_0 + 1,$$

and hence by (3.2), we have

$$r(v(P_{ij}^{(t)}) \uplus v(P_{ij}^{[n_0]})) = r(v(P_{ij}^{(t)}) \uplus rv(P_{ij}^{[n_0]}))$$

$$= r(v(P_{ij}^{(t)}) \uplus rv(P_{ij}^{[t-1]}))$$

$$= r(v(P_{ij}^{(t)}) \uplus v(P_{ij}^{[t-1]}))$$

$$= rv(P_{ij}^{[t]})$$

$$= rv(P_{ij}^{[n_0]})$$

Therefore, (4.10) is valid for all $k \geq n_0 + 1$.

Now by theorem 4.1, we have

$$v(P_{ij} \setminus P_{ij}^{[n_0]}) = v(\bigcup_{k=n_0+1}^{\infty} P_{ij}^{(k)}) = \biguplus_{k=n_0+1}^{\infty} v(P_{ij}^{(k)}) \quad ,$$

and hence it follows from the completeness assumption that

$$rv(P_{ij}) = r(v(P_{ij} \setminus P_{ij}^{[n_0]}) + v(P_{ij}^{[n_0]})) = rv(P_{ij}^{[n_0]})$$

as required. ∇

THEOREM 4.5 Every finite p-space is completely (n-1)-stable with respect to any network compatible with the p-space.

PROOF First, let us note from (iii) of corollary 4.1 above that a network $\mathcal{N}$ is compatible with a finite p-space iff $\mathcal{N}$ is acyclic. Now a network $\mathcal{N}$ is acyclic iff its graph $G(\mathcal{N})$ is acyclic. Since the order of each path in an acyclic graph is at most n-1, it follows that $P_{ij}^{(k)} = \phi$ for all $k \geq n$. Consequently, $P_{ij} = P_{ij}^{[s]}$ for all $s \geq n-1$ and hence the required result follows. ∇

THEOREM 4.6 Every q-stationary, idempotent and intensive p-space is completely n_0-stable with respect to any network compatible with the p-space.

PROOF Let $(X,o,\mathcal{V},r)$ be a q-stationary, idempotent and intensive p-space. Then we claim that $rv(P_{ij}) \in \mathcal{R}_X$ for all i,j. In order to justify this claim, let us first note the following consequence of the assumption that the p-space is intensive

(4.11) For any $A,B \in \mathcal{V}$ such that $r(A)$, $r(B) \in \mathcal{F}_X$, then both $r(A) \oplus r(B)$ and $r(A) \odot r(B)$ belong to $\mathcal{F}_X$ also

This follows because

(i) $r(A) \oplus r(B) = r(r(A) \uplus r(B)) \subseteq r(A) \uplus r(B) \in \mathcal{F}_X$, and

(ii) $r(A) \odot r(B) = r(r(A) \circ r(B)) \subseteq r(A) \circ r(B) \in \mathcal{F}_X$.

Now $rv(P_{ij}) = rv(P_{ij} \setminus P_{ij}^{(0)}) \oplus rv(P_{ij}^{(0)})$, and since $rv(P_{ij}^{(0)}) \subseteq v(P_{ij}^{(0)}) \in \mathcal{F}_X$ implies that $rv(P_{ij}^{(0)}) \in \mathcal{F}_X$, our claim can be justified by using (4.11) above if we can show that $rv(P_{ij} \setminus P_{ij}^{(0)}) \in \mathcal{F}_X$ also.

This we now do by showing that $rv(Q_{ij}^{\{n\}}) \in \mathcal{F}_X$, since $rv(Q_{ij}^{\{n\}}) = rv(P_{ij} \setminus P_{ij}^{(0)})$ by (ii) of theorem 4.3.

Now from (i) of theorem 4.3 and the assumption of intensitivity,

$$rv(Q_{ij}^{\{0\}}) = rv(P_{ij}^{(1)}) \subseteq v(P_{ij}^{(1)}) \in \mathcal{F}_X \text{ implies that}$$

$rv(Q_{ij}^{\{0\}}) \in \mathcal{F}_X$ for all i,j, and from (iii) of theorem 4.3, we have

$$rv(Q_{ij}^{\{1\}}) = rv(Q_{ij}^{\{0\}}) \oplus rv(Q_{i1}^{\{0\}}) \odot r(v(Q_{11}^{\{0\}})^*) \odot rv(Q_{1j}^{\{0\}})$$

for all i,j.

Consequently, it follows from (4.11) that $rv(Q_{ij}^{\{1\}}) \in \mathcal{F}_X$ if we can show that $r(v(Q_{11}^{\{0\}})^*) \in \mathcal{F}_X$. But the p-space is assumed to be q-stationary, and hence

$$r(v(Q_{11}^{\{0\}})^*) = r(e) \oplus rv(Q_{11}^{\{0\}}) \oplus \dots \oplus rv(Q_{11}^{\{0\}})^q \in \mathcal{F}_X \text{ as required.}$$

Similarly, one can then show that $rv(Q_{ij}^{\{2\}}) \in \mathcal{F}_X$ for all i,j, and so on for $rv(Q_{ij}^{\{3\}}),\ldots, rv(Q_{ij}^{\{n\}})$.

Therefore $rv(P_{ij}) \in \mathcal{F}_X$ as claimed. Now since $rv(P_{ij}) \subseteq v(P_{ij})$, it follows that there exists a finite subset $H \subsetneq P_{ij}$ such that for any $a \in rv(P_{ij})$, $a \in v(H)$ also. But then $a \in v(P_{ij}^{[n_0]})$ if n_0 is chosen to be the maximum order of paths in H. Therefore, $r(a) \prec rv(P_{ij}^{[n_0]})$ by corollary 3.1. Since this holds for all $a \in rv(P_{ij})$, it follows from property (0.5) of $\prec$ that $rv(P_{ij}) \prec rv(P_{ij}^{[n_0]})$.

But $v(P_{ij}^{[n_0]}) \subseteq v(P_{ij})$ implies that $rv(P_{ij}^{[n_0]}) \prec rv(P_{ij})$ by corollary 3.1 also, and hence

$$rv(P_{ij}) = rv(P_{ij}^{[n_0]}) \quad \text{as required.} \qquad \nabla$$

For the next four theorems, it is convenient to introduce the following concept analogous to that given by Roy (1975).

DEFINITION 4.5 A multiset $A \in \mathcal{V}$ of a p-space $(X,o,\mathcal{V},r)$ is said to be <u>q-absorptive</u> with respect to r iff for each q-tuple $(a_1,a_2,\ldots a_q)$ of elements in A, we have

$$r\{e,a_1,a_1 \circ a_2,\ldots,a_1 \circ a_2 \circ \ldots \circ a_q\} = r\{e,a_1,a_1 \circ a_2,\ldots,a_1 \circ a_2 \circ \ldots a_{q-1}\}$$

Note that by this definition, the set of all q-absorptive multisets of $\mathcal{V}$ (with respect to r) is hereditary, i.e. every submultiset of a q-absorptive multiset is also q-absorptive.

LEMMA 4.5 A multiset $A \in \mathcal{V}$ of a p-space $(X,o,\mathcal{V},r)$ is q-absorptive with respect to r iff for each s-tuple $(a_1,a_2,\ldots,a_s)$ of elements in A, we have

$$r(a_1 \circ a_2 \circ \ldots \circ a_s) \prec r\{e,a_1,a_1 \circ a_2,\ldots, a_1 \circ a_2 \circ \ldots \circ a_{q-1}\}$$

for all $s \geq q$,

where $\prec$ denotes the psuedo-ordering of the path algebra $(\mathcal{V}_r, \oplus, \odot)$

PROOF Its validity follows from an argument similar to that used in the beginning of the proof of lemma 1.4 above.

DEFINITION 4.6 Let $\mathcal{N}$ be a network over a monoid (X,o) and $(X,o,\mathcal{V},r)$ be a p-space. Then $\mathcal{N}$ is said to be q-absorptive with respect to the p-space iff the multiset $v(\Omega)$ is q-absorptive with respect to r, where Ω denotes the set of all elementary closed paths in $G(\mathcal{N})$.

DEFINITION 4.7 Let $\mathcal{N}$ be a network over a monoid (X,o) and $(X,o,\mathcal{V},r)$ be a p-space. Then $\mathcal{N}$ is said to be q-regular with respect to the p-space iff for every elementary closed path ω in $G(\mathcal{N})$, $v(\omega)$ is q-absorptive with respect to r.

We note here that since $v(\omega) \subseteq v(\Omega)$, it follows that $v(\omega)$ is q-absorptive whenever $v(\Omega)$ has this property, i.e. definition 4.7 is more general than definition 4.6 above.

LEMMA 4.6 Let $\mathcal{N}$ be a network over a monoid (X,o) which is q-absorptive with respect to a given p-space $(X,o,\mathcal{V},r)$ and M its label matrix. Then the graph $G(M_r)$ over the path algebra $(\mathcal{V}_r, \oplus, \odot)$ is a q-absorptive graph (i.e. (1.16) is satisfied). Here M_r is defined from M by (4.8) above.

PROOF We have to show that condition (1.16) is satisfied, i.e.

$$v(\omega_1) \odot v(\omega_2) \odot \dots \odot v(\omega_q) \prec r(e) \oplus v(\omega_1) \oplus \dots + v(\omega_1) \odot v(\omega_2) \odot \dots \odot v(\omega_{q-1})$$

for every q-tuple $(\omega_1,\omega_2,\dots,\omega_q)$ of elementary closed paths in $G(M_r)$

Let $v(\omega_i) = \{a_{i1},a_{i2},\dots,a_{im_i}\}$ for all $i \in \{1,2,\dots,q\}$

Then each $x \in v(\omega_1) \odot v(\omega_2) \odot \ldots \odot v(\omega_q)$ is of the form

$$x = r(a_{1k_1} \circ a_{2k_2} \circ \ldots \circ a_{qk_q}) \text{ , where } k_i \in \{1,2,\ldots,m_i\} \text{ for all } i \in \{1,2,\ldots,q\}$$

But clearly, $(a_{1k_1}, a_{2k_2}, \ldots, a_{qk_q})$ is a q-tuple of elements in $v(\Omega)$ where Ω is the set of elementary closed paths in $G(M_r)$. Since $v(\Omega)$ is q-absorptive, it follows that

$$x \prec r\{e, a_{1k_1}, a_{1k_1} \circ a_{2k_2}, \ldots, a_{1k_1} \circ a_{2k_2} \circ \ldots a_{(q-1)k_{(q-1)}}\}$$

It then follows from properties (0.2) and (0.4) of $\prec$ that

$$x \prec r(e) \oplus v(\omega_1) \oplus \ldots \oplus v(\omega_1) \odot v(\omega_2) \odot \ldots \odot v(\omega_{q-1}).$$

Since this holds for all $x \in v(\omega_1) \odot v(\omega_2) \odot \ldots \odot v(\omega_q)$, and there are only finitely many such x , it follows from property (0.5) of $\prec$ that

$$v(\omega_1) \odot v(\omega_2) \odot \ldots \odot v(\omega_q) \prec r(e) \oplus v(\omega_1) \oplus \ldots \oplus v(\omega_1) \odot v(\omega_2) \odot \ldots \odot v(\omega_{q-1})$$

as required. ∇

LEMMA 4.7 Let $\mathcal{N}$ be a network over a monoid $(X,\circ)$ which is q-regular with respect to a given p-space $(X,\circ,\mathcal{V},r)$ and M be its label matrix. Then the graph $G(M_r)$ over the path algebra $(\mathcal{V}_r,\oplus,\odot)$ is a q-regular graph , i.e. it satisfies condition (1.14) above.

PROOF Its validity follows from a similar argument used in the proof of lemma 4.6 above by taking

$$v(\omega_1) = v(\omega_2) = \ldots = v(\omega_q) \quad .$$

∇

THEOREM 4.7 Every p-space $(X,o,\mathcal{V},r)$ is (n-1)-stable with respect to any network $\mathcal{N}$ which is 1-absorptive with respect to the p-space. Moreover, if the p-space is complete and $\mathcal{N}$ is compatible with the p-space, then the p-space is completely (n-1)-stable with respect to $\mathcal{N}$.

PROOF By lemma 4.6, we know that the graph $G(M_r)$ over the path algebra $(\mathcal{V}_r, \oplus, \odot)$ is 1-absorptive, and therefore absorptive. Consequently, it follows from theorem 1.5 that the arc-value matrix M_r of the graph $G(M_r)$ is (n-1)-stable, i.e.

$$M_r^{[n-1]} = M_r^{[n]} \quad ,$$

and hence by lemma 4.3, the p-space is (n-1)-stable.
The rest of the proof follows from lemma 4.4 above. ∇

THEOREM 4.8 Every commutative p-space $(X,o,\mathcal{V},r)$ is (nq-1)-stable with respect to any network $\mathcal{N}$ which is q-absorptive with respect to the p-space. Moreover, if the p-space is complete and $\mathcal{N}$ is compatible with the p-space, then the p-space is completely (nq-1)-stable with respect to $\mathcal{N}$.

PROOF Its validity follows from lemma 4.6, theorem 1.7 and lemma 4.4 by an argument similar to that used in the proof of theorem 4.7 above. ∇

Note that theorem 4.8 only implies theorem 4.7 in the case where the p-space is commutative.

THEOREM 4.9 Every commutative p-space $(X,o,\mathcal{V},r)$ is n_0-stable with respect to any network $\mathcal{N}$ which is q-regular with respect to the p-space. Moreover, if $G(\mathcal{N})$ has t elementary closed paths, then

$$n_0 = nt(q-1) + (n-1) \quad .$$

Also, if the p-space is complete and $\mathcal{N}$ is compatible with the p-space, then the p-space is completely n_0-stable with respect to $\mathcal{N}$.

PROOF This follows from lemma 4.7, theorem 1.6 and lemma 4.4 above by an argument similar to that used in the proof of theorem 4.7.

The next theorem also requires the following

LEMMA 4.8 Let $(X,o,\mathcal{V},r)$ be a complete p-space. If A is a finite multiset which is closed in $\mathcal{V}$ and q-absorptive with respect to r , then $r(A^*)$ is finite and moreover, $r(A^*) = r(A^{[s]})$ for all $s \geq q-1$, where $A^{[s]} = \{e\} \uplus A \uplus A^2 \uplus \ldots \uplus A^s$.

PROOF Let $A = \{a_1,a_2,\ldots,a_k\}$. Then we can consider A to be the label matrix (of order 1×1) of the network $\mathcal{N}$ in figure 4.1 below. Therefore, the required result follows from theorem 4.2 if we can show that the p-space is completely $(q-1)$-stable with respect to $\mathcal{N}$. We note here that although the hypothesis of this theorem can be seen to be equivalent to the statement that $v(\Omega)$ is q-stable because $v(\Omega) = \{a_1,a_2,\ldots,a_k\}$, theorem 4.8 cannot be used to yield the required result because the p-space is not assumed to be commutative. So a proof of this result has to be obtained separately.

Let B be any finite subset such that $P_{11}^{[q-1]} \subseteq B \subseteq P_{11}$.

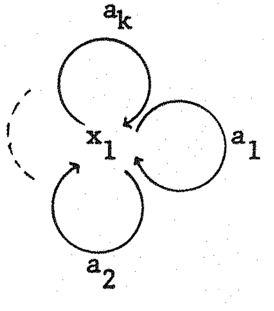

Figure 4.3

If $p \in B$ but $p \notin P_{11}^{[q-1]}$, then p must traverse more than $q - 1$ arcs in $\mathcal{N}$ and hence its label a can be written as $a = a_1 \circ a_2 \circ \ldots \circ a_s$ for some $s \geq q - 1$. Hence by lemma 4.5, we have

$$r(a) \prec r\{e, a_1, a_1 \circ a_2, \ldots, a_1 \circ a_2 \circ \ldots \circ a_{q-1}\} .$$

But $C = \{e, a_1, a_1 \circ a_2, \ldots, a_1 \circ a_2 \circ \ldots \circ a_{q-1}\} \subseteq v(P_{11}^{[q-1]})$, i.e. $C = v(H)$, say, for some $H \subseteq P_{11}^{[q-1]}$.

Therefore, by property (0.2) and (0.4) of $\prec$, we have

$$r(a) \prec rv(H) \oplus rv(P_{11}^{[q-1]} \setminus H) = rv(P_{11}^{[q-1]})$$

Since this holds for all $a \in v(B \setminus P_{11}^{[q-1]})$ and there are only finitely many such $r(a)$ in $rv(B \setminus P_{11}^{[q-1]})$ because B is assumed to be a finite set, it follows from property (0.5) of $\prec$ that $rv(B \setminus P_{11}^{[q-1]}) \prec rv(P_{11}^{[q-1]})$. But then

$$rv(B) = rv(B \setminus P_{11}^{[q-1]}) \oplus rv(P_{11}^{[q-1]}) = rv(P_{11}^{[q-1]}).$$

If we now choose $B = P_{11}^{[q]}$, we would then have

$rv(P_{11}^{[q]}) = rv(P_{11}^{[q-1]})$, which proves that the p-space is $(q-1)$ stable with respect to $\mathcal{N}$. The rest of the proof then follows from lemma 4.4 above. ∇

LEMMA 4.9 Let $(X, \circ, \mathcal{V}, r)$ be a complete and idempotent p-space. If $A \in \mathcal{V}$ is such that $r(A)$ is finite, closed in $\mathcal{V}$ and q-absorptive with respect to r, then $r(A^*)$ is finite, provided that $A^* \in \mathcal{V}$.

Moreover, $r(A^*) = r(r(A)^*) = r(A^{[q-1]})$.

PROOF By lemma 4.8 above, we have for all $s \geq q-1$,

$$r(r(A)^*) = r(r(A)^{[s]})$$

$$= r(A^{[s]}) \quad , \text{ by repeated use of (3.2) and (3.3)}$$

Hence by theorem 3.9, we have

$$r(A^*) = r(\biguplus_{k=0}^{\infty} A^k)$$

$$= \bigvee_{k=0}^{\infty} r(A^k)$$

$$= r(A^{[q-1]}) \quad , \text{ since } r(A^{[s]}) = r(A^{[q-1]}) \text{ for all } s \geq q-1.$$

$$= r(r(A)^*) \quad , \text{ as required.} \qquad \nabla$$

THEOREM 4.10 Let $(X,o,\mathscr{V},r)$ be a complete, idempotent and intensive p-space. Then this p-space is completely n_0-stable with respect to a network $\mathscr{N}$ which is compatible with the p-space and such that $v(P_{ii})$ is q-absorptive with respect to r for all i.

PROOF In view of the argument used at the end of the proof of theorem 4.6 above, it suffices to show that $rv(P_{ij})$ is a finite multiset or equivalently $rv(P_{ij} \setminus P_{ij}^{(0)})$ is a finite multiset. By (ii) of theorem 4.3, this is equivalent to showing that $rv(Q_{ij}^{\{n\}})$ is a finite multiset.

Now from (i) of theorem 4.3 and the assumption of intensitivity, $rv(Q_{ij}^{\{0\}}) = rv(P_{ij}^{(1)}) \subseteq v(P_{ij}^{(1)})$ implies that $rv(Q_{ij}^{\{0\}})$ is finite for all i,j, and from (iii) of theorem 4.3,

$$rv(Q_{ij}^{\{1\}}) = rv(Q_{ij}^{\{0\}}) \oplus rv(Q_{i1}^{\{0\}}) \odot r(v(Q_{11}^{\{0\}})^*) \odot rv(Q_{1j}^{\{0\}}) \quad \text{for all } i,j.$$

Since $v(Q_{11}^{\{0\}}) \subseteq v(P_{11})$, it follows that $v(Q_{11}^{\{0\}})$ is q-absorptive, and hence by lemma 4.8, $r(v(Q_{11}^{\{0\}})^*)$ is a finite multiset. Therefore by (4.11), $rv(Q_{ij}^{\{1\}})$ is finite for all i,j. In particular, $rv(Q_{22}^{\{1\}})$ is finite. But $rv(Q_{22}^{\{1\}}) \subseteq v(Q_{22}^{\{1\}}) \subseteq v(P_{22})$ (since the p-space is intensive) implies that $rv(Q_{22}^{\{1\}})$ is also q-absorptive . Therefore, by lemma 4.9, $r(v(Q_{22}^{\{1\}})^*)$ is a finite multiset. But then it follows from (4.11) and (iii) of theorem 4.3 that $rv(Q_{ij}^{\{2\}})$ is finite for all i,j.

Therefore, in particular, $rv(Q_{33}^{\{2\}})$ is finite. One can then use (4.11) and (iii) of theorem 4.3 to show in a similar manner as above that $rv(Q_{ij}^{\{3\}})$ is finite for all i,j. Likewise, we can therefore conclude that $rv(Q_{ij}^{\{4\}})$, ... ,$rv(Q_{ij}^{\{n\}})$ are all finite multisets. ∇

CHAPTER 5.

FUNDAMENTALS OF ALGEBRAIC METHODS FOR SOLVING PATH PROBLEMS

5.1 Path Problems Revisited

In chapter 1, we gave a number of problems which motivated their abstract study and considered several definitions of a path problem, each more general than the previous one. We also saw how all the given problems except problem 1.5 can be formulated as a path problem in accordance with definition 1.4 above. It was then suggested that a similar but more general approach to all these given problems is possible if one uses the concept of multisets. We then went on to investigate a number of concepts which were useful for the development of such an approach. In fact, all the results obtained in the preceding three chapters were carried out with this purpose in mind. Their relevance to the solution of path problems in accordance with the following definition will here be demonstrated in detail.

DEFINITION 5.1 Let $\mathcal{N}$ be a network over a monoid (X,o) which is compatible with a given p-space $(X,o,\mathcal{V},r)$. Then by a path problem, we mean the determination of $rv(P_{ij})$ or $rv(P_{ij} \setminus P_{ij}^{(0)})$ for one or more pairs (i,j), where P_{ij} and $P_{ij}^{(0)}$ are defined with respect to $G(\mathcal{N})$.

That this definition of a path problem is an extension of definition 1.4 can be easily seen to be a consequence of the following

THEOREM 5.1 Given a set $\mathcal{V}$ of subsets of a monoid (X,o) which has properties (1.26), (1.27) and is also closed with respect to union and complex product, and given a function r defined on $\mathcal{V}$ which satisfies (1.22) to (1.24) above, we can construct a p-space $(X, o, \mathcal{V}', r')$ such that $\mathcal{V} \subseteq \mathcal{V}'$ and $r'(A) = r(A)$ for all $A \in \mathcal{V}$.

PROOF Let $\mathcal{V}' = \{A \in N_\infty^X \mid d(A) \in \mathcal{V}\}$. Then $\mathcal{V} \subseteq \mathcal{V}'$ because $A = d(A) \in \mathcal{V}$ for any $A \in \mathcal{V}$. Now $\mathcal{V}'$ can be seen to be a hereditary semiring as follows

(i) $\mathcal{F}_X \subseteq \mathcal{V}'$ because for any $A \in \mathcal{F}_X$, $d(A)$ is a finite multiset and hence $d(A) \in \mathcal{V}$ by (1.26)

(ii) $\mathcal{V}'$ is a hereditary subset of N_∞^X. For let $A \in \mathcal{V}'$ and $B \subseteq A$. Then $d(B) \subseteq d(A)$, and hence by (1.27), $d(B) \in \mathcal{V}$ also.

(iii) $\mathcal{V}'$ is closed with respect to multisum and multiproduct because

$$d(A \uplus B) = d(A) \cup d(B) \in \mathcal{V}, \text{ and}$$
$$d(A \circ B) = d(A)\, d(B) \in \mathcal{V} \text{ whenever } d(A), d(B) \in \mathcal{V}$$

Now let $r' = rd$. Then clearly,

$$\begin{aligned} r'(A) &= r(d(A)) \\ &= r(A) \quad \text{for all } A \in \mathcal{V}. \end{aligned}$$

That r' is a reduction function can be seen as follows.

(i) $r'(\phi) = r(d(\phi)) = r(\phi) = \phi$ by (1.22)

(ii) $$\begin{aligned} r'(r'(A) \uplus B) &= rd(rd(A) \uplus B) \\ &= r(d(rd(A)) \cup d(B)) \\ &= r(rd(A) \cup d(B)) \quad \text{since } rd(A) \in \mathcal{V}. \end{aligned}$$

$= r(d(A) \cup d(B))$ by (1.23)

$= rd(A \uplus B)$

$= r'(A \uplus B)$

(iii) $r'(r'(A) \circ B) = rd(rd(A) \circ B)$

$= r(d(rd(A))\ d(B))$

$= r(rd(A)\ d(B))$ since $rd(A) \in \mathcal{V}$

$= r(d(A)\ d(B))$ by (1.24)

$= rd(A \circ B)$

$= r'(A \circ B)$

Similarly, $r'(A \circ r'(B)) = r'(A \circ B)$. ∇

Therefore, definition 1.4 is just a particular instance of definition 5.1. Moreover, definition 5.1 also includes problem 1.5 above. In fact, its corresponding p-space is $(N, +, \mathcal{F}_N, s)$ (see theorem 3.1).

Now from (iii) of theorem 4.2, we see that a path problem in a network is equivalent to the determination of $(M_r^*)_{ij}$ or $(M_r^+)_{ij}$, where $M_r^* = (M^*)_r$ and $M_r^+ = (M^+)_r$, using (4.8). Now since theorem 2.6 obviously holds for matrices over multisets, the matrix M^* is the least solution (with respect to $\subseteq$ as extended to matrices by (0.12) in terms of multiset-inclusion) of the equation

$$Z = M \circ Z \uplus I \quad \text{or} \quad Z = Z \circ M \uplus I \ .$$

Similarly, M^+ is the least solution of

$$Z = M \circ Z \uplus M \quad \text{or} \quad Z = Z \circ M \uplus M$$

Consequently, M_r^* is a solution of the matrix equation

$$Y = M_r \odot Y \oplus I_r \quad \text{or} \quad Y = Y \odot M_r \oplus I_r \quad ,$$

and M_r^+ is a solution of the matrix equation

$$Y = M_r \odot Y \oplus M_r \quad \text{or} \quad Y = Y \odot M_r \oplus M_r \ .$$

The above observation suggests that one can view the problem of determining M_r^* and M_r^+ as particular instances of a more general problem of solving the matrix equation

(5.1) $$Y = M_r \odot Y \oplus B \quad , \quad \text{or}$$

(5.2) $$Y = Y \odot M_r \oplus B \quad , \quad \text{where } B \in \mathcal{M}_n(\mathcal{V}_r)$$

In this way, a path problem can be seen to be a particular instance of a more general problem of determining $(M_r^* \odot B)_{ij}$ from (5.1) or $(B \odot M_r^*)_{ij}$ from (5.2) for one or more pairs (i,j).

For convenience, we shall here consider only the problem of determining $(M_r^* \odot B)_{ij}$ for one or more pairs (i,j) from (5.1), since the other can be considered in an analogous manner. Now this problem can be solved quite readily if the path algebra $(\mathcal{V}_r, \oplus, \odot)$ forms a field or is embeddable in a field (e.g. an integral domain), since in this case most of the available methods in linear algebra can be employed to solve (5.1) (see e.g. Fox(1964)). Since, there are relatively few examples of path problems whose corresponding path algebras form a field or are embeddable in a field, it is more significant to have available methods for solving (5.1) in a more general situation. This will be carried out in the rest of this section. However, it must be said at the outset that we shall not concern ourselves with the computational complexity of the methods to be discussed here since our aim is to present the fundamentals of algebraic methods for solving path problems. Moreover, our presentation will be limited to those methods which can be unified from the view point of solving a system of linear equations.

5.2 Elimination Methods

Let us first restrict ourselves to the problem of determining a particular column of the matrix $M_r^* \odot B$, say $M_r^* \odot b$ where b is a column of B. From (5.1), it follows that $M_r^* \odot b$ is a solution of the following system of equations

$$y = M_r \odot y \oplus b \tag{5.3}$$

Note that if b is the jth column of B, then y is the jth column of Y in (5.1). Thus (5.1) can be viewed as n systems of equations, each corresponding to a colum of B.

Perhaps the simplest and most well known method of solving (5.3) in the case where the path algebra $(\mathcal{V}_r, \oplus, \odot)$ forms a field or is embeddable in a field is the elimination method of Gauss in linear algebra (see e.g. Fox (1964)). It is therefore of interest to have a similar method developed to solve (5.3) in a more general situation. The most important example of this situation is where the path algebra is idempotent as a semiring and hence cannot be a non-trivial ring, let alone a field (see section 0.2). However, as will be seen below, methods analogous to those of Gauss and Jordan can be developed to solve (5.3) in this case. That these methods are available is a consequence of the pioneer work of Carré (1971).

For convenience of exposition, we shall first assume that the p-space $(X, o, \mathcal{V}, r)$ is idempotent and complete. Let us now write A for M_r in (5.3), i.e. $y = A \otimes y \oplus b$ which can also be written out in full as follows

$$(5.4) \quad \begin{cases} y_1 = A_{11} \otimes y_1 \oplus A_{12} \otimes y_2 \oplus \quad \cdots \quad \oplus A_{1n} \otimes y_n \oplus b_1 \\ y_2 = A_{21} \otimes y_1 \oplus A_{22} \otimes y_2 \oplus \quad \cdots \quad \oplus A_{2n} \otimes y_n \oplus b_2 \\ \vdots \\ y_n = A_{n1} \otimes y_1 \oplus A_{n2} \otimes y_2 \oplus \quad \cdots \quad \oplus A_{nn} \otimes y_n \oplus b_n \end{cases}$$

First let us assume that the matrix A is upper-triangular i.e. $A_{ij} = \theta$ for all $i > j$. Then (5.4) can be rewritten as

$$(5.5) \quad \begin{cases} y_1 = A_{11} \otimes y_1 \oplus A_{12} \otimes y_2 \oplus \quad \cdots \quad \oplus A_{1n} \otimes y_n \oplus b_1 \\ y_2 = \qquad\qquad A_{22} \otimes y_2 \oplus \quad \cdots \quad \oplus A_{2n} \otimes y_n \oplus b_2 \\ \vdots \\ y_n = \qquad\qquad\qquad\qquad\qquad\qquad A_{nn} \otimes y_n \oplus b_n \end{cases}$$

If we now write $H^{*r} = \bigvee_{k=0}^{\infty} H^k$, we see that $(A_{nn})^{*r} \otimes b_n$ is the least solution of the last equation of (5.5) because it obviously satisfies the last equation of (5.5) and moreover, we have
$y_n = A_{nn}^{k+1} \otimes y_n \oplus (r(e) \oplus A_{nn} \oplus \ldots \oplus A_{nn}^k) \otimes b_n$ for all $k \in N$, and hence

$$(r(e) \oplus A_{nn} \oplus \ldots \oplus A_{nn}^k) \otimes b_n < y_n \quad \text{for all} \quad k \in N$$

which in turn implies that $(A_{nn})^{*r} \otimes b_n < y_n$. Note that $(A_{nn})^{*r}$ exists because by theorem 4.4, $M^*_{nn} \in \mathcal{V}$ and

$$r(M^*_{nn}) = \bigvee_{k=0}^{\infty} r(M_{nn})^k \quad \text{by theorem 3.1}$$

$$= \bigvee_{k=0}^{\infty} (A_{nn})^k$$

$$= (A_{nn})^{*r}$$

We can now obtain the least solution of the $(n-1)^{th}$ equation in the same way by substituting $(A_{nn})^{*r} \otimes b$ for y_n in the equation. Similarly, we can obtain the least solution of the $(n-2)^{th}$ equation and so on until the least solution of the first equation is obtained. As will be proved below, these least solutions constitute the required least solution of (5.5). For convenience, we shall refer to the above method as the <u>generalized</u> <u>back-substitution</u> method, since it obviously resembles the well known back-substitution process in linear algebra (see e.g. Fox (1964))

THEOREM 5.2 The back-substitution method applied to (5.5) yields the least solution of (5.5).

PROOF From the above description of the back-substitution method, it is clear that we first obtain the least solution $\tilde{y}_n$ of the n^{th} equation, then the least solution $\tilde{y}_{n-1}$ of the $(n-1)^{th}$ equation given $\tilde{y}_n$, then the least solution $\tilde{y}_{n-2}$ of the $(n-2)^{th}$ equation given $\tilde{y}_n$, $\tilde{y}_{(n-1)}$ and similarly, we obtain $\tilde{y}_{n-3}, \ldots, \tilde{y}_1$ We have to show that if $(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n)$ is the least solution of (5.5), then $\tilde{y}_i = \bar{y}_i$ for all $i \in \{1,2,\ldots,n\}$.

Since $(\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_n)$ is obviously a solution of (5.5) we always have $(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n) \prec (\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_n)$, i.e. $\bar{y}_i \prec \tilde{y}_i$ for all $i \in \{1,2,\ldots,n\}$.

Now $(\bar{y}_1,\bar{y}_2,\ldots,\bar{y}_n)$ is a solution of (5.5), $\bar{y}_n$ must also satisfy the n^{th} equation of (5.5). But then $\tilde{y}_n \prec \bar{y}_n$, since $\tilde{y}_n$ was obtained as the least solution of the n^{th} equation of (5.5). Therefore, $\tilde{y}_n = \bar{y}_n$. But then $\bar{y}_{n-1}$ is also a solution of the $(n-1)^{th}$ equation given $\tilde{y}_n$, and hence $\tilde{y}_{n-1} \prec \bar{y}_{n-1}$. Consequently, $\tilde{y}_{n-1} = \bar{y}_{n-1}$. Similarly, we can use $\tilde{y}_n = \bar{y}_n$, $\tilde{y}_{n-1} = \bar{y}_{n-1}$ to show that $\tilde{y}_{n-2} = \bar{y}_{n-2}$ and so on, which therefore completes our proof. ∇

Let us define a matrix $M^{\{k\}}$ for all $k \in \{1,2,\ldots,n\}$ by

$$(5.6) \qquad (M^{\{k\}})_{ij} = M_{ij}^{\{k\}} = v(Q_{ij}^{\{k\}}) \quad \text{for all} \quad i,j,$$

where $Q_{ij}^{\{k\}}$ is as defined in theorem 4.3 above. Note that from theorem 4.3, we have $M^{\{0\}} = M$, $M^{+} = M^{\{n\}}$, and

$$(5.7) \qquad M_{ij}^{\{k\}} = M_{ij}^{\{k-1\}} \uplus M_{ik}^{\{k-1\}} \circ (M_{kk}^{\{k-1\}})^{*} \circ M_{kj}^{\{k-1\}} \quad \text{for all} \quad i,j.$$

Now consider the first equation of (5.4) above. This can be written as $y_1 = A_{11} \odot y_1 \oplus \hat{b}_1$, where

$$\hat{b}_1 = A_{12} \odot y_2 \oplus A_{13} \odot y_3 \oplus \ldots \oplus A_{1n} \odot y_n \oplus b_1$$

Then from our previous argument, we see that $(A_{11})^{*r} \odot \hat{b}_1$ is the least solution for the first equation of (5.4). Substituting this for y_1 in the second equation, we obtain the following "reduced" system.

$$(5.8) \qquad \begin{cases} y_1 = A_{11} \odot y_1 \oplus A_{12} \odot y_2 \oplus \ldots \oplus A_{1n} \odot y_n \oplus b_1 \\ y_2 = \qquad A_{22}^{\{1\}} \odot y_2 \oplus \ldots \oplus A_{2n}^{\{1\}} \odot y_n \oplus b_2^{\{1\}} \\ \vdots \\ y_n = \qquad A_{n2}^{\{1\}} \odot y_2 \oplus \ldots \oplus A_{nn}^{\{1\}} \odot y_n \oplus b_n^{\{1\}} \end{cases}$$

where $A_{ij}^{\{1\}} = A_{ij} \oplus A_{i1} \otimes (A_{11})^{*r} \otimes A_{1j}$ and

$$b_i^{\{1\}} = b_i \oplus A_{i1} \otimes (A_{11})^{*r} \otimes b_1 \quad \text{for all} \quad i \in \{2,3,\ldots,n\}$$

Note that from (5.7), it follows that $A_{ij}^{\{1\}} = r(M_{ij}^{\{1\}})$ for all $i \neq 1$.

We thus see that the above step essentially eliminates the variable y_1 from the last $n-1$ equations of the system which is therefore similar to the first step of the elimination process in the Gauss elimination method of solving ordinary linear equations. One can then use the same process to eliminate the variable y_2 from the last $(n-2)$ equations of the reduced system (5.8) and obtain another reduced system and so on. After the $(k-1)^{th}$ step, we obtain the following reduced system of equations

$$(5.9) \quad \begin{cases} y_1 = A_{11}\otimes y_1 \oplus A_{12}\otimes y_2 \oplus \ldots \oplus A_{1(k-1)}\otimes y_{k-1} \oplus A_{1k}\otimes y_k \oplus \ldots \oplus A_{1n}\otimes y_n \oplus b_1 \\ y_2 = \quad A_{22}^{\{1\}}\otimes y_2 \oplus \ldots \oplus A_{2(k-1)}^{\{1\}}\otimes y_{k-1} \oplus A_{2k}^{\{1\}}\otimes y_k \oplus \ldots \oplus A_{2n}^{\{1\}}\otimes y_n \oplus b_2^{\{1\}} \\ \vdots \\ y_k = \quad A_{kk}^{\{k-1\}}\otimes y_k \oplus \ldots \oplus A_{kn}^{\{k-1\}}\otimes y_n \oplus b_k^{\{k-1\}} \\ \vdots \\ y_n = \quad A_{nk}^{\{k-1\}}\otimes y_k \oplus \ldots \oplus A_{nn}^{\{k-1\}}\otimes y_n \oplus b_n^{\{k-1\}} \end{cases}$$

where

$$(5.10) \quad A_{ij}^{\{m\}} = A_{ij}^{\{m-1\}} \oplus A_{im}^{\{m-1\}} \otimes \left(A_{mm}^{\{m-1\}}\right)^{*r} \otimes A_{mj}^{\{m-1\}}, \quad A_{ij}^{\{0\}} = A_{ij} \quad \text{and}$$

$$(5.11) \quad b_i^{\{m\}} = b_i^{\{m-1\}} \oplus A_{im}^{\{m-1\}} \otimes \left(A_{mm}^{\{m-1\}}\right)^{*r} \otimes A_{mj}^{\{m-1\}}, \quad b_i^{\{0\}} = b_i$$

for all $i \in \{k,k+1,\ldots,n\}$ and all $m \in \{0,1,2,\ldots,k-1\}$.

Note again that from (5.7), we have $A_{ij}^{\{m\}} = r\left(M_{ij}^{\{m\}}\right)$ for all $i \in \{k, k-1,\ldots,n\}$ and also that $\left(A_{mm}^{\{m-1\}}\right)^{*r}$ exists because $\left(M_{mm}^{\{m-1\}}\right)^{*} \in \mathcal{V}$ by theorem 4.4, and

$$r\left(\left(M_{mm}^{\{m-1\}}\right)^{*}\right) = \bigvee_{q=0}^{\infty} r\left(M_{mm}^{\{m-1\}}\right)^{q}$$

$$= \bigvee_{q=0}^{\infty} \left(A_{mm}^{\{m-1\}}\right)^{q}$$

$$= \left(A_{mm}^{\{m-1\}}\right)^{*r}$$

Therefore, after the $(n-1)^{th}$ step, we obtain an upper triangular system which can then be solved by the generalized back-substitution method as described above. This method of solving (5.4) is similar to that of Gauss elimination in linear algebra and hence was called the generalized Gauss elimination by Carré (1971). At this point, it is of interest to obtain the following

THEOREM 5.3 If the matrix A of (5.4) is symmetric, i.e. $A_{ij} = A_{ji}$ for all i,j, and if the path algebra $(\mathcal{V}_r, \oplus, \odot)$ is a commutative semiring, then for all $k \in \{0,1,\ldots,n\}$, we have $A_{ij}^{\{k\}} = A_{ji}^{\{k\}}$ for all i,j, where $A_{ij}^{\{k\}}$ is given by (5.10)

PROOF We use mathematical induction on k as follows. For $k = 0$, $A_{ij}^{\{0\}} = A_{ij} = A_{ji} = A_{ji}^{\{0\}}$. So let us suppose that the result holds for $k = m-1$. But then

$$A_{ij}^{\{m\}} = A_{ij}^{\{m-1\}} \oplus A_{im}^{\{m-1\}} \odot \left(A_{mm}^{\{m-1\}}\right)^{*r} \odot A_{mj}^{\{m-1\}} \quad \text{by (5.10)}$$

$$= A_{ji}^{\{m-1\}} \oplus A_{jm}^{\{m-1\}} \odot \left(A_{mm}^{\{m-1\}}\right)^{*r} \odot A_{mi}^{\{m-1\}} \quad \text{by induction hypothesis}$$

$$= A_{ji}^{\{m\}}$$

Therefore, the result holds for all $k \in \{0,1,2,\ldots,n\}$. ∇

The above theorem means that the labour of calculating $A_{ij}^{\{k\}}$ for all $k \varepsilon \{0,1,2,\ldots,n\}$ when using the generalized Gauss elimination is almost halved in the case where A is symmetric and Θ `is commutative.

Now since the generalized Gauss elimination was obtained by analogy with the Gauss elimination method in linear algebra, this suggests that it can be modified to yield a method analogous to the Jordan elimination method which requires no back-substitution. In fact, this can be done by eliminating the relevant variable from all the equations in each step of the elimination process. For after the $(n-1)^{th}$ step of this modified elimination process, the matrix A of (5.4) will be transformed into a zero matrix, and hence no back-substitution is needed. Accordingly, this modified method was appropriately called the generalized Jordan elimination by Carré (1971).

It is now a logical question to ask whether the application of the generalized Gauss elimination to (5.3) actually yields the solution $M_r^* \otimes b$. That this is so can be easily seen to be the consequence of the following

THEOREM 5.4 The reduced systems obtained in the elimination process of the generalized Gauss elimination all have the same least solution.

PROOF From the above description of the generalized Gauss elimination, we see that the reduced system obtained after the variable $y_1,y_2,\ldots,y_{k-1}$ have been eliminated can be written as follows.

$$
(5.12)\quad \begin{cases}
y_1 = A_{11}y_1 + A_{12}y_2 + \dots + A_{1(k-1)}y_{k-1} + A_{1k}y_k + \dots + A_{1n}y_n + b_1 \\
y_2 = A_{21}\tilde{y}_1 + A_{22}y_2 + \dots + A_{2(k-1)}y_{k-1} + A_{2k}y_k + \dots + A_{2n}y_n + b_2 \\
-\,- \\
y_k = A_{k1}\tilde{y}_1 + A_{k2}\tilde{y}_2 + \dots + A_{k(k-1)}\tilde{y}_{k-1} + A_{kk}y_k + \dots + A_{kn}y_n + b_k \\
-\,- \\
y_n = A_{n1}\tilde{y}_1 + A_{n2}\tilde{y}_2 + \dots + A_{n(k-1)}\tilde{y}_{k-1} + A_{nk}y_k + \dots + A_{nn}y_n + b_n
\end{cases}
$$

where, for convenience, we have used $+$ and juxtaposition to denote respectively the addition $\oplus$ and multiplication $\odot$ of the path algebra $(\mathcal{V}_r, \oplus, \odot)$, and $\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_{k-1}$ are defined as follows.

(1) $\tilde{y}_1$ is the least solution of the first equation of (5.4) for all possible values of $y_2, y_3, \dots, y_n$.

(2) Given $\tilde{y}_1$, $(\tilde{y}_1, \tilde{y}_2)$ is the least solution of the first two equations of (5.4) for all possible values of $y_3, y_4, \dots, y_n$.

- -

(k-1) Given $\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_{k-2}$, $(\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_{k-1})$ is the least solution of the first $k-1$ equations of (5.4) for all possible values of $y_k, \dots, y_n$.

The elimination of the next variable y_k from all the equations below the k^{th} equation of (5.12) then yields the reduced system (5.13) below, where $\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_{k-1}$ are as before and $\tilde{y}_k$ is defined similarly.

$$
(5.13)\quad \begin{cases}
y_1 = A_{11}y_1 + A_{12}y_2 + \cdots + A_{1k}y_k + A_{1(k+1)}y_{k+1} + \cdots + A_{1n}y_n + b_1 \\
y_2 = A_{21}\tilde{y}_1 + A_{22}y_2 + \cdots + A_{2k}y_k + A_{2(k+1)}y_{k+1} + \cdots + A_{2n}y_n + b_2 \\
-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,- \\
y_{k+1} = A_{(k+1)1}\tilde{y}_1 + A_{(k+1)2}\tilde{y}_2 + \cdots + A_{(k+1)k}\tilde{y}_k + A_{(k+1)(k+1)}y_{k+1} + \cdots + A_{(k+1)n}y_n \\
\qquad\qquad + b_{k+1} \\
-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,-\,- \\
y_n = A_{n1}\tilde{y}_1 + A_{n2}\tilde{y}_2 + \cdots + A_{nk}\tilde{y}_k + A_{n(k+1)}y_{k+1} + \cdots + A_{nn}y_n + b_n
\end{cases}
$$

Now our task is to show that (5.12) and (5.13) have the same least solution, i.e. let $(\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n)$ be the least solution of (5.12) and $(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n)$ be the least solution of (5.13), then

$$\hat{y}_i = \bar{y}_i \qquad \text{for all} \quad i \in \{1, 2, \ldots, n\}$$

First, we show that $\hat{y}_i = \tilde{y}_i = \bar{y}_i$ for all $i \in \{1, 2, \ldots, k\}$. To this end, observe that the first equation of (5.13) is identical with that of (5.4). Consequently, $\tilde{y}_1 \prec \bar{y}_1$ by our definition of $\tilde{y}_1$. We then claim that $\tilde{y}_1 = \bar{y}_1$. For otherwise, $(\tilde{y}_1, \bar{y}_2, \ldots, \bar{y}_n)$ would be a "smaller" solution of (5.13) than $(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n)$, a contradiction.

Now $\tilde{y}_1 = \bar{y}_1$ implies that $(\tilde{y}_1, \bar{y}_2)$ is also a solution of the first two equations of (5.4). Consequently, $\tilde{y}_2 \prec \bar{y}_2$ by our definition of $\tilde{y}_2$. Again, we claim that $\tilde{y}_2 = \bar{y}_2$. For otherwise, $(\tilde{y}_1, \tilde{y}_2, \bar{y}_3, \ldots, \bar{y}_n)$ would be a "smaller" solution of (5.13) than $(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n)$, a contradiction.

Similarly, we can use $\tilde{y}_1 = \bar{y}_1$, $\tilde{y}_2 = \bar{y}_2$ to deduce that $\tilde{y}_3 = \bar{y}_3$ and so on until we have shown that $\tilde{y}_k = \bar{y}_k$.

By a similar argument, one can easily establish that $\tilde{y}_i = \hat{y}_i$ for all $i \in \{1, 2, \ldots, k-1\}$. That $\tilde{y}_k = \hat{y}_k$ also is a consequence of the following two arguments.

(i) $\tilde{y}_1 = \hat{y}_1, \tilde{y}_2 = \hat{y}_2, \ldots, \tilde{y}_{k-1} = \hat{y}_{k-1}$ together implies that $(\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_{k-1}, \hat{y}_k)$, satisfies the first k equations of (5.4) and hence $\tilde{y}_k \prec \hat{y}_k$ by our definition of $\tilde{y}_k$.

(ii) $\bar{y}_k = \tilde{y}_k$ implies that the solution $(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n)$ of (5.13) is also a solution of (5.12), and hence

$$(\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n) \prec (\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n) .$$

Thus in particular, $\hat{y}_k \prec \bar{y}_k = \tilde{y}_k$.

Now $\hat{y}_k = \tilde{y}_k$ implies that the solution $(\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n)$ of (5.12) is also a solution of (5.13), and hence

$$(\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n) \prec (\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n),$$

Therefore, $(\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n) = (\bar{y}_1, \bar{y}_2, \ldots, \bar{y}_n)$

as required. ∇

The validity of the generalized Jordan elimination can also be demonstrated in the same fashion as theorem 5.4 above and hence its detail will be omitted.

From the above proof of theorem 5.4, we see that the assumption of idempotency plays a fundamental role in that it enables us to make use of the fact that $M_r^* \ominus b$ is the least solution of (5.3) with respect to $\prec$ of the path algebra. Without this assumption, $\prec$ is not an ordering because it is not necessarily reflexive and hence one cannot speak of the least solution with respect to $\prec$. It thus appears that the generalized Gauss elimination may not be obtained without this assumption and hence is not applicable in the case where the path algebra is not idempotent. However, this would contradict our common sense, because after all, Gauss elimination was originally invented to solve a system of linear equations over a field and we know that for non-trivial fields, the idempotency of addition can never be satisfied. In fact, the recent work of Gondran (1975) was meant to substantiate this intuitive argument. But unfortunately, he overlooked the significance of having to prove the equivalence of the reduced systems (i.e. they all have the same required solution) in the above school-book approach to Gauss and Jordan elimination methods. Nevertheless, using the present formulation of a path problem, we are able to establish the validity of the generalized Gauss elimination for determining the solution $M_r^* \ominus b$ of (5.3) without assuming that the p-space is idempotent or complete. The proof of this result essentially rests upon the observation that $M_r^* \ominus b$ can be regarded as the image of $M^* \circ b$ under the function r, and since $M^* \circ b$ is the least solution of the system $z = M \circ z \uplus b$ with respect to $\subseteq$, intuitively it should be possible to define the generalized Gauss elimination for solving $y = M_r \ominus y \oplus b$ from the generalized Gauss elimination for solving $z = M \circ z \uplus b$ through the application of the function r. But can one develop the generalized Gauss elimination for solving $z = M \circ z \uplus b$, since the multisum operation is not idempotent ? The answer is made affirmative by the fact that here we can also speak of the least solution but with respect to multiset inclusion.

THEOREM 5.5 Let $\mathcal{N}$ be a network over a monoid (X,o) which is also compatible with a given p-space $(X,o,\mathcal{V},r)$, and $M^{\{k\}}$ be the matrix defined by (5.6) above. Then the generalized Gauss elimination as described above is valid for determing the solution $M_r^* \Theta b$ of (5.3) by defining

$$(5.14) \qquad \left(A_{mm}^{\{m-1\}}\right)^{*r} = r\left(\left(M_{mm}^{\{m-1\}}\right)^*\right) \quad \text{for all } m \in \{0,1,\ldots,n\}$$

PROOF We first show that the reduced systems all have the same solution $M_r^* \Theta b$. Now the reduced system obtained after the $(k-1)^{th}$ step as given by (5.9) can also be rewritten in terms of $r(z_i)$, where $y_i = r(z_i)$ for all $i \in \{1,2,\ldots,n\}$, as follows

$$(5.15)\begin{cases} r(z_1) = r(M_{11}oz_1 + M_{12}oz_2 + \ldots + M_{1(k-1)}oz_{k-1} + M_{1k}^{\{k\}}oz_k + \ldots + M_{1n}oz_n + b_1) \\ r(z_2) = \qquad r(M_{22}^{\{1\}}oz_2 + \ldots + M_{2(k-1)}^{\{1\}}oz_{k-1} + M_{2k}^{\{1\}}oz_k + \ldots + M_{2n}oz_n + b_2^{\{1\}}) \\ \vdots \\ r(z_k) = \qquad\qquad r(M_{kk}^{\{k-1\}}oz_k + \ldots + M_{kn}^{\{k-1\}}oz_n + b_k^{\{k-1\}}) \\ \vdots \\ r(z_n) = \qquad\qquad r(M_{nk}^{\{k-1\}}oz_k + \ldots + M_{nn}^{\{k-1\}}oz_n + b_n^{\{k-1\}}) \end{cases}$$

This is because by (5.14), we have $A_{ij}^{\{m\}} = r(M_{ij}^{\{m\}})$ for all i,j and for all $m \in \{0,1,2,\ldots,n\}$. This last result can be proved by mathematical induction as follows. For $m = 0$, $A_{ij}^{\{0\}} = r(M_{ij}) = r(M_{ij}^{\{0\}})$ as required. So let us suppose the result is true for $m = k-1$ and show that it is also true for $m = k$.

$$A_{ij}^{\{k\}} = A_{ij}^{\{k-1\}} \oplus A_{ik}^{\{k-1\}} \odot \left(A_{kk}^{\{k-1\}}\right)^{*r} \odot A_{kj}^{\{k-1\}}$$

$$= r\left(M_{ij}^{\{k-1\}}\right) \oplus r\left(M_{ik}^{\{k-1\}}\right) \odot r\left(\left(M_{kk}^{\{k-1\}}\right)^*\right) \odot r\left(M_{kj}^{\{k-1\}}\right)$$

by (5.14) and by induction hypothesis

$$= r\left(M_{ij}^{\{k-1\}} \uplus M_{ik}^{\{k-1\}} \circ \left(M_{kk}^{\{k-1\}}\right)^* \circ M_{kj}^{\{k-1\}}\right)$$

$$= r\left(M_{ij}^{\{k\}}\right) \quad \text{by (5.7)} .$$

Hence the results follows from the principle of mathematical induction.

Therefore, the reduced system obtained after the k th step can be written as

$$(5.16)\begin{cases} r(z_1) = r(M_{11}\circ z_1 + M_{12}\circ z_2 + \cdots + M_{1k}\circ z_k + M_{1(k+1)}\circ z_{k+1} + \cdots + M_{1n}\circ z_n + b_1) \\ r(z_2) = r(M_{22}^{\{1\}}\circ z_2 + \cdots + M_{2k}^{\{1\}}\circ z_k + M_{2(k+1)}^{\{1\}}\circ z_{k+1} + \cdots + M_{2n}\circ z_n + b_2^{\{1\}}) \\ \vdots \\ r(z_{k+1}) = r(M_{(k+1)(k+1)}^{\{k\}}\circ z_{k+1} + \cdots + M_{(k+1)n}^{\{k\}}\circ z_n + b_{k+1}^{\{k\}}) \\ \vdots \\ r(z_n) = r(M_{n(k+1)}^{\{k\}}\circ z_{k+1} + \cdots + M_{nn}^{\{k\}}\circ z_n + b_n^{\{k\}}) \end{cases}$$

From (5.15) and (5.16), we see that they have the same solution $M_r^* \odot b$ if we can show that the system of equations (over multisets) obtained from (5.15) by deleting r has the same least solution as the system of equations (over multisets) obtained from (5.16) by deleting r. Here the least solution is taken with respect to $\subseteq$. That this is so can be proved in exactly the same way as we did in theorem 5.3 above but in terms of the ordering $\subseteq$.

Since the back-substitution process can also be verified to yield the same solution $M_r^* \odot b$ through the application of the function r to each step of the back-substitution process in obtaining $M^* \circ b$, the proof is completed. ∇

The generalized Gauss elimination given by theorem 5.5 above can be easily seen to coincide with our previous description in the case where the p-space is both idempotent and complete. Moreover, it can also be seen to include the case where the path algebra $(\mathcal{V}_r, \oplus, \odot)$ forms a field or is embeddable in a field. For in such a case,

$$r\left(\left(M_{mm}^{\{m-1\}}\right)^*\right) = \left(r(e) \ominus r\left(M_{mm}^{\{m-1\}}\right)\right)^{-1}$$

$$= \left(r(e) \ominus A_{mm}^{\{m-1\}}\right)^{-1}$$

$$= \left(A_{mm}^{\{m-1\}}\right)^{*r} ,$$

where $\ominus A$ denotes the inverse of A with respect to $\oplus$, $B \ominus A = B \oplus (\ominus A)$ and A^{-1} denotes the inverse of A with respect to $\odot$ of the path algebra $(\mathcal{V}_r, \oplus, \odot)$.

Note that in these two cases $\left(A_{mm}^{\{m-1\}}\right)^{*r}$ is in fact expressible in terms of path algebraic operations, namely in terms of $\oplus$ and $\odot$ of $\mathcal{V}_r$ (or indirectly). Accordingly, the application of the generalized Gauss elimination to solve (5.3) for these two cases can be carried out entirely in terms of path algebraic operations. This point is significant in practice because otherwise, the description of the generalized Gauss elimination in theorem 5.5 above would be merely of theoretical interest. This is because the computation of $r\left(\left(M_{mm}^{\{m-1\}}\right)^*\right)$ by first computing $\left(M_{mm}^{\{m-1\}}\right)^*$ and then $r\left(\left(M_{mm}^{\{m-1\}}\right)^*\right)$ may not be feasible. For instance, consider the p-space $(R, +, \mathcal{W}_R, k\text{-min})$ and let $M_{mm}^{\{m-1\}} = \{2\}$. Then $\{2\}^* = \{0,2,4,6,8,\ldots\}$ and hence we cannot compute all the elements of $\{2\}^*$ in a finite number of steps. But from theoretical consideration, we know that

$$k\text{-min}\{2\}^* = k\text{-min}\{0\} \oplus k\text{-min}\{2\} \oplus \ldots \oplus k\text{-min}\{2^{k-1}\} .$$

and since the right-hand side of the above equality can be computed in a finite number of steps, $k\text{-min}\{2\}^*$ is therefore determined.

This example is in fact a special case of the more general situation where the p-space is q-stationary (see (iii) of definition 3.4 above). For in such a case, we always have

$$r\left(\left(M_{mm}^{\{m-1\}}\right)^*\right) = r(e) \oplus r\left(M_{mm}^{\{m-1\}}\right) \oplus \ldots \oplus r\left(M_{mm}^{\{m-1\}}\right)^q .$$

and hence the generalized Gauss elimination of theorem 5.5 can be applied by defining

$$\left(A_{mm}^{\{m-1\}}\right)^{*r} = r(e) \oplus A_{mm}^{\{m-1\}} \oplus \ldots \oplus \left(A_{mm}^{\{m-1\}}\right)^q$$

Another situation where the generalized Gauss elimination of theorem 5.5 can be applied by defining

$$(5.17) \qquad \left(A_{mm}^{\{m-1\}}\right)^{*r} = r(e) \oplus A_{mm}^{\{m-1\}} \oplus \ldots \oplus \left(A_{mm}^{\{m-1\}}\right)^{n_0}$$

for some positive integer n_0 is when the p-space is complete and the network $\mathcal{N}$ which is compatible with the p-space is 1-absorptive. In fact, it can be shown that in this case $n_0 = 0$. The necessary argument can be carried out in the same way that we proved (1.8) from (1.9) in chapter 1, and therefore will be omitted. Furthermore, (5.17) also holds when the p-space is not only complete but also commutative and its compatible network $\mathcal{N}$ is q-absorptive or q-regular with respect to the p-space. In fact, $n_0=n(q-1)$ for the first case and $n_0=nt(q-1)$ for the second, where t is the number of elementary closed paths in $G(\mathcal{N})$. These results can also be established in the same way that we proved theorems 1.6 and 1.7 by using lemma 1.3 and 1.4. Now from the proof of theorem 4.10, it is also clear that (5.17) holds whenever the p-space is complete, idempotent, intensive and its compatible network $\mathcal{N}$ is such that $v(P_{ii})$ is q-absorptive with respect to r for all i.

The generalized Jordan elimination can also be similarly defined as in theorem 5.5 by using (5.14) but its detail will be omitted here.

It is interesting to note that both the generalized Gauss and Jordan elimination methods can also be obtained through the derivation of product forms for M^*. In fact, the product forms of M^* can be obtained in the same way as presented by Backhouse and Carré (1975), since all the basic identities used by them in deriving these product forms are also valid in the complete semiring N_∞^X (Two such identities are given in theorem 2.7 above). Now from these product forms of M^*, one can derive methods of determining $M^* \circ b$ from the system $z = M \circ z \uplus b$. The corresponding methods for determining $M_r^* \odot b$ can then be obtained through the application of the function r to each step of the algorithms for determing $M^* \circ b$.

Thus far, we have restricted ourselves to the problem of determining a certain column of the matrix $M_r^* \otimes B$. Now if several columns of $M_r^* \otimes B$ are required, then we have to solve several systems of equations, one for each column of $M_r^* \otimes B$. However, since the matrix of those systems are the same, namely M_r, we can use the generalized Gauss elimination to solve these systems simultaneously by applying the elimination process as before, but all the different columns of B are now treated together and a separate back-substitution for each different column of B must now be carried out. The generalized Jordan elimination can also be applied to solve these systems simultaneously in the same fashion except that no back-substitution is required.

More generally, if we do not require complete columns of $M_r^* \otimes B$ but instead, we require a particular submatrix of $M_r^* \otimes B$, then an elimination method similar to that of Aitken (see e.g. Fox (1964)) can also be developed. As a matter of fact, the manner in which this can be done was well set out in the recent work of Backhouse and Carré (1975).

Finally, it is interesting to note here that when the path algebra $(\mathcal{V}_r, \oplus, \otimes)$ is a Q-semiring and its pseudo-ordering $\prec$ is also a total ordering, the method of Dijkstra (1959) for solving the shortest path problem (problem 1.1), which can be viewed as a form of elimination method, can also be generalized to solve (5.3) above. This generalization of Dijkstra's method was fully discussed by Carré (1976).

5.3 Iterative Methods

In this section, we shall consider variants of iterative methods in linear algebra (see e.g. Fox(1964)) for solving the system $y = A \circ y + b$ over a semiring $(X,+,\circ)$ in the case where the matrix $A \in \mathcal{M}_n(X)$ is n_0-stable, i.e. $A^{[n_0]} = A^{[n_0+1]}$ for some $n_0 \in N$, where $A^{[n]} = I + A + A^2 + \ldots + A^n$. Its application to the solution of path problems in the case where the corresponding p-space is completely n_0-stable (see definition 4.4) is a consequence of lemma 4.3 above.

Let us begin by making the following two assumptions which we shall relax later.

ASSUMPTION 5.1 The semiring $(X,+,o)$ is idempotent.

ASSUMPTION 5.2 The n_0-stable matrix A also has the property that whenever $B \prec A^*$, then B is n_1-stable for some $n_1 \in N$, where $A^* = A^{[n_0]}$.

Note that our use of the notation A^* for $A^{[n_0]}$ will not be confused with the closure of a multiset A or a matrix A of multisets as defined and used in previous sections since we shall not make any reference to them here. Also for convenience, we shall refer to the n_0-stable matrix A which satisfies assumption 5.2 as a <u>hereditary n_0-stable</u> matrix. The class of hereditary n_0-stable matrices is in fact quite extensive since it includes absorptive (and hence acyclic) matrices and matrices over commutative semirings which are q-regular or q-absorptive. That these matrices are n_0-stable can be seen respectively from theorems 1.5 to 1.7 above. That they satisfy assumption 5.2 can be easily seen to be consequences of the following two facts.

(i) $G(A^*)$ is q-absorptive or q-regular whenever $G(A)$ is.

(ii) From property (0.3) of the pseudo-ordering $\prec$ of a semiring $(X,+,o)$, we see that if A is q-absorptive or q-regular and $B \prec A$, then so is B.

Now an immediate consequence of assumption 5.1 is that the pseudo-ordering $\prec$ is also reflexive and hence an ordering. Thus the system $y = A \text{ o } y + b$ can be seen to have the least solution $A^* \text{ o } b$ with respect to $\prec$ as defined by (0.12) in terms of the pseudo-ordering of the semiring.

Our consideration of iterative methods of determining $A^* \circ b$ from the system $y = A \circ y + b$ can best be compared with a similar treatment in linear algebra. There, it is well known (Forsythe (1953)) that one can solve the system $y = A \circ y + b$ by considering an iterative scheme of the form

$$y^{(k+1)} = H \circ y^{(k)} + K \circ b,$$

where H and K are chosen so that the successive estimates $y^{(k)}$ for $k = 1,2,\ldots..$ ultimately yield the solution $y = (I - A)^{-1} \circ b$. In fact, the restriction imposed on the choice of H and K is obtained by noting that $y = y^{(k)} = y^{(k+1)} = \ldots$ must satisfy

$$y = H \circ y + K \circ b \quad , \quad \text{i.e.} \quad y = (I-H)^{-1} \circ K \circ b \quad ,$$

which then suggests that H and K should satisfy

$$(I - A)^{-1} = (I - H)^{-1} \circ K$$

so that the successive estimates would ultimately yield $y = (I - A)^{-1} \circ b$.

It is essentially from this consideration that Carré (1976) considered the iterative scheme $y^{(k+1)} = H \circ y^{(k)} + K \circ b$ in conjunction with the following restriction on the choice of H and K :

$$A^* = H^* \circ K$$

However, since H may not be k-stable for any $k \in N$, H^* may not be defined. One way of getting over this difficulty is to assume that $H \prec A^*$. For by assumption 5.2, it follows that H is n_1-stable for some $n_1 \in N$.

We note here that although $H \prec A^*$ can be deduced from $A^* = H^* \circ K$ (i.e. $H \prec H^* \circ A^* = H^* \circ H^* \circ K = H^* \circ K = A^*$), it must be assumed separately in order to guarantee that $A^* = H^* \circ K$ is meaningful in the first place. From these considerations, we obtain

THEOREM 5.6 Let $(X,+,\circ)$ be an idempotent semiring and $A \in \mathcal{M}_n(X)$ a hereditary n_0-stable matrix. Then the iterative scheme

$$y^{(k+1)} = H \circ y^{(k)} + K \circ b \ , \quad \text{where} \quad H \prec A^* \quad \text{and} \quad A^* = H^* \circ K$$

yields $y^{(n_1+1)} = A^* \circ b$ for some $n_1 \in N$ whenever $y^{(0)} \prec A^* \circ b$

If in addition, $y^{(k)} = y^{(k+1)}$ for some $k < n_1$, then $y^{(k)} = A^* \circ b$ also.

PROOF By assumption 5.2, H is n_1-stable for some $n_1 \in N$, i.e. $H^* = H^{[n_1]}$. Now from $y^{(k+1)} = H \circ y^{(k)} + K \circ b$, it follows that

$$\begin{aligned} y^{(n_1+1)} &= H^{n_1+1} \circ y^{(0)} + (I + H + \dots + H^{n_1}) \circ K \circ b \\ &= H^{n_1+1} \circ y^{(0)} + H^* \circ K \circ b \\ &= H^{n_1+1} \circ y^{(0)} + A^* \circ b \end{aligned}$$

But $y^{(0)} \prec A^* \circ b$ implies that

$$H^{n_1+1} \circ y^{(0)} \prec H^{n_1+1} \circ A^* \circ b = H^{n_1+1} \circ H^* \circ K \circ b \prec H^* \circ K \circ b = A^* \circ b,$$

and hence $y^{(n_1+1)} = A^* \circ b$ as required.

Now if $y^{(k+1)} = y^{(k)}$ for some $k < n_1$, then

$$y^{(k+2)} = H \circ y^{(k+1)} + K \circ b = H \circ y^{(k)} + K \circ b = y^{(k+1)} \quad .$$

Similarly, $y^{(k+3)} = y^{(k+2)}$ and so on.

Therefore, $y^{(k)} = y^{(k+1)} = \ldots = y^{(n_1+1)} = A^* \circ b$. ∇

Let us now examine some of the iterative methods which fit into the iterative scheme of theorem 5.5 above. The following possibilities were considered by Carré (1976).

(i) <u>The Jacobi Method</u> $(H = A\ ,\ K = I)$

The system $y = A \circ y + b$ immediately suggests the iterative method

$$y^{(k+1)} = A \circ y^{(k)} + b \qquad \text{for all } k = 0,1,2,\ldots. \tag{5.17}$$

Note that $A^* = H^* \circ K$ is here satisfied by $H = A,\ K = I$.

This method is credited to Jacobi because it is a counterpart of the Jacobi iterative method in linear algebra (see e.g. Fox (1964)).

(ii) <u>The Gauss-Siedel Method</u> $(H = L^* \circ U\ ,\ K = L^*)$

Here let us adopt the following assumption

ASSUMPTION 5.3 The n_0-stable matrix A is such that

$$A_{ii} = \theta \qquad \text{for all } i \in \{1,2,\ldots,n\}$$

We shall see later that this assumption does not cause any loss of generality whenever assumption 5.2 is satisfied.

If we now rewrite (5.17) as

$$y_i^{(k+1)} = \sum_{j=1}^{n} A_{ij} \circ y_j^{(k)} + b_i \qquad \text{for all } i \in \{1,2,\ldots,n\}$$

then we see that the Jacobi method makes use of all the elements $y_j^{(k)}$ of $y^{(k)}$ in calculating $y^{(k+1)}$. But intuitively, it would seem more reasonable to use always the last available estimates such as the following

(5.18) $$y_i^{(k+1)} = \sum_{j=1}^{i-1} A_{ij} \circ y_j^{(k+1)} + \sum_{j=i+1}^{n} A_{ij} \circ y_j^{(k)} + b_i \text{ for all}$$

$i \in \{1,2,\ldots,n\}$.

Note that we have here used assumption 5.3 above.

Since (5.18) can be calculated sequentially, it does not require the simultaneous storage of the two estimates $y_i^{(k)}$ and $y_i^{(k+1)}$ in the course of computation which is therefore an advantage over the Jacobi method. However, this advantage causes the method to be rather sensitive to the actual numbering used for the nodes.

If we now define two matrices L and U respectively by

(5.19) $$L_{ij} = \begin{cases} A_{ij} & \text{if } i < j \\ \theta & \text{, otherwise} \end{cases} \quad \text{and} \quad U_{ij} \begin{cases} A_{ij} & \text{if } i > j \\ \theta & \text{otherwise} \end{cases} ,$$

then we can rewrite (5.18) in matrix forms as follows

(5.20) $$y^{(k+1)} = L \circ y^{(k+1)} + U \circ y^{(k)} + b \quad .$$

Now since $G(L)$ is acyclic, $L^n = \theta$ and therefore (5.20) can be rewritten as

(5.21) $$y^{(k+1)} = L^* \circ U \circ y^{(k)} + L^* \circ b \quad , \quad \text{where } L^* = L^{[n-1]} \quad .$$

Note that $A^* = H^* \circ K$ is here satisfied by $H = L^* \circ U$, $K = L^*$ because $A = L + U$ by assumption 5.3, and $(L + U)^* = (L^* \circ U)^* \circ L^*$. The last equality can be proved either by comparison of terms in the expansion on both sides or by an argument similar to that used in theorem 2.7 above but in terms of $\prec$, which is an ordering whenever assumption 5.1 is satisfied.

This method is credited to Gauss and Siedel because it is a counterpart of the Gauss-Siedel iterative method in linear algebra (Fox (1964)).

(iii) The Double-Sweep Method $(H = L^* \circ U^* = K)$

Again assumption 5.3 will be adopted here, and just as in (ii), there is no loss of generality whenever assumption 5.2 is satisfied.

Now in view of our previous remark that the Gauss-Siedel method is rather sensitive to the actual numbering of the nodes, it might therefore prove useful to modify the procedure by alternately calculating the elements $y_j^{(k+1)}$ of $y^{(k+1)}$ first in a backward manner $(j = n, n-1, \ldots, 1)$ and then in a forward manner $(j = 1,2,\ldots,n)$. More precisely, we first obtain $y^{(k+\frac{1}{2})}$ by using

$$y_n^{(k+\frac{1}{2})} = y_n^{(k)} + b_n \ , \quad y_i^{(k+\frac{1}{2})} = \sum_{j=i+1}^{n} A_{ij} \circ y_j^{(k+\frac{1}{2})} + y_i^{(k)} + b_i$$

for all $i = n-1, n-2, \ldots, 1$

and then obtain $y^{(k+1)}$ by using

$$y_1^{(k+1)} = y_1^{(k+\frac{1}{2})} \quad , \quad y_i^{(k+1)} = \sum_{j=1}^{i-1} A_{ij} \circ y_j^{(k+1)} + y_i^{(k+\frac{1}{2})}$$

for all $i = 1,2,\ldots,n-1$.

In matrix forms, these can be written respectively as

(5.22) $\quad y^{(k+\frac{1}{2})} = U \circ y^{(k+\frac{1}{2})} + y^{(k)} + b \quad ,$ and

(5.23) $\quad y^{(k+1)} = L \circ y^{(k+1)} + y^{(k+\frac{1}{2})}$, where L and U are defined by (5.19).

Now since $G(L)$ and $G(U)$ are both acyclic, $L^n = \theta = U^n$, and hence (5.22) and (5.23) are respectively equivalent to

(5.24) $\quad y^{(k+\frac{1}{2})} = U^* \circ y^{(k)} + U^* \circ b \quad ,$ and

(5.25) $y^{(k+1)} = L^* \circ y^{(k+\frac{1}{2})}$

Therefore, combining (5.24) and (5.25), we obtain

$$y^{(k+1)} = L^* \circ U^* \circ y^{(k)} + L^* \circ U^* \circ b$$

Note that $A^* = H^* \circ K$ is here satisfied by $H = L^* \circ U^* = K$ because $A = L + U$ by assumption 5.3, and $(L + U)^* = (L^* \circ U^*)^*$, The last equality follows because using assumption 5.1, we always have

$$L + U \prec L^* \circ U^* \prec (L + U)^* \quad , \quad \text{and hence}$$

$$(L + U)^* \prec (L^* \circ U^*)^* \prec ((L + U)^*)^* = (L + U)^*$$

We note here that instead of (5.22) and (5.23), one can also consider the iterative method given by the following pair

(5.26) $y^{(k+\frac{1}{2})} = U \circ y^{(k+\frac{1}{2})} + y^{(k)}$, and

(5.27) $y^{(k+1)} = L \circ y^{(k+1)} + y^{(k+\frac{1}{2})} + b$,

which yields

$$y^{(k+1)} = L^* \circ U^* \circ y^{(k)} + L^* \circ b \quad ,$$

and $A^* = H^* \circ K$ is now satisfied by $H = L^* \circ U^*$ and $K = L^*$.

However, as will be apparent later, this method is inferior to the previous one, using (5.22) and (5.23), because $L^* \prec L^* \circ U^*$ always.

Moreover, whenever $b \prec y^{(0)}$ is used, both methods are equivalent to using (5.23) in conjunction with (5.26) because $y^{(0)} \prec y^{(\frac{1}{2})} \prec y^{(1)} \prec y^{(1+\frac{1}{2})} \prec \ldots\ldots$ always .

In fact, it was Yen(1970) who first used (5.23) in conjuction with (5.26), and $y^{(0)} = b$ to solve the shortest path problem. The more general method was called the Double-Sweep Method (cf. Shier (1974)) because of its similarity (though not a counterpart in a strict sense) to the Double-Sweep iterative method in linear algebra (Fox(1964), p.195)).

Let us now examine why assumption 5.3 does not cause any loss of generality in our presentation of the last two methods whenever assumption 5.2 is satisfied.

From the n_0-stable matrix A, let us define two matrices D and F by

$$D_{ij} = \begin{cases} A_{ij} & \text{if } i = j \\ \theta & \text{, otherwise} \end{cases} \quad \text{and} \quad F_{ij} = \begin{cases} A_{ij} & \text{if } i \neq j \\ \theta & \text{, otherwise} \end{cases} .$$

From this definition, we see that $A = D + F$. Accordingly, we can rewrite the system $y = A \circ y + b$ as

$$y = D \circ y + F \circ y + b \quad .$$

But this system has the same least solutions $A^* \circ b$ as the system $y = D^* \circ F \circ y + D^* \circ b$ because

$$A^* \circ b = (D + F)^* \circ b = (D^* \circ F)^* \circ D^* \circ b \quad .$$

Note that D^* is defined by assumption 5.2 because $D \prec A^*$, and similarly for $(D^* \circ F)^*$because

$$D^* \circ F \prec A^* \circ A \prec A^* \quad .$$

Hence from the system $y = A \circ y + b$, we can first obtain the equivalent system $y = A' \circ y + b'$, where $A' = D^* \circ F$ and $b' = D^* \circ b$. Since A' satisfies assumption 5.3, the last two methods can then be employed to solve $y = A' \circ y + b'$ to obtain the least solution $A^* \circ b$. This useful observation was made by Carré (1976).

A formal comparison of the above methods will now be discussed For this purpose, Carré (1976) obtained the following

THEOREM 5.7 Let $\tilde{y}^{(k+1)} = \tilde{H} \circ \tilde{y}^{(k)} + \tilde{K} \circ b$ and $\hat{y}^{(k+1)} = \hat{H} \circ \hat{y}^{(k)} + \hat{K} \circ b$ be two iterative methods for solving the system $y = A \circ y + b$ and let $\tilde{H} \prec \hat{H}$, $\tilde{K} \prec \hat{K}$. Then for all $k \in N$, $\tilde{y}^{(k)} \prec \hat{y}^{(k)}$ whenever $\tilde{y}^{(0)} = \hat{y}^{(0)}$.

PROOF We shall use mathematical induction on k. For $k = 0$, the result is true since $\tilde{y}^{(0)} = \hat{y}^{(0)}$ by assumption. So let us suppose that the result is true for $k = m$. But then

$$\tilde{y}^{(m+1)} = \tilde{H} \circ \tilde{y}^{(m)} + \tilde{K} \circ b$$

$$\prec \hat{H} \circ \hat{y}^{(m)} + \hat{K} \circ b$$

$$= \hat{y}^{(m+1)}$$

Therefore, the result is true for all $k \in N$. ∇

COROLLARY 5.1 The Double-Sweep method is superior to both the Jacobi and Gauss-Siedel methods in the sense that it yields the required solution in a number of iterations not exceeding those required by the other two, using the same initial estimate.

PROOF This follows because

(i) $A = L + U \prec L^* \circ U^*$, $I \prec L^* \circ U^*$, and

(ii) $L^* \circ U \prec L^* \circ U^*$, $L^* \prec L^* \circ U^*$. ∇

We note here that theorem 5.7 cannot be used to compare the

Gauss-Siedel method with the Jacobi method because in general A and $L^{*} o U$ are not comparable. For instance, let $A = \begin{bmatrix} \theta & e \\ e & \theta \end{bmatrix}$ be a matrix over the two element Q-semiring $X = \{\theta, e\}$.

$$\text{Then } L^{*} o U = \begin{bmatrix} \theta & e \\ \theta & e \end{bmatrix} \nprec A .$$

However, it would seem intuitively that the Gauss-Siedel method (which uses more "current" information than the Jacobi method) should be superior to the Jacobi method. In fact, this is so if one chooses the initial estimate $y^{(0)}$ such that the Gauss-Siedel method yields a better estimate at each iteration. This result was proved by Shier (1974) as follows.

THEOREM 5.8 Let $y_J^{(k+1)} = A \text{ o } y_J^{(k)} + b$ and

$$y_G^{(k+1)} = L \text{ o } y_G^{(k+1)} + U \text{ o } y_G^{(k)} + b, \qquad \text{where}$$

L and U are as defined by (5.19) above. Then $y_J^{(k)} \prec y_G^{(k)}$ for all $k \in N$ whenever $y_J^{(0)} = y_G^{(0)} \prec y_G^{(1)}$.

PROOF We first show by mathematical induction that $y_G^{(0)} \prec y_G^{(1)}$ implies $y_G^{(k)} \prec y_G^{(k+1)}$ for all $k \in N$.

For $k = 0$, this is true by assumption. So suppose that the result is true for $k = m$, i.e. $y_G^{(m)} \prec y_G^{(m+1)}$.

But then

$$y_G^{(m+1)} = L \text{ o } y_G^{(m+1)} + U \text{ o } y_G^{(m)} + b$$

$$= L^{*} \text{ o } U \text{ o } y_G^{(m)} + L^{*} \text{ o } b \quad , \quad \text{since } L^n = \theta$$

$$\prec \quad L^* \circ U \circ y_G^{(m+1)} + L^* \circ b$$

$$= \quad L \circ y_G^{(m+2)} + U \circ y_G^{(m+1)} + b$$

$$= \quad y_G^{(m+2)}$$

Hence $y_G^{(k)} \prec y_G^{(k+1)}$ for all $k \in N$.

We can now prove by mathematical induction that

$$y_J^{(k)} \prec y_G^{(k)} \quad . \qquad \text{for all } k \in N.$$

For $k = 0$, the result is true because $y_J^{(0)} = y_G^{(0)}$ by assumption. So suppose the result is true for $k = m$. But then

$$y_J^{(m+1)} = A \circ y_J^{(m)} + b$$

$$\prec A \circ y_G^{(m)} + b \qquad \text{by induction hypothesis}$$

$$= L \circ y_G^{(m)} + U \circ y_G^{(m)} + b \quad , \text{ since } A = L + U$$

$$\prec L \circ y_G^{(m+1)} + U \circ y_G^{(m)} + b, \text{ since } y_G^{(m)} \prec y_G^{(m+1)}$$

$$= y_G^{(m+1)} \quad .$$

Hence $y_J^{(k)} \prec y_G^{(k)}$ for all $k \in N$. ∇

Let us now return to theorem 5.6 above. From this theorem, we know that the number of iterations required for obtaining the least solution $A^* \circ b$ of $y = A \circ y + b$ by the above iterative methods will be at most $n_1 + 1$, where n_1 is such that $H^{[n_1]} = H^{[n_1+1]}$. But in fact, the number of iteration required by these methods will be at most $n_0 + 1$, where n_0 is such that $A^{[n_0]} = A^{[n_0+1]}$, as shown in the following

THEOREM 5.9 Let $(X,+,o)$ be an idempotent semiring and $A \varepsilon \mathcal{M}_n(X)$ a hereditary n_0 stable matrix. Then the iterative scheme

$$y^{(k+1)} = H \circ y^{(k)} + K \circ b \ , \quad \text{where} \quad H \prec A^* \text{ and } A^* = H^* \circ K$$

yields $y^{(n_0+1)} = A^* \circ b$ whenever $y^{(0)} \prec A^* \circ b$,

provided we also have in addition that

$$(5.28) \qquad K \circ A \prec (I + H) \circ K \quad \text{and} \quad I \prec K \quad .$$

Moreover, if $y^{(k)} = y^{(k+1)}$ for some $k < n_0$, then $y^{(k)} = A^* \circ b$ also.

PROOF Let us first show that $A^* = H^{[n_0]} \circ K$.

Since $H^{[k]} \circ K \prec H^* \circ K = A^*$ for all $k \varepsilon N$,

it follows that $H^{[n_0]} \circ K \prec A^*$.

Now we claim that (5.28) implies that

$$A^{[k]} \prec H^{[k]} \circ K \qquad \text{for all} \quad k \varepsilon N.$$

This claim can be proved by mathematical induction as follows. For $k = 0$, $A^{[0]} = I \prec K = H^{[0]} \circ K$. So let us suppose that the result holds for $k = m$.

But then

$$A^{[m+1]} = I + A^{[m]} \circ A$$

$$\prec I + H^{[m]} \circ K \circ A \quad \text{, by induction hypothesis}$$

$$\prec I + H^{[m]} \circ (I + H) \circ K \qquad \text{by } (5.28)$$

$$= I + (I + H)^{m+1} \circ K \text{, } \quad \text{since } H^{[m]} = (I + H)^m$$

(which uses assumption 5.1)

$$= I + H^{[m+1]} \circ K$$

$$= H^{[m+1]} \circ K \text{, } \qquad \text{since } I \prec H^{[m+1]} \text{ and } I \prec K \text{ .}$$

Hence the claimed is justified, and therefore

$$A^* = A^{[n_0]} \prec H^{[n_0]} \circ K \quad .$$

It then follows from the anti-symmetric property of $\prec$ that

$$A^* = H^{[n_0]} \circ K \quad .$$

The rest of the proof can be argued in exactly the same way as that of theorem 5.6 except that n_1 is now replaced by n_0. ∇

COROLLARY 5.2 The number of iterations required by the Jacobi, Gauss-Siedel and Double-Sweep methods to yield the least solution $A^* \circ b$ where A is a hereditary n_0-stable matrix is at most n_0+1 where $A^* = A^{[n_0]} = A^{[n_0+1]}$, provided that $y^{(0)} \prec A^* \circ b$ is used as the initial estimate.

PROOF It suffices to show that in all these methods (5.28) is satisfied as follows.

(i) If $H = A$ and $K = I$, then clearly $I \prec K$ and

$K \circ A = A \prec I + A = (I + H) \circ K$ as required

(ii) If $H = L^* \circ U$ and $K = L^*$, then clearly $I \prec K$ and

$$K \circ A = L^* \circ A$$

$$= L^* \circ (L + U) \quad , \quad \text{since } A = L + U$$

$$= L^* \circ L + L^* \circ U$$

$$\prec L^* + L^* \circ U \circ L^* \quad , \quad \text{since } L^* \circ L \prec L^* \text{ and } I \prec L^*$$

$$= (I + L^* \circ U) \circ L^*$$

$$= (I + H) \circ K \quad \text{as required.}$$

(iii) If $H = L^* \circ U^* = K$, then clearly $I \prec K$ and

$$K \circ A = (L^* \circ U^*) \circ (L + U) \ , \quad \text{since } A = L + U$$

$$= L^* \circ U^* \circ L + L^* \circ U^* \circ U$$

$$\prec L^* \circ U^* \circ L^* \circ U^* + L^* \circ U^* \ , \quad \text{since } L \prec L^* \circ U^* \text{ and } U^* \circ U \prec U^*$$

$$= (I + L^* \circ U^*) \circ L^* \circ U^*$$

$$= (I + H) \circ K \quad \text{as required.} \qquad \nabla$$

We note here that in fact the number of iterations required by the Double-Sweep method to yield $A^* \circ b$ was shown directly by Carré (1976) to be at most $\frac{1}{2}(n_0 + 1)$.

Let us now extend the result of Theorem 5.9 to n_0-stable matrices in general. The following theorem shows how this can be done.

THEOREM 5.10 Let $(X,+,o)$ be an idempotent semiring and $A \in \mathcal{M}_n(X)$ an n_0-stable matrix. Then the iterative scheme

$$y^{(k+1)} = H \circ y^{(k)} + K \circ b \quad , \quad \text{where}$$

$$H \circ A^* + K \prec A^* , \quad K \circ A \prec (I + H) \circ K , \quad I \prec K ,$$

yields $y^{(n_0+1)} = A^* \circ b$ whenever $y^{(0)} \prec A^* \circ b$.

If in addition, $y^{(k)} = y^{(k+1)}$ for some $k < n_0$, then $y^{(k)} = A^* \circ b$ also.

PROOF Let us show by mathematical induction that $H \circ A^* + K \prec A^*$ always implies $H^{k+1} \circ A^* + H^{[k]} \circ K \prec A^*$ for all $k \in N$. For $k = 0$, the result is true by assumption. So let us suppose that the result is true for $k = m$.

But then

$$\begin{aligned} & H^{m+2} \circ A^* + H^{[m+1]} \circ K \\ & = H \circ (H^{m+1} \circ A^* + H^{[m]} \circ K) + K \\ & \qquad \text{since } H^{[m+1]} = H \circ H^{[m]} + I \\ & \prec H \circ A^* + K \quad \text{by induction hypothesis} \\ & \prec A^* \end{aligned}$$

and hence the result is also true for $k = m+1$ whenever it is true for $k=m$. Therefore, the result is true for all $k \in N$. From this result, it follows that

$$H^{[k]} \circ K \prec H^{k+1} \circ A^* + H^{[k]} \circ K \prec A^* \quad \text{for all} \quad k \in N$$

The rest of the proof can be carried out in the same fashion as we did in the proof of theorem 5.9 above. ∇

COROLLARY 5.3 Corollary 5.2 above holds for any n_0-stable matrix A except that assumption 5.3 for the Gauss-Siedel and the Double-Sweep methods must be enforced. (because assumption 5.2 has been dropped).

PROOF It suffices to show that the inequality $H \circ A^* + K \prec A^*$ is satisfied by the H and K in these three methods, since the other two inequalities were already shown to be valid in the proof of corollary 5.2 above.

(i) If $H = A$ and $K = I$, then

$$H \circ A^* + K = A \circ A^* + I = A^* \prec A^* \text{ as required.}$$

(ii) If $H = L^* \circ U$ and $K = L^*$ then

$$\begin{aligned} H \circ A^* + K &= L^* \circ U \circ A^* + L^* \\ &\prec A^* \circ A^* + A^* \quad \text{since } L^* \prec A^* \text{ and } L^* \circ U \prec A^* \circ A \prec A^* \\ &= A^* \end{aligned}$$

(iii) If $H = L^* \circ U^* = K$, then

$$\begin{aligned} H \circ A^* + K &= (L^* \circ U^*) \circ A^* + L^* \circ U^* \\ &\prec (A^* \circ A^*) \circ A^* + A^* \circ A^* \quad \text{since } L^* \prec A^*,\ U^* \prec A^* \\ &= A^* \end{aligned}$$

∇

It is also interesting to note that both the Jacobi and Double-Sweep methods can also be fitted into the iterative scheme of the following theorem while the Gauss-Siedel method cannot.

THEOREM 5.11 Let $(X,+,o)$ be an idempotent semiring and $A \in \mathcal{M}_n(X)$ an n_0-stable matrix. Then the iterative scheme

$$y^{(k+1)} = H \circ y^{(k)} + K \circ b \quad , \quad \text{where}$$

$$H \circ A^* + K \prec A^* \; , \quad A \prec H \; , \quad I \prec K \; ,$$

yields $y^{(n_0+1)} = A^* \circ b$ whenever $y^{(0)} \prec A^* \circ b$.

If in addition, $y^{(k)} = y^{(k+1)}$ for some $k < n_0$, then $y^{(k)} = A^* \circ b$ also.

PROOF In view of the argument used in establishing theorem 5.10 above, it suffices to show that $A^{[k]} \prec H^{[k]} \circ K$ for all $k \in N$. This can be done by mathematical induction as follows.

For $k = 0$, $A^{[0]} = I \prec K = H^{[0]} \circ K$. So let us suppose that the result holds for $k = m$. But then

$$\begin{aligned} A^{[m+1]} &= I + A \circ A^{[m]} \\ &\prec I + A \circ H^{[m]} \circ K \quad , \quad \text{by induction hypothesis} \\ &\prec I + H^{[m+1]} \circ K \quad \text{since } A \prec H \text{ and } H \circ H^{[m]} \prec H^{[m+1]} \\ &= H^{[m+1]} \circ K \; , \quad \text{since } I \prec K \prec H^{[m+1]} \circ K \; . \end{aligned}$$

Hence the result holds for all $k \in N$. ∇

Since the Gauss-Siedel method fits into the iterative scheme of theorem 5.10 but does not fit into that of theorem 5.11, it follows that theorem 5.10 does not imply theorem 5.11. However, the converse is still an open question. We only know that if $K \circ H \prec (I + H) \circ K$ also holds, then $A \prec H$ implies $K \circ A \prec K \circ H \prec (I + H) \circ K$, and hence theorem 5.11 holds.

Let us now examine the possibility of dropping assumption 5.1. From the proof of theorem 5.6, we see that if one chooses $y^{(0)} = \theta$, then

$$y^{(k+1)} = H \circ y^{(k)} + K \circ b \quad \text{implies}$$

$$\begin{aligned} y^{(n_1+1)} &= H^{n_1+1} \circ y^{(0)} + (I + H + \dots + H^{n_1}) \circ K \circ b \\ &= H^* \circ K \circ b \quad , \text{where } H^* = H^{[n_1]} = H^{[n_1+1]} \\ &= A^* \circ b \end{aligned}$$

Therefore, theorem 5.6 remains valid without assumption 5.1 if $y^{(0)} = \theta$.

This means that the Jacobi method can be used to solve $y = A \circ y + b$ without assumption 5.1 whenever $y^{(0)} = \theta$ is used but A must still be a hereditary n_0-stable matrix. This result generalizes that given by Gondran (1975). The Gauss-Seidel method was also claimed by Gondran (1975) to be valid whenever A is absorptive. However, he used the identity that $(L + U)^* = (L^* \circ U)^* \circ L^*$ which he did not prove to be valid without assumption 5.1. Of course, if this result holds, then $A^* = H^* \circ K$ holds and hence Gauss-Siedel method could then be used without assumption 5.1 but with $y^{(0)} = \theta$ and A as a hereditary n_0-stable matrix. We think it is likely that the above identity holds without assumption 5.1, since this identity has its analogue in linear algebra, namely

$$(I - (L + U))^{-1} = (I - (I - L)^{-1}U)^{-1}(I - L)^{-1}$$

However, the proof appears to be difficult.

On the other hand, we do not think that the Double-Sweep method can be used without assumption 5.1. This is because $A^* = H^* \circ K$ is unlikely to be satisfied without assumption 5.1 since the identity $(L + U)^* = (L^* \circ U^*)^*$ does not have an analogue in linear algebra.

Finally, we note that theorems 5.9, 5.10 and 5.11 are not valid without assumption 5.1.

CHAPTER 6

SCHEDULE ALGEBRAS AND K-SHORTEST-PATHS PROBLEMS

6.1 Generalization of Giffler's Schedule Algebra

The p-space $(X,o,\mathcal{N}_X,r)$ of example 3.8 was first considered by Giffler (1963, 1968), although he did not define it formally in that way. In fact, he was directly concerned with the path algebra of this p-space which we shall henceforth refer to as Giffler's schedule algebra, since he referred to the ring of matrices over this path algebra by the name "schedule algebra". Such an algebra was noted in Giffler (1968) to have algebraic properties equivalent to those of ordinary integers including the fact that one can also define addition and multiplication of "quotients" or "ratios" of two elements. However, Giffler's investigation of these properties were rather informal and his definition of "quotients" or "ratios" were erroneous in much the same way as when one attempts to define rational numbers by a division algorithm of two integers which may or may not terminate. This pitfall can be avoided by adopting an algebraist's way of defining rational numbers, namely via the construction of the quotient field of integers. This approach will therefore be adopted below where we present a more generalized version of Giffler's schedule algebra.

Instead of the additive monoid $(N,+)$ of non-negative integers which we considered in example 3.8, let us consider a totally ordered commutative group $(S,\leq,+)$, see section 0.2 above. As in example 3.8, let us define $\hat{S}$ to be a set of arbitrary objects which can be put in one-to-one correspondence with S, i.e. to each $a \in S$, there exists $\hat{a} \in \hat{S}$ and vice versa. Also as before, the set $\{a,\hat{a}\}$ will be called a twin pair

DEFINITION 6.1 Let $X = S \cup \hat{S}$ and $|| \ || : X \to S$ be a function defined by $||a|| = a$ and $||\hat{a}|| = a$ for all $a \in S$. For convenience $||x||$ will be called the S-counterpart of $x \in X$.

From this definition, we can now define a binary operation o on X as follows.

$$(6.1) \qquad x \circ y = \begin{cases} ||x|| + ||y|| & \text{if } x,y \in S, \text{ or } x,y \in \hat{S} \\ \widehat{||x|| + ||y||} & , \text{ otherwise} \end{cases}$$

This definition of o may appear somewhat different from the rules (i) and (ii) in example 3.8, but in fact a second glance at definition 6.1 will reveal their equivalence.

LEMMA 6.1 For any $x,y \in X$, $||x \circ y|| = ||x|| + ||y||$

PROOF If both $x,y \in S$, then from (6.1), we have

$$x \circ y = ||x|| + ||y|| = x + y \quad ,$$

and hence

$$||x \circ y|| = ||x + y||$$

$$= x + y \quad , \text{ since } x + y \in S \text{ whenever } x,y \in S$$

$$= ||x|| + ||y||$$

If both $x,y \in \hat{S}$, then from (6.1), we have

$$x \circ y = ||x|| + ||y|| = a + b \quad ,$$

where a,b are elements of S such that $\hat{a} = x$, $\hat{b} = y$.

Consequently, $\|x \circ y\| = \|a + b\|$

$= a + b$, since $a + b \in S$ whenever $a, b \in S$

$= \|x\| + \|y\|$

Now if either $x \in S, y \in \hat{S}$ or $x \in \hat{S}$, $y \in S$, it follows from (6.1) that

$$x \circ y = \widehat{\|x\| + \|y\|} \quad ,$$

and hence $\|x \circ y\| = \|x\| + \|y\|$ by definition 6.1.

Therefore, in all cases, $\|x \circ y\| = \|x\| + \|y\|$ as required. ∇

THEOREM 6.1 $(X, \circ)$ forms a commutative group with the same identity element as in $(S,+)$. Moreover, if y^{-1} denotes the inverse of $y \in X$ with respect to $\circ$, and $-a$ denotes the inverse of $a \in S$ with respect to $+$, then

$$(6.2) \qquad x \circ y^{-1} = \begin{cases} \|x\| - \|y\| & \text{if } x,y \in S \text{ or } x,y \in \hat{S} \\ \widehat{\|x\| - \|y\|} & , \text{ otherwise} \end{cases}$$

PROOF From (6.1), it follows that X is closed with respect to the binary operation $\circ$. Thus it remains to verify that all the formal properties of a group are also satisfied.

This can be shown as follows.

(i) $x \circ \theta = x = \theta \circ x$, where θ is the identity for $+$ in S.

From lemma 6.1, we know that for all $x \in X$,

$$\|x \circ \theta\| = \|x\| + \|\theta\| = \|x\| + \theta = \|x\| \quad , \text{ and}$$

if $x \circ \theta \neq x$, then $\{x \circ \theta, x\}$ is a twin pair. Without loss of generality, let us suppose that $x \in S$, and hence $x \circ \theta \in \hat{S}$. But since $\theta \in S$, it follows from (6.1) that

$$x \circ \theta = ||x|| + ||\theta|| = x + \theta \in S, \text{ a contradiction.}$$

Therefore, $x \circ \theta = x$ for all $x \in X$. Similarly, one can show that $\theta \circ x = x$ for all $x \in X$.

(ii) $x \circ y = y \circ x$ for all $x,y \in X$.

From lemma 6.1, we know that for all $x,y \in X$,

$$||x \circ y|| = ||x|| + ||y|| = ||y|| + ||x|| = ||y \circ x|| ,$$

and if $x \circ y \neq y \circ x$, then $\{x \circ y, y \circ x\}$ is a twin pair. Without loss of generality. let us suppose that $x \circ y \in S$, and hence $y \circ x \in \hat{S}$. But from (6.1), we see that $x \circ y \in S$ iff both $x,y \in S$ or both $x,y \in \hat{S}$. In either case, $y \circ x \in S$, a contradiction.

(iii) $(x \circ y) \circ z = x \circ (y \circ z)$ for all $x,y,z \in X$.

From lemma 6.1, we know that for all $x,y,z \in X$,

$$\begin{aligned} ||(x \circ y) \circ z|| &= ||x \circ y|| + ||z|| \\ &= ||x|| + ||y|| + ||z|| \\ &= ||x|| + ||y \circ z|| \\ &= ||x \circ (y \circ z)|| , \end{aligned}$$

and if $(x \circ y) \circ z \neq x \circ (y \circ z)$, then $\{(x \circ y) \circ z, x \circ (y \circ z)\}$ forms a twin pair. Without loss of generality. let us suppose that $(x \circ y) \circ z \in S$, and hence $x \circ (y \circ z) \in \hat{S}$. But from (6.1), we see that the only way for $(x \circ y) \circ z$ to belong to S is that among the three elements x,y,z, none or exactly two elements belong to $\hat{S}$. If none, then $x \circ (y \circ z) \in S$, a contradiction. If two, then there are three possibilities, namely $x,y \in \hat{S}$, or $y,z \in \hat{S}$ or $x,z \in \hat{S}$.

If $x,y \in \hat{S}$, then from (6.1), we have

$$x \circ (y \circ z) = x \circ (\widehat{||y|| + ||z||}) = ||x|| + ||y|| + ||z|| \in S,$$

a contradiction. Similarly, one can easily check that the other two possibilities also lead to contradictions.

(iv) For each $y \in X$, there exists $y^{-1} \in X$ such that

$$y \circ y^{-1} = \theta = y^{-1} \circ y \quad .$$

If $y \in S$, then take $y^{-1} = -y \in S$ because from (6.1),

$$y \circ y^{-1} = y + (-y) = \theta$$

If $y \in \hat{S}$, say $y = \hat{a}$ for some $a \in S$, then take $y^{-1} = \hat{-a}$ because from (6.1),

$$y \circ y^{-1} = a + (-a) = \theta \quad .$$

Therefore from (i) to (iv) above, we conclude that $(X,\circ)$ forms a group with θ as identity for $\circ$, where θ is also the identity for $+$ in S. Moreover, from (iv) above, we see that $||y^{-1}|| = -||y||$ always, and hence (6.2) follows from (6.1). ∇

We now want to define a reduction function r as in example 3.8 above. There, the domain of r was chosen to be $\mathcal{N}_X$. The reason for this may not be obvious. However, if one considers the multiset $A = \{1,\hat{1},1,\hat{1},..\}$, then we see that there is more than one way of deleting all the twin pairs from A. Consequently, $r(A) = \phi$ and $r(A) = \{1\}$ are both possible, and hence r is not a function in accordance with our definition in section 0.1. It is for this reason that we have chosen $\mathcal{N}_X$ as the domain of r in example 3.8 above. But now in our present case, $\mathcal{N}_X$ alone won't do because (X,o) is not assumed to be locally finite, and hence $(\mathcal{N}_X, \uplus, o)$ may not be a semiring. For instance, let (S,+) be the additive group of real numbers and let

$$A = \{1,2, ...\} \qquad , \qquad B = \{-1,-2, ...\}$$

Then $A,B \in \mathcal{N}_X$, where $X = S \cup \hat{S}$. But

$$A \circ B = \{0, 0, ...\} \notin \mathcal{N}_X.$$

This example demonstrates that we must look for a subset of $\mathcal{N}_X$, but which ? In order to find an answer to this question, let us examine our final goal a little further.

Recall that Giffler's schedule algebra was noted to posses properties equivalent to those of integers including the fact that one can also define addition and multiplication for "quotients" or "ratios" of two elements. This suggests that our generalization of Giffler's schedule algebra must also retain these properties and we must also be able to define addition and multiplication for quotients of two elements. But any quotient $\frac{A}{B}$ is usually defined for non-zero B or in this case, $B \neq \phi$ and addition and multiplication of two quotients $\frac{A}{B}$, $\frac{C}{D}$, where $B,D \neq \phi$ are usually defined as follows.

(6.3) $$\frac{A}{B} \oplus \frac{C}{D} = \frac{A \odot D \oplus B \odot C}{B \odot D}$$

(6.4) $$\frac{A}{B} \ominus \frac{C}{D} = \frac{A \ominus C}{B \ominus D}$$

But for the right-hand sides of (6.3) and (6.4) to be consistent with the usual definition of quotients, we must have $B \ominus D \neq \phi$ as well. In order to guarantee that this restriction is satisfied the algebra to be constructed must therefore have the property that $B \ominus D = \phi$ implies either $B = \phi$ or $D = \phi$. Therefore, our choice of the domain of r must be such that the resulting path algebra has this property. Again, let us take $(S,+)$ to be the additive group of real numbers, and let

$$B = \{\ldots,-3,-2,-1,0,1,2,3,\ldots\}, \quad D = \{0,\hat{1}\} .$$

Then , $$B \circ D = \{\ldots,-3,-2,1,0,1,2,3,\ldots;\ldots-\hat{3},-\hat{2},-\hat{1},\hat{0},\hat{1},\hat{2},\hat{3},\ldots\}$$

and hence $B \ominus D = r(B \circ D) = \phi$. But clearly, $B,D \neq \phi$.

The acute reader will observe that the multiset B in this example and the one above does not have a least element and hence it might be for this reason that the difficulty arises. This suggests that our choice of the domain of r should be a subset of $\mathcal{W}_X$ which contains only well ordered multisets and ϕ (see (iii) of definition 2.2). But (X,o) may not be an ordered set, so how can one speak meaningfully of well ordered multisets ? Fortunately, the assumption that S is an ordered set is sufficient for our purpose, since we can use it to define some kind of pseudo-well-ordered multisets as follows.

DEFINITION 6.2 Let $X = S \cup \hat{S}$ as before. For any $A \in N_\infty^X$, let us define $\|A\| = \{ \|x\| \mid A(x) \neq 0\}$ which can be called the <u>$\| \|$ - support of A</u>. Then a multiset $A \in N_\infty^X$ is said to be <u>$\| \|$ - well ordered</u> iff $\|A\|$ is a well ordered set.

LEMMA 6.2 (i) $\|A\| \subseteq \|B\|$ whenever $A \subseteq B$

(ii) $\|A \uplus B\| = \|A\| \cup \|B\|$ or more generally,

$$\|\biguplus_{i\varepsilon I} A_i\| = \bigcup_{i\varepsilon I} \|A_i\|$$

(iii) $\|A \circ B\| = \|A\| \ \|B\|$, where juxtaposition denotes complex product induced by $+$ in S.

PROOF (i) Let $y \in \|A\|$. Then by definition, $y = \|x\|$ for some x such that $A(x) \neq 0$. But $A(x) \leq B(x)$ and therefore, $B(x) \neq 0$ also, i.e. $y \in \|B\|$.

(ii) Let $y \in \|\biguplus_{i\varepsilon I} A_i\|$. Then by definition, $y = \|x\|$ for some x such that $\sum_{i\varepsilon I} A_i(x) \neq 0$. But then $A_i(x) \neq 0$ for at least one $i \in I$. Consequently, $y \in \|A_i\|$ for at least one $i \in I$, and hence $\|\biguplus_{i\varepsilon I} A_i\| \subseteq \bigcup_{i\varepsilon I} \|A_i\|$.

But from (i) above, we have $\|A_i\| \subseteq \|\biguplus_{i\varepsilon I} A_i\|$ for all $i \in I$ since $A_i \subseteq \biguplus_{i\varepsilon I} A_i$ for all $i\varepsilon I$. Therefore,

$\bigcup_{i\varepsilon I} \|A_i\| \subseteq \|\biguplus_{i\varepsilon I} A_i\|$ also, and (ii) is therefore established.

(iii) Let $u \in \|A \circ B\|$. Then by definition, $u = \|y \circ z\|$ for some y,z such that $A(y), B(z) \neq 0$. But by lemma 6.1, $u = \|y\| + \|z\|$, and hence $u \in \|A\| \ \|B\|$, i.e. $\|A \circ B\| \subseteq \|A\| \ \|B\|$

Now let $u \in \|A\| \; \|B\|$, then $u = \|y\| + \|z\| = \|y \circ z\|$ for some y,z such that $A(y), B(z) \neq 0$, But then $\sum_{x=yoz} A(y)B(z) \neq 0$, i.e. $u = \|x\|$ for some $x \in \|A \circ B\|$, i.e. $\|A\| \; \|B\| \subseteq \|A \circ B\|$ also. ∇

LEMMA 6.3 Let $\mathcal{W}_{\|X\|}$ denote the set of all $\| \;\|$ - well ordered multisets of N_∞^X and ϕ, where $X = S \cup \hat{S}$ as before. Then $\mathcal{W}_{\|X\|}$ is a hereditary semiring .

PROOF First we show that $\mathcal{W}_{\|X\|}$ is a hereditary subset of N_∞^X. Let $A \in \mathcal{W}_{\|X\|}$, i.e. $\|A\|$ is a well ordered set, and suppose $B \subseteq A$. Then by (i) of lemma 6.2 above, we have $\|B\| \subseteq \|A\|$, and therefore $\|B\|$ is a well ordered set also, i.e. $B \in \mathcal{W}_{\|X\|}$.

That $\mathcal{W}_{\|X\|}$ is closed with respect to multisum and multiproduct is a consequence of (ii) and (iii) of lemma 6.2 and the fact that the union and complex product of two well ordered sets with elements in a totally ordered monoid are themselves well ordered sets (cf. the proof of (iv) of theorem 2.10 above).

Finally, let $A \in \mathcal{F}_X$ then $\|A\| = \{\, \|x\| \mid A(x) \neq 0\}$ is a finite subset because by definition of $\mathcal{F}_X$, $A(x) \neq 0$ for only a finite number of $x \in X$ and $A(x) \neq \infty$. Hence $\|A\|$ is a well ordered set, i.e. $A \in \mathcal{W}_{\|X\|}$. Consequently, $\mathcal{F}_X \subseteq \mathcal{W}_{\|X\|}$. ∇

Let us now resume our search for the domain of r. Thus far we know that it cannot be larger than $\mathcal{N}_X \cap \mathcal{W}_{\|X\|}$. However, we do not know whether $(\mathcal{N}_X \cap \mathcal{W}_{\|X\|}, \uplus, \circ)$ is a semiring but we do know from (vii) that the set of all countable and well ordered multisets with elements in a totally ordered group is. This suggests that we try the set $\mathcal{U}_{\|X\|} = \mathcal{C}_X \cap \mathcal{N}_X \cap \mathcal{W}_{\|X\|}$. The following theorems show that this is exactly what we are looking for.

THEOREM 6.2 Let $X = S \cup \hat{S}$ as before. Then the $(\mathcal{U}_{\|X\|}, \uplus, \circ)$ where $\mathcal{U}_{\|X\|} = \mathcal{G}_X \cap \mathcal{N}_X \cap \mathcal{W}_{\|X\|}$ is a hereditary semiring.

PROOF This result follows from an argument similar to that used in (vii) of theorem 2.10 by using lemma 6.1 and lemma 6.3 whenever necessary. ∇

THEOREM 6.3 Let $X = S \cup \hat{S}$ and $\mathcal{U}_{\|X\|}$ be as before, and $r : \mathcal{U}_{\|X\|} \rightarrow \mathcal{U}_{\|X\|}$ be defined by $r(A) = \phi$ if $A = \phi$, otherwise $r(A)$ is obtained from A by deleting all its twin pairs. Then $(X, \circ, \mathcal{U}_{\|X\|}, r)$ forms a commutative p-space.

PROOF $(X,\circ)$ is a commutative monoid by theorem 6.1, $(\mathcal{U}_{\|X\|}, \uplus, \circ)$ is a hereditary semiring by theorem 6.2, and hence $(X, \circ, \mathcal{U}_{\|X\|}, r)$ is a commutative p-space if we can show that r is a reduction function.

Since $r(\phi) = \phi$ by definition, it remains to show that r has properties (3.2) and (3.3) also. To do this, let us observe the following two immediate consequences of our definition of r.

(i) $r(A \uplus B) = r(B)$ whenever $r(A) = \phi$

(ii) $A = r(A) \uplus A'$, where $r(A') = \phi$.

Consequently, $r(A \uplus B) = r(r(A) \uplus A' \uplus B)$ by (ii)

$= r(r(A) \uplus B)$ by (i)

and

$$r(A \circ B) = r((r(A) \uplus A') \circ B) \quad \text{by (ii)}$$

$$= r(r(A) \circ B \uplus A' \circ B)$$

$$= r(r(A) \circ B) \quad \text{by (i), since } r(A' \circ B) = \phi .$$

That $r(A' \circ B) = \phi$ can be seen as follows.

Since $r(A') = \phi$, it must be either that $A' = \phi$ or A' contains only twin pairs. If the first, then clearly, $r(A' \circ B) = \phi$ also. If the second, let $x \in A' \circ B$, i.e.

$x = y \circ z$ for some $y \in A'$ and $z \in B$. Without loss of generality, let us suppose that $y \in S$. But then $\hat{y} \circ z$ must be an element of $A' \circ B$ as well. Since $\{y \circ z, \hat{y} \circ z\}$ forms a twin pair, it follows that $A' \circ B$ consists of only twin pairs, and hence $r(A' \circ B) = \phi$ as claimed. ∇

THEOREM 6.4 The path algebra of the p-space $(X, \circ, \mathcal{U}_{||X||}, r)$ in theorem 6.3 forms an integral domain (see section 0.2).

PROOF By theorem 3.2, we know that the path algebra $(\mathcal{V}_r, \oplus, \odot)$, where $\mathcal{V} = \mathcal{U}_{||X||}$, forms a commutative semiring. Now this path algebra is also a ring because for each $A \in \mathcal{V}_r$, there exists $\theta A = \{\hat{\theta}\} \odot A$ such that

$$\begin{aligned} A \ominus A &= A \oplus (\theta A) \quad \text{by definition} \\ &= A \oplus \{\hat{\theta}\} \odot A \\ &= r(A \uplus \{\hat{\theta}\} \circ A) \\ &= \phi \end{aligned}$$

It remains to show that for any $Y, Z \in \mathcal{V}_r$, $Y \odot Z = \phi$ implies either $Y = \phi$ or $Z = \phi$.

Suppose otherwise, and choose $y_0 \in Y$, $z_0 \in Z$ such that

$$||y_0|| = \min ||Y|| \quad \text{and} \quad ||z_0|| = \min ||Z|| .$$

We then claim that for all $x \in Y \circ Z$, $\{x, y_0 \circ z_0\}$ cannot be a twin pair which means that $Y \odot Z = r(Y \circ Z) \neq \phi$, a contradiction. This claim can be justified as follows

Let $x = y \circ z$ for some $y \in Y$ $z \in Z$. Then by our choice of y_0, z_0, we have $||y_0|| \leq ||y||$ and $||z_0|| \leq ||z||$.

Now if $\|y_0\| = \|y\|$ and $\|z_0\| = \|z\|$, then $y_0 = y$ and $z_0 = z$ because y_0 and z_0 cannot have their S-counterparts in Y and in Z respectively $(Y,Z \in \mathcal{V}_r)$, and hence $y_0 \circ z_0 = y \circ z$ means that $\{y \circ z, y_0 \circ z_0\}$ is not a twin pair.

So we may suppose that $\|y_0\| < \|y\|$, say. But then by lemma 6.1, we have

$$\|y_0 \circ z_0\| = \|y_0\| + \|z_0\| < \|y\| + \|z\| = \|y \circ z\| \ ,$$

and hence $\{y \circ z \ , y_0 \circ z_0\}$ is not a twin pair either. ∇

COROLLARY 6.1 The path algebra $(\mathcal{V}_r, \oplus, \odot)$ of the p-space in theorem 6.3 can be extended to the field $(\mathcal{Q}, \oplus, \odot)$, where $\mathcal{Q}$ is the set of equivalence subsets of $\mathcal{V}_r \times \mathcal{V}_r$ of the form

$$\frac{A}{B} = \{(Y,Z) \mid A \odot Z = B \odot Y, \ B,Z \neq \phi\} \quad ,$$

and $\oplus$ is defined by (6.3) whereas $\odot$ is defined by (6.4). Moreover, $\frac{A}{A}$ and $\frac{\phi}{A}$, where $A \neq \phi$, are respectively the unit and zero of $\mathcal{Q}$.

PROOF The construction of $\mathcal{Q}$ from the integral domain $\mathcal{V}_r$ is similar to the construction of rational numbers from integers. Its proof will therefore be omitted. Note that $\mathcal{Q}$ is usually called the <u>quotient field</u> of the integral domain $\mathcal{V}_r$. ∇

The path algebra $(\mathcal{V}_r, \oplus, \odot)$, where $\mathcal{V} = \mathcal{U}_{\|X\|}$, can be seen as a generalization of Giffer's schedule algebra as follows.

When $X = N \cup \hat{N}$, X is always a countable set because $\hat{N}$ is a countable set whenever N is, and so is $N \cup \hat{N}$. Therefore, any multiset $A \in N_\infty^X$, where $X = N \cup \hat{N}$,is always a quasi-countable multiset since $d(A) \subseteq X$ must be a countable set also. Consequently, $\mathcal{Q}_X = N_\infty^X$ when $X = N \cup \hat{N}$.

Now since N is also a well ordered set, and hence $X = N \cup \hat{N}$ is a $\| \ \|$ - well ordered set. But this means that any multiset $A \in N_\infty^X$, where $X = N \cup \hat{N}$ is always a $\| \ \|$ - well ordered multiset. Consequently, $\mathcal{W}_{\|X\|} = N_\infty^X$ when $X = N \cup \hat{N}$.

Therefore, $\mathcal{U}_{\|X\|} = \mathcal{Q}_X \cap \mathcal{N}_X \cap \mathcal{W}_{\|X\|}$

$$= N_\infty^X \cap \mathcal{N}_X \cap N_\infty^X \qquad \text{when } X = N \cup \hat{N}$$

$$= \mathcal{N}_X .$$

Therefore, the p-space $(X, \circ, \mathcal{N}_X, r)$ where $X = N \cup \hat{N}$ is a particular instance of the p-space $(X, \circ, \mathcal{U}_{\|X\|}, r)$, and hence we can regard the path algebra $(\mathcal{V}_r, \oplus, \odot)$ where $\mathcal{V} = \mathcal{U}_{\|X\|}$ as a generalization of Giffler's schedule algebra. For convenience, this path algebra will be referred to simply as a <u>schedule algebra over the totally ordered commutative group $(S, \leq, +)$</u>.

6.2 <u>Schedule Algebraic Division Algorithm</u>

Since the quotient-field extension of a schedule algebra was obtained by an analogy with the construction of rational numbers, it might be fruitful to carry the same analogy a little further. Since it is well known that any rational number or a quotient of two integers can always be expressed as decimals (though it might be just approximately) by a division algorithm, it might be useful to have a similar way of expressing the element $\frac{A}{B}$ in the quotient field of a schedule algebra. To this end, Giffler (1968) invents a procedure for "dividing" A by B. Unfortunately, his presentation of the procedure is not explicitly defined since he only gives an example of the procedure in a tableau form, and that is all (see Giffler (1968), p.269). Nevertheless, it is clear from his example that a strong analogy with the ordinary long-division procedure for two integers can be fruitfully exploited in carrying out his method. The result to be presented below will confirm that this is so.

We shall say that a multiset B of $\mathcal{V}_r$ <u>divides</u> a multiset A of $\mathcal{V}_r$ in the integral domain $\mathcal{V}_r$, where $\mathcal{V} = \mathcal{U}_{\|X\|}$, iff there exists a multiset C of $\mathcal{V}_r$ such that

$$A = B \ominus C,$$

or equivalently,

$$\frac{A}{B} = \frac{C}{\{\theta\}} \text{ or simply } \frac{A}{B} = C \quad .$$

Let $A,B \in \mathcal{V}_r$ and $n_0 \in \{1,2, \ldots .\}$ be given. Then the following algorithm yields a multiset $C \in \mathcal{V}_r$ such that $A = B \ominus C$ or an approximation for $\frac{A}{B}$ to be specified below.

DIVISION ALGORITHM

Step 0. Specify n_0 and set $k = 1$, $A_1 = A$

Step 1. Choose $b_0 \in B$ such that $\|b_0\| = \min \|B\|$

Step 2. Choose $a_k \in A_k$ such that $\|a_k\| = \min \|A_k\|$

and compute $c_k = a_k \circ b_0^{-1}$ by (6.2)

Step 3. Set $A_{k+1} = A_k \ominus \{c_k\} \ominus B$ Terminate when either of the following hold

(i) $A_{k+1} = \phi$ (In this case the division result is exactly $\{c_1,c_2, \ldots ,c_k\}$) .

(ii) $k = n_0$ (In this case, the division result $\{c_1, c_2, \ldots ,c_{n_0}\}$ differs from $\frac{A}{B}$ by $\frac{A_{n_0+1}}{B}$.

Step 4. Increase the value of k by 1 and return to step 2.

In order to illustrate the steps of the above algorithm, let us consider an example below.

Let (S, +) be the additive group of integers and take

$A = \{-1, \hat{2}, \hat{3}, 4\}$, $B = \{\hat{3}, 5\}$. Clearly, $A, B \in \mathcal{V}_r$. Now all the steps involved in computing $\frac{A}{B}$ by the above algorithm can be conveniently expressed in the "long-division" tableau form as used by Giffler (1968) as follows

$$
\begin{array}{r|llll}
 & -\hat{4}, & -\hat{2}, & -1 & \\
\hline
\hat{3},\ 5 & -1, & \hat{2}, & \hat{3}, & 4 \\
 & -1, & \hat{1} & & \\
\hline
 & 1, & \hat{2}, & \hat{3}, & 4 \\
 & 1, & \hat{3} & & \\
\hline
 & \hat{2}, & 4 & & \\
 & \hat{2}, & 4 & & \\
\hline
 & \phi & & &
\end{array}
\qquad\qquad
\begin{array}{r|l}
 & \{c_1, c_2, c_3\} \\
\hline
B & A_1 \\
 & \{c_1\} \ominus B \\
\hline
 & A_2 \\
 & \{c_2\} \ominus B \\
\hline
 & A_3 \\
 & \{c_3\} \ominus B \\
\hline
 & A_4
\end{array}
$$

Note that in the above example, the algorithm terminates with $A_4 = \phi$. However, in general, it is possible that $A_k \neq \phi$ for all k, and hence the algorithm will not terminate if n_0 is not specified.

As an illustration, take $A = \{2, 2, 3, 4\}$ and $B = \{1, 1\}$. Then we have for $k > 1$,

$$
A_k = \begin{cases} \{3, 4\} & \text{if } k \text{ is even} \\ \{\hat{3}, 4\} & \text{otherwise} \end{cases}
$$

Consequently, the algorithm <u>cannot</u> be used to decide whether or not a given multiset B of $\mathcal{V}_r$ divides a given multiset A of $\mathcal{V}_r$. All we know is that if $A_{k+1} = \phi$ for some $k \leq n_0$, then B divides A, but we do not know otherwise, because the algorithm is made to terminate at $k = n_0$ while it is still possible that $A_{k+1} = \phi$ for some $k > n_0$.

In order to prove the validity of the above algorithm, let us first obtain the following

LEMMA 6.4 In the above algorithm,

$$\|c_k\| \leq \|c_{k+1}\| \quad \text{whenever} \quad A_{k+1} \neq \phi .$$

PROOF Since $c_k = a_k \circ b_0^{-1}$, it follows from (6.2) that

$$\|c_k\| = \|a_k \circ b_0^{-1}\| = \|a_k\| - \|b_0\|$$

Similarly, $\|c_{k+1}\| = \|a_{k+1}\| - \|b_0\|$.

Therefore, $\|c_k\| \leq \|c_{k+1}\|$ whenever $\|a_k\| \leq \|a_{k+1}\|$.

But $\|a_{k+1}\| = \min \|A_{k+1}\|$, and hence the required result follows if we can show that

$$\|a_k\| \leq \|x\| \qquad \text{for all} \quad x \in A_{k+1} .$$

Now since $A_{k+1} = A_k \ominus \{c_k\} \odot B$, we must have either $x \in A_k$ or $x \in \theta \{c_k\} \odot B$ for any $x \in A_{k+1}$.

If the former, then $\|a_k\| \leq \|x\|$ because $\|a_k\| = \min \|A_k\|$.

If the latter, then $x = \hat{\theta} \circ c_k \circ b$ for some $b \in B$.

But then
$$\begin{aligned}
\|x\| &= \|\hat{\theta} \circ c_k \circ b\| \\
&= \|\hat{\theta}\| + \|c_k\| + \|b\| && \text{by lemma 6.1} \\
&= \|c_k\| + \|b\| && \text{since } \|\hat{\theta}\| = \theta \in S. \\
&= \|c_k\| + \|b_0\| + (\|b\| - \|b_0\|) . \\
&\geq \|c_k\| + \|b_0\| && \text{since } \|b_0\| \leq \|b\| \\
&= \|c_k \circ b_0\| && \text{by lemma 6.1} \\
&= \|a_k\| && \text{as required.}
\end{aligned}$$

∇

LEMMA 6.5 In the above algorithm, if k is the first index such that $A_{k+1} = \phi$, then the multiset $\{c_1, c_2, \ldots, c_k\}$ does not contain any twin pairs.

PROOF By lemma 6.4 above, it suffices to show that $\{c_i, c_{i+1}\}$ cannot be a twin pair for any $i \varepsilon \{1, 2, \ldots, k-1\}$

Suppose otherwise, and let i_0 be the first index such that $\{c_{i_0}, c_{i_0+1}\}$ forms a twin pair. Without loss of generality, we may assume that $\hat{c}_{i_0+1} = c_{i_0}$.

But then

$$
\begin{aligned}
A_{i_0+2} &= A_{i_0+1} \ominus \{c_{i_0+1}\} \ominus B \\
&= A_{i_0+1} \oplus \{\hat{\theta}\} \ominus \{c_{i_0+1}\} \ominus B \\
&= A_{i_0+1} \oplus \{\hat{c}_{i_0+1}\} \ominus B \\
&= A_{i_0+1} \oplus \{c_{i_0}\} \ominus B \\
&= A_{i_0}, \quad \text{since } A_{i_0+1} = A_{i_0} \ominus \{c_{i_0}\} \ominus B .
\end{aligned}
$$

Hence $\|a_{i_0+2}\| = \min \|A_{i_0+2}\| = \min \|A_{i_0}\| = \|a_{i_0}\|$, which implies that $a_{i_0+2} = a_{i_0}$, since $A \varepsilon \mathcal{V}_r$, i.e. A does not contain any twin pair.

Therefore,

$$
\begin{aligned}
c_{i_0+2} &= a_{i_0+2} \circ b_0^{-1} \\
&= a_{i_0} \circ b_0^{-1} \\
&= c_{i_0},
\end{aligned}
$$

and hence $A_{i_0+3} = A_{i_0+2} \ominus \{c_{i_0+2}\} \circleddash B$

$$= A_{i_0} \ominus \{c_{i_0}\} \circleddash B \qquad \text{by the above result}$$

$$= A_{i_0+1}$$

This result can then be used to show in a similar manner as above that $A_{i_0+4} = A_{i_0+2}$ and so on.

Thus in general, we have

$$A_{i_0+j} = \begin{cases} A_{i_0} & \text{if } j \text{ is even} \\ A_{i_0+1} & \text{if } j \text{ is odd.} \end{cases}$$

Now $i_0 \in \{1,2,\ldots,k-1\}$ implies that $i_0 < i_0 + 1 \leq k$, and hence both A_{i_0} and A_{i_0+1} are non-empty by our assumption that k is the first index such that $A_{k+1} = \phi$.

But then $A_{i_0+j} \neq \phi$ for all j, i.e. $A_{k+1} \neq \phi$ for all k, a condradiction. ∇

PROOF OF THE DIVISION ALGORITHM.

We can now demonstrate the validity of the division algorithm as follows.

First let us use mathematical induction to show that at the end of the k^{th} iteration, we have

$$(6.5) \qquad \frac{A}{B} = C \oplus \frac{A_{k+1}}{B} \quad , \text{ where } C = \{c_1,c_2,\ldots,c_k\} .$$

Now for $k = 1$, if $A_2 = \phi$, then $A_1 = \{c_1\} \circleddash B$ because by definition $A_2 = A_1 \ominus \{c_1\} \ominus B$, and hence

$$\frac{A}{B} = \frac{A_1}{B} = \frac{\{c_1\} \ominus B}{B} = \{c_1\} = \{c_1\} \oplus \frac{A_2}{B} \quad \text{as required.}$$

If $A_2 \neq \phi$, then $\{c_1\} \oplus \frac{A_2}{B} = \{c_1\} \oplus \frac{A_1 \ominus \{c_1\} \ominus B}{B}$

$$= \frac{\{c_1\} \ominus B \oplus A_1 \ominus \{c_1\} \ominus B}{B}$$

$$= \frac{A_1}{B}$$

$$= \frac{A}{B}$$

Hence (6.5) holds for $k = 1$. So let us suppose it holds for all k such that $1 \leqslant k < m$.

Now if $\{c_m, c_{m+1}\}$ forms a twin pair, then

$$\begin{aligned}\{c_1, c_2, \ldots, c_{m+1}\} \ominus B &= r(\{c_1, \ldots, c_{m+1}\} \circ B) \\ &= r(\{c_1, \ldots, c_{m-1}\} \circ B \uplus \{c_m, c_{m+1}\} \circ B) \\ &= r(\{c_1, \ldots, c_{m-1}\} \circ B), \text{ since } r(\{c_m, c_{m+1}\} \circ B) = \phi \\ &= \{c_1, c_2, \ldots, c_{m-1}\} \ominus B .\end{aligned}$$

Also if $\{c_m, c_{m+1}\}$ forms a twin pair, then $A_{m+2} = A_m$ follows from an argument similar to that used in the proof of lemma 6.5 above.

Consequently,

$$\{c_1,\dots,c_{m+1}\} \oplus \frac{A_{m+2}}{B} = \frac{\{c_1,\dots,c_{m+1}\} \ominus B \oplus A_{m+2}}{B}$$

$$= \frac{\{c_1,\dots,c_{m-1}\} \ominus B \oplus A_m}{B}$$

$$= \{c_1,\dots,c_{m-1}\} \oplus \frac{A_m}{B}$$

$$= \frac{A}{B} \quad \text{by induction hypothesis.}$$

So we may suppose that $\{c_m, c_{m+1}\}$ does not form a twin pair. But then $\{c_1,c_2,\dots,c_{m+1}\} = \{c_1,c_2,\dots,c_m\} \oplus \{c_{m+1}\}$, and hence

$$\{c_1,\dots,c_{m+1}\} \oplus \frac{A_{m+2}}{B} = \{c_1,\dots,c_m\} \oplus \{c_{m+1}\} \oplus \frac{A_{m+1} \ominus \{c_{m+1}\} \ominus B}{B}$$

$$= \{c_1,\dots,c_m\} \oplus \frac{\{c_{m+1}\} \ominus B \oplus A_{m+1} \ominus \{c_{m+1}\} \ominus B}{B}$$

$$= \{c_1,\dots,c_m\} \oplus \frac{A_{m+1}}{B}$$

$$= \frac{A}{B} \quad \text{by induction hypothesis.}$$

Therefore, (6.5) holds for all $k \geq 1$.

Therefore, if $A_{k+1} \neq \phi$ for all $k \leq n_0$, then the algorithm must terminate at $k = n_0$, and hence by (6.5),

$$\frac{A}{B} = \{c_1,\dots,c_{n_0}\} \oplus \frac{A_{n_0+1}}{B} \quad .$$

If $A_{k+1} = \phi$ for some $k \leq n_0$, then k must be the first index such that $A_{k+1} = \phi$, and hence by lemma 6.5, $\{c_1,\ldots,c_k\} \in \mathcal{V}_r$, and by (6.5)

$$\frac{A}{B} = \{c_1,\ldots,c_k\} \oplus \frac{A_{k+1}}{B}$$

$$= \{c_1,\ldots,c_k\}$$

which completes the proof.

6.3 Application of Schedule Algebra to K-Shortest-Paths Problems.

For convenience of exposition, we shall assume through out that $\mathcal{N}$ is a network in which the labels are elements of the additive group $(R,+)$ of real numbers. However, the results to be given below remain valid if the labels of $\mathcal{N}$ belong to any Archimedean totally ordered commutative group.

We shall first consider the problem of determining all the numerical labels of paths in $\mathcal{N}$ and then show how one can modify this result to solve k-shortest-paths problems. Using the terminology introduced in section 4.1, our immediate problem is just the computation of $v(P_{ij})$ or $(M^*)_{ij}$, where M is the label matrix of $\mathcal{N}$, for all $i,j \in \{1,2,\ldots,n\}$. Our strategy for solving this problem is to embed $(R,+)$ in the commutative group (X,o) of theorem 6.1, where $X = R \cup \hat{R}$, and to consider the p-space $(X,o,\mathcal{U}_{\|X\|},r)$ for solving this problem. The problem now becomes a path problem in accordance with definition 5.1, where $\mathcal{N}$ is now considered to be a network over (X,o) which is also compatible with the p-space $(X,o,\mathcal{U}_{\|X\|},r)$. This is so because by the definition of r (see theorem 6.3), it follows that $rv(P_{ij}) = v(P_{ij})$ for all i,j or equivalently $M_r^* = M^*$. Now from an argument similar to that used in establishing corollary 4.5 (plus the use of lemma 6.1 and 6.3 where necessary), it follows that $\mathcal{N}$ is compatible with the p-space $(X, o, \mathcal{U}_{\|X\|}, r)$ iff the numerical label of any elementary closed path in $\mathcal{N}$ must be strictly positive. When this condition is satisfied $(M^*)_{ij} \in \mathcal{U}_{\|X\|}$ for all i,j. Now from our discussion in section 5.1, we know that $M^* = M_r^*$ always satisfies the matrix equation $Y = M_r \odot Y \oplus I_r$ over the path algebra $(\mathcal{V}_r, \oplus, \odot)$ where $\mathcal{V} = \mathcal{U}_{\|X\|}$. Since $M_r = M$ and $I_r = I$, it follows that $M^* = M \odot M^* \oplus I$. But then by theorem 6.4, we have

$$(6.6) \qquad M^* = (I \ominus M)^{-1} \quad ,$$

where $\ominus M$ denotes the additive inverse of M, $I \ominus M = I \oplus (\ominus M)$ and $(I \ominus M)^{-1}$ denotes the multiplicative inverse of $I \ominus M$.

Therefore, our immediate problem is reduced to the computation of $(I \ominus M)^{-1}$. Now it is well known that $(I \ominus M)^{-1}$ can be obtained by using the following formula.

$$(6.7) \qquad ((I \ominus M)^{-1})_{ij} = \frac{s(i+j) \odot \det([I \ominus M]_{ji})}{\det(I \ominus M)} \quad \text{for all } i,j \quad ,$$

where $\det(A)$ denotes the determinant of the matrix A (for the definition of a determinant, see e.g. Fox (1964)), $[A]_{ji}$ denotes the matrix obtained from A by deleting its jth row and ith column, and

$$s(k) = \begin{cases} \{0\} & \text{if } k \text{ is even} \\ \{\hat{0}\}, & \text{otherwise} \end{cases} \quad .$$

However, a more efficient way of computing $(I \ominus M)^{-1}$ is to solve the equation $Y = M \odot Y \oplus I$ by Gauss or Jordan elimination in linear algebra (or as discussed in the section 4.2). We note that these methods will yield solutions of the form (6.7) or its equivalence.

The required numerical labels of paths in $\mathcal{N}$ can then be obtained from the application of the division algorithm to (6.7) or its equivalence. Moreover, by lemma 6.4, the labels so obtained will be non-decreasing and are guaranteed to be elements of R by (6.6) above. Note also that any labels of paths in $\mathcal{N}$ can always be obtained by continuing the division algorithm long enough.

We turn now to consider the use of the above result for solving the k-shortest-paths problems. But first, let us note that there are in general two types of k-shortest-paths problems. The first is to find k paths in $\mathcal{N}$ from a given node x_i to a node x_j such that their labels can be ranked as 1^{st}, $2^{nd}, \ldots, k^{th}$ smallest among the labels in $v(P_{ij})$. The second differs from the first in that the required k shortest paths must also be elementary. For $k = 1$, both problems coincide and are better known as the shortest paths problem, and in this case the problem can be more efficiently solved by using the p-space $(R, +, \mathcal{W}_R, \min)$ of example 3.2 above (see also problem 1.1). For $k > 1$, the first type of k-shortest-paths problem can also be solved by using the p-space $(R, +, \mathcal{W}_R, k\text{-min})$ of example 3.3 above (see also problem 1.6). Of course, the second type of k-shortest-paths problems can also be solved in this way except that in tracing the actual paths corresponding to k smallest labels given by this method, a large number of non-elementary paths may have to be traced before the required elementary paths are found. This remark also applies if one uses the above method for computing all the labels of paths in $\mathcal{N}$ by terminating the division algorithm as soon as the required k shortest elementary paths are obtained. Therefore, it is clearly more efficient if one can obtain as small a set of labels as possible which contains those of the k shortest elementary paths. In fact, the method to be presented below is actually based on this principle and the set obtained in fact contains all the labels of elementary paths for any given pair of nodes in $\mathcal{N}$. The elements of this set which are labels of the required k shortest paths are then identified by a certain path tracing algorithm which we modified from the method of "backward subtraction" described by Pollack (1961, p.558).

K-SHORTEST-ELEMENTARY-PATHS ALGORITHM.

The following steps yield k shortest elementary paths from x_i to x_j

Step 0. Set M equal to the label matrix of $\mathcal{N}$.

Step 1. Compute $(M^*)_{hj}$ for all $h \in \{1,2,\ldots,n\}$ by applying Gauss or Jordan elimination to the system $y = M \ominus y \oplus I_j$, where I_j denotes the j^{th} columm of the unit matrix I.

Step 2. If $(M^*)_{ij} = \phi$ or $(M^*)_{ij} = \{0\}$ when $i = j$, terminate; there are no paths from x_i to x_j. If $(M^*)_{ij}$ is free of denominator, set $Y' = (M^*)_{ij}$ and go to step 5. Otherwise, set A equal to the numerator of $(M^*)_{ij}$ and if $i \neq j$ go to step 4, else

Step 3 Set B equal to the denominator of $(M^*)_{ij}$ and B' the multiset obtained from B by deleting all the elements in B which are also in A as well as those which appear without hats (^). Choose $b_0 = \max\{x | x \text{ in } \ominus B'\}$ and apply the division algorithm to $(M^*)_{ij}$ until an element not less than b_0 is obtained. Set Y' equal to the multiset obtained from $\theta B'$ by deleting all those elements which do not also appear in the division result and go to step 5.

Step 4. Set A' equal to the multiset obtained from A by deleting all those elements with hats (^). Choose $a_0 = \max\{x | x \in A'\}$ and apply the division algorithm to $(M^*)_{ij}$ until we obtain an element not less than a_0. Set Y' equal to the multiset obtained from A' by deleting all the elements which do not also appear in the division result.

Step 5. Set Y equal to the set of distinct elements of Y' and trace the paths corresponding to each of the elements of Y in the order of increasing magnitude by using the path-tracing algorithm given below. Terminate when either k elementary paths have been obtained or Y has been exhausted.

Note that the computation of $(M^*)_{hj}$ for all h other than i in step 1 above is a prerequisite for the use of the path-tracing algorithm in step 5. Note also that when more than k shortest elementary paths are subsequently required, we need only trace additional paths corresponding to those remaining elements of Y. The justification of this algorithm will be given at the end of this section.

<u>PATH-TRACING ALGORITHM</u> (cf. Pollack (1961)).

Here we still assume that the label of any elementary closed path in $\mathcal{N}$ is strictly positive. Suppose we wish to trace an elementary path corresponding to the k^{th} smallest label b_k of $(M^*)_{ij}$. Then the following algorithm which presupposes the knowledge of k-min $(M^*)_{hj}$ for all $h \in \{1,2,..,n\}$ (see example 3.3 for the definition of k-min(A)) will trace an elementary path whose label is b_k if at least one such path exists, otherwise it will terminate with a negative answer.

Step 0. Set $s = 1$, $A = \{i\}$ and $B_s = \phi$.

Step 1. Find $a_{ih} \in M_{ih}$ such that $b_k - a_{ih} = b_t$ for some $b_t \in (M^*)_{hj}$. Put each such h in B_s

Step 2. If $j \in B_s$ and $b_k - a_{ij} = 0$, set $i_s = j$ and terminate; an elementary path corresponding to b_k is

$$x_i \xrightarrow{a_{ii_1}} x_{i_1} \xrightarrow{a_{i_1 i_2}} x_{i_2} \;\cdots\cdots\; x_{i_{s-1}} \xrightarrow{a_{i_{s-1} i_s}} x_{i_s}$$

Step 3. If $B_s = \phi$, go to step 6, otherwise choose an $i_s \in B_s$

Step 4. If $b_k - a_{ii_s} \neq 0$ and $i_s \in A$, delete i_s from B_s and return to step 3. Otherwise, put i_s in A.

Step 5. Increase the value of s by 1 and return to step 1 with $i = i_{s-1}$, $b_k = b_t$.

Step 6. Decrease the value of s by 1 and delete i_s from B_s. If $B_1 = \phi$, terminate; there are no elementary paths corresponding to b_k. Otherwise return to step 3.

Note that the assumption of having all the elementary paths in the network carry strictly positive labels ensures that the above algorithm terminates in all cases and its justification is in fact based on the observation that each subpath from x_h to x_j of a k^{th} shortest path from x_i to x_j is always a t^{th} shortest path for some $t \leq k$. We note also that a similar algorithm can also be devised from the knowledge of k-min $(M^*)_{ih}$ for all $h \in \{1,2,\ldots,n\}$.

Now let us illustrate the above two algorithms by considering the following example.

Suppose we wish to find 2 shortest elementary closed paths from x_5 to itself and 3 shortest elementary paths from x_2 to x_6 in the following network.

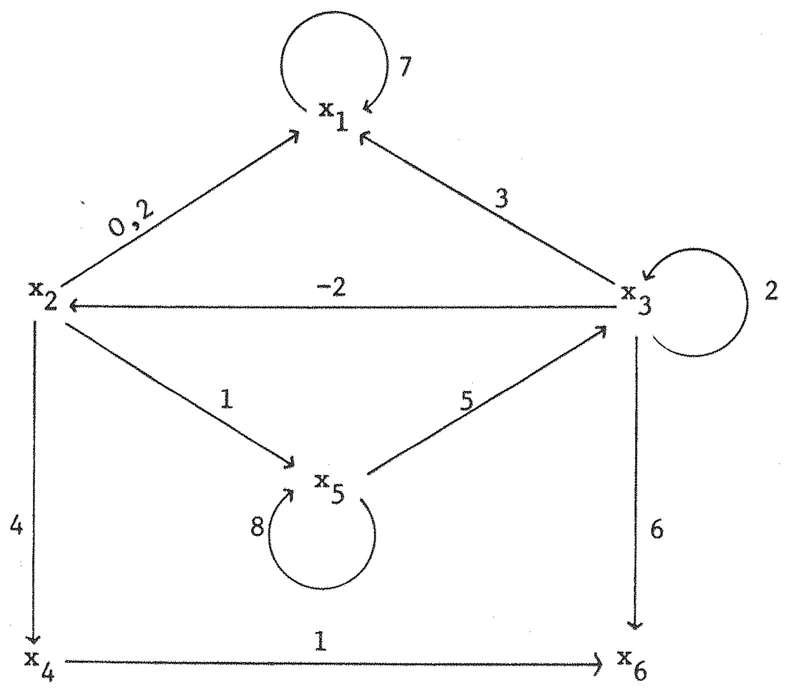

Figure 6.1

The label matrix of the above network is

$$M = \begin{bmatrix} \{7\} & \phi & \phi & \phi & \phi & \phi \\ \{0,2\} & \phi & \phi & \{4\} & \{1\} & \phi \\ \{3\} & \{-2\} & \{2\} & \phi & \phi & \{6\} \\ \phi & \phi & \phi & \phi & \phi & \{1\} \\ \phi & \phi & \{5\} & \phi & \{8\} & \phi \\ \phi & \phi & \phi & \phi & \phi & \phi \end{bmatrix}$$

(i) To find 2 shortest elementary paths from x_5 to itself, we first obtain $(M^*)_{h5}$ for all $h \in \{1,2,\ldots,6\}$ by solving the following system of equations

$$\begin{bmatrix} \{0,\hat{7}\} & \phi & \phi & \phi & \phi & \phi \\ \{\hat{0},\hat{2}\} & \{0\} & \phi & \{\hat{4}\} & \{\hat{1}\} & \phi \\ \{\hat{3}\} & \{-\hat{2}\} & \{0,\hat{2}\} & \phi & \phi & \{\hat{6}\} \\ \phi & \phi & \phi & \{0\} & \phi & \{\hat{1}\} \\ \phi & \phi & \{\hat{5}\} & \phi & \{0,\hat{8}\} & \phi \\ \phi & \phi & \phi & \phi & \phi & \{0\} \end{bmatrix} \odot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} = \begin{bmatrix} \phi \\ \phi \\ \phi \\ \phi \\ \{0\} \\ \phi \end{bmatrix}$$

Using Gauss elimination to solve this system, we obtain, after the completion of the elimination phrase, the following "triangulated" system

$$\begin{bmatrix} \{0,\hat{7}\} & \phi & \phi & \phi & \phi & \phi \\ & \{0\} & \phi & \{\hat{4}\} & \{\hat{1}\} & \phi \\ & & \{0,\hat{2}\} & \{\hat{2}\} & \{-\hat{1}\} & \{\hat{6}\} \\ & & & \{0\} & \phi & \{\hat{1}\} \\ & \phi & & & \dfrac{\{0,\hat{2},\hat{4},\hat{8},10\}}{\{0,\hat{2}\}} & \dfrac{\{\hat{8},\hat{1}1\}}{\{0,\hat{2}\}} \\ & & & & & \{0\} \end{bmatrix} \odot \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{bmatrix} = \begin{bmatrix} \phi \\ \phi \\ \phi \\ \phi \\ \{0\} \\ \phi \end{bmatrix}$$

The back substitution then yields successively

$$y_6 = \phi\ ,\ y_5 = \frac{\{0,\hat{2}\}}{\{0,\hat{2},\hat{4},\hat{8},10\}}\ ,\ y_4 = \phi,\ y_3 = \frac{\{-1\}}{\{0,\hat{2},\hat{4},\hat{8},10\}},$$

$$y_2 = \frac{\{1,\hat{3}\}}{\{0,\hat{2},\hat{4},\hat{8},10\}}\ ,\ y_1 = \phi\ .$$

The next step is to set $A = \{0,\hat{2}\}$, $B = \{0,\hat{2},\hat{4},\hat{8},10\}$ which are respectively the numerator and denominator of y_5 above.

Hence $B' = \{\hat{4},\hat{8}\}$ and $b_0 = 8$. So we apply the division algorithm to y_5 until we obtain $y_5 = \{0,4,6,8, \ldots\}$ which contains all the elements in $\theta B'$. It remains to trace the actual path corresponding to each element of $Y = Y' = \theta B'$. Using the path-tracing algorithm we obtain the required elementary closed paths, namely

$$x_5 \xrightarrow{5} x_3 \xrightarrow{-2} x_2 \xrightarrow{1} x_5 \quad \text{and} \quad x_5 \xrightarrow{8} x_5$$

(ii) To find $(M^*)_{h6}$, we need only replace the right-hand side of the above system, which was obtained after the completion of the elimination phrase, by the 6^{th} column of the unit matrix I. The back substitution will then give successively

$$y_6 = \{0\},\ y_5 = \frac{\{8,11\}}{\{0,\hat{2},\hat{4},\hat{8},10\}},\ y_4 = \{1\},\ y_3 = \frac{\{3,\hat{5},6,\hat{8},\hat{11},13,\hat{14},16\}}{\{0,\hat{2},\hat{2},6,\hat{8},10,10,\hat{12}\}},$$

$$y_2 = \frac{\{5,\hat{7},12,\hat{13},15\}}{\{0,\hat{2},\hat{4},\hat{8},10\}},\quad y_1 = \phi .$$

Let us now set $A = \{5,\hat{7},12,\hat{13},15\}$ which is the numerator of y_2. Hence $A'=\{5,12,15\}$ and $a_0 = 15$. So $y_2 = \{5,9,11,12,13,13,14,15,...\}$ is obtained by using the division algorithm. Since all the elements of A' are contained in y_2, we have $Y' = A'$ and hence $Y = A'$ also. Finally, using the path-tracing algorithm, we find the elementary path $x_2 \xrightarrow{4} x_4 \xrightarrow{1} x_6$ corresponds to 5 in Y, the elementary path $x_2 \xrightarrow{1} x_5 \xrightarrow{5} x_3 \xrightarrow{6} x_6$ corresponds to 12, and no elementary paths has label 15.

As an illustration, let us show how to trace an elementary path corresponding to 12 in Y above. First, we note that 12 is the 4^{th} smallest element in $4\text{-min}(M^*)_{26} = \{5,9,11,12\}$. Thus we need also compute $4\text{-min}(M^*)_{h6}$ for all $h \neq 2$ by using the division algorithm where necessary.

These are as follows.

$$4\text{-min}(M^*)_{16} = \phi \quad , \quad 4\text{-min}(M^*)_{36} = \{3,5,6,7\} \quad ,$$

$$4\text{-min}(M^*)_{46} = \{1\} \ , \quad 4\text{-min}(M^*)_{56} = \{8,10,11,12\} \ , \ 4\text{-min}(M^*)_{66} = \{0\}.$$

Initially, $A = \{2\}$, $B_1 = \phi$. From the 2nd row of the label matrix M, we find that $a_{25} = 1$ is the only element satisfying $12 - a_{25} = 11$, where 11 is the 3rd smallest element in $4\text{-min}(M^*)_{56}$. Hence $B_1 = \{5\}$ and $A = \{2,5\}$. Applying the same procedure to 11, we get $a_{53} = 5$; $11 - a_{53} = 6$, where 6 is the 3rd smallest element in $4\text{-min}(M^*)_{36}$. Since a_{53} is the only such candidate, $B_2 = \{3\}$ and hence $A = \{2,5,3\}$. Applying the same procedure once more, we obtain $a_{36} = 6$; $6 - a_{36} = 0$, where 0 is the first smallest element in $4\text{-min}(M^*)_{66}$. Again, this is the only candidate and hence $B_3 = \{6\}$. Therefore, the algorithm terminates and yields the elementary path

$$x_2 \xrightarrow{1} x_5 \xrightarrow{5} x_3 \xrightarrow{6} x_6$$

PROOF OF THE K-SHORTEST-ELEMENTARY-PATHS ALGORITHM

The validity of the k-shortest elementary-paths algorithm will be verified with the help of the following well-known result relating the elementary paths in a graph G over a field $(X,+,o)$ and the arc-value matrix A.

$$\text{(6.8)} \qquad \det(I-A) = e + \sum_{h=1}^{n} \Big(\sum_{m_1+m_2+..+m_k=h} (-e)^k \circ v(\omega_{m_1}) \circ v(\omega_{m_2}) \circ\ldots\circ v(\omega_{m_k}) \Big)$$

$$\text{(6.9)} \quad (-e)^{(i+j)} \circ \det([I-A]_{ji}) = \sum_{h=2}^{n} \Big(\sum_{m_0+m_1+..+m_k=h-1} (-e)^k \circ v(p_{m_0}) \circ v(\omega_{m_1}) \circ..\circ v(\omega_{m_k}) \Big)$$

if $i \neq j$.

Here $[I-A]_{ji}$ denotes the matrix obtained from I-A by deleting its j^{th} row and i^{th} column; p_{m_0} is an elementary open path from x_i to x_j of order $m_0 \geq 1$; ω_{m_t} is the elementary closed path of order $m_t \geq 1$, for all $t \in \{1,2,\ldots,k\}$; all the paths $p_{m_0}, \omega_{m_1}, \ldots, \omega_{m_k}$ are disjoint, i.e. they do not have any nodes in common. For a proof of (6.8) and (6.9), we shall refer the reader to Ponstein (1966).

Since the label matrix M of $\mathcal{N}$ can also be regarded as the arc-value matrix of the graph G(M) over the quotient-field of the schedule algebra , it follows respectively from (6.8) and (6.9) that

$$(6.10)\quad \det(I\ominus M) = \{0\} \oplus \left(\bigoplus_{h=1}^{n} \left(\bigoplus_{m_1+m_2+\ldots+m_k=h} s(k) \ominus \{c_{m_1} + c_{m_2} + \ldots + c_{m_k}\}\right)\right)$$

$$(6.11)\quad s(i+j)\ominus \det([I\ominus M]_{ji}) = \bigoplus_{h=2}^{n} \left(\bigoplus_{m_0+m_1+\ldots+m_k=h-1} s(k) \ominus\{d_{m_0}+c_{m_1}+\ldots+c_{m_k}\}\right)$$

if $i \neq j$,

where $d_{m_0} \in v(p_{m_0})$, $c_{m_t} \in v(\omega_{m_t})$ for all $t \in \{1,2,\ldots,k\}$,

and $$s(k) = \begin{cases} \{0\} & \text{if } k \text{ is even} \\ \{\hat{0}\} & \text{, otherwise.} \end{cases}$$

Now let us recall our earlier note that the use of Gauss or Jordan elimination in solving $Y = M \odot Y \oplus I$ will always yield $(M^*)_{ij}$ as a quotient of two multisets equivalent to (6.7). In fact, from (6.10) and (6.11) above, it is easily seen that these two multisets are also finite because the network is assumed to have only a finite number of nodes and arcs, and hence a finite number of elementary paths in the network. The proof can now be conveniently divided into two cases.

(i) For $i \neq j$, we have three possibilities, namely

(a) $(M^*)_{ij} = \phi$. There is nothing to prove in this case

(b) $(M^*)_{ij} \neq \phi$ but free of denominator.

Since $(M^*)_{ij}$ so obtained is equivalent to (6.7), it follows that in this case, the numerator must contain a factor equal to $\det(I \ominus M)$. This occurs when all the elementary closed paths are disjoint from the elementary paths from x_i to x_j. Hence $(M^*)_{ij}$ so obtained is already a multiset of elements which are labels of elementary paths from x_i to x_j.

(c) $(M^*)_{ij} = \frac{A}{B}$, where A and B are non-empty finite multisets. Let us assume first that

$$A = s(i+j) \odot \det([I \ominus M]_{ji}) \text{ and } B = \det(I \ominus M).$$

It suffices to show that all the labels of elementary paths from x_i to x_j are contained in A and the elements deleted from A to yield A' in step 4 of the algorithm are not labels of elementary paths from x_i to x_j. Let d_{m_0} be a label of an elementary path in $\mathcal{N}$ which begins at x_i ends at x_j, and has order m_0. Then $d_{m_0} \in v(p_{m_0})$, where p_{m_0} is as defined above. Since $m_0 \geq 1$ (because $i \neq j$), there is always a positive integer $h \geq 2$ such that $m_0 = h-1$, and hence $s(0) \ominus \{d_{m_0}\}$ is a term in the sum on the righthand side of (6.11) above. Since $s(0) = \{0\}$, it follows that $d_{m_0} \in A$. It remains to show that all the elements in A which appear with hats cannot be labels of elementary paths from x_i to x_j. Since $d_{m_0}, c_{m_1}, \ldots, c_{m_k}$ are elements of R, the only elements in A with hats are those corresponding to the terms $s(k) \ominus \{d_{m_0} + c_{m_1} + .. + c_{m_k}\}$ in the sum on the right-hand side of (6.11) where k is odd, and hence they cannot be labels of elementary paths in $\mathcal{N}$. Now consider the case where

$$A \neq s(i+j) \ominus \det([I\ominus M]_{ji}), \quad B \neq \det(I\ominus M) \quad \text{but}$$

$$\frac{A}{B} = \frac{s(i+j)\ominus \det([I\ominus M]_{ji})}{\det(I\ominus M)}$$

This implies that a common factor has been cancelled out. In view of (6.10) and (6.11) above, this can only occur when some elementary closed paths in the network are disjoint from the elementary paths from x_i to x_j. In fact, the common factor which has been cancelled out comprises precisely the labels or sum of labels (disregarding the hats) of these closed paths. Hence it cannot contain labels of elementary paths from x_i to x_j. Thus the above argument for A' remains valid in this case.

(ii) For $i = j$, we have two possibilities:

(a) $(M^*)_{ii} = \{0\}$. Again, there is nothing to prove in this case.

(b) $(M^*)_{ii} = \frac{A}{B}$, where A and B are finite non-empty multisets.

Let M_i be the label matrix of a network obtained from the given network by deleting the node x_i and all the arcs beginning and ending at x_i. It is easily seen that

$$M_i = [M]_{ii} \quad \text{and} \quad I \ominus M_{ii} = I \ominus [M]_{ii} = [I\ominus M]_{ii} .$$

Let us assume first that $A = \det([I\Theta M]_{ii})$ and $B = \det(I\Theta M)$. Again, we have to show that B contains all the labels (disregarding the hats) of elementary closed paths from x_i to itself and the elements deleted from B to yield B' in step 3 of the algorithm are not labels of those elementary closed paths. The argument is similar to that used in (i) above except that we need also show that the elements in B which also appear in A cannot be labels (disregarding the hats) of elementary closed paths from x_i to itself. Indeed, by the definition of M_i and (6.10) above, $A = \det([I\Theta M]_{ii}) = \det(I\Theta M_i)$ cannot contain such labels. Now if $A \neq \det([I\Theta M]_{ii})$ and $B \neq \det(I\Theta M)$ but $\frac{A}{B} = \frac{\det([I\Theta M]_{ii})}{\det(I\Theta M)}$,

then $\det([I\Theta M]_{ii})$ and $\det(I\Theta M)$ must have a common factor. But by the definition of M_i, this common factor cannot contain labels of any elementary closed paths from x_i to itself. Hence the above argument for B' applies.

REFERENCES

BACKHOUSE, R.C. and CARRÉ, B.A. (1975) "Regular Algebra Applied to Path-Finding Problems", J. Inst. Math. Applics., Vol. 15, No.2, pp. 161-186.

BENZAKEN, C. (1968) "Structures Algébriques des Cheminements: Pseudo-Treillis, Gerbiers de Carré Nul", in BIORCI, G.(Ed.), Network and Switching Theory, Academic Press, New York, pp. 40-57.

BIRKHOFF, G. (1967) Lattice Theory, American Mathematical Society Colloquim Publications.

BRUCKER, P. (1974) Theory of Matrix Algorithms, Verlag Anton Hain, Meisenheim Am Glam, Germany. (Mathematical Systems in Economics No. 13)

CARRÉ, B.A. (1971) "An Algebra for Network Routing Problems", J. Inst. Math. Applics., Vol. 7, No.3, pp. 273-294.

CARRÉ, B.A. (1976) To be published by Oxford University Press

DERNIAME, J.C. and PAIR, C. (1971) Problèmes de Cheminement dans las Graphes, Dunod, Paris.

DIJKSTRA, E.W. (1959) "A Note on Two Problems in Connexion with Graphs", Numer. Math., Vol. 1, No.5, pp. 269-271.

EILENBERG, S. (1974) Automata, Languages, and Machines: Volume A, Academic Press, New York.

FLOYD, R.W. (1962) "Algorithm 97: Shortest Path", Comm. ACM, Vol. 5, No. 6, p.345.

FORSYTHE, G.E. (1953) "Solving Linear Algebraic Equations Can Be Interesting", Bull. Amer. Math. Soc., Vol.59, pp.299-329.

FOX, L. (1964) An Introduction to Numerical Linear Algebra, Clarendon Press, Oxford.

GIFFLER, B. (1963) "Scheduling General Production System Using Schedule Algebra", Nav. Res. Log. Quart., Vol. 10, No.3, pp. 237-255.

GIFFLER, B. (1968) "Schedule Algebra: A Progress Report", Nav. Res. Log. Quart., Vol. 15, No.2, pp. 255-280.

GONDRAN, M. (1975) "Algèbre Linéaire et Cheminement dans un Graphe", Rev. Fr. Aut. Inf. Rech. Opér., 9^{e} annee, V-1, pp. 77-99.

IRI, M. (1962) "Fundamentals of the Algebraic and Topological Treatments of General Information Networks", Artcle G-V in RAAG Memoirs, Vol.3, (Research Association of Applied Geometry, Japan) pp. 418-450.

KNUTH, D.E. (1969) The Art of Computer Programming; Volume 2: Seminumerical Algorithms, Addison-Wesley.

LUNTS, A.G. (1950) "The Application of Boolean Matrix Algebra to the Analysis and Synthesis of Relay-Contact Networks" (in Russian), Dokl. Akad. Nauk SSSR, Vol. 70, No.3, pp. 421-423.

McNAUGHTON, R. and YAMADA, H. (1960) "Regular Expressions and State Graphs for Automata", IRE Trans. Electron. Comput. Vol. 9, No.1, pp. 39-47.

MINIEKA, E. and SHIER, D.R. (1973) "Note on an Algebra for the K Best Routes in a Network", J. Inst. Math. Applics., Vol. 11, No.2, pp. 145-149.

MOISIL, G.C. (1960) "Asupra unor Reprezentări ale Grafurilor ce Intervin în Probleme de Economia Transporturilor", Com. Acad. R.P. Romîne, Vol.10, No. 8, pp. 647-652.

MURCHLAND, J.D. (1965) "A New Method for Finding All Elementary Paths in a Complete Directed Graph", Report LSE-TNT-22, London School of Economics.

PAIR, C. (1967) "Sur des algorithmes pour les problemes de cheminement dans les graphes finis" P. Rosenstiehl (ed.), Theory of Graphs, pp. 271-300, Dunod, Gordon and Breach.

PETEANU, V. (1969) "Optimal Paths in Networks and Generalizations (I) " Mathematica (Cluj), Vol. 11, No.2, pp. 311-327.

PETEANU (1970) "Optimal Paths in Networks and Generalizations (II) " Mathematica (Cluj), Vol. 12, No. 1, pp. 159 - 186.

POLLACK, M. (1961) "Solutions of the K Best-Routes through a network - A Review", J. of Math.Anal. and Applics., Vol. 13, No. 3, pp. 547-559.

PONSTEIN, J.(1966) "Self-Avoiding Paths and the Adjacency Matrix of a Graph", J. SIAM Appl. Maths. Vol. 14, No. 3, pp. 600-609.

ROBERT, P. and FERLAND, J. (1968) "Généralisation de l' Algorithme de Warshall", Rev. Fr. Inf. Rech. Opér., Vol.2, No. 7, pp. 71 - 85.

ROY, B. (1959) "Transitivité et Connexité", C.R. Acad. Sci. Paris, Vol. 249, No.2, pp. 216-218.

ROY, B. (1975) "Chemins et Circuits: Enumération et Optimisation", in Combinatorial Programming: Methods and Applications, ROY, B. et al., D. Reidel Publishing Co., Dordrecht, Hollande.

SHIER, D.R. (1973) "Network Optimality Problems: Theory and Algorithms" Ph.D. Thesis (London).

SHIER, D.R. (1974) "Iterative Methods for Determining the k Shortest Paths in a Network", submitted for publication.

TOMESCU, I (1968) "Sur l'algorithme matricel de B.Roy", *Rev.Fr. Inf. Rech.Opér.* Vol. 2, No. 7, pp. 87-91.

VAZSONYI, A. (1954) "The Use of Mathematics in Production and Inventory Control", *Management Science*, Vol.1, No. 1, pp. 70-85.

WARSHALL, S. (1962) "A Theorem on Boolean Matrices", *J.ACM*, Vol. 9, No. 1, pp. 11-12.

WONGSEELASHOTE, A. (1976) "An Algebra for Determining All Path-Values in a Network with Application to K-Shortest-Paths Problems", *Networks*, Vol. 6, No.3, pp.

YEN, J.Y. (1970) "An Algorithm for Finding Shortest Routes from All Source Nodes to a Given Destination in General Networks", *Quart. Appl. Math.*, Vol. 27, pp. 526-530.

YOELI, M. (1961) "A Note on a Generalization of Boolean Matrix Theory", *Amer. Math. Monthly*, Vol. 68, pp. 552-557.