Absorbing Lexicographic Products in Metarouting

Eric Parsonage, Hung X. Nguyen, Matthew Roughan
University of Adelaide, Australia
{eric.parsonage,hung.nguyen,matthew.roughan}@adelaide.edu.au

***Abstract*—Modern treatments of routing protocols use algebraic techniques to derive the protocol's properties, permitting a semantic richness more flexible than simple numerical "shortest paths". Many such routing protocols make preference decisions based on multiple criteria. This fits well with an algebraic formulation with each strata in the decision process modeled as an algebraic structure, that are combined to create the full routing protocol. Routing protocols constructed in this manner are the focus of this paper. To implement such a routing protocol we must understand the properties needed on each of the algebraic formulations representing a strata. In this paper we examine a stratified routing algebra based on a recently suggested absorbing product and provide the necessary and sufficient conditions required by each of the operands to guarantee that such a routing language ensures globally optimal paths will be found.**

I. Introduction

Convergence of a routing protocol to an optimum is crucial for stability and efficiency. Proofs of convergence have their roots in the work of Carré [1] who in 1971 showed that several different kinds of pathfinding problems could be solved by variants of classical methods in linear algebra using matrices over semirings. Since then it has been possible to view a routing protocol as comprising two distinct components:

$$\text{routing protocol} = \text{routing language} + \text{algorithm},$$

where a protocol's routing language is used to configure a network and the protocol's algorithm computes routing solutions using the network's configuration. This abstraction is very useful as we know that if a routing language has certain properties then the associated protocol will converge.

However, more recently, it was discovered that the Border Gateway Protocol (BGP), that is used universally for inter-domain routing in the Internet, does not always converge [2], [3], despite the fact that the properties needed for such convergence were known. It might reasonably be concluded that although the theory was available, it was too hard to apply to the complex routing protocols needed to implement the flexible policies of modern inter-domain routing.

Carré's formulation has been extended in various ways in the intervening years [4]–[6]. Griffin and Sobrinho [6], in particular, extended the standard formulation to allow for combinations of simple building blocks into larger, more complex routing protocols. The process, that is referred to as *metarouting*, allows one to derive the properties of the combined algebra from its components without the tedious or difficult process of deriving these properties *ab initio*. The most prominent example of such a construction is the lexicographic product, which has been used to model BGP-like protocols where the choice of best routes descends through a series of metrics, using those after the first only as a tie-break if the preceding metrics are equal.

In this paper we examine a structure based on an *absorbing lexicographic product* proposed by Gurney [7] and further elaborated upon by Griffin [8], specifically designed to overcome practical difficulties that arise when implementing a routing language using a simple lexicographic product.

Several examples of the use of the absorbing lexicographic product to model subsets of BGP policy can be found in [8]. One of these considers a simplified model of BGP with only two components. The first component is used to model the business relationships between Autonomous Systems (ASs) (call it local preference to match the conventions of BGP) and the second models AS path length. The path selection decision process is based on policy first (via the local preference metric) with tiebreaks then decided by the shortest AS path. The classification of business relationships between ASs into customer-provider, peer-peer, and downstream leads to import/export rules such as: (1) an AS does not export to a provider or peer routes that it has learned from other providers and other peers; and (2) an AS can export to its customers any route it knows of. The model implements these policies by assigning functions to links in the network that represent the combined import and export policies of the routers at each end of a link. Thus as paths are propagated these functions combine to cause invalid paths (paths that are filtered out by the import/export rules) to be marked with an absorbing zero thus excluding them from further consideration.

A less conventional example is the calculation of shared-risk groups across paths using Martelli's algebra [9]. Here the size of the state information that needs be conveyed increases as path lengths increase. Thus, eliminating routes of no interest from consideration would reduce the computational and communications overheads of the protocol. This could be achieved by using an absorbing lexicographic product where the first component is used to absorb paths of no interest and the second component is an instance of Martelli's algebra.

Unfortunately the increased flexibility of the absorbing product comes at the cost of disallowing the use of several powerful theorems for the proof of its properties. Sufficient conditions for a routing algebra formed from the absorbing product always to converge to a globally optimal solution were previously known [8]. Our main result is the provision of *necessary and sufficient* conditions.

Discovery of necessary conditions is crucial if we want to create these routing algebras with all the possible flexibility that is allowed given the requirement for convergence.

Moreover, necessary conditions are required for the analysis of existing routing protocols that may be convergent, despite not meeting the existing sufficient conditions.

II. Definitions and Simple Results

As an *aide-mémoire* this section contains pertinent definitions and propositions. These can be found in textbooks such as [4].

Definition 1. *A semigroup consists of a set S and a binary operator $\oplus$ under which S is closed, and which is associative:*

$$\forall\, a, b, c \in S : (a \oplus b) \oplus c = a \oplus (b \oplus c).$$

Definition 2. *A semigroup $(S, \oplus)$ is* commutative *if*

$$\forall\, a, b \in S : a \oplus b = b \oplus a.$$

Definition 3. *A semigroup $(S, \oplus)$ has an* identity $\bar{0}_S$ *if*

$$\exists\, \bar{0}_S \in S \text{ such that } \forall\, a \in S : \bar{0}_S \oplus a = a \oplus \bar{0}_S = a.$$

A semigroup with an identity is called a *monoid*.

Definition 4. *A semigroup $(S, \oplus)$ is* selective *if*

$$\forall\, a, b \in S : a \oplus b = a \text{ or } a \oplus b = b.$$

Definition 5. *A semigroup $(S, \oplus)$ is* idempotent *if $\forall\, a \in S$: $a = a \oplus a$. A selective semigroup is also idempotent.*

Definition 6. *An algebra of monoid endomorphisms (AME) is defined as a quadtuple $(S, F, \oplus, \bar{0})$ where $(S, \oplus, \bar{0})$ forms a monoid and F is the set of mappings $S \to S$ with:*

$$\forall\, f \in F, a \in S, b \in S : f(a \oplus b) = f(a) \oplus f(b)$$

and $\forall\, f \in F : f(\bar{0}) = \bar{0}$. Thus F is the set of endomorphisms of $(S, \oplus)$ that satisfy $f(\bar{0}) = \bar{0}$.

Definition 7. *A semiring denoted $(R, \oplus, \otimes, \bar{0}, \bar{1})$ is a set R with two binary operators $\oplus$ and $\otimes$, called addition and multiplication, such that the following four properties hold:*

1) *$(R, \oplus)$ is a commutative monoid with identity $\bar{0}$.*
2) *$(R, \otimes)$ is a monoid with identity $\bar{1}$.*
3) *Multiplication distributes over addition from the left and the right.*
4) *$\forall a \in R : \bar{0} \otimes a = a \otimes \bar{0} = \bar{0}$.*

Proposition 1. *Given a commutative monoid $(S, \oplus)$, it is always possible to define a reflexive, transitive binary operation, denoted $\leq$, as either:*

$$a \leq b \iff \exists\, c \in S \text{ such that } b = a \oplus c, \text{ or} \tag{1}$$
$$a \leq b \iff \exists\, c \in S \text{ such that } a = b \oplus c. \tag{2}$$

For $\leq$ to be a total order we need it to be antisymmetric (that is if $a \leq b$ and $b \leq a$ then $a = b$). Since the antisymmetry of $\leq$ is not guaranteed, we call $\leq$ a *preorder relation* of $(S, \oplus)$.

Proposition 2. *If $(S, \oplus)$ is a commutative and idempotent monoid, then either of the above preorder relations is a* partial order relation.

Proposition 3. *If $(S, \oplus)$ is a commutative, selective (thus idempotent) monoid then $\leq$ is a* total order relation.

Proposition 4. *If $(S, \oplus)$ is a commutative, idempotent monoid then either*

$$a \leq b \iff a = a \oplus b \text{ or } a \leq b \iff b = a \oplus b.$$

Proposition 5. *If $(S, \oplus)$ is a commutative, selective monoid then $\bar{0}$ is either the largest or smallest element in $(S, \oplus)$.*

III. The Algebraic Formulation of Routing

Simple shortest-path algorithms aim to minimize the sum of the weights of the links forming a path. Carré's key observation was that this decision process could be encoded in the semiring $(\mathbb{R}^+ \cup \infty, \min, +, \infty, 1)$ where $\oplus$ chooses shortest paths using the min operator, and $\otimes$ extends paths using the $+$ operator. Thus linear algebra could be used to find shortest paths without concern about the underlying routing protocol. The idea extends to any semiring, though guarantees of convergence and path optimality require the set and the operators to satisfy specific properties.

We can replace the $\otimes$ operator with a set of functions F: in the simplest case, consider f_b, where $f_b(a) = b \otimes a$, with $F = f_b : b \in R$. The use of arbitrary functions allows decision processes with more semantic richness than simple shortest-path algorithms provide. Gondran and Minoux [4] investigate the requirements for such a construction to provide globally optimal solutions, and Griffin uses these types of algebraic structures as building blocks [8].

A. Lexicographic Products

Some routing protocols (*e.g.*, BGP) make preference decisions based on multiple criteria in a way that can be implemented as lexicographic choice. That is, the criteria are evaluated in order of importance with criteria of lower importance used to break ties arising from more important criteria. Evaluation ends when a criterion has yielded a decisive result. This is exactly the type of decision process that results from constructing a routing algebra using a lexicographic product.

The absorbing lexicographic product presented by Griffin [8] builds upon a specific lexicographic product defined thus:

Definition 8. *Let $(S, \oplus_S, \bar{0}_S, F_S)$ be a commutative and selective monoid augmented with a set of functions $F_S : S \mapsto S$ which have the property that $\forall f \in F_S : f(\bar{0}_S) = \bar{0}_S$ and let $(T, \oplus_T, \bar{0}_T, F_T)$ be a monoid also augmented with a set of functions $F_T : T \mapsto T$ such that $\forall f \in F_T : f(\bar{0}_T) = \bar{0}_T$, since S is commutative and selective by Proposition 3 we can define $<$ on S, which allows the operation of $\oplus_{S\times T}$ in the following manner:*

$$(s_1, t_1) \oplus_{S\times T} (s_2, t_2) = \begin{cases} (s_1, t_1 \oplus_T t_2) & \text{if } s_1 = s_2, \\ (s_1, t_1) & \text{if } s_1 < s_2, \\ (s_2, t_2) & \text{if } s_2 < s_1. \end{cases}$$

The multiplicative component is constructed using the direct product, where a pair $(f, g) \in F_S \times F_T$ is taken to represent a new function $h : (S \times T) \to (S \times T)$ where $h(s, t) =$

Property	Definition	Remarks
DIST	$\forall\, a, b \in S,\ f \in F : f(a \oplus b) = f(a) \oplus f(b)$	distributive, or additive
INFL	$\forall\, a \in S,\ f \in F : a \leq f(a)$	inflationary
S.INFL	$\forall\, a \in S,\ f \in F : a < f(a)$	strictly inflationary
K	$\forall\, a, b \in S,\ f \in F : f(a) = f(b) \implies a = b$	cancellative, or injective
$\mathrm{K}_{\bar{0}}$	$\forall\, a, b \in S,\ f \in F : f(a) = f(b) \implies (a = b \vee f(a) = \bar{0})$	almost cancellative, or 0-cancellative
C	$\forall\, a, b \in S,\ f \in F : f(a) = f(b)$	constant
$\mathrm{C}_{\bar{0}}$	$\forall\, a, b \in S,\ f \in F : f(a) \neq f(b) \implies (f(a) = \bar{0} \vee f(b) = \bar{0})$	almost constant

TABLE I
PROPERTIES OF $(S, \oplus, F, \bar{0})$ DISCUSSED. TAKEN FROM [8].

$(f(s), g(t))$. *Now define the lexicographic product of* S *and* T*, denoted* $S \vec{\times} T$*, as* $(S \times T, \oplus_{S \times T}, (\bar{0}_S, \bar{0}_T), F_S \times F_T)$.

In an exhaustive examination of the various types of structure that can be combined using a lexicographic product [10] Gurney and Griffin prove that the main property required of an algebra of this type in order to ensure globally optimal paths are found is that it should be distributive. This leads to an important question: what are the necessary and sufficient properties of the components of such an algebra that will ensure the algebra is distributive? They answered this emphatically with the following (the properties required are described in Table I):

$$\mathrm{DIST}(S \vec{\times} T) \iff \mathrm{DIST}(S) \wedge \mathrm{DIST}(T) \wedge (\mathrm{K}(S) \vee \mathrm{C}(T)). \quad (3)$$

B. Motivation for an absorbing product

There is a problem with the above definition when $(S, \oplus, \bar{0}, F_S)$ represents a finite algebra. To ensure that $S \vec{\times} T$ causes only globally optimal solutions to be found S must be distributive and cancellative (properties DIST and K in Table I). The combination of these conditions force F_S to contain only the identity function (consider Lemma 1 below, along with property K and the order on S). This is very restrictive.

Further motivation for modifying the lexicographic product of Definition 8 can be found in [7]. These centre around the overloading of the semantics of infinite or error elements. Given our chosen preorder we have $\bar{0}$ as the largest element in S. This could be taken to be either a largest element that is not an error condition; an erroneous route encoded as a maximal element to be preferred less than a valid route; or as a place keeper in the operation of the algorithm, such as an initial value that will be overwritten by better information.

The solution proposed by Griffin in [8] is to define the absorbing lexicographic product denoted $S \vec{\times}_{\bar{0}} T$ by building upon Definition 8 thus:

Definition 9. *Using the definitions for* $(S, \oplus_S, \bar{0}_S, F_S)$ *and* $(T, \oplus_T, \bar{0}_T, F_T)$ *from Definition 8, take the set*

$$S \times_{\bar{0}} T = \{S \backslash \bar{0}_S \times T\} \cup \{\bar{0}\} \text{ where } \bar{0} \notin S \text{ and } \bar{0} \notin T,$$

Now, rather than pairs of the form (s, t)*, we use pairs of the form* $\langle s, t \rangle$*, which denote elements of* $S \times_{\bar{0}} T$*,*

$$\langle s, t \rangle = \begin{cases} \bar{0} & \text{if } s = \bar{0}_S, \\ (s, t) & \text{otherwise.} \end{cases}$$

Now redefine the direct product so that every pair $(f, g) \in F_S \times F_T$ *is taken to represent a function* $h : (S \times_{\bar{0}} T) \to (S \times_{\bar{0}} T)$ *where* $h(\langle s, t \rangle) = \langle f(s), g(t) \rangle$ *and* $h(\bar{0}) = \bar{0}$.

Finally $\oplus_{S \vec{\times}_{\bar{0}} T}$ *extends* $\oplus_{S \times T}$ *in Definition 8 to handle the absorbing element* $\bar{0}$ *thus:*

$$\bar{0} \oplus_{S \vec{\times}_{\bar{0}} T} \langle s, t \rangle = \langle s, t \rangle \oplus_{S \vec{\times}_{\bar{0}} T} \bar{0} = \langle s, t \rangle.$$

In [10] Gurney and Griffin provide proofs that give the properties required of an algebra created with the standard lexicographic product that will ensure globally optimal paths are found. Examination of these reveals that the same properties are required of algebras created with the absorbing lexicographic product if the same end is to be achieved. This leaves the task of proving a statement similar to (3) that will supply the necessary and sufficient properties of the components of an algebra creating using the absorbing lexicographic product to ensure the algebra is distributive. It is this result which is the main contribution of this paper as proved in Section IV.

C. Locally vs globally optimal paths

It has been shown that algebraic structures which satisfy all of the semiring axioms except distributivity will provide solutions when used as a routing algebra; however, rather than calculating globally optimal paths between each pair of nodes, the result is a stable solution among path assignments with the properties of a Nash Equilibrium [11]. That is, if we view each of the nodes in the graph as having their own preferences amongst possible paths and they are only allowed to choose paths that are consistent with the choices made by their neighbours, then the solution found is one where the assignment of paths to nodes is such that no node can choose a better path from the candidates available as a result of the choices of the other nodes. In the literature this is referred to as "the stable paths problem" [12] and is the type of solution that BGP attempts to find. In [8], Griffin is interested in the properties that are required of an algebra created using the absorbing lexicographic product such that these "locally optimal" solutions can be found. The result proved in Section IV forms an "upper bound" in the sense that we expect weaker conditions to ensure local optimality.

IV. CONDITIONS FOR $S \vec{\times}_{\bar{0}} T$ TO BE DISTRIBUTIVE

We will need a set of conditions on $(S, \oplus_S, \bar{0}_S, F_S)$ and $(T, \oplus_T, \bar{0}_T, F_T)$ for $S \vec{\times}_{\bar{0}} T$ to be distributive and vice-versa.

These properties are defined in Table I and are discussed in the following lemmas.

Lemma 1. *If $(S, \oplus)$ is a commutative, selective monoid with an associated order $\leq$ and $F : S \mapsto S$ a set of functions mapping elements of S to elements of S, then $f(a \oplus b) = f(a) \oplus f(b)$ for all $a, b \in S$ and $f \in F$ is equivalent to $a \leq b \implies f(a) \leq f(b)$ for all $a, b \in S$ and $f \in F$* (i.e., *distibutivity/additivity is equivalent to monotonicity).*

Proof: By Proposition 3, the associated order $\leq$ on $(S, \oplus)$ is total. Without loss of generality we assume that (2) is the preorder; hence $a \leq b \iff a = a \oplus b$.

We first prove that if $(S, \oplus)$ is distributive then $a \leq b$ implies $f(a) \leq f(b)$ for all $a, b \in S$ and $f \in F$. Assume $a \leq b$ (total order and $\oplus$ commutative means we can make this assumption without loss of generality) then:

$$\begin{aligned} a \leq b &\iff a = a \oplus b \\ &\implies f(a) = f(a \oplus b) \\ &\implies f(a) = f(a) \oplus f(b), \text{ (as } (S, \oplus) \text{ is distributive)} \\ &\implies f(a) \leq f(b). \end{aligned}$$

We now prove the reverse direction that if $\forall\, a, b \in S$ and $\forall\, f \in F$, $a \leq b \implies f(a) \leq f(b)$, then $(S, \oplus)$ is distributive. For $f \in F$ and $a, b \in S$ assume that $a \leq b$ (total order and $\oplus$ commutative means we can make this assumption without loss of generality) thus:

$$a \leq b \implies f(a) \leq f(b) \implies f(a) = f(a) \oplus f(b).$$

and

$$a \leq b \implies a = a \oplus b \implies f(a) = f(a \oplus b).$$

Thus, $\forall\, a, b \in S$ and $\forall\, f \in F$, $f(a \oplus b) = f(a) \oplus f(b) = f(a)$. Hence, $(S, \oplus)$ is distributive. ■

By design $(S, \oplus)$ is a commutative, selective monoid thus by Proposition 3 it has total order. Therefore, $\forall s_1, s_2 \in S$ we can assume (without loss of generality) that $s_1 \leq s_2$ is defined by the preorder relation described in (2). Note that under this relation, $\bar{0}_S \geq s$ for all $s \in S$ and that there are four cases for the relationship between s_1, s_2 and the special elements $\bar{0}_S$: (1) $s_1 = s_2 = \bar{0}_S$, (2) $s_1 = s_2 \neq \bar{0}_S$, (3) $s_1 < s_2 = \bar{0}_S$, and (4) $s_1 < s_2 \neq \bar{0}_S$. Our proofs rely on enumeration of these cases.

Even though the zero element and the operator $\oplus$ are defined on all the sets S, T and $S \vec{\times}_{\bar{0}} T$, for convenience, in the rest of this paper when the context is clear we will drop the subscript from the notation $\bar{0}$ and $\oplus$.

Lemma 2. $\text{C}(S) \iff \text{C}(S \vec{\times}_{\bar{0}} T) \iff h(u) = \bar{0}$ *for all $h \in F_{S \vec{\times}_{\bar{0}} T}$ and for all $u \in S \vec{\times}_{\bar{0}} T$.*

We omit the proof as it follows from the definitions of $\text{C}(S)$: $\text{C}(S)$ is a necessary and sufficient condition for $S \vec{\times}_{\bar{0}} T$ to be trivially distributive.

Lemma 3. $\text{DIST}(S) \wedge \text{C}(T) \implies \text{DIST}(S \vec{\times}_{\bar{0}} T)$

Proof: We need to show that if $f(s_1 \oplus s_2) = f(s_1) \oplus f(s_2)$ and $g(t) = \bar{0}$ for all $s_1, s_2 \in S$, $t \in T$, $f \in F_S$ and $g \in F_T$, then $h(\langle s_1, t_1 \rangle \oplus \langle s_2, t_2 \rangle) = h(\langle s_1, t_1 \rangle) \oplus h(\langle s_2, t_2 \rangle)$.

We consider the four cases for the relationship between s_1 and s_2 and $\bar{0}_S$ separately.

1) $s_1 = s_2 = \bar{0}$:
$$h(\langle s_1, t_1 \rangle \oplus \langle s_2, t_2 \rangle) = h(\bar{0} \oplus \bar{0}) = h(\bar{0}) = \bar{0},$$
and
$$h(\langle s_1, t_1 \rangle) \oplus h(\langle s_2, t_2 \rangle) = h(\bar{0}) \oplus h(\bar{0}) = \bar{0}.$$
2) $s_1 = s_2 \neq \bar{0}$ giving rise to the following two sub-cases.
   a) $f(s_1) = f(s_2) = \bar{0}$:
   $$h(\langle s_1, t_1 \rangle \oplus \langle s_2, t_2 \rangle) = \langle f(s_1), g(t_1 \oplus t_2) \rangle = \bar{0},$$
   and
   $$\begin{aligned} &h(\langle s_1, t_1 \rangle) \oplus h(\langle s_2, t_2 \rangle) \\ &= \langle f(s_1), g(t_1) \rangle \oplus \langle f(s_2), g(t_2) \rangle = \bar{0} \oplus \bar{0} = \bar{0}. \end{aligned}$$
   b) $f(s_1) = f(s_2) \neq \bar{0}$: since $\text{C}(T)$ implies $g(t) = 0,\ \forall t \in T$, we get
   $$h(\langle s_1, t_1 \rangle \oplus \langle s_2, t_2 \rangle) = h(\langle s_1, t_1 \oplus t_2 \rangle) = \langle f(s_1), \bar{0} \rangle,$$
   and
   $$\begin{aligned} &h(\langle s_1, t_1 \rangle) \oplus h(\langle s_2, t_2 \rangle) \\ &= \langle f(s_1), g(t_1) \oplus g(t_2) \rangle = \langle f(s_1), \bar{0} \rangle. \end{aligned}$$
3) $s_1 < s_2 = \bar{0}$:
$$h(\langle s_1, t_1 \rangle \oplus \langle s_2, t_2 \rangle) = h(\langle s_1, t_1 \rangle \oplus \langle \bar{0}, t_2 \rangle) = h(\langle s_1, t_1 \rangle),$$
and
$$h(\langle s_1, t_1 \rangle) \oplus h(\langle s_2, t_2 \rangle) = h(\langle s_1, t_1 \rangle) \oplus \bar{0} = h(\langle s_1, t_1 \rangle).$$
4) $s_1 < s_2 \neq \bar{0}$: from Lemma 1, $\text{DIST}(S)$ implies that $f(s_1) \leq f(s_2)$ if $s_1 < s_2$. Furthermore, as $s_1 < s_2$, in cases 4c and 4d we can use
$$h(\langle s_1, t_1 \rangle \oplus \langle s_2, t_2 \rangle) = h(\langle s_1, t_1 \rangle) = \langle f(s_1), g(t_1) \rangle.$$
We need to consider the following subcases for the relationship between $f(s_1)$ and $f(s_2)$.
   a) $f(s_1) = f(s_2) = \bar{0}$: thus
   $$h(\langle s_1, t_1 \rangle \oplus \langle s_2, t_2 \rangle) = \langle f(s_1), g(t_1) \rangle = \bar{0},$$
   and
   $$\begin{aligned} &h(\langle s_1, t_1 \rangle) \oplus h(\langle s_2, t_2 \rangle) \\ &= \langle f(s_1), g(t_1) \rangle \oplus \langle f(s_2), g(t_2) \rangle = \bar{0} \oplus \bar{0} = \bar{0}. \end{aligned}$$
   b) $f(s_1) = f(s_2) \neq \bar{0}$: since $\text{C}(T)$ implies $g(t) = 0,\ \forall t \in T$,
   $$h(\langle s_1, t_1 \rangle \oplus \langle s_2, t_2 \rangle) = h(\langle s_1, t_1 \rangle) = \langle f(s_1), \bar{0} \rangle,$$
   and as shown in case 2b
   $$h(\langle s_1, t_1 \rangle) \oplus h(\langle s_2, t_2 \rangle) = \langle f(s_1), \bar{0} \rangle.$$

c) $f(s_1) < f(s_2) = \bar{0}$:

$$h(\langle s_1, t_1\rangle) \oplus h(\langle s_2, t_2\rangle) \\ = \langle f(s_1), g(t_1)\rangle \oplus \bar{0} = \langle f(s_1), g(t_1)\rangle.$$

d) $f(s_1) < f(s_2) \neq \bar{0}$:

$$h(\langle s_1, t_1\rangle) \oplus h(\langle s_2, t_2\rangle) = \langle f(s_1), g(t_1)\rangle,$$

as $f(s_1) < f(s_2)$.

■

Lemma 4. $\text{DIST}(S) \wedge \text{DIST}(T) \wedge \text{K}_{\bar{0}}(S) \implies \text{DIST}(S \vec{\times}_{\bar{0}} T)$

Proof: We again consider the four cases for the relationship between s_1, s_2 and $\bar{0}_S$.

1) $s_1 = s_2 = \bar{0}$: the same result as case 1 of Lemma 3, $h(\langle s_1, t_1\rangle \oplus \langle s_2, t_2\rangle) = h(\langle s_1, t_1\rangle) \oplus h(\langle s_2, t_2\rangle) = \bar{0}$.
2) $s_1 = s_2 \neq \bar{0}$: giving rise to two cases for the relationship between $f(s_1)$ and $f(s_2)$:
   a) $f(s_1) = \bar{0}$ and $f(s_2) = \bar{0}$ as shown in case 2a of Lemma 3, $h(\langle s_1, t_1\rangle \oplus \langle s_2, t_2\rangle) = h(\langle s_1, t_1\rangle) \oplus h(\langle s_2, t_2\rangle) = \bar{0}$.
   b) $f(s_1) \neq \bar{0}$ and $f(s_2) \neq \bar{0}$ thus

   $$h(\langle s_1, t_1\rangle \oplus \langle s_2, t_2\rangle) = \langle f(s_1), g(t_1 \oplus t_2)\rangle,$$

   and

   $$h(\langle s_1, t_1\rangle) \oplus h(\langle s_2, t_2\rangle) = \langle f(s_1), g(t_1) \oplus g(t_2)\rangle \\ = \langle f(s_1), g(t_1 \oplus t_2)\rangle \text{ (as DIST}(T)).$$
3) $s_1 < s_2 = \bar{0}$: again similarly to case 3 in Lemma 3 we have $h(\langle s_1, t_1\rangle \oplus \langle s_2, t_2\rangle) = h(\langle s_1, t_1\rangle) \oplus h(\langle s_2, t_2\rangle) = \langle f(s_1), g(t_1)\rangle$.
4) $s_1 < s_2 \neq \bar{0}$, using $\text{DIST}(S)$, $\text{K}_{\bar{0}}(S)$ and Lemma 1, we get $f(s_1) < f(s_2)$, giving the following two subcases.
   a) $f(s_1) < f(s_2) = \bar{0}$ thus $h(\langle s_1, t_1\rangle \oplus \langle s_2, t_2\rangle) = \langle f(s_1), g(t_1)\rangle = h(\langle s_1, t_1\rangle) \oplus h(\langle s_2, t_2\rangle)$ as shown in case 4c of Lemma 3.
   b) $f(s_1) < f(s_2) \neq \bar{0}$ thus $h(\langle s_1, t_1\rangle \oplus \langle s_2, t_2\rangle) = \langle f(s_1), g(t_1)\rangle = h(\langle s_1, t_1\rangle) \oplus h(\langle s_2, t_2\rangle)$ as shown in case 4d of Lemma 3.

■

An alternative method for proving the previous two lemmas is to show that the equivalence relation defined by $(s_1, t_1) \sim (s_2, t_2) \iff ((s_1 = s_2 \text{ and } t_1 = t_2) \text{ or } s_1 = s_2 = \bar{0})$ is a congruence. This provides a connection between the ordinary lexicographic product and the absorbing lexicographic product, allowing the theorems of [7] and [10] to be applied. The approach above is used here because (i) it uses only basic steps, suitable for implementation through theorem provers, and (ii) it can be extended to the case where both S and T can throw an absorbing state, a case that we intend to tackle in future work.

Lemma 5. $\text{DIST}(S \vec{\times}_{\bar{0}} T) \implies \text{DIST}(S)$

Proof: Lemma 1 gives $\text{DIST}(S) \iff s_1 \leq s_2 \implies f(s_1) \leq f(s_2)$ for all $s_1, s_2 \in S$ and $f \in F$, thus we only need to show that $\text{DIST}(S \vec{\times}_{\bar{0}} T)$ implies $s_1 \leq s_2 \implies f(s_1) \leq f(s_2)$ for all $s_1, s_2 \in S$.

Again, we consider four cases for the relationship between s_1, s_2 and $\bar{0}$:

1) $s_1 = s_2 = \bar{0}$ thus $f(s_1) = f(s_2)$
2) $s_1 = s_2 \neq \bar{0}$ thus $f(s_1) = f(s_2)$
3) $s_1 < s_2 = \bar{0}$ giving rise to two cases for the relationship between $f(s_1)$ and $f(s_2)$:
   a) $f(s_1) = \bar{0}_S$ and $f(s_2) = \bar{0}_S$ thus $f(s_1) = f(s_2)$
   b) $f(s_1) \neq \bar{0}_S$ and $f(s_2) = \bar{0}_S$ as $\bar{0}_S \geq s$ for all $s \in S$, $f(s_1) \leq f(s_2) = \bar{0}_S$.
4) $s_1 < s_2 \neq \bar{0}$ we prove that $f(s_1) \leq f(s_2)$ by contradiction. Assume that $f(s_1) > f(s_2)$ as $\text{DIST}(S \vec{\times}_{\bar{0}} T)$ implies

$$h(\langle s_1, t\rangle \oplus \langle s_2, t\rangle) = h(\langle s_1, t\rangle) \oplus h(\langle s_2, t\rangle) \\ \iff \langle f(s_1), g(t)\rangle = \langle f(s_2), g(t)\rangle \\ \iff f(s_1) = f(s_2)$$

$\Rightarrow\Leftarrow$ thus $f(s_1) \leq f(s_2)$. ■

Lemma 6. $\neg\text{C}(S) \wedge \text{DIST}(S \vec{\times}_{\bar{0}} T) \implies \text{DIST}(T)$

Proof: As $\neg\text{C}(S)$ there exists $s \in S$, $f \in F_S$ such that $f(s) \neq \bar{0}_S$. From $\text{DIST}(S \vec{\times}_{\bar{0}} T)$ for all $t_1, t_2 \in T$, $g \in F_T$ with $h = \langle f, g\rangle \in F_{S \vec{\times}_{\bar{0}} T}$ we have

$$h(\langle s, t_1\rangle \oplus \langle s, t_2\rangle) = h(\langle s, t_1\rangle) \oplus h(\langle s, t_2\rangle)$$

which implies

$$\langle f(s), g(t_1 \oplus t_2)\rangle = \langle f(s), g(t_1)\rangle \oplus \langle f(s), g(t_2)\rangle.$$

Now $f(s) \neq \bar{0}_S$ thus $\langle f(s), g(t_1 \oplus t_2)\rangle \neq \bar{0}$ which implies $\forall\, t_1, t_2 \in T,\ g \in F_T\ g(t_1 \oplus t_2) = g(t_1) \oplus g(t_2)$ hence $\text{DIST}(T)$. ■

Lemma 7. $\neg\text{C}(S) \wedge \text{DIST}(S \vec{\times}_{\bar{0}} T) \implies \text{K}_{\bar{0}}(S) \vee \text{C}(T)$

Proof: Proof by contradiction. Assume $\neg\text{C}(S) \wedge \text{DIST}(S \vec{\times}_{\bar{0}} T) \wedge \neg\text{K}_{\bar{0}}(S) \wedge \neg\text{C}(T)$ thus

$$\neg\text{C}(T) \implies \exists t \in T, g \in F_T : g(t) \neq \bar{0}_T,$$

and

$$\neg\text{K}_{\bar{0}}(S) \implies \exists s_1, s_2 \in S, f \in F_S : f(s_1) = f(s_2) \\ \text{with } s_1 < s_2 \text{ and } f(s_1) \neq \bar{0}_S.$$

For the above t, s_1, s_2,

$$h(\langle s_1, \bar{0}_T\rangle \oplus \langle s_2, t\rangle) = h(\langle s_1, \bar{0}_T\rangle) = \langle f(s_1), \bar{0}_T\rangle,$$

and

$$h(\langle s_1, \bar{0}_T\rangle) \oplus h(\langle s_2, t\rangle) = \langle f(s_1), g(\bar{0}_T)\rangle \oplus \langle f(s_2), g(t)\rangle \\ = \langle f(s_1), \bar{0}_T \oplus g(t)\rangle.$$

Now $f(s_1) \neq \bar{0}_S$ and $\text{DIST}(S \vec{\times}_{\bar{0}} T)$ thus

$$\langle f(s_1), \bar{0}_T\rangle = \langle f(s_1), \bar{0}_T \oplus g(t)\rangle \implies \bar{0}_T = g(t).$$

$\Rightarrow\Leftarrow$ thus $\neg\text{C}(S) \wedge \text{DIST}(S \vec{\times}_{\bar{0}} T) \implies \text{K}_{\bar{0}}(S) \vee \text{C}(T)$ ■

Using the above lemmas, we now state our main theoretical contribution, the properties of $(S, \oplus, \bar{0}_S, F_S)$ and $(T, \oplus, \bar{0}_T, F_T)$ that make $S \vec{\times}_{\bar{0}} T$ distributive.

Theorem 1. *The necessary and sufficient conditions for $S\vec{\times}_{\bar{0}}T$ to be distributive are*

$$\mathrm{C}(S) \vee (\mathrm{DIST}(S) \wedge \mathrm{C}(T)) \vee (\mathrm{DIST}(S) \wedge \mathrm{DIST}(T) \wedge \mathrm{K}_{\bar{0}}(S)) \iff \mathrm{DIST}(S\vec{\times}_{\bar{0}}T), \tag{4}$$

where the properties $\mathrm{C}, \mathrm{DIST}$ *and* $\mathrm{K}_{\bar{0}}$ *are defined in Table I.*

Proof: Lemmas 2, 3 and 4 provide

$$\mathrm{C}(S) \vee (\mathrm{DIST}(S) \wedge \mathrm{C}(T)) \vee (\mathrm{DIST}(S) \wedge \mathrm{DIST}(T) \wedge \mathrm{K}_{\bar{0}}(S)) \implies \mathrm{DIST}(S\vec{\times}_{\bar{0}}T).$$

Lemmas 2, 5, 6 and 7 along with the fact that $\mathrm{DIST}(S\vec{\times}_{\bar{0}}T)$ can be partitioned into $(\mathrm{C}(S) \wedge \mathrm{DIST}(S\vec{\times}_{\bar{0}}T))$ and $(\neg\mathrm{C}(S) \wedge \mathrm{DIST}(S\vec{\times}_{\bar{0}}T))$ provide

$$\mathrm{DIST}(S\vec{\times}_{\bar{0}}T) \implies \mathrm{C}(S) \vee (\mathrm{DIST}(S) \wedge \mathrm{C}(T)) \vee (\mathrm{DIST}(S) \wedge \mathrm{DIST}(T) \wedge \mathrm{K}_{\bar{0}}(S)).$$

■

Remarks: Theorem 4 divides condition (4) into three parts: the first two involve the requirement that either $C(S)$ or $C(T)$, *i.e.*, that the component algebra is constant. These make no sense in context – why use strata if one level is constant? The interesting condition for distributivity of the overall algebra is that the components be distributive, and that S be almost cancellative. This was stated without proof as a sufficient condition for distributivity in [8]. Here we have shown it is also a necessary condition for useful algebras of this type.

V. Further results

By assuming $\exists f \in F_S, a \in S : f(a) < a$ as a base case for an induction that leads to a contradiction, we can show that $\mathrm{DIST}(S) \wedge \mathrm{K}_{\bar{0}}(S) \implies \mathrm{INFL}(S)$.

This leads to a way of counting the number of possible functions in F_S when we have $\mathrm{DIST}(S) \wedge \mathrm{K}_{\bar{0}}(S)$ and S is finite. We consider paths on a $|S|-1 \times |S|-1$ square grid. The constraints of $\mathrm{DIST}(S) \wedge \mathrm{K}_{\bar{0}}(S) \wedge \mathrm{INFL}(S)$ mean that we can move only up or to the right, so the set of paths we are interested in can be written as a word consisting of the letters $\mathcal{U}$ and $\mathcal{R}$. Suppose that the value that $f(a)$ takes is the highest point of the path at a. The words can be written as $|S|-1$ $\mathcal{U}$s each of which can be followed by 0 or 1 $\mathcal{R}$s, thus the number of distributive functions possible is $2^{|S|-1}$.

VI. Conclusion

The main result of this paper is proof of the necessary and sufficient properties of $(S, \oplus_S, \bar{0}_S, F_S)$ and $(T, \oplus_T, \bar{0}_T, F_T)$ such that $S\vec{\times}_{\bar{0}}T$ is distributive. It is this distributivity which ensures $S\vec{\times}_{\bar{0}}T$ will find globally optimal paths. The required properties are given in Theorem 1. They are remarkably similar to the those for $S\vec{\times}T$ to be distributive given in (3); however, when S is finite, comparison of the number of possible functions in F_s (Sections III-B and V) demonstrates how much more flexible the absorbing lexicographic product is in comparison to the standard lexicographic product.

The minimal assumption that the functions in F_S and F_T map zero onto itself was made, but the requirement for $S\vec{\times}_{\bar{0}}T$ to be distributive forces these functions to be endomorphisms and thus S and T must be AMEs.

We considered an absorbing lexicographic product where only the highest level could generate an absorbing zero. Ideally we could create algebras using an n-ary product, where every level might be able to create an absorbing zero. The first step towards this end would be to consider the simple case of two levels, both of which can generate an absorbing zero. This could then be used recursively to generate an n-ary product. By defining a new property $\mathrm{N} : f(a) = \bar{0} \implies a = \bar{0}$ on T we can provide the following sufficient conditions for such an algebra to be distributive:

$$\mathrm{DIST}(S) \wedge \mathrm{DIST}(T) \wedge \mathrm{K}_{\bar{0}}(S) \wedge \mathrm{N}(T) \implies \mathrm{DIST}(S\vec{\times}_{\bar{0}}T).$$

If these are also necessary conditions, we suspect that the requirement for $\mathrm{N}(T)$ could be quite restrictive in the case where T was finite. This would also have implications in an n-ary product as with each recursive definition the property would be needed on larger and larger constructs.

The addition of the conditions that generate an absorbing zero means that many powerful theorems can no longer be applied to this construct to provide proof of various properties. Elementary methods were applied in this case. Their success gives an indication that automated theorem provers might be useful in proving the properties of these constructs and would be a necessity if n-ary products with a mix of levels capable of generating an absorbing zero were required.

References

[1] B. A. Carré, "An Algebra for Network Routing Problems," *IMA Journal of Applied Mathematics*, vol. 7, no. 3, pp. 273–294, 1971.

[2] K. Varadhan, R. Govindan, and D. Estrin, "Persistent route oscillations in inter-domain routing," *Computer Networks*, vol. 32, no. 1, pp. 1–16, 2000.

[3] T. Griffin. BGP "Wedgies", Request for Comments: 4264: http://www.ietf.org/rfc/rfc4264.txt, November 2005.

[4] M. Gondran and M. Minoux, *Graphs, Dioids and Semirings: New Models and Algorithms (Operations Research/Computer Science Interfaces Series).* Springer Publishing Company, Incorporated, 1st ed., 2008.

[5] J. Sobrinho, "Algebra and algorithms for QoS path computation and hop-by-hop routing in the Internet," *Networking, IEEE/ACM Transactions on*, vol. 10, pp. 541–550, August 2002.

[6] T. G. Griffin and J. L. Sobrinho, "Metarouting," *SIGCOMM Comput. Commun. Rev.*, vol. 35, pp. 1–12, August 2005.

[7] A. Gurney, *Construction and verification of routing algebras*. PhD thesis, University of Cambridge, 2009.

[8] T. G. Griffin, "The stratified shortest-paths problem," in *Proceedings of the 2nd international conference on COMmunication systems and NETworks*, COMSNETS'10, (Piscataway, NJ, USA), pp. 268–277, IEEE Press, 2010.

[9] A. Martelli, "An application of regular algebra to the enumeration of the cut sets of a graph," *Information Processing*, vol. 74, pp. 511–515, 1976.

[10] A. Gurney and T. Griffin, "Lexicographic products in metarouting," in *Network Protocols, 2007. ICNP 2007. IEEE International Conference on*, pp. 113–122, October 2007.

[11] T. Griffin and A. Gurney, "Increasing Bisemigroups and Algebraic Routing," in *Relations and Kleene Algebra in Computer Science* (R. Berghammer, B. Möller, and G. Struth, eds.), vol. 4988 of *Lecture Notes in Computer Science*, pp. 123–137, Springer Berlin / Heidelberg, 2008. 10.1007/978-3-540-78913-0-11.

[12] T. G. Griffin, F. B. Shepherd, and G. Wilfong, "The stable paths problem and interdomain routing," *IEEE/ACM Trans. Netw.*, vol. 10, pp. 232–243, April 2002.