The convergence of best-path algorithms over non-distributive algebras

L11 guest lecture

Matthew Daggitt (mld46)

29th of November 2016

Motivation

From previous lectures (hopefully!):

- Most existing best-path theory has been concerned with the semiring world.
- BGP is definitely not distributive and furthermore its extension of routes cannot be described by a single operator.
- Therefore the routing problems solved in the internet cannot be modelled by semirings.

Q: Where does this leave us?
A: In a new, exciting world of more general structures!

Questions?

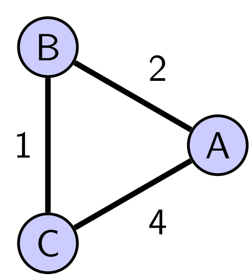
The questions we might want to ask about these new general structures are:

- When do our algorithms converge?
- When do they reconverge?
- How long does convergence/reconvergence take?

Quick semiring recap

We can represent a routing problem as a semiring $(S, \oplus, \otimes, 0, 1)$ where:

- S is the carrier set
- $\oplus : S \to S \to S$ is the choice operator
- $\otimes : S \to S \to S$ is the extension operator
- $0 \in S$ is the invalid path
- $1 \in S$ is the zero length path



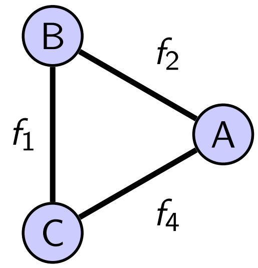
For example:

- Shortest paths: $(\mathbb{N}^{\infty}, \min, +, \infty, 0)$

Algebras of monoid endomorphisms

We can generalise such structures to algebras of monoid endomorphisms $(S, \oplus, F, 0, \bar{1})$ where:

- S is the carrier set
- $\oplus : S \rightarrow S \rightarrow S$ is the choice operator
- $F \subseteq S \rightarrow S$ is a set of edge functions
- $0 \in S$ is the invalid route
- $\bar{1} \subseteq S$ is the set of zero length routes



For example:

- Shortest paths: $(\mathbb{N}^{\infty}, \min, F_{+}, \infty, \{0\})$ where $F_{+} = \{f_n \mid n \in \mathbb{N}\}$ and $f_n(x) = n + x$.

Generalising properties

Property	AME formulation
$\oplus$ associative	-
$\oplus$ commutative	-
$\oplus$ selective	-
F associative	n/a
0 identity for $\oplus$	-
0 annihilator for F	$\forall f : f(0) = 0$
$\bar{1}$ identity for F	n/a
$\bar{1}$ annihilator for $\oplus$	$\forall f : \forall 1 \in \bar{1} : \forall a \in S : f(a) \oplus 1 = 1$
F distributes over $\oplus$	$\forall f : \forall ab \in S : f(a + b) = f(a) \oplus f(b)$
F increasing over $\oplus$	$\forall f : \forall a \in S : a = a \oplus f(a)$
F s. increasing over $\oplus$	$\forall f : \forall a \in S : a = a \oplus f(a) \neq f(a)$

Bellman-Ford and convergence

Generalising the distributed Bellman-Ford algorithm is simple:

$$\sigma(X)_{ij} = \left(\bigoplus_k A_{ik}(X_{kj}) \right) \oplus I_{ij}$$

and so once again we have that

$$\sigma(X) = AX \oplus I$$

The algorithm converges from a starting state X iff:

$$\exists n \in \mathbb{N}.\, \sigma^n(X) = \sigma^{n+1}(X)$$

To prove

Ideally we would like to guarantee that we converge from all states rather than just I, the identity matrix.

This means that the algorithm will be able to cope with:

- Node failure (servers going down)
- Edge failure (connections going down)
- Change in edge weights (due to congestion etc.)

Hence we should find conditions on the algebra such that:

$$\forall X : \exists n \in N : \sigma^n(X) = \sigma^{n+1}(X)$$

Assumptions for proof

Property	SRs	AMEs
$\oplus$ associative	Yes	Yes
$\oplus$ commutative	Yes	Yes
$\oplus$ selective	Yes	Yes
$\otimes / F$ associative	Yes	-
0 identity for $\oplus$	Yes	Yes
0 annihilator for F	Yes	Yes
$\bar{1}$ identity for F	Yes	-
$\bar{1}$ annihilator for $\oplus$	Yes	Yes
F distributes over $\oplus$	Yes	No
F increasing over $\oplus$	No	Yes
F s. increasing over $\oplus$	No	Yes

Additional assumptions

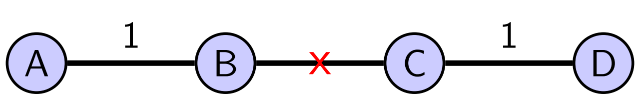
We're also going assume that set S is finite...

This is a *very* strong assumption, but so too is our claim that:

$$\forall X.\, \exists n \in \mathbb{N}.\, \sigma^n(X) = \sigma^{n+1}(X)$$

For instance this doesn't hold even for the shortest-paths algebra.

Shortest-paths only converges when starting from I. For general starting states we get count-to-convergence/infinity problems.



Proof invariants

In the semiring world we can prove that with each iteration every entry in the routing matrix improves monotonically.

In the non-distributive world this is no longer true. We need to find a new invariant...

Creating an invariant

As S is finite then we can define the height function:

$$h(x) = |\{a \mid a \leq_L^{\oplus} x\}|$$

Using this we can then define the similarity between $x, y \in S$ as:

$$s(x,y) = \begin{cases} h(0)+1 & \text{if } x = y \\ \min(h(x), h(y)) & \text{otherwise} \end{cases}$$

If x and y differ than they're only as similar as the height of the best of the pair. Intuitively this makes sense as the better the element is the more likely it is to be incorporated into other routes.

Far enough?

Is this enough? Does the similarity between X_{ij} and $\sigma(X)_{ij}$ increase each iteration?

Sadly not as the lack of distributivity means that individual entries may get worse.

Matrix invariants

Instead of trying to reason about the properties of individual entries in the routing matrices, we can instead reason about the matrices themselves.

For all X and Y define the similarity between them as:

$$S(X, Y) = \min_{ij}(s(X_{ij}, Y_{ij}))$$

Finally an invariant!

We will now prove that for all X and Y such that $\sigma(X) \neq \sigma(Y)$ then

$$S(X, Y) < S(\sigma(X), \sigma(Y))$$

i.e. an iteration must strictly improve the *best* entry that X and Y disagree on.

Proof 1/3

As $\sigma(X) \neq \sigma(Y)$ there must be distinct i and j such that $\sigma(X)_{ij} \neq \sigma(Y)_{ij}$ and $S(X, Y) = h(\sigma(X)_{ij})$ or $h(\sigma(Y)_{ij})$.

Without loss of generality let us assume the former and so $h(\sigma(X)_{ij}) \leq h(\sigma(Y)_{ij})$.

If $\sigma(X)_{ij} = I_{ij}$ then we have a contradiction as:

$$\begin{aligned} h(\sigma(Y)_{ij}) &\leq h(0) \\ &= h(I_{ij}) \\ &= h(\sigma(X)_{ij}) \end{aligned}$$

and by anti-symmetry of $\leq$ we have that $\sigma(X)_{ij} = \sigma(Y)_{ij}$.

Proof 2/3

Therefore $\sigma(X)_{ij} = A_{ik}(X_{kj})$ for some k.

If $X_{kj} = Y_{kj}$ then we have a similar contradiction in that:

$$\begin{aligned}\sigma(Y)_{ij} &\leq A_{ik}(Y_{kj}) \\ &= A_{ik}(X_{kj}) \\ &= \sigma(X)_{ij}\end{aligned}$$

and by anti-symmetry of $\leq$ we have that $\sigma(X)_{ij} = \sigma(Y)_{ij}$.

Proof 3/3

Therefore $X_{kj} \neq Y_{kj}$ and hence we have that:

$$\begin{aligned}S(X, Y) &\leq h(X_{kj}) \\ &< h(A_{ik}(X_{kj})) \\ &= h(\sigma(X)_{ij}) \\ &= S(\sigma(X), \sigma(Y))\end{aligned}$$

as required.

Convergence

Using this lemma we can therefore form a chain:

$$S(X, \sigma(X)) < S(\sigma(X), \sigma^2(X)) < S(\sigma^2(X), \sigma^3(X)) < \ldots$$

This chain is bounded above by $h(0) + 1$ and therefore there must be an n such that $\sigma^n(X) = \sigma^{n+1}(X)$.

Interestingly now that we have the existence of such n it is easy to show that $S(X, \sigma^n(X)) < S(\sigma(X), \sigma^n(X))$.

Uniqueness

This fixed point is also necessarily unique.

Suppose we had two distinct fixed points X and Y then by applying the lemma we could get:

$$\begin{aligned} S(X, Y) &< S(\sigma(X), \sigma(Y)) \\ &= S(X, Y) \end{aligned}$$

which is clearly a contradiction.

Too strong by half?

We now have a set of sufficient conditions for convergence from any state, but are they too strong? Even shortest-distance doesn't fulfil them!

But how do we fix count-to-convergence problems? By introducing paths.

Adding paths

Given an algebra $(S, \oplus, F, 1)$ and a graph G we can form an augmented algebra $(S_P, \oplus_P, F_P, 0_P, \bar{1}_P)$ that tracks and removes paths.

- Let $\mathcal{P}_S(G)$ be the set of simple paths.

$$S_P = (S \times \mathcal{P}_S(G)) \cup \{null\}$$

- $0_P = null$
- $\bar{1}_P = \{(1, i) \mid i \in G\}$

- Choice operator

  Let $\oplus_L$ be a selective operator over paths that selects the shortest path and breaks ties by the lexicographic order of the nodes in the path.

$$
\begin{aligned}
null \oplus_P (w, q) &= (w, q) \\
(v, p) \oplus_P null &= (v, p) \\
(v, p) \oplus_P (w, q) &= \begin{cases} (v, p) & \text{if } v = v \oplus w \neq w \\ (w, q) & \text{if } v \neq v \oplus w = w \\ (v \oplus w, p \oplus_L q) & \text{otherwise} \end{cases}
\end{aligned}
$$

- Extension functions

  Let $e_{ij} \in F$ be the edge weight in G between i and j.

  Then we define f_{ij} as:

$$
\begin{aligned}
&f_{ij}(null) = null \\
&f_{ij}(x, p) = \begin{cases} null & \text{if } i \in p \text{ or } j \neq src(p) \\ (e_{ij}(x), i :: p) & \text{otherwise} \end{cases}
\end{aligned}
$$

  We then have that

$$F_P = \{f_{ij} \mid i, j \in G\}$$

Paths and sufficient conditions

It is easy to show that given an $(S, \oplus, F, 1)$ satisfying:

- $\oplus$ associative
- $\oplus$ commutative
- $\oplus$ selective
- F increasing over $\oplus$ (note: not strictly increasing!)
- 1 is an annihilator for $\oplus$

that all the required conditions for our previous theorem hold.

The one exception is the finiteness of S.

Consistent pairs

But we've added simple paths... surely we must be finite!

Unfortunately not. For any given entry (x, p), the weight of p may not be x, and therefore the cardinality of S_P is the same as the cardinality of S.

Instead take the sub-algebra of only consistent pairs where $x = w(p)$ where the weight function w obviously depends on the underlying graph. Our theorem then applies to this consistent algebra.

Inconsistent pairs

We can then prove by induction on path lengths that any routing matrix containing inconsistent information eventually has such information flushed.

Therefore we eventually reach a consistent state from which we then converge!

Inconsistent pairs

We can then prove by induction on path lengths that any routing matrix containing inconsistent information eventually has such information flushed within n iterations.

Therefore we eventually reach a consistent state from which we then converge!

Convergence time

So we have some (relatively) weak conditions for which we converge. How long does such convergence take?

For k-stable semirings we know that convergence is guaranteed in $nk - 1$ iterations.

Can we come up with a similar bound for non-distributive algebras?

Linear convergence time

Following on from our intuition about the shortest-widest paths algebra, we first tried to attempt to prove that if l_{ij} is the length of the longest simple path from i to j then:

$$\forall ij : \forall t \in \mathbb{N} : \sigma^{l_{ij}+t}(I)_{ij} = \sigma^{l_{ij}+t+1}(I)_{ij}$$

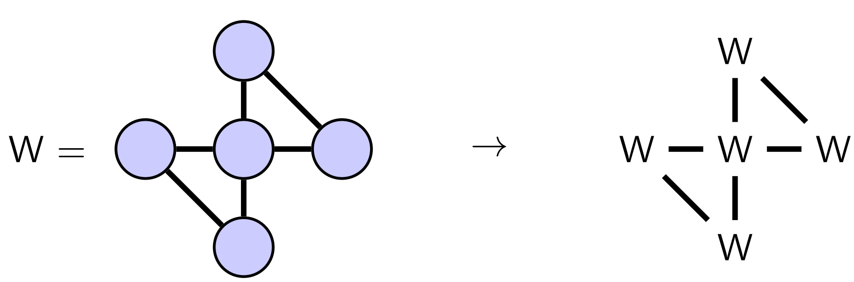
Although this holds for the shortest-widest paths, in general you cannot make the proof go through.

Non-linear convergence time

The hole in the proof lead us to the following counter-example graph...

Chaining widgets

Which we can then chain to form an exponential counter-example.



Chaining widgets

Consider a widget with n nodes in it and which takes c iterations to converge. Let a_k be the number of iterations the kth level meta-widget takes to converge with $a_0 = c$.

$$\begin{aligned} a_k &= (c+1)a_{k-1} + c - 1 \\ &= a_0(c+1)^k + \ldots \\ &= c(c+1)^k + \ldots \end{aligned}$$

Exponential convergence time

Let n_k be the number of nodes in meta-widget at level k then:

$$n_k = n^{k+1}$$

Therefore under the assumption that $c \geq n$ and so $c + 1 > n$ the ratio between $c(c+1)^k$ and n^{k+1} grows exponentially.