Demonic composition and inverse semigroups

Tim Stokes (University of Waikato)

Let $Rel_X = 2^{X \times X}$ be the set of binary relations on the set X.

For example, if $X = \{a, b, c\}$, let

$$s = \{(a, a), (a, b), (b, c), (c, c)\}, t = \{((a, c), (b, b)\}.$$

We say the domain of s is $dom(s) = \{a, b, c\}$ and $dom(t) = \{a, b\}$.

Neither s nor t is a function:

s is multiply defined at $a \in X$, and $dom(t)$ is not all of X.

However, $t \in PT_X$, the set of partial functions on X.

Ordinary or “angelic” composition of binary relations is well-known.

For binary operations $s, t \in Rel_X$,

$$s; t = \{(x, y) \in X \times X \mid \exists z \in X : (x, z) \in s, (z, y) \in t\}.$$

This generalises composition of (partial) functions and shares some of its nice features: composition of binary relations is associative.

So Rel_X is a semigroup, having PT_X as subsemigroup.

And the full transformation semigroup T_X is a subsemigroup of PT_X.

However, there is a less familiar way to compose binary relations: "demonic composition".

For $s, t \in Rel_X$, define

$$s * t = \{(x, y) \in s; t \mid (x, z) \in s \Rightarrow z \in dom(t)\}.$$

The terminology comes from computer science thought experiments when thinking about "nondeterministic programs":

angelic composition does whatever can be done, while demonic composition assumes a "demon" will mess things up if possible.

In our example with s, t as before:

$$s = \{(a, a), (a, b), (b, c), (c, c)\}, t = \{((a, c), (c, b)\},$$

we have that

$$s; t = \{(a, c), (b, b), (c, b)\}$$

while

$$s * t = \{(b, b), (c, b)\}.$$

But in general, $s * t$ need not be a partial function.

Angelic composition is important when thinking about partial correctness of “non-deterministic” programs, demonic for total correctness.

Demonic composition also reduces to composition of partial functions when applied to them.

For binary relations with domain all of X, it agrees with usual relational composition too.

And it is still an associative operation (as is non-obvious).

So arguably, “demonic” composition is just as good a generalisation of functional composition as is “angelic” composition.

How is $(Rel_X, *)$ different to $(Rel_X, ;)$?

To help see how, we define the operation of domain D on Rel_X;

$$D(s) = \{(x, x) \mid x \in dom(s)\}.$$

So $D(s)$ is the restriction of the identity function to $dom(s)$.

So D is also well-defined on PT_X (but rather uninteresting for T_X).

In fact $(PT_X, ;, D)$ is what's called a *left restriction semigroup*.

Laws for left restriction semigroups are those for semigroups, plus

- $D(x)x = x$; $D(x)D(y) = D(y)D(x)$
- $D(D(x)y) = D(D(x)D(y)) = D(x)D(y)$
- $xD(y) = D(xy)x$.

Every left restriction semigroup embeds in one of the form PT_X (a "Cayley theorem").

However, $(Rel_X, ;, D)$ is not a left restriction semigroup, since the final law fails.

It does satisfy some weaker laws, but there is no Cayley theorem...

...unless infinitely many laws are used!

BUT... $(Rel_X, *, D)$ *is* a left restriction semigroup!

So there is a Cayley theorem again (the same as for PT_X).

How does one go from $(Rel_X, ;, D)$ to $(Rel_X, *, D)$?

In fact it can be shown that for all $s, t \in Rel_X$,

$$s * t = A(s; A(t)); s; t,$$

where $A(s) = \{(x, x) \in X \times X \mid x \notin dom(s)\}$, the so-called antidomain of s.

Why does this work – is something more general going on?

Indeed there is!

Suppose S is a unary semigroup with unary operation D, and we define

$$s \circ t = st \text{ only if } sD(t) = s.$$

For the case of Rel_X, this is saying that the image of s is contained in $dom(t)$ (regardless of which type of composition is used!).

If we do this for a left restriction semigroup S, Gould and Hollings showed that we get a so-called *inductive constellation* $(S, \circ, D)$

(a special kind of partial algebra defined by a finite set of laws).

Conversely, every inductive constellation arises in this way from a left restriction semigroup.

So given an inductive constellation $(S, \circ, D)$, it's possible to "add in" all the "missing" products and make a left restriction semigroup out of it.

These constructions are mutually inverse, and indeed the two categories are isomorphic. (An "ESN Theorem" variant.)

But not only left restriction semigroups can give inductive constellations in this way!

First, when do we get a constellation? And what is a constellation?!

Given a set P equipped with a partial binary operation $\circ$, we say $e \in P$ is a *right identity* if for all $a \in P$, if $a \circ e$ exists then it equals a.

We say $(P, \circ, D)$ is a constellation if, for all $x, y, z \in P$:

- if $x \circ (y \circ z)$ exists then so does $(x \circ y) \circ z$, and then the two are equal;
- if $x \circ y$ and $y \circ z$ exist then $x \circ (y \circ z)$ exists;
- $D(x)$ is the unique right identity in P such that $D(x) \circ x = x$.

On a constellation, there is the natural quasiorder given by $s \leq t$ iff $s = D(s) \circ t$ (which is assumed to exist).

To be inductive, the constellation must have a notion of “co-restriction” satisfying some further laws.

These are:

(O4) If $e \in D(P)$ and $a \in P$, then there is a maximum $x \in P$ with respect to the natural quasiorder $\leq$ on P, such that $x \leq a$ and $x \circ e$ exists, the *co-restriction* of a to e, denoted $a|e$; and

(O5) for $x, y \in P$ and $e \in D(P)$, if $x \circ y$ exists then

$$D((x \circ y)|e) = D(x|(D(y|e)).$$

Then we may extend $\circ$ to a total operation by setting

$$s \odot t := (s|D(t)) \circ t.$$

And $(P, \odot, D)$ is a left restriction semigroup!

Given a unary semigroup S equipped with unary D, when is $(S, \circ, D)$ at least a constellation?

(Recall $s \circ t = st$ but only when $sD(t) = s$.)

Exactly when it is a “demigroup”. (Name to be confirmed!)

Laws for these are fairly simple.

- $D(x)^2 = D(x)$;
- $D(x)x = x$;
- $D(xy) = D(xD(y))$.

In this case, $D(D(x)) = D(x)$ for all $x \in S$.

And it follows that $D(S) = \{D(x) \mid x \in S\}$ consists of idempotents.

Let's call a demigroup $(S, \times, D)$ an *inductive demigroup* if $(S, \circ, D)$ is an inductive constellation.

These may be characterised by just two further conditions.

The first is that for all $e, f \in D(S)$, the quasiorder given by $e \leq_r f$ iff $e = ef$ is a partial order;

equivalently, for all $e, f \in D(S)$, $e = ef, f = fe \Rightarrow e = f$.

Aside from that, we require that

for all $s \in S$, $e \in D(S)$, there exists (unique) $s \cdot e \in D(S)$

such that $s \cdot e \leq_r D(s)$, and for all $t \in S$,

$$tse = ts \Leftrightarrow t(s \cdot e) = t,$$

that is, the set of left equalizers of se, s is generated as a left ideal by $(s \cdot e) \in D(S)$.

Then the co-restriction in the derived inductive constellation is

$$s|e = (s \cdot e)s.$$

It follows that if S is an inductive demigroup, $D(S)$ is a meet-semilattice under $\leq_r$,

with $e \wedge f = e \cdot f$ for all $e, f \in D(S)$.

And S can be turned into a left restriction semigroup $(S, \odot, D)$ by setting

$s \odot t = (s \cdot D(t))st$ for all $s, t \in S$.

Digression: one can give a purely equational characterisation of inductive demigroups.

The demigroup S is inductive if and only if,

for all $s \in S$ and $e \in D(S)$, there is $s \cdot e \in S$ satisfying,

for all $s, t \in S$ and $e, f \in D(S)$:

- $D(s \cdot e) = s \cdot e$,
- $se \cdot e = D(se)$,
- $e \cdot f = f \cdot e$,
- $(t \cdot e)t = (t \cdot e)te$ and
- $(sD(t)) \cdot (t \cdot e) = (st) \cdot e$.

(Extend $\cdot$ to S by adding $s \cdot t := s \cdot D(t)$ for all $s, t \in S$.)

Note that $(Rel_X, ;, D)$ is an inductive demigroup!

That is, $(Rel_X, \circ, D)$ is an inductive constellation.

So in theory we can rebuild it to be a left restriction semigroup instead.

The action of Rel_X on $D(Rel_X)$ is given by $s \cdot D(t) = A(s; A(t)); D(t)$...

so $s \odot t = A(s; A(t)); D(t); s; t = A(s; A(t)); s; t$,

which is just the demonic composition $s * t$ as defined before!

This "explains" why $(Rel_X, *, D)$ is a left restriction semigroup, and in particular, why $*$ is associative.

But are there any other interesting special cases?

What follows applies to any $*$-regular ring or indeed any $*$-regular Baer $*$-semigroup,

but I'm here going to concentrate on matrices.

Consider the multiplicative semigroup of $n \times n$ matrices $M_n(R)$, where R is the real or complex number field.

Although not every matrix has an inverse, it does always have a (necessarily unique) *Moore-Penrose inverse*.

Given matrix M, this is an M' such that

$$MM'M = M, M'MM' = M',$$

and both $MM', M'M$ are symmetric ($M^H = M$).

If M is non-singular then $M' = M^{-1}$, and $MM' = M'M = I$.

The Moore-Penrose inverse is important in statistics and beyond.

In general semigroup theory, we say x has an inverse y if

$$xyx = x,\ yxy = y.$$

(In a group, this is just the usual notion of inverse.)

So the Moore-Penrose inverse of a matrix is an inverse satisfying some additional properties.

But in general a given matrix has many other semigroup inverses.

By contrast, an *inverse semigroup* is a semigroup in which each element has a unique semigroup inverse.

They are “closer” to being groups than any other class of semigroups.

Inverse semigroups model algebras of 1:1 partial functions equipped with inversion.

In an inverse semigroup, the idempotents form a subsemigroup which is a semilattice (*cf.* the idempotents in $M_n(R)$).

Every inverse semigroup is a left restriction semigroup

if we define $D(s) = ss'$, where s' is the inverse of s.

And one can characterise those left restriction semigroups arising in this way:

for all s there is s' such that $D(s) = ss', D(s') = s's$.

As we've noted, $M_n(R)$ is not an inverse semigroup.

However, it is an inductive demigroup with $D(S) = SS'$, $S \in M_n(R)$.

Define $A(S) = I - D(S)$; this behaves like antidomain in Rel_X.

Then for $S, T \in M_n(R)$, $S \odot T = A(S \times A(T))ST$, like in Rel_X.

Hence $(M_n(R), \odot, D)$ is a left restriction semigroup.

But it turns out that $D(S) = S \odot S'$ and $D(S') = S' \odot S$,

so it is one arising from an inverse semigroup, in which inverse is Moore-Penrose inverse.

So $M_n(R)$ is an inverse semigroup with respect to $\odot$ and Moore-Penrose inverse!

$S' = S^{-1}$ when it exists, and "almost always" $S \odot T = ST$ (e.g. if T^{-1} exists).

But it's not easy to write down an explicit formula for $S \odot T$, even when $n = 2$!

(It splits into cases, like the "formula" for Moore-Penrose inverse.)

There is much interest in the structure of these inverse semigroups...

but that's another story.