Weighted Automata and CospanSpan

R.F.C. Walters
L. de Franceso Albasini N. Sabadini
Università dell'Insubria

25 June, CT2010, Genova

Problem

If someone tells you they have two children, (at least) one a boy born on Tuesday, what is the probability that the other child is a boy?

Weighted Automata (Transducer) (Schützenberger 1961)

A *weighted automaton* with left parallel interface A, right parallel interface B, top sequential interface X and bottom sequential interface Y consists of

- a set Q of states;
- an $A \times B$ family of $Q \times Q$ matrices of non-negative reals;
- two functions $\gamma_0 : X \to Q$, $\gamma_1 : Y \to Q$ (usually in the literature, inclusions).

M. Droste, W. Kuich, H. Vogler, (Eds.), *Handbook of Weighted Automata*, EATCS Monographs in Theoretical Computer Science: 2009.

Weighted Automata (Transducer) (Schützenberger 1961)

A *weighted automaton* with left parallel interface A, right parallel interface B, top sequential interface X and bottom sequential interface Y consists of

a set Q of states;

an $A \times B$ family of $Q \times Q$ matrices of non-negative reals;

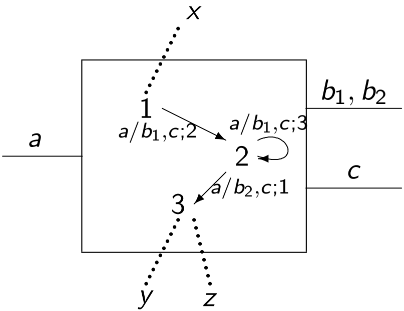
two functions $\gamma_0 : X \to Q$, $\gamma_1 : Y \to Q$ (usually in the literature, inclusions).

M. Droste, W. Kuich, H. Vogler, (Eds.), *Handbook of Weighted Automata*, EATCS Monographs in Theoretical Computer Science: 2009.

Weighted Automata

Consider the automaton with parallel interfaces $\{a\}$, $\{b_1, b_2\} \times \{c\}$, sequential interfaces $\{x\}$, $\{y, z\}$; with states $\{1, 2, 3\}$ sequential interface functions $x \mapsto 1$ and $y, z \mapsto 3$; and transition matrices

$$Q_{a,(b_1,c)} = \begin{bmatrix} 0 & 2 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \; Q_{a,(b_2,c)} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$



The algebra of weighted automata

Eilenberg and others gave various operations on transducers, which may be grouped into sequential (Σ operations) and parallel operations (Π operations). The sequential operations correspond to Kleene operations on languages, but in addition there is a *composition* of automata, which we interpret as communicating parallel.

Samuel Eilenberg, Automata, Languages, and Machines, vol. A. Academic Press, New York, 1974.

Rather than defining these operations I will describe a slightly different algebra in which the main operations and many of their properties are obvious.

The algebra of weighted automata

Eilenberg and others gave various operations on transducers, which may be grouped into sequential (Σ operations) and parallel operations (Π operations). The sequential operations correspond to Kleene operations on languages, but in addition there is a *composition* of automata, which we interpret as communicating parallel.

Samuel Eilenberg, Automata, Languages, and Machines, vol. A. Academic Press, New York, 1974.

Rather than defining these operations I will describe a slightly different algebra in which the main operations and many of their properties are obvious.

CospanSpan(Graph): first attempt

A square matrix may be thought of as a *graph* with vertex set the number of rows=number of columns. An $A \times B$ family of square matrices of the same size may be thought of as an edge-doubly-labelled graph. Hence a weighted automaton contains the following information: four graphs (X, Y graphs without edges, A, B graphs with one vertex) and four graph morphisms as shown:

$$\begin{array}{ccccc} & & X & & \\ & & \downarrow{\scriptstyle \gamma_0} & & \\ A & \xleftarrow{\partial_0} & G & \xrightarrow{\partial_1} & B \\ & & \uparrow{\scriptstyle \gamma_1} & & \\ & & Y & & \end{array} .$$

Idea: G is the graph of states and transitions of the system, X and Y are the graphs of the sequential interfaces, A and B are graphs of the parallel interfaces.

CospanSpan(Graph): definition

But the sequential interfaces are systems and need parallel interfaces, and the parallel interfaces are systems and need sequential interfaces. Hence a *system with sequential and parallel interfaces* is a *cospan of spans of graphs* and consists of a commutative diagram of graphs and graph morphisms

$$
\begin{array}{ccccc}
G_0 & \xleftarrow{\partial_0} & X & \xrightarrow{\partial_1} & G_1 \\
\downarrow{\scriptstyle \gamma_0} & & \downarrow{\scriptstyle \gamma_0} & & \downarrow{\scriptstyle \gamma_0} \\
A & \xleftarrow{\partial_0} & G & \xrightarrow{\partial_1} & B \\
\uparrow{\scriptstyle \gamma_1} & & \uparrow{\scriptstyle \gamma_1} & & \uparrow{\scriptstyle \gamma_1} \\
G_2 & \xleftarrow{\partial_0} & Y & \xrightarrow{\partial_1} & G_3
\end{array}
$$

Notice a system is also a *span of cospans of graphs*.
We denote such a system very briefly as $G^X_{Y;A,B}$.

The algebra CospanSpan: compositions

Now there are two obvious compositions:

Sequential composition: Two systems $G^X_{Y;A,B}$ and $K^Y_{Z;D,E}$ admit a compositions by pushout, the *sequential* (or vertical) *composition*, denoted $G^X_{Y;A,B} \circ_{seq} K^Y_{Z;D,E}$.

Parallel composition: Two systems $G^X_{Y;A,B}$ and $H^Z_{W;B,C}$ admit a compositions by pullback, the *parallel* (or horizontal) *composition*, denoted $G^X_{Y;A,B} \circ_{par} H^Z_{W;B,C}$.

These operations make sense for cospans of spans in any category with finite limits and colimits.

The algebra CospanSpan(Graph): tensor products

And there are two obvious tensor products:

Sequential tensor product: The *sum* $G^X_{Y;A,B} + H^Z_{W;C,D}$ of two systems is formed as

$$\begin{array}{ccccc} \bullet+\bullet & \longleftarrow & X+Z & \longrightarrow & \bullet+\bullet \\ \downarrow & & \downarrow & & \downarrow \\ A+C & \longleftarrow & G+H & \longrightarrow & B+D \\ \uparrow & & \uparrow & & \uparrow \\ \bullet+\bullet & \longleftarrow & Y+W & \longrightarrow & \bullet+\bullet \end{array}$$

Parallel tensor product: The *product* $G^X_{Y;A,B} \times H^Z_{W;C,D}$ of two systems is formed as

$$\begin{array}{ccccc} \bullet\times\bullet & \longleftarrow & X\times Z & \longrightarrow & \bullet\times\bullet \\ \downarrow & & \downarrow & & \downarrow \\ A\times C & \longleftarrow & G\times H & \longrightarrow & B\times D \\ \uparrow & & \uparrow & & \uparrow \\ \bullet\times\bullet & \longleftarrow & Y\times W & \longrightarrow & \bullet\times\bullet \end{array}$$

So we have two symmetric monoidal structures on *CospanSpan* which we denote $\mathit{CospanSpan}_{seq}$ and $\mathit{CospanSpan}_{par}$.

Lax double category?

Given four systems $G^X_{Y;A,B}$, $H^U_{V;B,C}$, $K^Y_{Z;D,E}$, $L^V_{W;E,F}$ in the following configuration

$$
\begin{array}{ccccccccc}
\bullet & \leftarrow & X & \rightarrow & \bullet & \leftarrow & U & \rightarrow & \bullet \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
A & \leftarrow & G & \rightarrow & B & \leftarrow & H & \rightarrow & C \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow \\
\bullet & \leftarrow & Y & \rightarrow & \bullet & \leftarrow & V & \rightarrow & \bullet \\
\downarrow & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\
D & \leftarrow & K & \rightarrow & E & \leftarrow & L & \rightarrow & F \\
\uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow \\
\bullet & \leftarrow & Z & \rightarrow & \bullet & \leftarrow & W & \rightarrow & \bullet
\end{array}
$$

there is a comparison map

$$
\begin{array}{c}
(G^X_{Y;A,B} \circ_{par} H^U_{V;B,C}) \circ_{seq} (K^Y_{Z;D,E} \circ_{par} L^V_{W;E,F}) \\
\downarrow \\
(G^X_{Y;A,B} \circ_{seq} K^Y_{Z;D,E}) \circ_{par} (H^U_{V;B,C} \circ_{seq} L^V_{W;E,F})
\end{array}
$$

satisfying appropriate (lax monoidal) coherence equations, which however is *not in general* an isomorphism. This reflects the fact that the top expression involves more synchronization than the bottom.

Remarks

Transferring these operations to the weighted automata setting we get appropriate operations on weighted automata.

Alternating sequential and parallel operations allow the description of hierarchical systems. For example,

Σ expressions describe automata,

$\Pi\Sigma$ expressions describe nets of automata,

$\Sigma\Pi\Sigma$ expressions describe evolving nets of automata,

$\Pi\Sigma\Pi\Sigma$ expressions describe nets of evolving nets of automata.

Sequential and parallel constants

The following constants are not used explicitly in the classical theory of weighted automata.

As a category of spans *CospanSpan* has for each horizontal object a diagonal Δ, an opposite of diagonal Δ^{op} giving the object a commutative separable algebra structure with respect to $\times$, and

As a category of cospans has each vertical object a codiagonal ∇, an opposite of diagonal ∇^{op} giving the object a commutative separable algebra structure with respect to $+$.

These constants make both $CospanSpan_{seq}$ and $CospanSpan_{par}$ into well-supported compact closed (Wscc) categories - symmetric monoidal categories, each object is equipped with a commutative separable (Frobenius) algebra, which structure is compatible with the tensor.

Sequential and parallel constants

The following constants are not used explicitly in the classical theory of weighted automata.

- As a category of spans *CospanSpan* has for each horizontal object a diagonal Δ, an opposite of diagonal Δ^{op} giving the object a commutative separable algebra structure with respect to $\times$, and
- As a category of cospans has each vertical object a codiagonal ∇, an opposite of diagonal ∇^{op} giving the object a commutative separable algebra structure with respect to $+$.

These constants make both $CospanSpan_{seq}$ and $CospanSpan_{par}$ into well-supported compact closed (Wscc) categories - symmetric monoidal categories, each object is equipped with a commutative separable (Frobenius) algebra, which structure is compatible with the tensor.

Sequential and parallel constants

The following constants are not used explicitly in the classical theory of weighted automata.

- As a category of spans *CospanSpan* has for each horizontal object a diagonal Δ, an opposite of diagonal Δ^{op} giving the object a commutative separable algebra structure with respect to $\times$, and
- As a category of cospans has each vertical object a codiagonal ∇, an opposite of diagonal ∇^{op} giving the object a commutative separable algebra structure with respect to $+$.
- These constants make both $CospanSpan_{seq}$ and $CospanSpan_{par}$ into well-supported compact closed (Wscc) categories - symmetric monoidal categories, each object is equipped with a commutative separable (Frobenius) algebra, which structure is compatible with the tensor.

Remark

We began with the fact that automata are cospans of spans of graphs.

The algebra we have described exists for any category C with finite limits and colimits.

However the next theorem shows that even for $CospanSpan(C)$ graphs (automata) emerge.

Remark

We began with the fact that automata are cospans of spans of graphs.

The algebra we have described exists for any category C with finite limits and colimits.

However the next theorem shows that even for $CospanSpan(C)$ graphs (automata) emerge.

Remark

We began with the fact that automata are cospans of spans of graphs.

The algebra we have described exists for any category C with finite limits and colimits.

However the next theorem shows that even for $CospanSpan(C)$ graphs (automata) emerge.

Theorem

A multigraph consists of edges and vertices, such that the source and target of each edge is a word in the vertices. Multigraphs form a presheaf category.

Theorem (Rosebrugh, Sabadini, Walters)

The free symmetric monoidal category with Wscc structure on a multigraph G, $FreeWscc(G)$, is the full subcategory of

$$Cospan(Multigraph_{finite}/G)$$

with the objects are restricted to discrete multigraphs.

The induced functors

$FreeWscc(CospanSpan_{seq}) \to CospanSpan_{seq}$ is colimit.

$FreeWscc(CospanSpan_{par}) \to CospanSpan_{par}$ is limit.

Theorem

A multigraph consists of edges and vertices, such that the source and target of each edge is a word in the vertices. Multigraphs form a presheaf category.

Theorem (Rosebrugh, Sabadini, Walters)

The free symmetric monoidal category with Wscc structure on a multigraph G, $FreeWscc(G)$, is the full subcategory of

$$Cospan(Multigraph_{finite}/G)$$

with the objects are restricted to discrete multigraphs.

The induced functors

$FreeWscc(CospanSpan_{seq}) \to CospanSpan_{seq}$ is colimit.

$FreeWscc(CospanSpan_{par}) \to CospanSpan_{par}$ is limit.

Theorem

A multigraph consists of edges and vertices, such that the source and target of each edge is a word in the vertices. Multigraphs form a presheaf category.

Theorem (Rosebrugh, Sabadini, Walters)

The free symmetric monoidal category with Wscc structure on a multigraph G, $FreeWscc(G)$, is the full subcategory of

$$Cospan(Multigraph_{finite}/G)$$

with the objects are restricted to discrete multigraphs.

The induced functors

$FreeWscc(CospanSpan_{seq}) \to CospanSpan_{seq}$ is colimit.

$FreeWscc(CospanSpan_{par}) \to CospanSpan_{par}$ is limit.

Proof

The proof consists in mild generalizations of the results in the papers

R. Rosebrugh, N. Sabadini, R.F.C. Walters, *Generic commutative separable algebras and cospans of graphs*, Theory and Applications of Categories, 15, 264-177, 2005.

R. Rosebrugh, N. Sabadini, R.F.C. Walters, *Calculating colimits compositionally*, Montanari Festschrift, LNCS 5065, pp. 581-592, 2008.

These results were presented at CT's in Vancouver and Faro. The first depended on

S. Lack, Composing PROPs, Theory and Applications of Categories, Vol. 13, No. 9, 147-163, 2004.

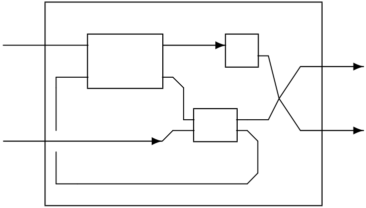
Geometrical notation

A direct result of this theorem is that expressions in the operations of *CospanSpan*$_{seq}$ and of *CospanSpan*$_{par}$ can be pictured as cospans of multigraphs (labelled in *CospanSpan*).

The fact that only the multigraph is relevant, and that the evaluation in the parallel case is a limit is how engineers think of circuits.

Geometrical notation

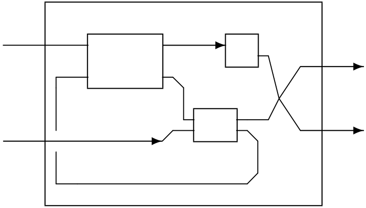
A direct result of this theorem is that expressions in the operations of $CospanSpan_{seq}$ and of $CospanSpan_{par}$ can be pictured as cospans of multigraphs (labelled in *CospanSpan*).



The fact that only the multigraph is relevant, and that the evaluation in the parallel case is a limit is how engineers think of circuits.

Geometrical notation

A direct result of this theorem is that expressions in the operations of $\mathit{CospanSpan}_{seq}$ and of $\mathit{CospanSpan}_{par}$ can be pictured as cospans of multigraphs (labelled in *CospanSpan*).



The fact that only the multigraph is relevant, and that the evaluation in the parallel case is a limit is how engineers think of circuits.

Geometric notation

We have two types of multigraph pictures, corresponding to sequential expressions and parallel ones.

We distinguish the parallel pictures from the sequential ones by drawing dotted 'wires' in the sequential expressions, and sequential operations are vertical, parallel operations are horizontal.

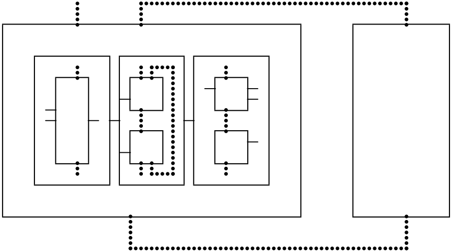
A general expression may be seen as a nested diagram of sequential and parallel multigraphs.

Geometric notation

We have two types of multigraph pictures, corresponding to sequential expressions and parallel ones.

We distinguish the parallel pictures from the sequential ones by drawing dotted 'wires' in the sequential expressions, and sequential operations are vertical, parallel operations are horizontal.

A general expression may be seen as a nested diagram of sequential and parallel multigraphs.



Distribution of limits over colimits

The interesting cases are when C is a topos. Exactness is important in such things as flattening of hierarchy,

$$\Pi\Sigma\Pi\Sigma\cdots\Sigma \to \Sigma\Pi;$$

for example, using the Kronecker product in calculating $\Pi\Sigma$ etc.

Probability problems

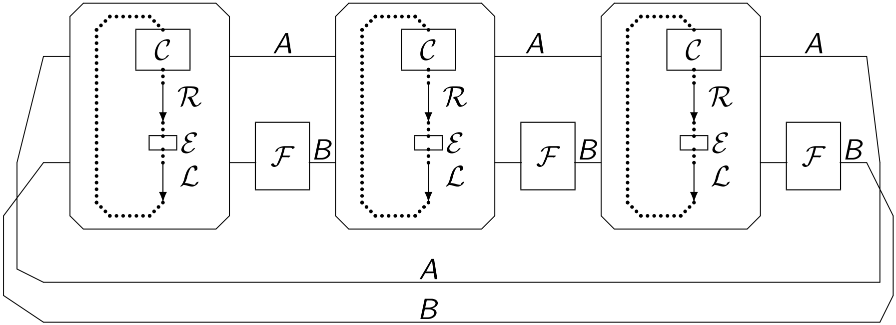
In the following papers we study probability problems in the algebra of weighted automata we have described. Another issue which arises is normalization, which behaves well with respect to parallel operations, but not sequential ones.

L. de Francesco Albasini, N. Sabadini, R.F.C. Walters, *The compositional construction of Markov processes*, arXiv:0901.2434.

L. de Francesco Albasini, N. Sabadini, R.F.C. Walters, The compositional construction of Markov processes II, arXiv:1005.0949.

A $\Pi\Sigma\Sigma$ example we study is a version of the Dining Philosophers system. The idea is a table of children who may move around the table, as well as eat: we call it Sofia's Birthday Party. For details see the second paper.

Sofia's Birthday Party



What is the probability of a child eating after n steps. The probability after 1 step from initial state $(5, 1, 1, 1, 1, 1)$ is 0, after 2 steps is $\frac{19}{60}$, after 3 steps is $\frac{98}{225}$, after 4 steps is $\frac{49133}{108000}$, after 5 steps is $\frac{1473023}{3240000}$ and after 100 steps is 0.3768058221.

Problem

If someone tells you they have two children, (at least) one a boy born on Tuesday, what is the probability that the other child is a boy?

Two answers have been proposed $\frac{13}{27}$ or $\frac{1}{2}$.

The first answer is the orthodox one, and comes from a simple counting argument. If you consider all the possible states, BoyMonGirlTue, BoyTueBoyWed, ..., there are $2 \times 7 \times 2 \times 7$ possibilities; the number which contain a pair BoyTue is $7 + 7 + 7 + 7 - 1 = 27$ (since BoyTueBoyTue is counted only once); the number of these which which are both boys is $7 + 7 - 1 = 13$ and hence the probability is $\frac{13}{27}$.

Problem

If someone tells you they have two children, (at least) one a boy born on Tuesday, what is the probability that the other child is a boy?

Two answers have been proposed $\frac{13}{27}$ or $\frac{1}{2}$.

The first answer is the orthodox one, and comes from a simple counting argument. If you consider all the possible states, BoyMonGirlTue, BoyTueBoyWed, ..., there are $2 \times 7 \times 2 \times 7$ possibilities; the number which contain a pair BoyTue is $7 + 7 + 7 + 7 - 1 = 27$ (since BoyTueBoyTue is counted only once); the number of these which which are both boys is $7 + 7 - 1 = 13$ and hence the probability is $\frac{13}{27}$.

Problem

If someone tells you they have two children, (at least) one a boy born on Tuesday, what is the probability that the other child is a boy?

Two answers have been proposed $\frac{13}{27}$ or $\frac{1}{2}$.

The second answer comes from modelling the protocol involved (a very simple automaton). In addition to the states there is also the action of the informant. The weight behind declaring that he has a boy born on Tuesday is different if he has two boys born on Tuesday, than if for example he has a boy born on Tuesday and another born on Wednesday, or for example a boy born on Tuesday and girl born on Tuesday. The declaration in the case BoyTueBoyTue should carry twice the weight and hence the calculation gives $\frac{14}{28} = \frac{1}{2}$.

In our view to estimate the likelihood it is necessary to describe precisely a protocol or automaton, as we did with Sofia's Birthday Party.

Problem

If someone tells you they have two children, (at least) one a boy born on Tuesday, what is the probability that the other child is a boy?

Two answers have been proposed $\frac{13}{27}$ or $\frac{1}{2}$.

The second answer comes from modelling the protocol involved (a very simple automaton). In addition to the states there is also the action of the informant. The weight behind declaring that he has a boy born on Tuesday is different if he has two boys born on Tuesday, than if for example he has a boy born on Tuesday and another born on Wednesday, or for example a boy born on Tuesday and girl born on Tuesday. The declaration in the case BoyTueBoyTue should carry twice the weight and hence the calculation gives $\frac{14}{28} = \frac{1}{2}$.

In our view to estimate the likelihood it is necessary to describe precisely a protocol or automaton, as we did with Sofia's Birthday Party.

Other papers on CospanSpan

P. Katis, N. Sabadini, R.F.C. Walters, A formalisation of the IWIM Model, in: Proc. COORDINATION 2000, (Eds.) A. Porto, G.-C. Roman, LNCS 1906, 267–283, Springer Verlag, 2000.

L. de Francesco Albasini, N. Sabadini, R.F.C. Walters, Cospans and spans of graphs: a categorical algebra for the sequential and parallel composition of discrete systems, arXiv:0909.4136.

S. Eilenberg [1974] This work is addressed both to pure mathematicians and to computer scientists. To the pure mathematician, I have tried to reveal a body of new algebra, which, despite its external motivation (or perhaps because of it) contains methods and results that are deep and elegant. I believe that eventually some of them will be regarded as a standard part of algebra. To the computer scientist I tried to show the correct conceptual setting for many of the results known to him (and some new ones). This should help him to obtain a better and sharper mathematical perspective on the theoretical aspects of his researches.