

Coarse-graining open Markov processes

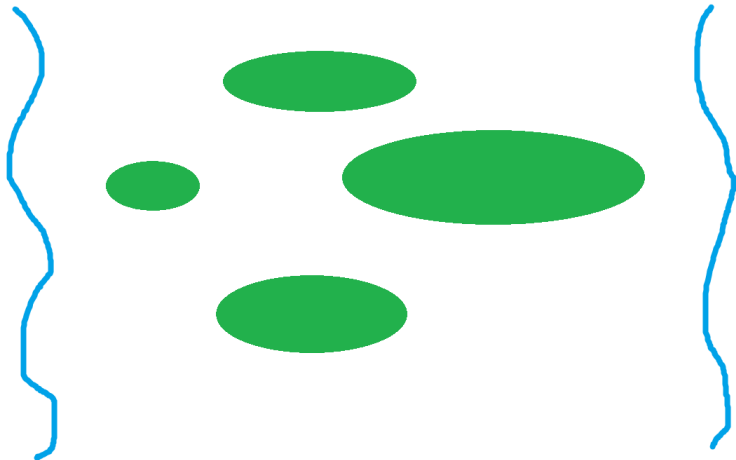
John Baez and Kenny Courser

University of California, Riverside

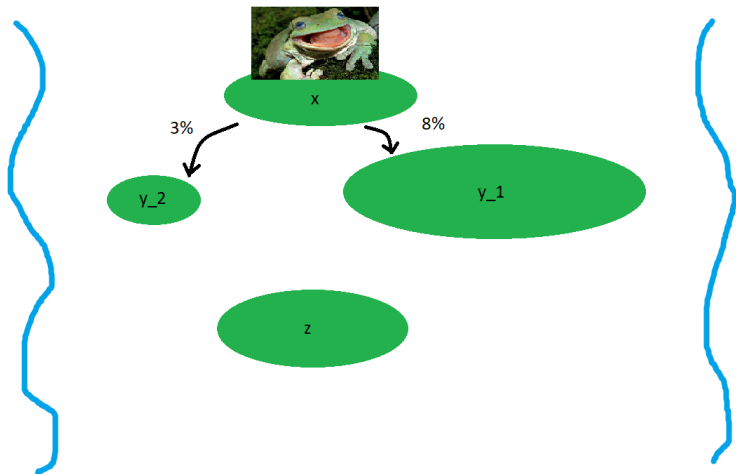
June 12, 2019

Suppose frogs are hopping down a stream...

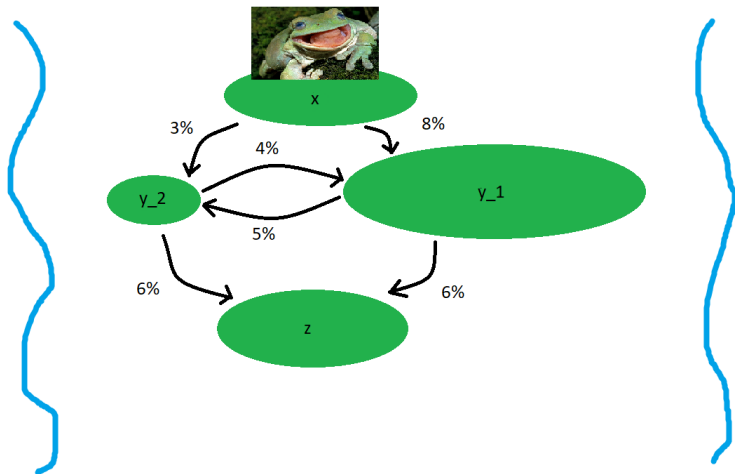
Suppose frogs are hopping down a stream...



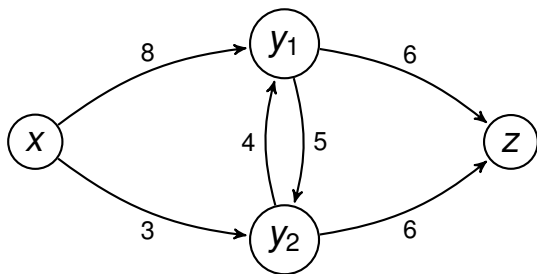
Suppose frogs are hopping down a stream...



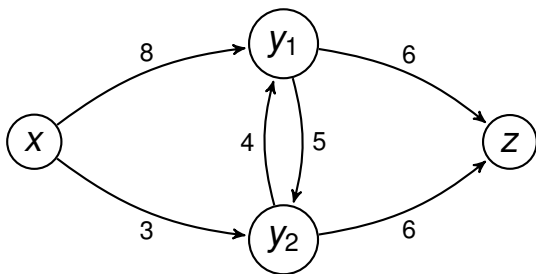
Suppose frogs are hopping down a stream...



This is an example of a Markov process, which we can represent more formally with this diagram:



This is an example of a Markov process, which we can represent more formally with this diagram:



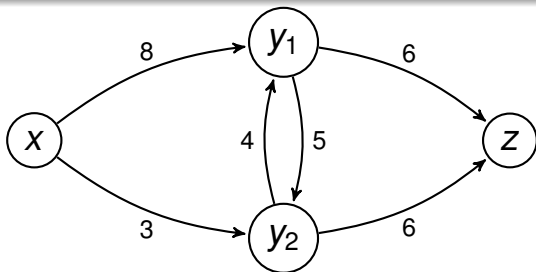
$$H = \begin{bmatrix} -11 & 0 & 0 & 0 \\ 8 & -11 & 4 & 0 \\ 3 & 5 & -10 & 0 \\ 0 & 6 & 6 & 0 \end{bmatrix}$$

Definition

A square real-valued matrix is **infinitesimal stochastic** if each non-diagonal entry is non-negative and each column sums to 0.

Definition

A square real-valued matrix is **infinitesimal stochastic** if each non-diagonal entry is non-negative and each column sums to 0.



$$X = \{x, y_1, y_2, z\}$$

$$H = \begin{bmatrix} -11 & 0 & 0 & 0 \\ 8 & -11 & 4 & 0 \\ 3 & 5 & -10 & 0 \\ 0 & 6 & 6 & 0 \end{bmatrix}$$

We can also view this infinitesimal matrix as a map $H: X \times X \rightarrow \mathbb{R}$.

Definition

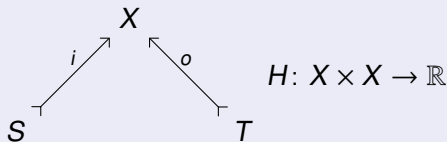
A **Markov process** on a finite set X is given by an infinitesimal stochastic matrix $H: X \times X \rightarrow \mathbb{R}$ where $H(i, j) = H_{j,i}$.

Definition

A **Markov process** on a finite set X is given by an infinitesimal stochastic matrix $H: X \times X \rightarrow \mathbb{R}$ where $H(i, j) = H_{j,i}$.

Definition

An **open Markov process** is a monic cospan of finite sets where the apex is equipped with a Markov process.



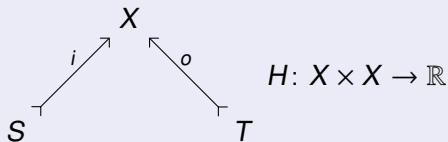
We call S and T the **inputs** and **outputs**.

Definition

A **Markov process** on a finite set X is given by an infinitesimal stochastic matrix $H: X \times X \rightarrow \mathbb{R}$ where $H(i, j) = H_{j,i}$.

Definition

An **open Markov process** is a monic cospan of finite sets where the apex is equipped with a Markov process.

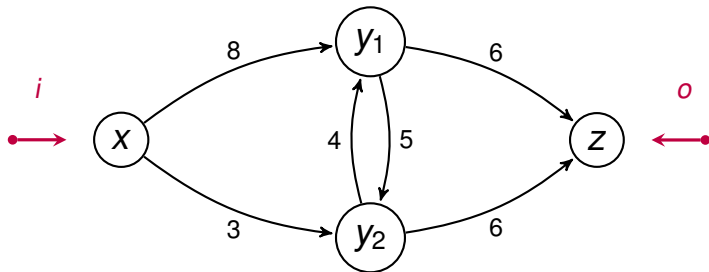


We call S and T the **inputs** and **outputs**.

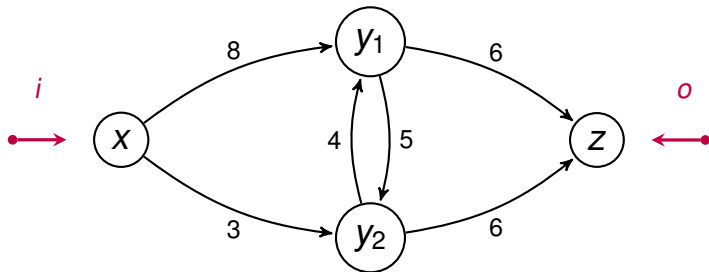
For brevity, we will denote an open Markov process as:

$$S \xrightarrow{i} (X, H) \xleftarrow{o} T$$

Here's an example of an open Markov process:

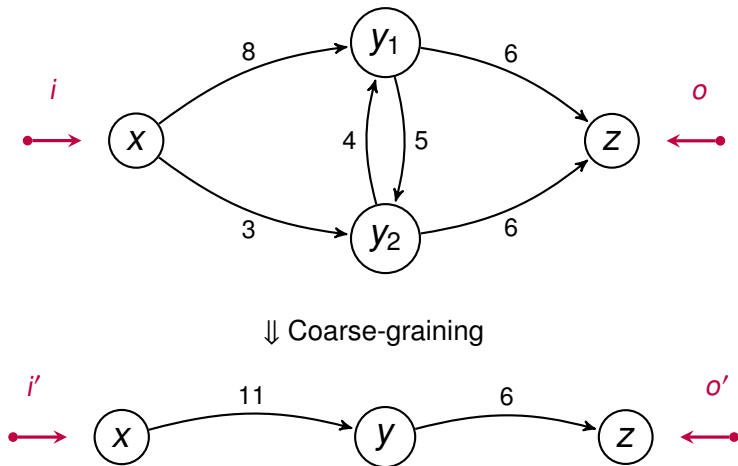


Here's an example of an open Markov process:

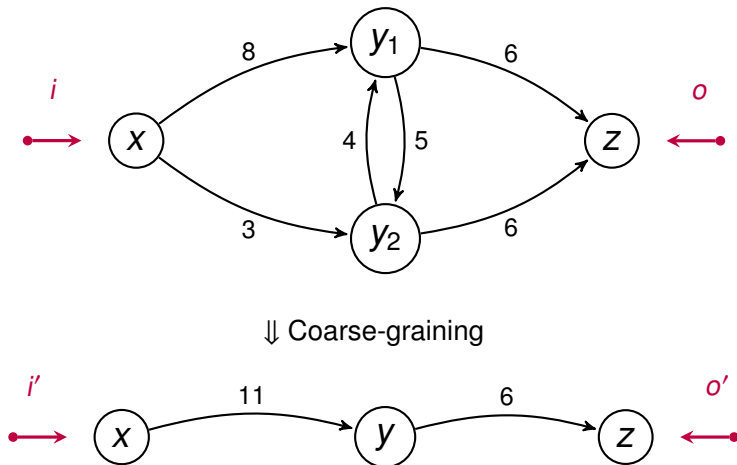


Open Markov processes will constitute the 'horizontal 1-cells' in a double category.

A coarse-graining looks like this:



A coarse-graining looks like this:

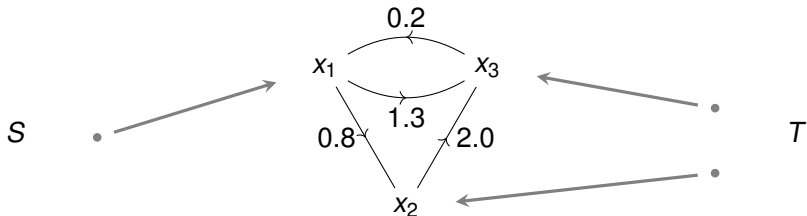


Coarse-grainings can be seen as 2-morphisms in a 'double category'.

What can we do with open Markov processes?

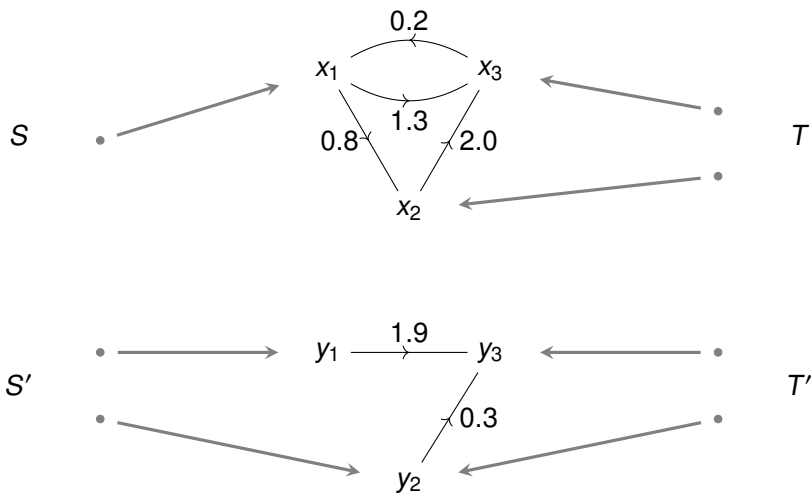
What can we do with open Markov processes?

Given two open Markov processes:

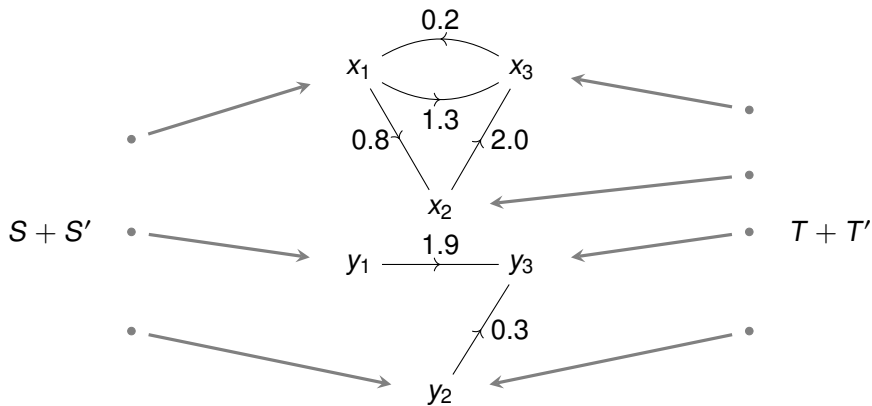


What can we do with open Markov processes?

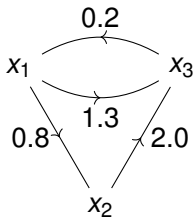
Given two open Markov processes:



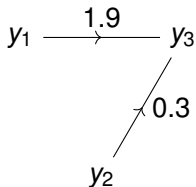
We can tensor them:



Each of these Markov processes has an associated infinitesimal stochastic matrix:

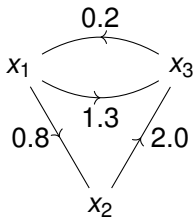


$$H = \begin{bmatrix} -2.1 & 0 & 0.2 \\ 0.8 & -2 & 0 \\ 1.3 & 2 & -0.2 \end{bmatrix}$$

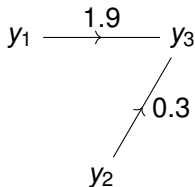


$$H' = \begin{bmatrix} -1.9 & 0 & 0 \\ 0 & -0.3 & 0 \\ 1.9 & 0.3 & 0 \end{bmatrix}$$

Each of these Markov processes has an associated infinitesimal stochastic matrix:

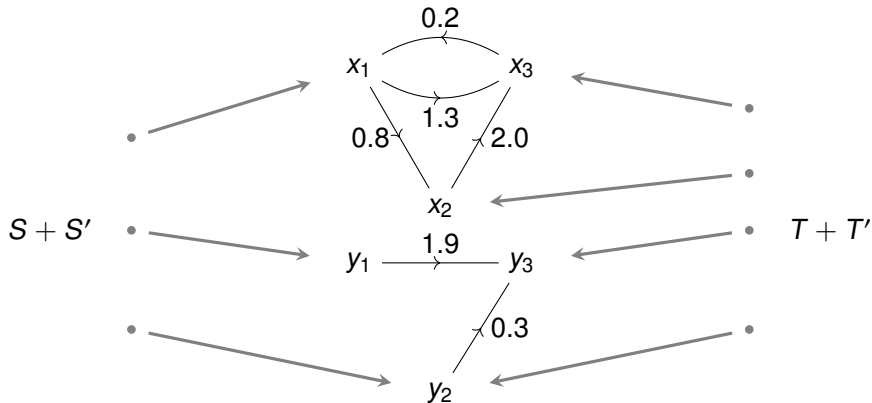


$$H = \begin{bmatrix} -2.1 & 0 & 0.2 \\ 0.8 & -2 & 0 \\ 1.3 & 2 & -0.2 \end{bmatrix}$$



$$H' = \begin{bmatrix} -1.9 & 0 & 0 \\ 0 & -0.3 & 0 \\ 1.9 & 0.3 & 0 \end{bmatrix}$$

The tensor product will have the direct sum $H \oplus H'$ as its associated infinitesimal stochastic matrix.

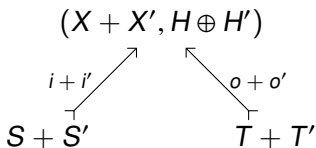


$$H \oplus H' = \begin{bmatrix} -2.1 & 0 & 0.2 & 0 & 0 & 0 \\ 0.8 & -2 & 0 & 0 & 0 & 0 \\ 1.3 & 2 & -0.2 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1.9 & 0 & 0 \\ 0 & 0 & 0 & 0 & -0.3 & 0 \\ 0 & 0 & 0 & 1.9 & 0.3 & 0 \end{bmatrix}$$

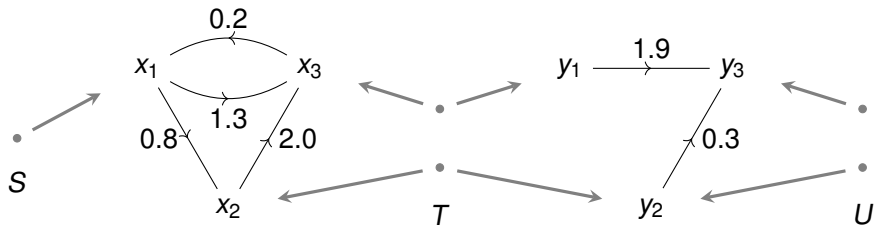
More formally, given two open Markov processes:



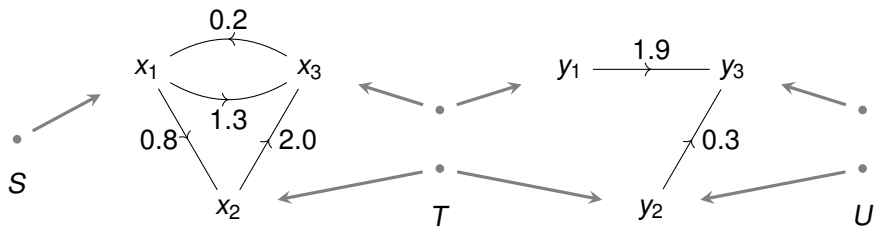
their tensor product is given by:



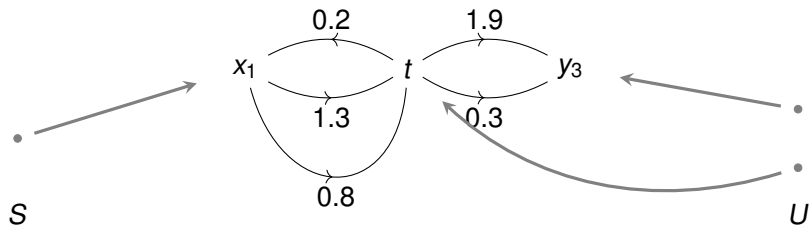
Given two open Markov processes such that the outputs of the first coincide with the inputs of the second:



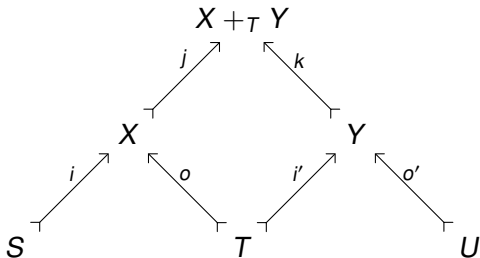
Given two open Markov processes such that the outputs of the first coincide with the inputs of the second:



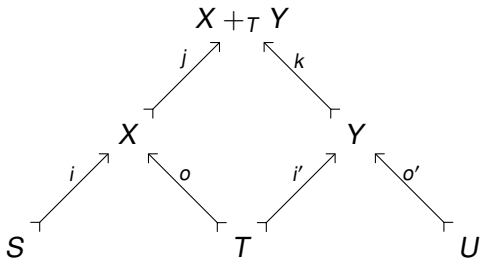
we can compose them:



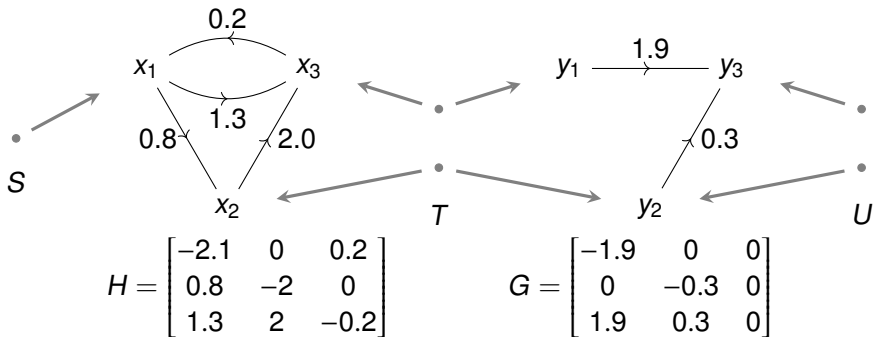
This is done by taking a pushout of the underlying cospans:

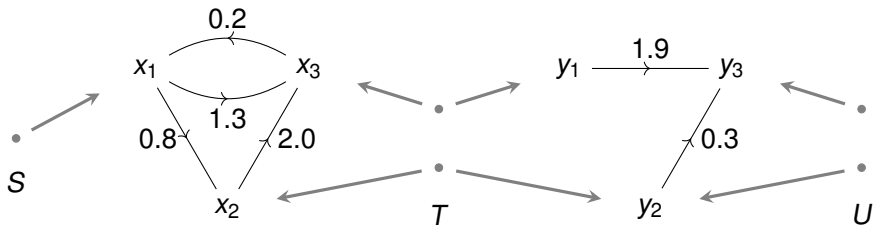


This is done by taking a pushout of the underlying cospans:



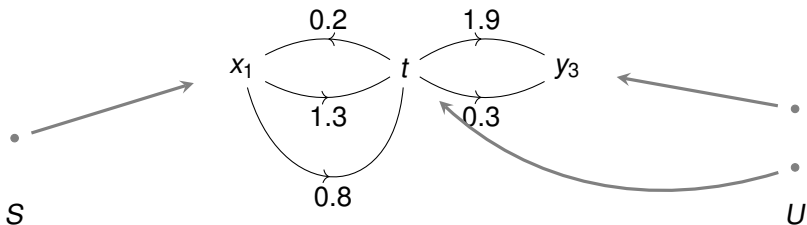
But what about the infinitesimal stochastic matrices?





$$H = \begin{bmatrix} -2.1 & 0 & 0.2 \\ 0.8 & -2 & 0 \\ 1.3 & 2 & -0.2 \end{bmatrix}$$

$$G = \begin{bmatrix} -1.9 & 0 & 0 \\ 0 & -0.3 & 0 \\ 1.9 & 0.3 & 0 \end{bmatrix}$$



$$H \odot G = \begin{bmatrix} -2.1 & 0.2 & 0 \\ 2.1 & -2.4 & 0 \\ 0 & 2.2 & 0 \end{bmatrix}$$

How do we obtain $H \odot G$?

How do we obtain $H \odot G$?

Definition

Let $f: A \rightarrow B$ be a map between finite sets. The linear map $f^*: \mathbb{R}^B \rightarrow \mathbb{R}^A$ sends any vector $v \in \mathbb{R}^B$ to its **pullback** along f , given by

$$f^*(v) = v \circ f.$$

How do we obtain $H \odot G$?

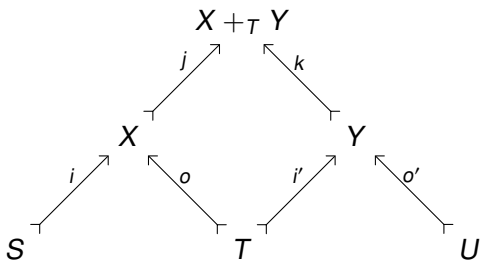
Definition

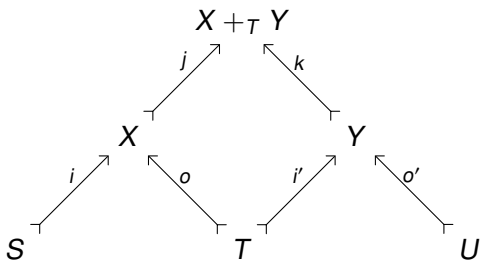
Let $f: A \rightarrow B$ be a map between finite sets. The linear map $f^*: \mathbb{R}^B \rightarrow \mathbb{R}^A$ sends any vector $v \in \mathbb{R}^B$ to its **pullback** along f , given by

$$f^*(v) = v \circ f.$$

The linear map $f_*: \mathbb{R}^A \rightarrow \mathbb{R}^B$ sends any vector $v \in \mathbb{R}^A$ to its **pushforward** along f , given by

$$(f_*(v))(b) = \sum_{\{a: f(a)=b\}} v(a).$$





$$j^*: \mathbb{R}^{X+TY} \rightarrow \mathbb{R}^X$$

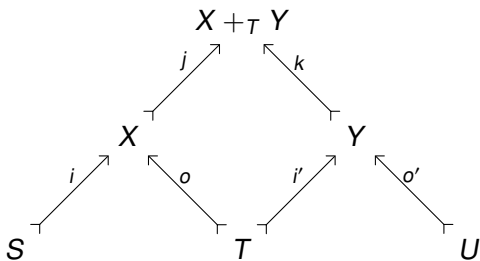
$$H: \mathbb{R}^X \rightarrow \mathbb{R}^X$$

$$j_*: \mathbb{R}^X \rightarrow \mathbb{R}^{X+TY}$$

$$k^*: \mathbb{R}^{X+TY} \rightarrow \mathbb{R}^Y$$

$$G: \mathbb{R}^Y \rightarrow \mathbb{R}^Y$$

$$k_*: \mathbb{R}^Y \rightarrow \mathbb{R}^{X+TY}$$



$$j^*: \mathbb{R}^{X+TY} \rightarrow \mathbb{R}^X$$

$$k^*: \mathbb{R}^{X+TY} \rightarrow \mathbb{R}^Y$$

$$H: \mathbb{R}^X \rightarrow \mathbb{R}^X$$

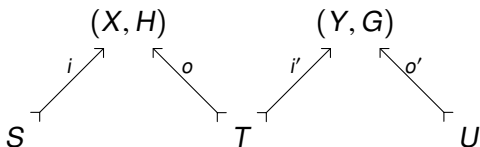
$$G: \mathbb{R}^Y \rightarrow \mathbb{R}^Y$$

$$j_*: \mathbb{R}^X \rightarrow \mathbb{R}^{X+TY}$$

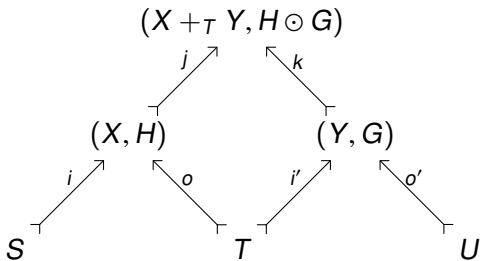
$$k_*: \mathbb{R}^Y \rightarrow \mathbb{R}^{X+TY}$$

$$H \odot G = j_* H j^* + k_* G k^*$$

And so the composite of these two open Markov processes:



is given by:



Theorem (Baez, C.)

There exists a symmetric monoidal category *Mark* with:

- finite sets as objects, and
- isomorphism classes of open Markov processes as morphisms...

...where two open Markov processes are in the same isomorphism class if (f, p, g) is a triple of bijections that make the two underlying squares of the following diagram commute.

$$\begin{array}{ccccc} S & \xrightarrow{i} & (X, H) & \xleftarrow{o} & T \\ f \downarrow & & p \downarrow & & \downarrow g \\ S' & \xrightarrow{i'} & (X', H') & \xleftarrow{o'} & T' \end{array}$$

Now let's bring coarse-grainings into the picture.

Now let's bring coarse-grainings into the picture.

Given two open Markov processes, we want 'coarse-grainings' to act as morphisms between open Markov processes.

$$S \succrightarrow^i (X, H) \leftarrow^o T$$

⇓ Coarse-graining

$$S' \succrightarrow^{i'} (X', H') \leftarrow^{o'} T'$$

Now let's bring coarse-grainings into the picture.

Given two open Markov processes, we want 'coarse-grainings' to act as morphisms between open Markov processes.

$$S \xrightarrow{i} (X, H) \xleftarrow{o} T$$

↓ Coarse-graining

$$S' \xrightarrow{i'} (X', H') \xleftarrow{o'} T'$$

To do this, we use 'double categories'.

A double category has figures like this:

$$\begin{array}{ccc} A & \xrightarrow{M} & B \\ f \downarrow & \Downarrow \alpha & \downarrow g \\ C & \xrightarrow{N} & D \end{array}$$

A double category has figures like this:

$$\begin{array}{ccc} A & \xrightarrow{M} & B \\ f \downarrow & \Downarrow \alpha & \downarrow g \\ C & \xrightarrow{N} & D \end{array}$$

We have objects, here denoted as A , B , C and D .

Vertical 1-morphisms between objects, here denoted as f and g .

Also, horizontal 1-cells between objects, here denoted as M and N ,

and morphisms between horizontal 1-cells, called 2-morphisms, here denoted as α .

These 2-morphisms can be composed both vertically and horizontally.

$$\begin{array}{ccc}
 A & \xrightarrow{M} & B \\
 f \downarrow & \Downarrow \alpha & \downarrow g \\
 C & \xrightarrow{N} & D
 \end{array}
 \qquad
 \begin{array}{ccc}
 B & \xrightarrow{M'} & E \\
 g \downarrow & \Downarrow \beta & \downarrow h \\
 D & \xrightarrow{N'} & F
 \end{array}$$

$$\begin{array}{ccc}
 C & \xrightarrow{N} & D \\
 f' \downarrow & \Downarrow \alpha' & \downarrow g' \\
 G & \xrightarrow{O} & H
 \end{array}
 \qquad
 \begin{array}{ccc}
 D & \xrightarrow{N'} & F \\
 g' \downarrow & \Downarrow \beta' & \downarrow h' \\
 I & \xrightarrow{P} & J
 \end{array}$$

These 2-morphisms can be composed both vertically and horizontally.

$$\begin{array}{ccc}
 A & \xrightarrow{M} & B \\
 \downarrow f & \Downarrow \alpha & \downarrow g \\
 C & \xrightarrow{N} & D
 \end{array}
 \qquad
 \begin{array}{ccc}
 B & \xrightarrow{M'} & E \\
 \downarrow g & \Downarrow \beta & \downarrow h \\
 D & \xrightarrow{N'} & F
 \end{array}$$

$$\begin{array}{ccc}
 C & \xrightarrow{N} & D \\
 \downarrow f' & \Downarrow \alpha' & \downarrow g' \\
 G & \xrightarrow{O} & H
 \end{array}
 \qquad
 \begin{array}{ccc}
 D & \xrightarrow{N'} & F \\
 \downarrow g' & \Downarrow \beta' & \downarrow h' \\
 I & \xrightarrow{P} & J
 \end{array}$$

$$(\alpha \circ \beta)(\alpha' \circ \beta') = (\alpha\alpha') \circ (\beta\beta')$$

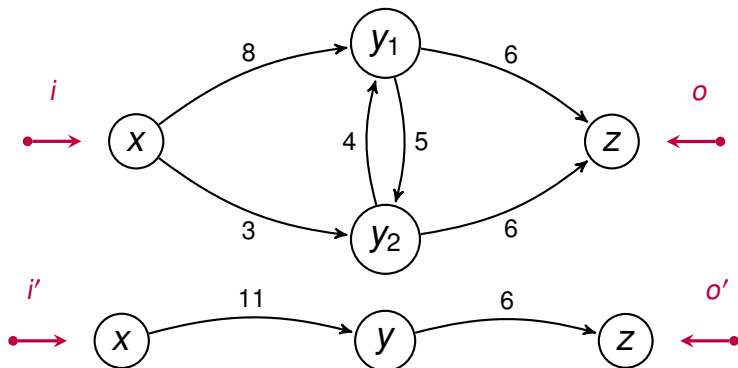
Definition

A **stochastic matrix** is a real-valued matrix such that each entry is non-negative and each column sums to 1.

Definition

A **stochastic matrix** is a real-valued matrix such that each entry is non-negative and each column sums to 1.

We can think of a surjection $p: X \rightarrow X'$ as a $|X'| \times |X|$ stochastic matrix. E.g.



$$p_* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Definition

Given a surjection $p: X \rightarrow X'$, a **stochastic section of p** is a stochastic matrix $s: X' \rightarrow X$ such that $p_*s = 1_{X'}$.

E.g.

$$p_* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad s = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 2/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Definition

Given a surjection $p: X \rightarrow X'$, a **stochastic section of p** is a stochastic matrix $s: X' \rightarrow X$ such that $p_*s = 1_{X'}$.

E.g.

$$p_* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad s = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 2/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Theorem

Let H be an infinitesimal stochastic matrix on a finite set X and $p: X \rightarrow X'$ a surjection with a stochastic section $s: X' \rightarrow X$. Then $H' := p_*Hs$ is an infinitesimal stochastic matrix.

Definition

Given a surjection $p: X \rightarrow X'$, a **stochastic section of p** is a stochastic matrix $s: X' \rightarrow X$ such that $p_*s = 1_{X'}$.

E.g.

$$p_* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad s = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 2/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Theorem

Let H be an infinitesimal stochastic matrix on a finite set X and $p: X \rightarrow X'$ a surjection with a stochastic section $s: X' \rightarrow X$. Then $H' := p_*Hs$ is an infinitesimal stochastic matrix.

Definition

We say a Markov process (infinitesimal stochastic matrix) H on a finite set X is **lumpable** with respect to a surjection $p: X \rightarrow X'$ if $H' := p_*Hs$ is independent of the choice of stochastic section $s: X' \rightarrow X$.

Definition

Given open Markov processes

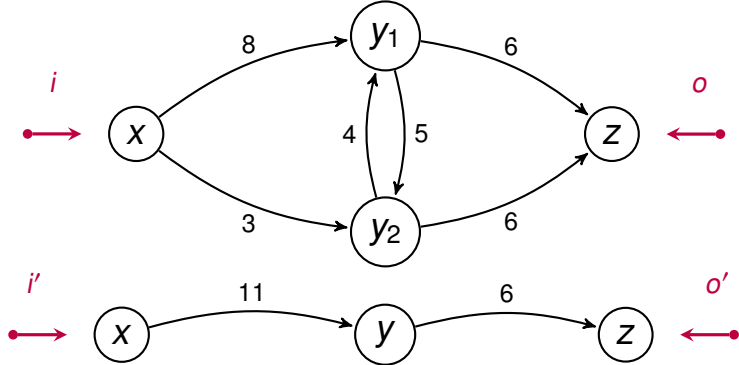
$$S \xrightarrow{i} (X, H) \xleftarrow{o} T$$

and

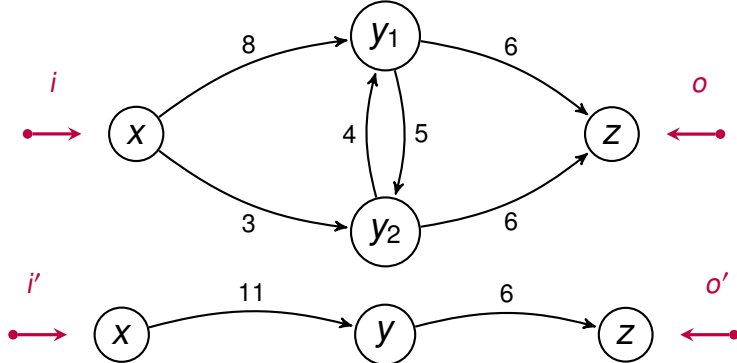
$$S' \xrightarrow{i'} (X', H') \xleftarrow{o'} T'$$

a **coarse-graining of open Markov processes** is a triple of functions (f, p, g) such that $H' = pHs$ for any stochastic section s of p and that the two following squares are pullbacks.

$$\begin{array}{ccccc} S & \xrightarrow{i} & (X, H) & \xleftarrow{o} & T \\ f \downarrow & & p \downarrow & & \downarrow g \\ S' & \xrightarrow{i'} & (X', H') & \xleftarrow{o'} & T' \end{array}$$



$$p_* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} H = \begin{bmatrix} -11 & 0 & 0 & 0 \\ 8 & -11 & 4 & 0 \\ 3 & 5 & -10 & 0 \\ 0 & 6 & 6 & 0 \end{bmatrix} s = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 2/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$



$$p_* = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad H = \begin{bmatrix} -11 & 0 & 0 & 0 \\ 8 & -11 & 4 & 0 \\ 3 & 5 & -10 & 0 \\ 0 & 6 & 6 & 0 \end{bmatrix} \quad s = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 2/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$p_* H = \begin{bmatrix} -11 & 0 & 0 & 0 \\ 11 & -6 & -6 & 0 \\ 0 & 6 & 6 & 0 \end{bmatrix} \implies H' = p_* H s = \begin{bmatrix} -11 & 0 & 0 \\ 11 & -6 & 0 \\ 0 & 6 & 0 \end{bmatrix}$$

A key result in our paper is the following:

Theorem (Baez, C.)

There exists a symmetric monoidal double category $\mathbb{M}ark$ with:

- finite sets as objects,
- functions as vertical 1-morphisms,
- open Markov processes as horizontal 1-cells, and
- coarse-grainings of open Markov processes as 2-morphisms.

$$\begin{array}{ccccc} S & \xrightarrow{i} & (X, H) & \xleftarrow{o} & T \\ f \downarrow & & p \downarrow & & \downarrow g \\ S' & \xrightarrow{i'} & (X', H') & \xleftarrow{o'} & T' \end{array}$$

A key result in our paper is the following:

Theorem (Baez, C.)

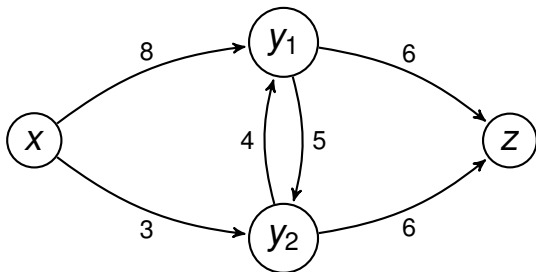
There exists a symmetric monoidal double category \mathbf{Mark} with:

- finite sets as objects,
- functions as vertical 1-morphisms,
- open Markov processes as horizontal 1-cells, and
- coarse-grainings of open Markov processes as 2-morphisms.

$$\begin{array}{ccccc} S & \xrightarrow{i} & (X, H) & \xleftarrow{o} & T \\ f \downarrow & & p \downarrow & & \downarrow g \\ S' & \xrightarrow{i'} & (X', H') & \xleftarrow{o'} & T' \end{array}$$

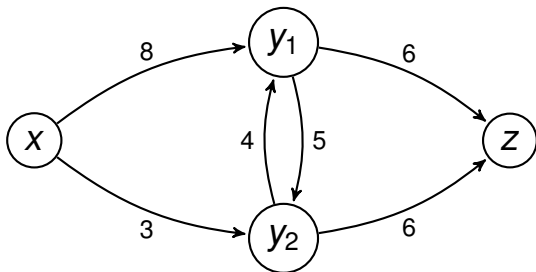
We assign semantics to this double category via a black-box ‘double functor’

$$\mathbf{■} : \mathbf{Mark} \rightarrow \mathbf{LinRel}.$$



The dynamics of this Markov process are described by the **master equation**:

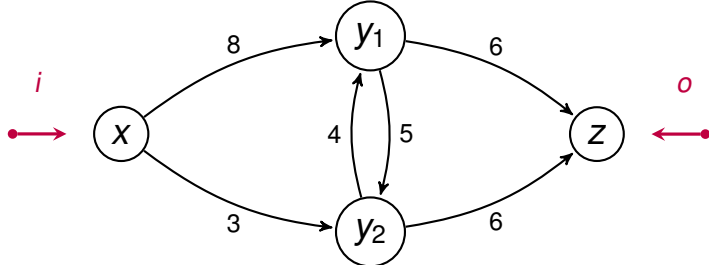
$$\frac{d}{dt}v(t) = Hv(t)$$



The dynamics of this Markov process are described by the **master equation**:

$$\frac{d}{dt}v(t) = Hv(t)$$

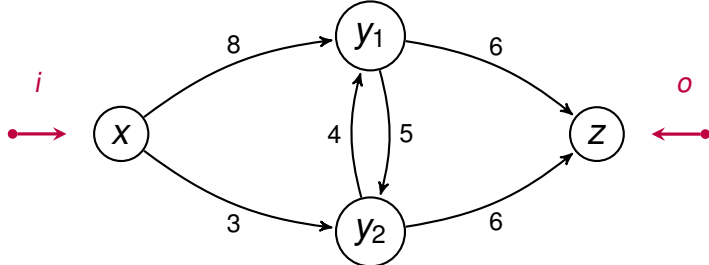
$$\frac{d}{dt} \begin{bmatrix} v_x(t) \\ v_{y_1}(t) \\ v_{y_2}(t) \\ v_z(t) \end{bmatrix} = \begin{bmatrix} -11 & 0 & 0 & 0 \\ 8 & -11 & 4 & 0 \\ 3 & 5 & -10 & 0 \\ 0 & 6 & 6 & 0 \end{bmatrix} \begin{bmatrix} v_x(t) \\ v_{y_1}(t) \\ v_{y_2}(t) \\ v_z(t) \end{bmatrix}$$



The dynamics of this *open* Markov process are described by the **open master equation**:

$$\frac{d}{dt}v(t) = Hv(t) + i_*(I(t)) - o_*(O(t))$$

$$\frac{d}{dt} \begin{bmatrix} v_X(t) \\ v_{y_1}(t) \\ v_{y_2}(t) \\ v_Z(t) \end{bmatrix} = \begin{bmatrix} -11 & 0 & 0 & 0 \\ 8 & -11 & 4 & 0 \\ 3 & 5 & -10 & 0 \\ 0 & 6 & 6 & 0 \end{bmatrix} \begin{bmatrix} v_X(t) \\ v_{y_1}(t) \\ v_{y_2}(t) \\ v_Z(t) \end{bmatrix} + \begin{bmatrix} i_*(I(t)) \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ o_*(O(t)) \end{bmatrix}$$



The dynamics of this *open* Markov process are described by the **open master equation**:

$$\frac{d}{dt}v(t) = Hv(t) + i_*(I(t)) - o_*(O(t))$$

$$\frac{d}{dt} \begin{bmatrix} v_X(t) \\ v_{y_1}(t) \\ v_{y_2}(t) \\ v_Z(t) \end{bmatrix} = \begin{bmatrix} -11 & 0 & 0 & 0 \\ 8 & -11 & 4 & 0 \\ 3 & 5 & -10 & 0 \\ 0 & 6 & 6 & 0 \end{bmatrix} \begin{bmatrix} v_X(t) \\ v_{y_1}(t) \\ v_{y_2}(t) \\ v_Z(t) \end{bmatrix} + \begin{bmatrix} i_*(I(t)) \\ 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \\ o_*(O(t)) \end{bmatrix}$$

An open Markov process is in **steady state** if the open master equation is equal to zero:

$$\frac{d}{dt}v = Hv + i_*(I) - o_*(O) = 0$$

Our main result is the following:

Theorem (Baez, C.)

There exists a black-boxing functor of symmetric monoidal double categories $\blacksquare: \mathbf{Mark} \rightarrow \mathbf{LinRel}$. This double functor is defined by:

For a finite set S , $\blacksquare(S) = \mathbb{R}^S \oplus \mathbb{R}^S$.

For a function $f: S \rightarrow S'$, $\blacksquare(f) = f_ \oplus f_*: \mathbb{R}^S \oplus \mathbb{R}^S \rightarrow \mathbb{R}^{S'} \oplus \mathbb{R}^{S'}$.*

For an open Markov process $S \rightarrow (X, H) \leftarrow T$:

$$\blacksquare(S \rightarrow (X, H) \leftarrow T) \subseteq \mathbb{R}^S \oplus \mathbb{R}^S \oplus \mathbb{R}^T \oplus \mathbb{R}^T$$

consisting of all 4-tuples which describe a steady state, and

For a coarse-graining given by (f, p, g) ,

$$\blacksquare(f, p, g): \mathbb{R}^S \oplus \mathbb{R}^S \oplus \mathbb{R}^T \oplus \mathbb{R}^T \rightarrow \mathbb{R}^{S'} \oplus \mathbb{R}^{S'} \oplus \mathbb{R}^{T'} \oplus \mathbb{R}^{T'}$$

which pushes forward a steady state of one open Markov process onto another.

For more details, see our paper in *Theory and Applications of Categories*:

- J. Baez and K. Courser, Coarse-graining open Markov processes. Available as <http://www.tac.mta.ca/tac/volumes/33/39/33-39abs.html>.