### Directed Algebraic Topology and Concurrency

Emmanuel Haucourt

MPRI : Concurrency (2.3)

Monday, the 30<sup>th</sup> of January 2012

Discrete vs Topological The algebraic topologist way

## Control Flow Graph (CFG)



#### done

Discrete vs Topological The algebraic topologist way

## Control Flow Graph (CFG)



done

Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way





Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way



Discrete vs Topological The algebraic topologist way

# Execution traces of a program as paths over its CFG

- Any execution trace induces a path
- Some paths do not come from an execution trace

Discrete vs Topological The algebraic topologist way

# Execution traces of a program as paths over its CFG

- Any execution trace induces a path
- Some paths do not come from an execution trace

Therefore the collection of path provides a (strict) overapproximation of the collection of execution traces

Discrete vs Topological The algebraic topologist way

# Execution traces of a program as paths over its CFG

- Any execution trace induces a path
- Some paths do not come from an execution trace

Therefore the collection of path provides a (strict) overapproximation of the collection of execution traces

The (infinite) collection of paths is entirely determined by the (finite) CFG

Discrete vs Topological The algebraic topologist way

The model of a program should be the finite representation of an overapproximation of the collection of all its execution traces.

Discrete vs Topological The algebraic topologist way

#### Category CDefinition (the "underlying graph" part)

Ob(C) : collection of objects

Mo(C) : collection of morphisms

s, t : mappings source, target as follows

$$\mathsf{Mo}(\mathcal{C}) \xrightarrow[t]{s} \mathsf{Ob}(\mathcal{C})$$
  
We define the homset  $\mathcal{C}[x, y] := \left\{ \gamma \in \mathsf{Mo}(\mathcal{C}) \mid \mathsf{s}(\gamma) = x \text{ and } \mathsf{t}(\gamma) = y \right\}$ 

Discrete vs Topological The algebraic topologist way

#### Category CDefinition (the "underlying local monoid" part)

id : provides each object with an identity

$$\mathsf{Mo}(\mathcal{C}) \xrightarrow[t]{s} \mathsf{Ob}(\mathcal{C})$$

The (local) composition is a partially defined binary operation often denoted by  $\circ$ 

$$\left\{(\gamma, \delta) \mid \gamma, \delta \text{ morphisms of } \mathcal{C} \text{ s.t. } \mathsf{s}(\gamma) = \mathsf{t}(\delta) \right\} \xrightarrow{\text{composition}} \mathsf{Mo}(\mathcal{C})$$

Discrete vs Topological The algebraic topologist way

#### Category CDefinition (the axioms)



The composition law is associative For all morphisms  $\gamma$  one has  $id_{t(\gamma)} \circ \gamma = \gamma = \gamma \circ id_{s(\gamma)}$ 



For all objects x one has  $s(id_x) = x = t(id_x)$
Discrete vs Topological The algebraic topologist way

Category of paths  $(1)^{l}$ freely generated by a graph

-  $I_n$  is the *finite linear order* with n + 1 elements

$$\underbrace{\begin{array}{c}0\\\bullet\end{array}}_{\bullet} \underbrace{\begin{array}{c}1\\\bullet\end{array}}_{\bullet} \underbrace{\begin{array}{c}2\\\bullet\end{array}}_{\bullet} \underbrace{\begin{array}{c}3\\\bullet\end{array}}_{\bullet} \underbrace{\begin{array}{c}n\\\bullet\end{array}}_{\bullet} \underbrace{\begin{array}{c}n\\\bullet} \underbrace{\begin{array}{c}n\\\bullet}\\ \underbrace{\begin{array}{c}n\\\bullet\end{array}}_{\bullet} \underbrace{\begin{array}{c}n\\\bullet} \underbrace{\begin{array}{c}n\\\bullet}\\ \underbrace{\begin{array}{c}n\\\bullet\end{array}}_{\bullet} \underbrace{\begin{array}{c}n\\\bullet} \underbrace{\begin{array}{c}n\\\bullet}\\ \underbrace{\begin{array}{c}n\\\bullet} \underbrace{\begin{array}{c}n\\\bullet}\\ \underbrace{\begin{array}{c}n\\\bullet} \underbrace{\begin{array}{c}n\\\bullet} \underbrace{\begin{array}{c}n\\\bullet} \underbrace{n\\\bullet} \underbrace{\begin{array}{c}n\\\bullet} \underbrace{n\\\bullet} \underbrace$$

- A path  $\gamma$  on G is a morphism of graphs from  $I_n$  to G
- The source and the target of  $\gamma$  are  $\gamma(0)$  and  $\gamma(n)$

Discrete vs Topological The algebraic topologist way

Category of paths (2) freely generated by a graph

- Given two paths  $\gamma$  (over  $I_n$ ) and  $\delta$  (over  $I_m$ ) such that  $tgt(\delta) = src(\gamma)$  we can define the *concatenation*  $\delta \cdot \gamma$  as the following path

$$I_{n+m} \longrightarrow G$$

$$k \longmapsto \begin{cases} \delta(\vec{k}) & \text{if } 0 \leq k < n \\ \gamma(\vec{k-n}) & \text{if } n \leq k < n+m \end{cases}$$

where  $\overrightarrow{k}$  stands for the arrow (k, k+1) of  $I_n$ 

- The concatenation is associative
- If  $\gamma$  (resp.  $\delta$ ) is defined over  $I_0$  then  $\delta \cdot \gamma = \delta$  (resp.  $\gamma$ )

We defined F(G) also called the *Free Category over G*.

Discrete vs Topological The algebraic topologist way

#### Model of a sequential program Pwith $G_P$ the control flow graph of P

# The model of the program is defined as the *category of paths* over its control flow graph

$$\llbracket P \rrbracket := F(G_p)$$

Discrete vs Topological The algebraic topologist way

# Cartesian product in Set

$$A imes B := ig\{(a,b) ig| a \in A ext{ and } b \in Big\}$$

There exist two mappings  $\pi_A$  and  $\pi_B$ 

 $\pi_{A}: A \times B \longrightarrow A \qquad \pi_{B}: A \times B \longrightarrow B$  $(a, b) \longmapsto a \qquad (a, b) \longmapsto b$ 

such that for all sets X the following map is a bijection

$$\operatorname{Set}[X, A \times B] \longrightarrow \operatorname{Set}[X, A] \times \operatorname{Set}[X, B]$$
$$h \longmapsto (\pi_{A^{\circ}}h, \pi_{B^{\circ}}h)$$

Discrete vs Topological The algebraic topologist way

# Cartesian product in a category C

The object *c* is the Cartesian product (in C) of *a* and *b* when there exist two morphisms  $\pi_a : c \to a$  and  $\pi_b : c \to b$  such that for all objects *x* of *C* the following map is a bijection

$$C[x, c] \longrightarrow C[x, a] \times C[x, b]$$
$$h \longmapsto (\pi_a \circ h, \pi_b \circ h)$$

When such an object c exists we write  $c = a \times b$ 

Discrete vs Topological The algebraic topologist way

Cartesian products modelling Concurrency

#### A family $P_1, \ldots, P_n$ of programs is independent if

$$\llbracket P_1 | \cdots | P_n \rrbracket \cong \llbracket P_1 \rrbracket \times \cdots \times \llbracket P_n \rrbracket$$

Discrete vs Topological The algebraic topologist way

#### Example Cartesian product in the category of graphs (*Grph*)

#### The elements of V are the vertices and those of A are the arrows In particular A and V are sets



Discrete vs Topological The algebraic topologist way

#### Example Cartesian product in the category of graphs (*Grph*)

The elements of V are the vertices and those of A are the arrows In particular A and V are sets



The Cartesian product in Grph is deduced form the Cartesian product in Set

Discrete vs Topological The algebraic topologist way

### Two simple sequential programs



Discrete vs Topological The algebraic topologist way

Two simple sequential processes running concurrently What goes wrong with the graphs



Discrete vs Topological The algebraic topologist way

Two simple sequential processes running concurrently What goes wrong with the graphs



Discrete vs Topological The algebraic topologist way

# Topological spaces reminder

A topological space is a set X and a collection  $\Omega_X \subseteq \mathcal{P}(X)$  s.t.

Discrete vs Topological The algebraic topologist way

# Topological spaces reminder

A topological space is a set X and a collection  $\Omega_X \subseteq \mathcal{P}(X)$  s.t.

1)  $\emptyset \in \Omega_X$  and  $X \in \Omega_X$ 

Discrete vs Topological The algebraic topologist way

# Topological spaces

A topological space is a set X and a collection  $\Omega_X \subseteq \mathcal{P}(X)$  s.t.

- 1)  $\emptyset \in \Omega_X$  and  $X \in \Omega_X$
- 2)  $\Omega_X$  is stable under union

Discrete vs Topological The algebraic topologist way

# Topological spaces

A topological space is a set X and a collection  $\Omega_X \subseteq \mathcal{P}(X)$  s.t.

- 1)  $\emptyset \in \Omega_X$  and  $X \in \Omega_X$
- 2)  $\Omega_X$  is stable under union
- 3)  $\Omega_X$  is stable under finite intersection

Discrete vs Topological The algebraic topologist way

#### Topological spaces reminder

A topological space is a set X and a collection  $\Omega_X \subseteq \mathcal{P}(X)$  s.t.

- 1)  $\emptyset \in \Omega_X$  and  $X \in \Omega_X$
- 2)  $\Omega_X$  is stable under union
- 3)  $\Omega_X$  is stable under finite intersection

A continuous map  $f: (X, \Omega_X) \to (Y, \Omega_Y)$  is a map  $f: X \to Y$  s.t.

 $\forall U \in \Omega_Y \ f^{-1}(U) \in \Omega_X$ 

Discrete vs Topological The algebraic topologist way

#### Topological spaces reminder

A topological space is a set X and a collection  $\Omega_X \subseteq \mathcal{P}(X)$  s.t.

- 1)  $\emptyset \in \Omega_X$  and  $X \in \Omega_X$
- 2)  $\Omega_X$  is stable under union
- 3)  $\Omega_X$  is stable under finite intersection

A continuous map  $f: (X, \Omega_X) \to (Y, \Omega_Y)$  is a map  $f: X \to Y$  s.t.

$$\forall U \in \Omega_Y \ f^{-1}(U) \in \Omega_X$$

Topological spaces and continuous maps form the category Top

Discrete vs Topological The algebraic topologist way

Two simple sequential processes running concurrently The topological model

#### Working in Top instead of Grph we have



Discrete vs Topological The algebraic topologist way

#### Functors f from C to DDefinition (preserving the "underlying graph")

A functor  $f : C \to D$  is defined by two "mappings" Ob(f) and Mo(f) such that



with  $s'(Mo(f)(\alpha)) = Ob(f)(s(\alpha))$  and  $t'(Mo(f)(\alpha)) = Ob(f)(t(\alpha))$ 

Hence it is in particular a morphism of graphs.

Discrete vs Topological The algebraic topologist way

#### Functors f from C to DDefinition (preserving the "underlying local monoid")

The "mappings" Ob(f) and Mo(f) also make the following diagram commute



and satisfies  $Mo(f)(\gamma \circ \delta) = Mo(f)(\gamma) \circ Mo(f)(\delta)$ 



Discrete *vs* Topological The algebraic topologist way

### Functors compose as morphisms of graphs do



Hence the functors should be thought of as the morphisms of categories The small categories and their funtors form a (large) category denoted by *Cat* 

Discrete *vs* Topological The algebraic topologist way

A morphism  $\gamma \in C[x, y]$  is an isomorphism when there exists  $\delta \in C[y, x]$  s.t.

$$\gamma \circ \delta = \mathrm{id}_{\mathsf{t}(\gamma)}$$
 and  $\delta \circ \gamma = \mathrm{id}_{\mathsf{s}(\gamma)}$ 

In this case  $\delta$  is unique and we write  $\delta=\gamma^{\text{-}1}$ 

$$x \xrightarrow[\delta=\gamma^{-1}]{\gamma=\delta^{-1}} y$$

We also say that x and y are isomorphic which is denoted by  $x \cong y$ A category in which every morphism is an isomorphism is called a groupoid

Discrete *vs* Topological The algebraic topologist way

### The overall idea of Algebraic Topology

#### Any functor preserve the isomorphisms

Problem: prove the topological spaces X and Y are *not* the same Strategy: find a functor F defined over Top such that  $F(X) \ncong F(Y)$ 

In this case, if  $X = \llbracket P \rrbracket$  and  $Y = \llbracket Q \rrbracket$  then the programs P and Q do not have the same behaviour.

Discrete *vs* Topological The algebraic topologist way

### The connected component functor

from Top to Set

- 1) A topologial space X is the disjoint sum of its *connected* components
- 2) Any connected subset of X is contained in a connected component of X
- 3) Any continuous direct image of a connected subset of X is connected





Moreover we have  $\pi_0(g \circ f) = \pi_0(g) \circ \pi_0(f)$ 

Discrete vs Topological The algebraic topologist way

An application involving basic (algebraic) topology The continuous image of a connected space is connected

The image of the space B is entirely contained in a *connected component* of the space V.



Discrete *vs* Topological The algebraic topologist way

## The set of connected components is a functorial construction

This situation is abstracted by classifying continuous maps from *B* to *V* according to which connected component ( $V_1$  or  $V_2$ ) the single connected components of *B* (namely *B* itself) is sent to. There are exactly two set theoretic maps from the singleton {*B*} to the pair { $V_1, V_2$ } hence there is at most (in fact exactly) two kinds of continuous maps from *B* to *V*.

$$\{B\} \longrightarrow \{V_1, V_2\}$$

In particular B and V are not homeomorphic.

Discrete vs Topological The algebraic topologist way

#### Application The compact interval and the circle are not homeomorphic

Let  $\mathbb{S}^1 := ig\{ z \in \mathbb{C} ig| |z| = 1 ig\}$  be the Euclidean circle.

Suppose  $arphi:[0,1] o \mathbb{S}^1$  is a homeomorphism. Then arphi induces a homeomorphism

 $[0, rac{1}{2}[ \ \cup \ ]rac{1}{2}, 1] \ o \ \mathbb{S}^1 ackslash \{ \varphi(rac{1}{2}) \}$ 

which does not exist!



Discrete *vs* Topological The algebraic topologist way

#### Generalization Bouquets of circles

These topological spaces are pairwise not homeomorphic. Why ?



Discrete vs Topological The algebraic topologist way

## Examples of large categories used in (directed) algebraic topology

- Set : sets and mappings
- $\mathcal{T} op$  : topological spaces and continuous maps
- $\ensuremath{\mathit{\mathfrak{Pre}}}$  : preordered sets and preorder preserving maps
- $\mathcal{P}\textit{os}$  : partially ordered sets and order preserving maps
- Mon : Monoids and their morphisms
- Cmon : Commutative monoids and their morphisms
- Gr : Groups and their morphisms
- $\mathcal{A} \mathcal{b}$  : Abelian groups and their morphisms

Discrete vs Topological The algebraic topologist way

### Example of small categories

(N, ≤): set of objects N (guess the remaining)

 $(\mathbb{N}, +, 0)$ : set of morphisms  $\mathbb{N}$  (guess the remaining)

The same way, any poset or monoid can be seen as a small category

F(G): The category freely generated by the graph G

Syntax Semantics

# The *Prolaag-Verhogen* language Edsger Wybe Dijkstra (1968)

 $\begin{array}{l} \mathcal{S}: \text{ set of semaphores and } \alpha: \mathcal{S} \to \mathbb{N} \backslash \{0,1\} \text{ associates each semaphore s with its arity } \alpha_{s} \geq 2. \\ \text{Hypothesis}: \text{ For all } \alpha \geq 2, \text{ there exist infinitely many semaphores whose arity is } \alpha. \\ \mathbb{P}(s) \text{ and } \mathbb{V}(s) \text{ are the only instructions (where } s \in \mathcal{S}) \text{ of the language.} \\ \text{A processes } P \text{ is a finite sequence of instructions, } P(j) \text{ the } j^{\text{th}} \text{ instruction with } j \geqslant 1. \end{array}$ 

P(a).V(a) and P(a).P(b).V(a).V(b)

A PV program is a finite sequence of processes

P(a).V(a) | P(a).V(a)

P(a).P(b).V(a).V(b) | P(b).P(a).V(b).V(a)

Therefore a PV program can be seen as a matrix of instructions each line of which being a process. The operator . bounds tighter that the operator |

Syntax Semantics

PV Programs as heterogeneous matrices of instructions

PV program = vector of processes 
$$\overrightarrow{P}$$
  
 $\overrightarrow{P_i} = i^{\text{th}}$  process of the program  
 $\overrightarrow{P_i}(j) = j^{\text{th}}$  instruction of the  $i^{\text{th}}$  process.  
 $l_i = \text{number of instructions of the process } \overrightarrow{P_i}$  (indexed from 1 to  $l_i$ )

$$\mathsf{dom}(\overrightarrow{P}) := \{0, \ldots, l_1\} \times \cdots \times \{0, \ldots, l_n\}$$

One has intentionally included 0

### Intuition

- P stands for "prolaag" (short for "probeer te verlagen" i.e. "try to reduce" in Dutch) and P(s) means: take an occurence of the semaphore s from the pool of resources, but wait if none is available.

- V stands for "verhogen" ("increase" in Dutch) and V(s) means: release an occurence of the semaphore s, if the process trying to perform this action does not hold any occurence of s then the instruction is just ignored and the process keeps on running.

Syntax Semantics

### An example of trace



Syntax Semantics

### Another example of trace



Syntax Semantics

### Semaphore held by a process

The real positive half-line is  $\mathbb{R}_+ = [0, +\infty[$ For each process *P*, each semaphore s and each point  $x \in \mathbb{R}_+$ , we define

$$\begin{array}{l} a_{\mathsf{x}} := \max \left\{ k \in \mathbb{N} \mid k \leq \mathsf{x} \text{ et } P(k) = \mathsf{P}(\mathsf{s}) \right\} \\ \text{and} \\ b_{\mathsf{x}} := \min \left\{ k \in \mathbb{N} \mid a_{\mathsf{x}} \leq k \text{ et } P(k) = \mathtt{V}(\mathsf{s}) \right\} \end{array}$$

with the (unusual) convention that  $\max \emptyset = \min \emptyset = \infty$ . The occupied/held part of s is

$$B_{\mathrm{s}}(P) := \left\{ x \in \mathbb{R}_{+} \mid x \in [a_{x}, b_{x}] \right\}$$
Syntax Semantics

# Example



 $a_x = 1$  and  $b_x = 4$ 

Syntax Semantics

## The forbidden area generated by a semaphore s of a PV program $\vec{P} = P_1 | ... | P_n$

The indicator function of the set  $B_{s}(P)$ 

 $\chi_{p}^{s}: \mathbb{R}_{+} \longrightarrow \{0,1\}$ 

 $\begin{array}{rcl} x & \longmapsto & \begin{cases} 1 & \text{if } x \in B_{\mathtt{B}}(P) \\ 0 & \text{otherwise} \end{cases} \\ \\ \text{For } \overrightarrow{f} := (f_1, \dots, f_n) \text{ $n$-uple of functions $\mathbb{R}_+ \to \mathbb{R}$ and $\overrightarrow{x} := (x_1, \dots, x_n) \in \mathbb{R}_+^n$} \\ \\ & \overrightarrow{f} \cdot \overrightarrow{x} := \sum_{i=1}^n f_i(x_i) \end{array}$ 

Let  $\overrightarrow{\chi}$  be the *n*-uple  $(\chi_{P_1}^s, \dots, \chi_{P_n}^s)$  of indicators of the sets  $B_s(P_1), \dots, B_s(P_n)$ The forbidden region generated by s of arity  $\alpha$  is

$$F_{s} := \left\{ \overrightarrow{x} \in \mathbb{R}^{n}_{+} \mid \overrightarrow{\chi} \cdot \overrightarrow{x} \ge \alpha \right\}$$

Syntax Semantics

Forbidden area and Model of a PV program  $\vec{P} = P_1 | ... | P_n$ 

The forbidden area of the program  $\overrightarrow{P}$  is

$$F := \bigcup_{s \in S} F_s$$

The model is the set theoretic complement (relatively to  $\mathbb{R}^n_+$ ) of its forbidden area.

 $\llbracket P_1 | \dots | P_n \rrbracket := \mathbb{R}^n_+ \backslash F$ 

Syntax Semantics

### Example P(a).V(a) | P(a).V(a)



If we were working in  $\mathbb{R}_+\times\mathbb{R}_+$  then all the points of the grey square should be removed.

Why topology is not enough The convenient structure Results

# Taking direction into account

The following spaces are homeomorphic though the first one has no local maximum



Why topology is not enough The convenient structure Results

### Topological spaces are not enough

The homset  $\mathcal{T}\!\mathit{op}\big[[0,1],[\![\overrightarrow{P}]\!]\big]$  still contains elements which do not correspond to any execution trace



Why topology is not enough The convenient structure Results

## Partially Ordered Spaces or Pospaces Leopoldo Nachbin (1968)

A pospace is a topological space X and a partial order  $\sqsubseteq$  over the underling set of X s.t.

 $\{(x, y) \in X \times X \mid x \sqsubseteq y\}$  is closed in  $X \times X$ 

The morphisms of pospaces are the continuous order preserving maps

The pospaces and their morphisms form the category Po

The real line  $\mathbb{R}$  (with its standard topology and order) provides a pospace

The products  $\mathbb{R}^n$  with the product topology and product order are pospaces

Any subset of a pospace inherits a pospace structure

Why topology is not enough The convenient structure Results

# Models of PV program $\vec{P}$

The previous topological model  $[\![\vec{P}]\!]$  inherits a pospace structures as a subset of  $\mathbb{R}^n$ 

The paths on a pospace X are the elements of the homset

 $\mathcal{P}o\big[[0,1],X\big]$ 

 $\forall \gamma \in \mathscr{Po}ig[[0,1],Xig]$ ,  $\gamma$  is constant iff  $\gamma(0) = \gamma(1)$ 

 $\forall \gamma \in \mathscr{Po}[[0,1],[0,1]], \ \gamma \text{ is onto (surjective) iff } \gamma(0) = 0 \text{ and } \gamma(1) = 1$ 

Why topology is not enough The convenient structure Results

# Paths and Execution traces

### Theorem

Given a PV program  $\overrightarrow{P}$ , any element of  $\mathcal{P}_{\mathcal{D}}[[0,1], [\![\overrightarrow{P}]\!]]$  induces an execution trace of  $\overrightarrow{P}$  and conversely, any execution trace of  $\overrightarrow{P}$  is induced by some element of  $\mathcal{P}_{\mathcal{D}}[[0,1], [\![\overrightarrow{P}]\!]]$ 



Introduction Why topology is not end The PV language The convenient structur Introducing Direction Results

### Direct image of paths on a pospaces Characterization

The direct image im(f) of any  $f \in \mathcal{P}o[X, Y]$  inherits a pospace struture still denoted by im(f)

#### Theorem

Given a pospace X and  $\gamma \in \mathcal{P}o[[0,1], X]$ , we have  $im(\gamma) \cong \{*\}$  or  $im(\gamma) \cong [0,1]$