Card Shuffles & Cantor space: an inverse semigroup perspective

Peter M. Hines
Y.C.C.S.A. , Univ. York

N.B.S.A.N. – York – 2019

The intuition / motivation

Card shuffles are well-studied in *combinatorics*, *probability*, *representation theory*, *statistics*, &c.

Credit: Johnny Blood Photography

There are also applications in theoretical computer science.

These are best understood via (inverse) semigroup theory.

Why Computer Science?

One example: Race Conditions

"A parent process spawns several child processes, each of which competes for the parent's resources. These requests must be dealt with in order, one at a time. The outcome varies depending on the order in which these are processed"

This is the motivation. However, today's talk is about the **semigroup theory**. Any applications are side-effects!

How to shuffle two (possibly infinite) decks of cards

The Riffle Shuffle

- Cards from Deck A and Deck B are merged into a single stack.
- At each step, a single card is taken from the bottom of either A or B, and placed on top of the stack.

Some important conventions:

- The ordering of cards is preserved.
- Every card from each deck ends up in the stack.

Modeling Riffle Shuffles

We model a **deck** of a cards by the well-ordered set

$$[0, a) \;= \{n \in \mathbb{N} : n < a\}$$

(We allow for $a = \infty$ in this definition, giving $[0, \infty) = \mathbb{N}$).

A **pair of decks** is modeled by the disjoint union

$$[0, a) \uplus [0, b) \;=\; [0, a) \times \{0\} \;\cup\; [0, b) \times \{1\}$$

equipped with the induced partial order

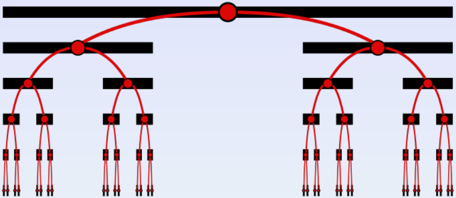
$$(x, i) \leqslant (y, j) \;\text{ iff } i = j \;\text{ and } x \leqslant y$$

A **riffle shuffle** is then an order-preserving bijection:

$$\phi : [0, a) \uplus [0, b) \to [0, a + b)$$

The infinitary setting

Every riffle shuffle of two infinite decks uniquely determines & is determined by a point of the Cantor set $\mathcal{C}$.



Formally, one-sided countably infinite strings over $\{0,1\}$, or equivalently, $\mathcal{C} = \mathbf{Set}(\mathbb{N}, \{0,1\})$.

Computational motivation

Infinite shuffles model potentially *non-terminating* processes.

The correspondence:

Given a shuffle of two infinite decks

$$\phi : [0, \infty) \uplus [0, \infty) \to [0, \infty)$$

we define the corresponding Cantor point $p_\phi \in \mathcal{C}$ by

$$p_\phi(n) = \begin{cases} 0 & n = \phi(x, 0) \text{ for some } x \in \mathbb{N} \\ 1 & n = \phi(x, 1) \text{ for some } x \in \mathbb{N} \end{cases}$$

Operationally: p_ϕ is an instruction

At the n^{th} step, take the next card from:

- The first deck, when $p_\phi(n) = 0$
- The second deck, when $p_\phi(n) = 1$

An illustrative example

The **perfect riffle shuffle**:

Cards are alternately taken from each deck

corresponds to the **alternating Cantor point** $a(n) = n \ (mod\ 2)$**.**

$$a = 0101010101\ldots$$

Not all Cantor points determine shuffles:

We require a Cantor point $c \in \mathcal{C}$ to satisfy:

$$\sum_{j=0}^{\infty} c(j) = \infty = \sum_{j=0}^{\infty} 1 - c(j)$$

For the condition, "every card is played at some point".

An inverse semigroup approach ...

Some notation ...

We will be mixing order theory & partiality.

By analogy with Kleene equality

In a poset we write $f(a) \precsim g(b)$ for

"$f(a) \leqslant g(b)$ provided both $f(a)$ and $g(b)$ are defined".

A partial injection $f : (P, \leqslant) \rightarrow (Q, \leqslant)$ is

- **monotone (mono.)** when $a \leqslant b \Rightarrow f(a) \precsim f(b)$,
- **anti-monotone (anti)** when $a \leqslant b \Rightarrow f(b) \precsim f(a)$.

Monos, antis. and composition

Notation:

Denote the monotone partial injections from P to Q by $\mathbf{mono}(P,Q)$, and the anti-monotone partial injections from P to Q by $\mathbf{anti}(P,Q)$.

Given partial injections:

- $m \in \mathbf{mono}(P,Q)$ and $n \in \mathbf{mono}(Q,R)$
- $a \in \mathbf{anti}(P,Q)$ and $b \in \mathbf{anti}(Q,R)$

then $ba, nm \in \mathbf{mono}(P,R)$ and $na, bm \in \mathbf{anti}(P,R)$.

Composing monotone & anti-monotone partial injections

	mono.	anti.
mono.	mono.	anti.
anti.	anti.	mono.

A word of warning!

Monotone partial injections form categories / monoids

— *these are not generally inverse categories / monoids.*

A simple counterexample:

Consider the successor function

$$\mathit{succ} \in \mathbf{mono}((\mathbb{N}, =), (\mathbb{N}, \leqslant))$$

This is monotone, but its generalised inverse certainly is not!

A simple and relevant setting:

Denote the set of monotone partial injections on $\mathbb{N}$ by

$$\mathbf{mono}(\mathbb{N},\mathbb{N}) \leqslant \mathbf{pInj}(\mathbb{N},\mathbb{N})$$

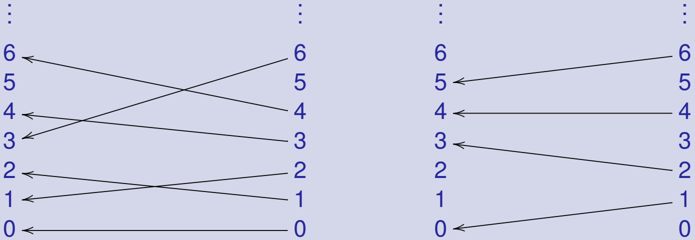
This set is closed under composition and generalised inverse and so forms an inverse monoid.

Some intuition ...

Think of monotone partial injections on $\mathbb{N}$ as 'planar diagrams'.

Monotonicity as planarity for partial injections on $\mathbb{N}$

non-planar | planar



non-monotone | monotone

Planarity is a **big deal** in many areas of C.S.

Why planarity?

1. The **quantum Jones polynomial** algorithm (Aharanov, Jones, Landau)
   - A QM algorithm for computing Jones polynomials at $e^{\frac{2k\pi i}{5}}$
   - Classically, a (presumably) $P\#$ problem.
   - Based on the *Temperley-Lieb algebra* "Knot theory without crossings" – L. Kaufmann.
2. **Lambek pregroups** (From categorical linguistics)
   - Becoming used Natural Language Processing
   - Diagrams determined by *planarity & acyclicity*.
3. **Complexity theory** (Planarity provides bounds to complexity).
   - *Matchgates and classical simulation of quantum circuits* – R. Jozsa, A. Miyake
   - Restricting swap gates allows for *efficient classical simulation* of QM circuits.
4. ...

First ... some simple theory!

Characterising monotone partial injections on $\mathbb{N}$

Graphically, or otherwise, the following is straightforward:

Every $f \in mono(\mathbb{N}, \mathbb{N})$ is uniquely determined by its initial & final idempotents.

planar

⋮	⋮
6	**6**
5	5
4	**4**
3	3
2	**2**
1	**1**
0	0

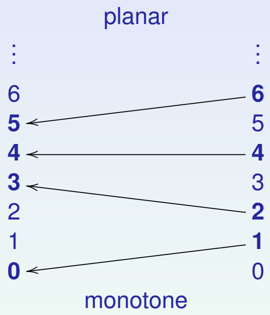
monotone

$f^{-1}f$ and ff^{-1} are partial identities on $\mathbb{N}$.

Characterising monotone partial injections on $\mathbb{N}$

Graphically, or otherwise, the following is straightforward:

Every $f \in mono(\mathbb{N}, \mathbb{N})$ is uniquely determined by its initial & final idempotents on the well-ordered set $\mathbb{N}$.



$f^{-1}f$ and ff^{-1} are partial identities on $\mathbb{N}$.

Idempotents as Cantor points

Indicator functions for subsets of $\mathbb{N}$ are points of the Cantor set.

Abusing notation slightly: given $e^2 = e = 1_E \in mono(\mathbb{N},\mathbb{N})$,

we write $Ind_e : \mathbb{N} \to \{0,1\}$, or $Ind_e \in \mathcal{C}$.

A trivial observation:

For a monotone partial injection $f \in \mathbf{mono}(\mathbb{N},\mathbb{N})$,

$$\sum_{n=0}^{\infty} ind_{ff^{-1}}(n) = \sum_{n=0}^{\infty} ind_{f^{-1}f}(n) \in \mathbb{N} \cup \{\infty\}$$

A few simple definitions

A pair of Cantor points $(d, c) \in \mathcal{C} \times \mathcal{C}$ is **balanced** when

$$\sum_{j=0}^{\infty} d(j) \; = \; \sum_{j=0}^{\infty} c(j) \; \in \; \mathbb{N} \cup \{\infty\}$$

We denote the set of balanced Cantor pairs by $\mathfrak{B} \subseteq \mathcal{C} \times \mathcal{C}$.

There is a $1:1$ correspondence $\mathfrak{B} \; \equiv \; \mathbf{mono}(\mathbb{N}, \mathbb{N})$.

Giving this explicitly:

Balanced Cantor pairs $\equiv$ monotone partial injections

- $\Leftarrow$ Given $f \in \mathbf{mono}(\mathbb{N}, \mathbb{N})$, the balanced pair is:

$$(Ind_{ff^{-1}}, Ind_{f^{-1}f}) \in \mathfrak{B}$$

- $\Rightarrow$ Given $(t, s) \in \mathfrak{B}$, define $m_{(t,s)} \in \mathbf{mono}(\mathbb{N}, \mathbb{N})$ by

$$m_{(t,s)}(n) = \begin{cases} \perp & s(n) = 0 \\ min_{x \in \mathbb{N}} \left\{ \sum_{j=0}^{x} t(j) = \sum_{j=0}^{n} s(j) \right\} & s(n) = 1 \end{cases}$$

An illustration:

A balanced pair of Cantor points:

$$t = 1001110\ ,\ \ s = 0110101\ldots$$

$n =$	0	1	2	3	4	5	6	...
$s(n) =$	0	1	1	0	1	0	1	...
$t(n) =$	1	0	0	1	1	1	0	...

An illustration:

A balanced pair of Cantor points:

$$t = 1001110\ldots \quad , \quad s = 0110101\ldots$$

$n =$	0	1	2	3	4	5	6	...
$s(n) =$	0	1	1	0	1	0	1	...
$\sum_{j\leqslant n} s(j) =$	0	1	2	2	3	3	4	...
$\sum_{j\leqslant n} t(j) =$	1	1	1	2	3	4	4	...
$t(n) =$	1	0	0	1	1	1	0	...

A monoid operation on balanced pairs?

There exists some composition operation $\cdot : \mathfrak{B} \times \mathfrak{B} \rightarrow \mathfrak{B}$ such that $(\mathfrak{B}, \cdot) \;\cong\; \textbf{mono}(\mathbb{N}, \mathbb{N})$.

What does this look like?

Normal forms (I)

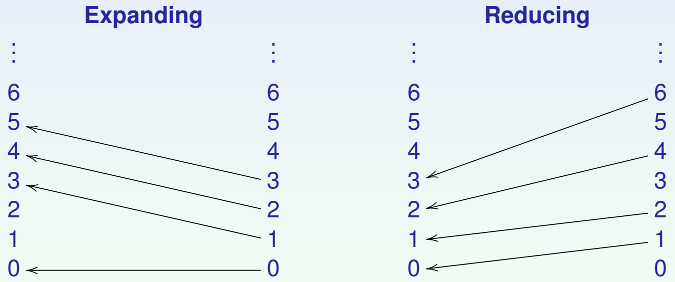
A monotone partial injection is **reducing** when

$$ff^{-1} = 1_{[0,x)} \quad \text{for some } x \in \mathbb{N} \cup \{\infty\}$$

Dually, it is **expanding** when

$$ff^{-1}f = 1_{[0,x)} \quad \text{for some } x \in \mathbb{N} \cup \{\infty\}$$

An illustrative example:



Cantor points as reducing / expanding arrows

Reducing (resp. expanding) arrows are uniquely determined by their initial (resp. final) idempotents.

Given $c \in \mathfrak{C}$, define $Red_c \in Mono(\mathbb{N}, \mathbb{N})$ by

$$Red_c(n) = \begin{cases} \perp & n = 0 \\ \sum_{j=0}^{n} c(n) - 1 & n = 1 \end{cases}$$

Dually, define $Exp_c \in Mono(\mathbb{N}, \mathbb{N})$ by

$$Exp_c = Red_c^{-1}$$

Normal forms (II)

Given arbitrary $f \neq 0 \in \textbf{mono}(\mathbb{N}, \mathbb{N})$,
then the balanced pair $(t, s) = (Ind_{ff^{-1}}, Ind_{f^{-1}f}) \in \mathfrak{B}$
is the unique balanced pair satisfying

$$f = Exp_t Red_s$$

(The only non-trivial point is uniqueness, which follows since (t, s) is required to be balanced).

By considering normal forms (or directly)

Given $(v,u),(t,s) \in \mathfrak{B}$, define a composition by:

$$(v,u) \cdot (t,s) = \begin{cases} (\mathbf{0},\mathbf{0}) & t(n)u(n) = 0 \ \ \forall n \in \mathbb{N} \\ (x,w) & \text{otherwise} \end{cases}$$

where $w(n) = s(n).u(j).t(j) \in \{0,1\}$,

$$j = min_{j \in \mathbb{N}} \left\{ \sum_{\alpha=0}^{j} t(\alpha) = \sum_{\alpha=0}^{n} s(\alpha) \right\}$$

and similarly, $x(n) = v(n).u(k).t(k) \in \{0,1\}$,

$$k = min_{k \in \mathbb{N}} \left\{ \sum_{\alpha=0}^{k} u(\alpha) = \sum_{\alpha=0}^{n} v(\alpha) \right\}$$

The generalised inverse is immediate: $(t,s)^{-1} = (s.t)$.

This gives$(\mathfrak{B},\cdot) \cong$ **mono**$(\mathbb{N},\mathbb{N})$ as required.

Duals and self-encodings

Recall the complement / dual operation on the Cantor set:

$$c^{\perp}(n) = c(n) + 1 \ (mod\ 2) \ \ \forall\ c \in \mathcal{C}$$

(e.g. $c = 0100101\ldots$ has complement $c^{\perp} = 1011010\ldots$).

A key definition

An element $(b, a) \in \mathfrak{B}$ is **complemented** when $(b^{\perp}, a^{\perp}) \in \mathfrak{B}$, and is **dual-inverse** when $(b, a)^{-1} = (b^{\perp}, a^{\perp})$.

Duals, inverses, and shuffles

Fairly simply, $(b, a) \in \mathfrak{B}$ is dual-inverse iff

$$b = a^{\perp} \text{ and } \sum_{n=0}^{\infty} a(n) = \infty = \sum_{n=0}^{\infty} 1 - a(n)$$

There is then a bijective correspondence between dual-inverse arrows of $\mathfrak{B}$, and riffle shuffles of two infinite decks of cards.

These are both determined by Cantor points $a \in \mathcal{C}$ satisfying

$$\sum_{n=0}^{\infty} a(n) = \infty = \sum_{n=0}^{\infty} 1 - a(n)$$

We call these **dual-balanced Cantor points.**

From D.-B. Cantor points to Young Tableaux

There is an correspondence between

1. D.-B. Cantor points,
2. Shuffles of infinite decks of cards,
3. (∞, ∞) Young tableaux.

$$c = 1\ 0\ 1\ 0\ 1\ 0\ 1\ 1\ 0\ 0\ \ldots \in \mathcal{C}$$

$c(n) = 0$	1	3	5	8	9	$\ldots$
$c(n) = 1$	0	2	4	6	7	$\ldots$

The obvious question:

What about **standard** Young tableaux?

This is where we start to need the inverse semigroup theory.

D.-B. Cantor points as inverse monoids

Proposition: There is a 1:1 correspondence between:

- Dual-balanced Cantor points,
- Effective representations of the 2-generator polycyclic monoid within **mono**$(\mathbb{N}, \mathbb{N})$.

Recall – the polycyclic monoid P_2

- Two generators, $\{p, q\}$
- Relations:

$$pq^{-1} = 0 = q^{-1}p \text{ and } pp^{-1} = 1 = qq^{-1}$$

Useful fact: polycyclic monoids are *congruence-free*.

Monotone representations of P_2

Let $c \in \mathcal{C}$ be dual-balanced. Looking at normal forms,

$$Exp_c Red_{c^\perp} \; = \; (c, c^\perp) \quad \text{and} \quad (c^\perp, c) = Exp_{c^\perp} Red_c$$

By construction, $Red_c Exp_c \; = \; 1_{\mathbb{N}} \; = \; Red_{c^\perp} Exp_{c^\perp}$.

By definition, of $(\ )^\perp : \mathcal{C} \to \mathcal{C}$,

$$c(n) = 0 \iff c^\perp(n) = 1$$

and so

$$Red_c Exp_{c^\perp} \; = \; 0_{\mathbb{N}} \; = \; Red_{c^\perp} Exp_c$$

The assignment $p \mapsto Red_c$, $q \to Red_{c^\perp}$ gives an effective monotone representation of P_2.

— all effective monotone representations arise in this way.

As always ... an example

For the *alternating Cantor point*, or *perfect riffle shuffle*,

$$a(n) = n \ (mod\ 2) \quad \text{or } a = 010101010101\ldots$$

we derive the representation of P_2 corresponding to the Cantor pairing:

$$p^{-1}(x) = 2x \quad \text{and} \quad q^{-1}(x) = 2x + 1$$

On to standard Young tableaux

In **standard** Young tableaux, the cells are well-ordered both *horizontally* and *vertically*.

x	a
y	b

$$\begin{array}{ccc} x & \leqslant & a \\ \leqslant & & \leqslant \\ b & \leqslant & y \end{array}$$

Horizonal ordering corresponds to monotonicity.

What about the vertical ordering?

Some standard(?) semigroup theory

A (binary) **ballot sequence** is an element $w \in \{0,1\}^*$ where, for *every* prefix u of w,

$$\#1s \text{ in } u \quad \leqslant \quad \#0s \text{ in } u$$

Denote the set of all finite ballot sequences by *Ballot*
— this forms a submonoid of $\{0,1\}^*$.

By contradiction: Consider $v, w \in Ballot$ such that $vw \notin Ballot$. Then there exists some prefix u of vw satisfying $\#0s$ in $u < \#1s$ in u. As $v \in Ballot$, u is not a prefix of v, so $u = vl$, for some prefix l of w. However, $\#0s$ in $v \geqslant \#1s$ in v. Therefore, $\#1s$ in $l \geqslant \#0s$ in l, contradicting the assumption that $w \in Ballot$.

Ballot sequences are *well-studied* in combinatorics – but also make for interesting monoids!

Proposition The monoid of binary ballot sequences is not finitely generated.

By contradiction: Assume a finite generating set G for $Ballot \leqslant \{0, 1\}^*$. As G is finite, the longest contiguous string of $1s$ in any member of G is bounded by some finite $K \in \mathbb{N}$. No composite of members of G can account for the ballot sequence $0^{K+1}1^{K+1}$.

From the finite to the infinite:

A Cantor point $c \in \mathcal{C}$ is **ballot** when every prefix is a member of the Ballot monoid.

$$\sum_{j=0}^{N} c(j) \leqslant \sum_{j=0}^{N} c^{\perp}(j) \ \forall N \in \mathbb{N}$$

Denote the ballot Cantor points by $\mathcal{S} \subseteq \mathcal{C}$.

Question:

How do such Cantor points behave under the point-wise partial order:

$$a \leqslant b \text{ iff } a(n) \leqslant b(n) \ \forall n \in \mathbb{N}$$

The ballot Scott domain

Key properties:

- There is no top element & they are **not** closed under joins $(c \vee d)(n) \; = \; max\{c(n), d(n)\}$.
- They **are** closed under the meet, $(c \wedge d)(n) \; = \; c(n)d(n)$
- There is a bottom element $\bot(n) = 0$, for all $n \in \mathbb{N}$.
- The supremum of every chain $c_0 \leqslant c_1 \leqslant c_2 \leqslant \ldots$ is also in $\mathcal{B}$

  – chain-completeness $\Rightarrow$ directed completeness, assuming the axiom of choice (Iwamura's Lemma).
- There is a notion of **finite support** / **compactness**: $c \in \mathcal{B}$ is "**compact**" iff $\sum_{j=0}^{\infty} c(j) < \infty$, and every element is the supremum of a chain of such elements.

Scott Domains ...

- Introduced by Dana Scott (early 1970s) to model pure untyped λ calculus

  — and hence **computational universality**.

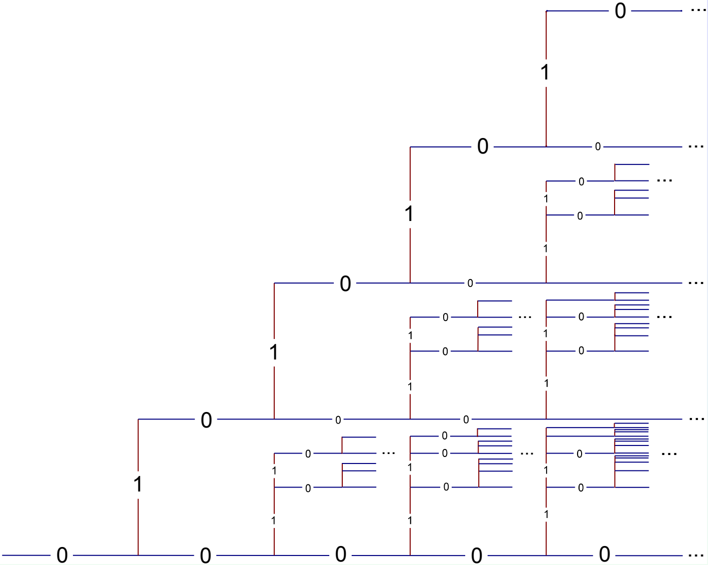
- Also used for semantics of **functional programming** languages, due to the existence of solutions of arbitrary **fixed-point equations**.

This particular Scott domain is

a subset of Cantor space.

We can draw a picture.

The ballot Cantor points



Combining two properties:

A **dual-balanced ballot** Cantor point $c \in \mathcal{C}$ satisfies:

- $\sum_{j=0}^{\infty} c(j) = \sum_{j=0}^{\infty} c^{\perp}(j)$
- $\sum_{j=0}^{N} c(j) \leqslant \sum_{j=0}^{N} c^{\perp}(j)$.

There is a 1:1 correspondence:

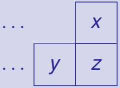
DBB Cantor points $\equiv$ Standard (∞, ∞) Young tableaux

These are given by:

Removing the 'compact' points from the ballot Scott domain.

The motivation

As card-shuffling:

The only way we can see: 

with $z \leqslant x$ is when

"More cards have been laid from Deck B than from Deck A"

As DBB Cantor points are dual-balanced, they uniquely determine representations of P_2, as monotone partial injections on $\mathbb{N}$.

Call these **standard monotone representations.**

Some computer science motivation

Recall the motivation for studying Shuffles, from *race conditions*.

- Operations from Process A push data onto a stack.
- Operations from Process B pop data off a stack.
- The Ballot condition prevents us from trying to *read data from an empty stack*.

A very standard result **N. & P. (1970)**

There exists an embedding of P_∞ into P_2.

Recall:

The infinite-generator polycyclic monoid P_∞ has generating set $\{p_j\}_{j\in\mathbb{N}}$, with relations $p_j p_k^{-1} = \begin{cases} 1 & j = k \\ 0 & j \neq k \end{cases}$

The embedding is given by

$$p_j \mapsto pq^j \quad , \quad p_j^{-1} \mapsto q^{-j}p^{-1}$$

Straightforward to check that the required relations are satisfied!

Polycyclic monoids as bijections

A slightly lesser-known result **PH & MVL (... a while back)**

Representations of P_∞ within **pInj**$(\mathbb{N}, \mathbb{N})$ correspond to injections

$$\mathbb{N} \times \mathbb{N} \hookrightarrow \mathbb{N}$$

which are bijections when the representation is effective

A very simple construction

For a given representation, we define

$$\Psi(x, y) \ = \ p_x^{-1}(y) \ \ \forall (x, y) \in \mathbb{N} \times \mathbb{N}$$

A worked example:

Let's do this for the *standard monotone representation* determined by the *alternating Cantor point* $a \in \mathfrak{C}$.

$$p^{-1}(n) = 2n \quad \text{and} \quad q^{-1}(n) = 2n + 1$$

Expanding out, we get

$$\Psi_a(x, y) = q^{-x}p^{-1}(y) = 2^{x+1}y + 2^x - 1$$

A (Hilbert-hotel style) bijection from $\mathbb{N} \times \mathbb{N}$ to $\mathbb{N}$.

View this as ‘shuffling countably infinitely many decks of cards’.

$x =$ \ $y =$	0	1	2	3	4	5	$\dots$
0	0	2	4	6	8	10	$\dots$
1	1	5	9	13	17	21	$\dots$
2	3	11	19	27	35	43	$\dots$
3	7	23	39	55	71	87	$\dots$
4	15	47	79	111	143	175	$\dots$
5	31	94	159	223	287	351	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

A few observations:

Simple observations

- This table contains every natural number.
- Both rows and columns appear to be well-ordered

  — an $(\infty, \infty, \infty, \ldots)$ standard Young tableau?
- There seems to be some 'underlying fractal structure' ...

More practically — how easy is it to perform this shuffle?

Deep fractal structure ??

On the n^{th} step, we play from Deck x:

	$\cdots$ $Deck_4$	$Deck_3$	$Deck_2$	$Deck_1$	$Deck_0$
$n = 1$					•
$n = 2$				•	
$n = 3$					•
$n = 4$			•		
$n = 5$					•
$n = 6$				•	
$n = 7$					•
$n = 8$		•			
$n = 9$					•
$n = 10$				•	
$n = 11$					•
$n = 12$			•		
$n = 13$					•
$n = 14$				•	
$n = 15$					•
$n = 16$	•				

This looks kind of familiar!

	…	2^4	2^3	2^2	2^1	2^0
$n=1$						1
$n=2$					1	0
$n=3$					1	1
$n=4$				1	0	0
$n=5$				1	0	1
$n=6$				1	1	0
$n=7$				1	1	1
$n=8$			1	0	0	0
$n=9$			1	0	0	1
$n=10$			1	0	1	0
$n=11$			1	0	1	1
$n=12$			1	1	0	0
$n=13$			1	1	0	1
$n=14$			1	1	1	0
$n=15$			1	1	1	1
$n=16$		1	0	0	0	0

Performing the perfect infinite riffle

A very simple rule

1. Count in binary ...
2. Which bit has changed from 0 to 1?
3. Play a card from that deck!

The standard Young property

It is straightforward that **rows** and **columns** are well-ordered:

k	m
l	

$k = \Psi_a(x, y)$ for some $(x, y) \in \mathbb{N} \times \mathbb{N}$.

- $l = 2k + 1 \ > \ k$
- $m = k + 2^{y+1} \ > \ k$.

They also contain all natural numbers.

Claim These properties follow generally from:

1. The fact that representations of P_2 are *monotone* (since they are derived from DB Cantor points).
2. The ballot property on these Cantor points.

A quick outline

Let $c \in \mathfrak{C}$ be a dual-balanced ballot Cantor point. This determines an effective monotone representation $P_2 \overset{c}{\hookrightarrow} \mathbf{pInj}(\mathbb{N}, \mathbb{N})$ which corresponds to an (∞, ∞) Young tableau:

$p^{-1}(0)$	$p^{-1}(1)$	$p^{-1}(2)$	$p^{-1}(3)$	$p^{-1}(4)$	$\dots$
$q^{-1}(0)$	$q^{-1}(1)$	$q^{-1}(2)$	$q^{-1}(3)$	$q^{-1}(4)$	$\dots$

By the ballot property, $p^{-1}(n) \leqslant q^{-1}(n)$, so this is *standard*.

A quick outline (cont.)

By the same properties, $q^{-k}(n) < q^{-(k+1)}(n)$, so the following table is has well-ordered rows and columns:

$p^{-1}(0)$	$p^{-1}(1)$	$p^{-1}(2)$	$p^{-1}(3)$	$p^{-1}(4)$	$\dots$
$q^{-1}p^{-1}(0)$	$q^{-1}p^{-1}(1)$	$q^{-1}p^{-1}(2)$	$q^{-1}p^{-1}(3)$	$q^{-1}p^{-1}(4)$	$\dots$
$q^{-2}p^{-1}(0)$	$q^{-2}p^{-1}(1)$	$q^{-2}p^{-1}(2)$	$q^{-2}p^{-1}(3)$	$q^{-2}p^{-1}(4)$	$\dots$
$q^{-3}p^{-1}(0)$	$q^{-3}p^{-1}(1)$	$q^{-3}p^{-1}(2)$	$q^{-3}p^{-1}(3)$	$q^{-3}p^{-1}(4)$	$\dots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\ddots$

Finally, as q^{-1} is monotone and $q^{-1}(x) > p^{-1}(x)$, we deduce that

$$\bigcap_{j=0}^{\infty} q^{-j}(\mathbb{N}) = \varnothing$$

and so the embedding of P_∞ is effective.

From the infinite to the finite

Every DBB Cantor point determines an $(\infty, \infty, \infty, \ldots)$ standard Young tableau. These can be written as sequences of finite standard Young tableaux.

For the alternating Cantor point:

n=0

0	

n=1

0
1

n=2

0	2
1	

n=3

0	2
1	
3	

n=4

0	2	4
1		
3		

n=5

0	2	4
1	5	
3		

. . . just a complicated way of counting in binary!

From Sets to Spaces

Adding in Topology & Category Theory

The clopen topology

The **Cantor space** $\mathfrak{C}$ is the Cantor set $\mathcal{C}$ together with the **clopen topology**.
This is generated by the **clopen basis**

$$\{w\mathcal{C} \quad : \quad w \in \{0,1\}^*\}$$

Basic open covers

These are determined by some $R \in \{0,1\}^*$ where

$$\bigcup_{r \in R} r\mathfrak{C} \ = \ \mathfrak{C}$$

A **minimal** cover is a basic open cover satisfying

$$r\mathfrak{C} \cap r'\mathfrak{C} = \emptyset \ \ \forall r \neq r' \in R$$

From open covers to prefix codes

Given some $R \subseteq \{0,1\}^*$, then

$R\mathfrak{C}$ is a minimal open cover
iff
R is a maximal prefix code.

A relevant fact ...

The set of maximal prefix codes is closed under the induced subset composition.

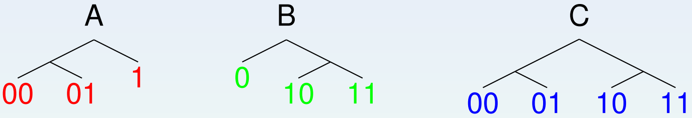
We will concern ourselves with **finite** covers.

A picture is worth a thousand words ...

There is a well-established bijection between

- (Finite) prefix codes over $\{0,1\}^*$
- (Finite) complete binary trees.

  $A = \{00, 01, 1\}$, $B = \{0, 10, 11\}$, $C = \{00, 01, 10, 11\}$



An uninteresting(?) groupoid

Define the groupoid $\mathcal{P}$ as follows:

Objects All finite maximal prefix codes over $\{0, 1\}^*$

Arrows Bijections of prefix codes that are **monotone** w.r.t. the lexicographic ordering.

This is fairly uninteresting:

There is precisely one arrow between any two prefix codes of the same size.

What is interesting about $\mathcal{W}$?

The groupoid $\mathcal{P}$ has two distinct categorical tensors.

Given finite, maximal prefix codes $R, S \subseteq \{0,1\}^*$,

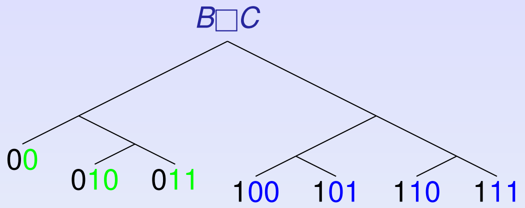
The Multiplicative tensor $R \boxtimes S = \{rs\}_{r \in R, s \in S}$

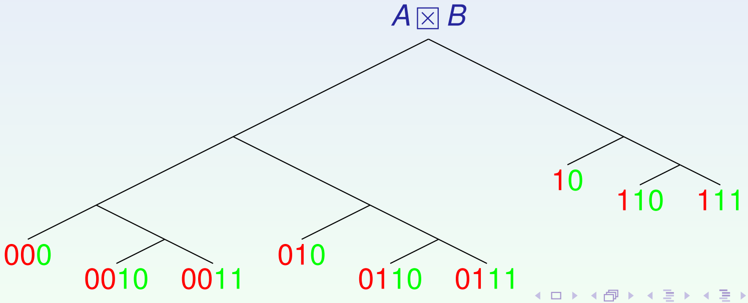
The additive tensor $R \square S = \{0\} \times R \cup \{1\} \times S$

On arrows ...

The tensor on arrows is determined by uniqueness(!)

As binary trees ...





A categorical reminder ...

A key structure from the foundations of category theory (MacLane's Theorem):

MacLane's $(W, \_\square\_)$

Objects	All finite complete binary trees.
Arrows	Unique arrow between any two trees of the same rank.
Tensor	Paste two trees together at their root!

We have an equivalence of categories $(\mathcal{P}, \_\square\_) \equiv (\mathcal{W}, \_\square\_)$.

Question: What about the 'other tensor' & categorical distributivity?

Back to Cantor space

Each arrow of MacLane's $\mathcal{W}$ uniquely determines a homeomorphism of Cantor space:

- Given $\phi : R \to S$, a monotone bijection of finite maximal prefix codes,
- define $T(\phi) : \mathfrak{C} \to \mathfrak{C}$ by:

$$T(\phi)(rw) = \phi(r)w \ \ \forall\, r \in R,\ \ w \in \mathfrak{C}$$

$T(\phi)$ is:

- Injective, by construction.
- Globally defined, as $R\mathfrak{C}$ is an open cover.
- Surjective, as $S\mathfrak{C}$ is an open cover.
- Continuous — basic open sets map to basic open sets.

What the F. is this group?

$T(\ )$ is a **faithful functor** from a groupoid to a group.

Its image is closed under *composition* and *inverses*, and contains the identity.

The obvious question:

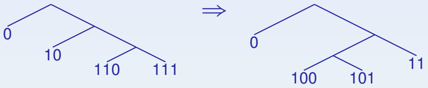
What is this group of homeomorphisms of $\mathfrak{C}$?

Some explicit calculations ...

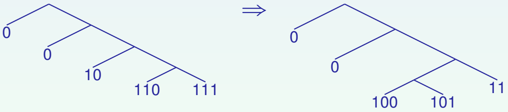
Within the groupoid $\mathcal{W}$ (equivalently, $\mathcal{P}$)

Let X_0 be the unique arrow

Let X_1 be the unique arrow



Let X_2 be the unique arrow



Let X_3 be . . .

Mapping things down to Cantor space ...

Let us then define $x_j = T(X_j) : \mathfrak{C} \to \mathfrak{C}$ for all $n \in \mathbb{N}$

Simple direct calculation gives:

$$x_i^{-1} x_j x_i \ = \ x_{j+1} \ \ \forall i < j \in \mathbb{N}$$

We have the *generators* and *relations* of Thompson's group $\mathcal{F}$

Appealing to the fact that F has no non-abelian quotients,

The image of $T(\_)$ contains a copy of $\mathcal{F}$.

With a little more work ...

The image of $T(\ )$ is precisely Thompson's $\mathcal{F}$.

More questions than answers

There is a close connection between:

1. Minimal basic open covers of Cantor space
2. MacLane's W & the foundations of category theory
3. Thompson's group $\mathcal{F}$.

By varying assumptions (finiteness, monotonicity, maximality, &c.) we recover many interesting & familiar structures!

What structures do we recover when we look at

minimal basic open covers of:

1. The Ballot Scott domain?
2. Dual-Balanced Ballot Cantor points?