Lecture 1
Background: Discrete mathematical structures and automata

MIEFis - Quantum Computation, L. S. Barbosa, 2019-20

Summary
(1) Sets, functions, relations. Isomorphism and cardinality.
(2) Ordered structures: preorders, partial orders, lattices. Complete lattices. Ideals and filters. Boolean algebras. Application: the theorem of Knaster-Tarski. Lattices as algebraic structures.
(3) Reversibility. Groups as a prototypical algebraic structure. Groups of permutations. Action of a group. Application: Cayley theorem.
(4) Automata as a basic computational model

1 Sets, relations, cardinality

Sets, relations, functions

(cf recall your HASKELL experience)

- Function $f : A \longrightarrow B$. Composition. Injective, surjective and bijective functions
- Set. Cardinality. Powerset (2^A).
- Binary relations; $2^{A\times B} \cong 2^{A^B}$.
- Equivalence relations. Partition. Quotient set as a partition.

Finite and infinite sets.

- equicardinality *vs* isomorphism.
- finite *vs* infinite
- countable *vs* uncountable:
  A set A is countable iff there is an injective function $f : A \longrightarrow \mathbb{N}$.
  It is infinite countable if f is a bijection.

Ranking cardinality.

$$|A| \leq |B| \ \text{ iff } \ \text{there is an injection } f : A \longrightarrow B$$

Relation $\leq$ above is a total order. Note that proving antisymmetry (i.e. the Cantor-Bernstein-Schroeder theorem) and totality (which requires the axiom of choice) is extremely hard.

Some applications:

Theorem

$\mathbb{N}$ and $\mathbb{Z}$ have the same cardinal

Proof (hint).
Consider $h : \mathbb{Z} \longrightarrow \mathbb{N}$ defined as follows and show it is a bijection

$$h(x) \;=\; \begin{cases} 2x & \Longleftarrow x \geq 0 \\ -2x+1 & \Longleftarrow \text{otherwise} \end{cases}$$

□

Theorem

$\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$ have the same cardinal

Look for a bijection between $\mathbb{N}$ and $\mathbb{N} \times \mathbb{N}$. Let's see some (of several) possibilities:

Proof (1).
Enumerate all pairs of numbers which sum $0, 1, 2, \cdots$:

$$\begin{array}{lllll}
(0,0) & & & & \\
(0,1) & (1,0) & & & \\
(0,2) & (1,1) & (2,0) & & \\
(0,3) & (1,2) & (2,1) & (3,0) & \cdots \\
\vdots & & & &
\end{array}$$

For every sum n there are only finitely many, actually $n+1$ pais (i, j) that sum n. I.e. for every number n, one gets to all the pairs which sum n. On the other hand, since every pair of numbers (i, j) has a finite sum it will appear somewhere on this list. This defines a bijection

$$\begin{array}{rcl}
(0,0) & \mapsto & 0 \\
(0,1) & \mapsto & 1 \\
(1,0) & \mapsto & 2 \\
(0,2) & \mapsto & 3 \\
(1,1) & \mapsto & 4 \\
 & \cdots\cdots &
\end{array}$$

□

Proof (2).

Consider the following correspondence:

$$n \;\mapsto\; (i, j)$$

such that $n = 2i + j$ (i.e. basically factoring number n in its odd and even parts). Clearly the pair (i, j) is unique, thus establishing a bijection

$$f : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}_{odd}$$

Remains to prove that $\mathbb{N} \cong \mathbb{N}_{odd}$.

□

Theorem

The union of a finite number of countably infinite sets is countably infinite.

Proof (hint).

$$\begin{tikzcd}
X & x_0 \arrow[d] & x_1 \arrow[d] & x_3 \arrow[d] & \cdots \\
Y & y_0 \arrow[d] & y_1 \arrow[d] & y_3 \arrow[d] & \cdots \\
Z & z_0 \arrow[uur] & z_1 \arrow[uur] & z_3 \arrow[uur] & \cdots
\end{tikzcd}$$

□

$\boxed{\textbf{Theorem}}$

$|\mathbb{N}| < |\mathbb{R}|$

Proof.
To show that $|\mathbb{N}| \leq |\mathbb{R}|$ is trivial: function $h(n) = n$ is injective.

The difficult part is to prove that $\mathbb{N} \neq \mathbb{R}$. Let us prove an even stronger statement: that there is no surjection from $\mathbb{N}$ to $[0, 1[$. Consider an arbitrary function $h : \mathbb{N} \longrightarrow [0, 1[$ with which one may enumerate an infinite sequence of real numbers

$$r_0, r_1, r_2, \cdots$$

making $r_i = h(i)$.

To show that h in not surjective, we have to find a real x such that $r_n \neq x$ for all $n \in \mathbb{N}$. Let us build x as an infinite dizime

$$0.x[0]x[1]x[2] \cdots$$

such that

$$x[i] \ = \ \begin{cases} 1 \ \Leftarrow \ r_i[i] = 0 \\ 0 \ \Leftarrow \ \text{otherwise} \end{cases}$$

Observe that any real $h(n)$ differs from number x exactly in position n, and conclude that x does not belong to the image of h.

□

$\boxed{\textbf{Theorem}}$

$2^{\mathbb{N}}$ is uncountable.

Proof.
If this is not the case, and $2^{\mathbb{N}}$ is countably infinite, there is an enumeration of sets such that

$$2^{\mathbb{N}} \ = \ \{R_1, R_2, \cdots\}$$

Let $D = \{n \in \mathbb{N} \mid n \notin R_n\}$. Set D is a set of natural numbers and thus should appear somewhere in the enumeration $R_1, R_2, \cdots$. Suppose $D = R_j$ for some value j. Does $j \in R_j$? If yes, by definition of D, $j \notin D$, which contradicts $D = R_j$. If, alternatively, $j \notin R_j$ then $j \in D$ which is again a contradiction.

□

$\boxed{\text{Exercise (Cantor Theorem)}}$

Generalise the previous proof to show that, for any set X, $|X| < |\mathbf{2}^X|$.

Proof (Cantor Theorem).

Function

$$\eta : x \longrightarrow \{x\}$$

is trivially injective. However there is no surjection $h : X \longrightarrow \mathbf{2}^X$.

To argue by contradiction, suppose such a function h exists and consider a set

$$W = \{x \in X \mid x \notin h(x)\}$$

If h is indeed surjective it must exist an element $w \in X$ such that $h(w) = W$ and w may or may not belong to $h(w)$.
These two cases are as follows as both lead to a contradiction:

- $w \in h(w)$ but then $w \notin W$,
- $w \notin h(w)$ but then $w \in W$

which invalidates our assumption that h is a surjection.

□

Note. This theorem sheds light on the limits of computability: *there are more problems that we might want to solve than there are programs to solve them, even though both are infinite*. To see this restrict your attention to one type of problem: deciding whether a string has some property (e.g. having even length, being a palindrome, or a legal Haskell program). A property can be identified with the set of strings that happen to share it. Clearly, the number of possible programs is no bigger than the number of strings, while the number of sets of strings is strictly greater. This shows the existence of unsolvable problems, i.e. problems that can be formulated but not possibly solved.

Theorem: the pigeonhole principle

Let m objects be distributed into n containers. If $m > n$, then some container contains at least two objects

Proof.
By contrapositive: let us show that if every container contains at most one object, then $m \leq n$.

If c_i is the number of objects in container i, then

$$m = \sum_{i=1}^{n} c_i$$

but, every container contains at most one object, we get

$$m \leq \sum_{i=1}^{n} c_i \leq \sum_{i=1}^{n} 1 = n$$

□

Applications: Given a large enough number of objects with a bounded number of properties, eventually at least two of them will share a property.

Ex. 1 - Theorem
Suppose that every point in the real plane is coloured either red or blue. Then for any distance $d > 0$, there are two points exactly distance d from one another that are the same color.

Ex. 2 - Theorem
For any natural number n, there is a nonzero multiple of n whose digits are all 0s and 1s.

Note. The proofs applying the pigeonhole proofs are non-constructive, i.e. although they prove the existence of something, they fail to provide an explicit example that proves the theorem. Constructive proofs are the ones more suitable for Computer Science (why?).

2 Orders

- pre-order
- partial order
- lattice, bounded lattice and complete lattice

Exercise 1

In a poset $(P, \sqsubseteq)$ define the supremum of S, represented by $\sqcup S$, as the least upper bound (lub) of S. The *dual* notion of infimum, $\sqcap S$ is defined as the greatest lower bound (glb) of S.

Characterise lub and glb in $(\mathcal{P}(X), \subseteq)$ and $(\mathbb{N}, div)$, where div is integer division. Suppose P is a poset with a top and a bottom element $\top$ and $\bot$, respectively, i.e. $\sqcap P = \bot$ and $\sqcup P = \top$. Explain why $\sqcap\emptyset = \top$ and $\sqcup\emptyset = \bot$.

Exercise 2

A morphism between posets $(P, \sqsubseteq)$ and $(Q, \subseteq)$ is a function $f : P \longrightarrow Q$ such that

$$x \sqsubseteq y \Rightarrow f(x) \subseteq f(y)$$

i.e. a monotonic function. What extra structure must a morphism between lattices, bounded lattices or complete lattice preserve?

Exercise 3

A lattice is complete if infimum and supremum are defined for arbitrary subsets. Characterise as complete lattices i) the set of all sub-spaces of a vectorial space; ii) the set of sub-groups of a group; iii) any finite lattice.

The Knaster-Tarski theorem.

A most relevant result about complete lattices for the semantics of computation is the theorem of Knaster-Tarski [5] on the existence of fixed points of a monotonic function. Such special points, for which $x = f(x)$, give meaning to recursive functions.

Theorem

Let $(U, \sqsubseteq)$ be a complete lattice and $f : U \longrightarrow U$ a monotonic function. The least and the greatest fixed points of f are given by

$$\begin{aligned} m &= \bigsqcap \{x \in U \mid f(x) \sqsubseteq x\} \\ M &= \bigsqcup \{x \in U \mid x \sqsubseteq f(x)\} \end{aligned}$$

respectively.

Proof.
Let us show that m is the least fixed point of f. Let $X = \{x \in U \mid f(x) \sqsubseteq x\}$ and choose $x \in X$ arbitrarily. Clearly, $m \sqsubseteq x$ and, f being monotonic, $f(m) \sqsubseteq f(x)$. On the other hand, $f(x) \sqsubseteq x$, because $x \in X$. Thus, we may conclude that, for all $x \in X$, $f(m) \sqsubseteq x$. In particular, $f(m) \sqsubseteq m$, which leads us to $f(f(m)) \sqsubseteq f(m)$. We conclude that $f(m) \in X$ and, thus, $m \sqsubseteq f(m)$. But then $f(m) = m$ as expected. The second part of the theorem comes from this one; if f is monotonic in $(U, \sqsubseteq)$, then it is also monotonic in the complete lattice formed by the inverse order $(U, \sqsupseteq)$. If M is the least fixed point of f in $(U, \sqsupseteq)$, it will be the *greatest* fixed point of the same function in $(U, \sqsubseteq)$.

□

Lattices as algebraic structures

Lattices can as well be seem as algebraic structures taking $\sqcup$ and $\sqcap$ as binary operations satisfying the axioms for commutativity, associativity, idempotence, and the following absorption laws:

$$\begin{aligned} a \sqcup (a \sqcap b) = a \\ a \sqcap (a \sqcup b) = a \end{aligned}$$

Clearly, $a \leq b \Leftrightarrow a \sqcup b = b \Leftrightarrow a \sqcap b = a$.

3 Groups

A group (G, θ, u) is a set G with a binary operation θ which is associative, and equipped with an identity element u and an inverse:

$$a^{-1}\theta a = u = a\theta a^{-1}$$

Note that *monoid* lacks inverse, and a *semigroup* also drops the identity element.

Exercise 4

Show that $(\mathbb{R}^+, \times, 1)$ and $(\mathbb{R}^+, +, 0)$ are groups. Prove that a bijection between them is obtained by functions ln_e and e^-.

Exercise 5

Show that $S_n = (\{\sigma : n \longrightarrow n \mid \sigma \text{ is a permutation}\}, \cdot, id)$ is a group. This is usually called the *symmetry group of degree* n.

Exercise 6

Prove the following properties:

1. $a\theta b = a\theta c \Rightarrow b = c$ (dually, $b\theta a = c\theta a \Rightarrow b = c$).
2. $a^{-1^{-1}} = a$.
3. $(a\theta b)^{-1} = b^{-1}\theta a^{-1}$.
4. The equation $a\theta x = b$ has a unique solution $x = a^{-1}\theta b$.

Action of a group

A group (G, θ, u) acts over a set X through a function (the action) $\tau : G \times X \longrightarrow X$ which satisfies the following properties: $\tau(u, x) = x$ and $\tau(g\theta f, x) = \tau(g, (\tau(f, x)))$.

Exercise 7

Show that i) the group S_n acts over set n (initial fragment of $\mathbb{N}$ with n numbers), and ii) that every group (G, θ, u) acts over itself through the map $(g, x) \mapsto g\theta x\theta g^{-1}$.

Cayley Theorem.

The set of bijections $f : X \longrightarrow X$ over a set X with functional composition forms a group of *transformations* (which is the identity? And the inverse?). The following is a main result in the theory of groups:

Theorem

Every group is isomorphic to a group of transformations

Proof.

Let (G, θ, u) be a group. For each element a of G define a map $f_a : G \longrightarrow G$ such that $f_a(x) = a\theta x$.

Let us show that a new group T can be defined over the set of transformations above:

1. The (functional) composition of two elements of T is in T:

$$(f_a \cdot f_b)(x) = f_a(f_b(x)) = f_a(b\theta x) = a\theta(b\theta x) = (a\theta b)\theta x = f_{a\theta b}(x)$$

2. For identity,

$$f_u(x) = u\theta x = x$$

3. For inverse,

$$f_a \cdot f_{a^{-1}}(x) = a\theta(a^{-1}\theta x) = (a\theta a^{-1})\theta x = u\theta x = x$$

We have proved that T is a group (note that axioms are inherited from the properties of function composition restricted to bijections). It remains to show that T is *isomorphic* to G. Let $h : G \longrightarrow T$ be defined by $h\,(a) = f_a$.

- Clearly $h(a\theta b) = f_{a\theta b} = f_a \cdot f_b = h(a) \cdot h(b)$, $h(u) = f_u$ (which is the identity) and $ha^{-1} = f_{a^{-1}} = h(a^{-1})$. Thus h is a homomorphism between both groups.
- T is entirely composed of bijections f_a for every element $a \in G$, thus h is a surjection.
- If $a \neq b$, then $h(a) = f_a \neq f_b = h(b)$; thus h is injective.

□

4 Automata

Classical Automata

An automaton over a set N of names is a tuple $\langle \mathsf{S}, \mathsf{s}_0, \mathsf{N}, \downarrow, \longrightarrow \rangle$ where

- $\mathsf{S} = \{\mathsf{s}_0, \mathsf{s}_1, \mathsf{s}_2, ...\}$ is a set of states, with a distinguished initial state s_0
- $\downarrow \subseteq \mathsf{S}$ is the set of terminating or final states

$$\downarrow \mathsf{s} \;\equiv\; \mathsf{s} \in \downarrow$$

- $\longrightarrow \subseteq \mathsf{S} \times \mathsf{N} \times \mathsf{S}$ is the transition relation, often given as an N-indexed family of binary relations

$$\mathsf{s} \xrightarrow{\mathsf{a}} \mathsf{s}' \;\equiv\; \langle \mathsf{s}', \mathsf{a}, \mathsf{s} \rangle \in \longrightarrow$$

Variants

- deterministic
- non deterministic
- finite
- image finite
- ...

Morphism

A morphism relating two automata over N, $\langle \mathsf{S}, \mathsf{N}, \mathsf{s}_0, \downarrow, \longrightarrow \rangle$ and $\langle \mathsf{T}, \mathsf{N}, \mathsf{t}_0, \downarrow', \longrightarrow' \rangle$, is a function $\mathsf{h} : \mathsf{S} \longrightarrow \mathsf{S}'$ st

$$\begin{array}{rcl} \mathsf{s} \xrightarrow{\mathsf{a}} \mathsf{s}' & \Rightarrow & \mathsf{h}\,\mathsf{s} \xrightarrow{\mathsf{a}}{}' \mathsf{h}\,\mathsf{s}' \\ \mathsf{s} \downarrow & \Rightarrow & \mathsf{h}\,\mathsf{s} \downarrow' \\ \mathsf{t}_0 & = & \mathsf{h}(\mathsf{s}_0) \end{array}$$

Morphisms preserve transitions, initial state and termination

Reachability

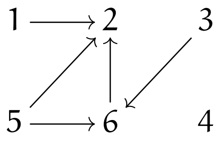
The reachability relation, $\longrightarrow^* \subseteq \mathsf{S} \times \mathsf{N}^* \times \mathsf{S}$, is defined inductively

- $\mathsf{s} \xrightarrow{\epsilon}{}^{*} \mathsf{s}$ for each $\mathsf{s} \in \mathsf{S}$, where $\epsilon \in \mathsf{N}^*$ denotes the empty word;
- if $\mathsf{s} \xrightarrow{\mathsf{a}} \mathsf{s}''$ and $\mathsf{s}'' \xrightarrow{\sigma}{}^{*} \mathsf{s}'$ then $\mathsf{s} \xrightarrow{\mathsf{a}\sigma}{}^{*} \mathsf{s}'$, for $\mathsf{a} \in \mathsf{N}, \sigma \in \mathsf{N}^*$

A state $t \in S$ is reachable from $s \in S$ iff there is a word $\sigma \in N^*$ st $s \xrightarrow{\sigma}^* t$

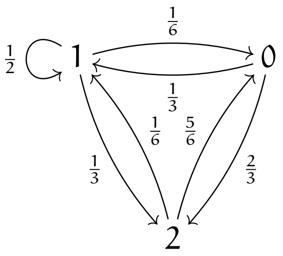
Matricial representation

(with $N = \emptyset$)



- The states of a system correspond to column vectors;
- The automata dynamics is encoded in Boolean matrices: $M[i, j] = 1$ if and only if there is an edge (path of length 1) from vertex j to vertex i;
- Multiplying the current state vector by matrix M yields progress from one state to another in one time step;
- Multiple step dynamics are obtained via matrix multiplication.

Probabilistic Automata

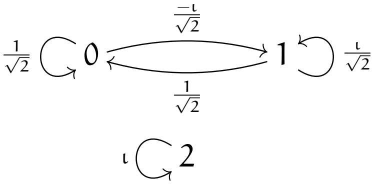


- The vectors that represent states of a probabilistic physical system express indeterminacy about the exact physical state of the system;
- The matrices that represent the dynamics express indeterminacy about the way the physical system will change over time;
- The entries of the evolution matrix enable the computation of the likelihood of transitioning from one state to the next: $M[i, j]$ gives the probability of a transition from vertex j to vertex i;
- Typically, but not necessarily, matrices encoding automata dynamics are *double stochastic*, which encodes the following two conditions:
  - the sum of all the weights leaving a vertex is 1 and
  - the sum of all the weights entering a vertex is 1.

i.e. the dynamics is time symmetric.

- The way in which the indeterminacy progresses is simulated by matrix multiplication.

Quantum automata



- States in a quantum automaton are represented by column vectors of complex numbers whose sum of moduli squared is 1.
- The dynamics is represented by unitary matrices and is therefore reversible. $M^\dagger$ takes a state from time t to $t-1$. Note that the modulus squared of a unitary matrix forms a doubly stochastic matrix. Actually, the probabilities of quantum mechanics are always given as the modulus square of complex numbers.
- The weights on a quantum automaton are complex numbers whose modulus square is a real number between 0 and 1. Actually, if real number probabilities can only increase when added, complex numbers can cancel each other and lower their probability (therefore capturing *interference*): $|c_0 + c_2|^2$ need not be bigger than $|c_0|^2$ or $|c_0|^2$.
- Quantum states can be superposed, i.e. a physical system can be in more than one basic state simultaneously.

Notes.
There are several introductory textbooks on the mathematical background stuff discussed in this lecture. I would recommend Paul Halmos' very well written introductions to set theory [2] and to modern logic from an algebraic perspective [3]. Davey and Priestley textbook [1] on ordered structures is recommended for the second topic in the summary. For a very pleasant and solid, although not elementary, reading on algebraic structures I can't but recommend *the* book [4].

References

[1] D. A. Davey and H. A. Priestley. *Introduction to lattices and order (Second Edition)*. Cambridge University Press, 2002.

[2] P. Halmos. *Naive Set Theory*. Springer (Undergraduate texts in Mathematics), 1974.

[3] P. Halmos and S. Givant. *Logic as Algebra*. The Mathematical Association of America (Dolciani Mathematical Expositions, 21), 1998.

[4] S. Mac Lane and G. Birkhoff. *Algebra (Third Edition)*. Cambridge University Press, 1988.

[5] A. Tarski. A lattice–theoretic fixpoint theorem and its applications. *Pacific Journal of Mathematics*, 5:285–309, 1955.