Polishness of some topologies related to automata

Olivier Finkel

Joint work with Olivier Carton and Dominique Lecomte

IMJ-PRG
Institut de Mathématiques de Jussieu-Paris Rive Gauche

IRIF
INSTITUT DE RECHERCHE EN INFORMATIQUE FONDAMENTALE

JAF 2018 – Florence

Outline

The Cantor space of infinite words

The set $\Sigma^{\mathbb{N}}$ of infinite words over some finite alphabet Σ can be endowed with the distance d defined for words $x = x_0 x_1 x_2 \cdots$ and $y = y_0 y_1 y_2 \cdots$ by

$$d(x,y) = \begin{cases} 0 & \text{if } x = y \\ 2^{-\min\{i \,:\, x_i \neq y_i\}} & \text{otherwise} \end{cases}$$

Two words x and y are close if they coincide on a long prefix.
A base of the topology is the family of basic clopen sets of the form $N_w = w\Sigma^{\mathbb{N}} = \{x : x_0 \cdots x_{|w|-1} = w\}$.

Polish spaces

A topological space is called a Polish space if it is a separable completely metrizable topological space, that is

- It has a dense countable subset
- Its topology can be defined by a distance which makes it complete

These spaces are intensively studied in Descriptive Set Theory.
Examples:

- The real line $\mathbb{R}$ and $\mathbb{R}^k$ for $k \geq 2$,
- Intervals $[0; 1]$ and $(0; 1)$ (not with the usual distance for the latter one),
- The Cantor space $\Sigma^{\mathbb{N}}$ for each finite alphabet Σ,
- The Baire space $\mathbb{N}^{\mathbb{N}}$.

Borel hierarchy

$$
\begin{tikzcd}
 & \mathbf{\Sigma}^0_1 & & \mathbf{\Sigma}^0_2 & & \cdots & & \mathbf{\Sigma}^0_\alpha & & \cdots \\
\mathbf{\Delta}^0_1 \arrow[ur, dash] \arrow[dr, dash] & & \mathbf{\Delta}^0_2 \arrow[ul, dash] \arrow[ur, dash] \arrow[dl, dash] \arrow[dr, dash] & & \mathbf{\Delta}^0_3 \arrow[ul, dash] \arrow[dl, dash] & \cdots & \mathbf{\Delta}^0_\alpha \arrow[ur, dash] \arrow[dr, dash] & & \mathbf{\Delta}^0_{\alpha+1} \arrow[ul, dash] \arrow[dl, dash] & \cdots \\
 & \mathbf{\Pi}^0_1 & & \mathbf{\Pi}^0_2 & & \cdots & & \mathbf{\Pi}^0_\alpha & & \cdots
\end{tikzcd}
$$

where

- $\mathbf{\Delta}^0_1$ is the family of clopen (closed and open) sets
- $\mathbf{\Sigma}^0_1$ is the family of open sets
- $\mathbf{\Pi}^0_1$ is the family of closed sets
- $\mathbf{\Sigma}^0_2$ is the family of F_σ sets
- $\mathbf{\Pi}^0_2$ is the family of G_δ sets

Changing the topology

It is sometimes needed to consider other topologies by changing the base of open sets:

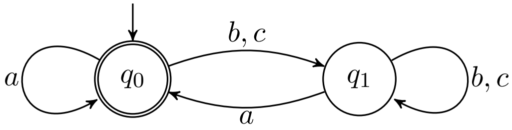
- the alphabetic topology:
  $wA^{\mathbb{N}}$ for some word $w \in \Sigma^*$ and some alphabet $A \subseteq \Sigma$
- the strictly alphabetic topology: $wA^{\mathbb{N}} \setminus \bigcup_{B \subsetneq A} wB^{\mathbb{N}}$ for some word $w \in \Sigma^*$ and some alphabet $A \subseteq \Sigma$
- the automatic topology: all closed (for the Cantor topology) ω-regular sets.
- the Büchi topology: all ω-regular sets.

All these topologies, considered by S. Schwartz and L. Staiger in 2010, are finer than the Cantor topology because the cylinders are always included in the base of open sets.
In the classical Cantor topology, the set $P = (0^*1)^{\mathbb{N}}$ is a complete $\mathbf{\Pi}^0_2$ set. In the Büchi topology, it becomes an open set.

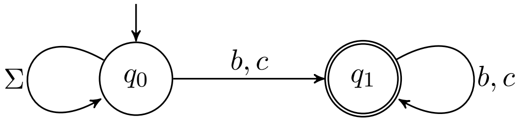
Regular sets

A subset $X \subseteq \Sigma^{\mathbb{N}}$ is ω-regular if it is the set of infinite words accepted by a Büchi automaton, or equivalently, accepted by a deterministic Muller automaton.

Example: Deterministic Büchi automaton accepting the set $(\Sigma^* a)^{\mathbb{N}}$ of words having infinitely many a.



Non-deterministic Büchi automaton accepting the complement $\Sigma^*(b+c)^{\mathbb{N}}$



First attempt

A Büchi automaton separates two infinite words x and y if it accepts one of the two and rejects the other one.
Let define the distance d_B by

$$d_B(x,y) = \begin{cases} 0 & \text{if } x = y \\ 2^{-\min\{|\mathcal{B}| \,:\, \mathcal{B} \text{ separates } x \text{ and } y\}} & \text{otherwise} \end{cases}$$

Two words x and y are close if a big automaton is needed to separate them.
The space $\Sigma^{\mathbb{N}}$ endowed with the distance d_B is not complete.
The sequence $(a^{n!}b^{\mathbb{N}})_{n\geq 0}$ is a Cauchy sequence but it does not converge.
The topology induced by the distance d on $\Sigma^{\mathbb{N}}$ is the Büchi topology.

Main Results

Theorem

All the four topologies introduced before are Polish.

The main tool: Choquet games

The Choquet games is played by two players 1 and 2 in a topological space. At each turn i,

- Player 1 chooses an open set $U_i \subseteq V_{i-1}$ and a point $x_i \in U_i$,
- Player 2 chooses an open set $V_i \subseteq U_i$ such that $x_i \in V_i$.

Player 2 wins the play if $\bigcap_{i \geq 0} V_i \neq \varnothing$.
The topological space is strong Choquet if player 2 wins the game (that is, has a winning strategy).

Theorem (Choquet)

A nonempty, second countable (countable basis) topological space is Polish if and only if it is T1 (singleton sets are closed), regular (for each open neighborhood U, there is a open neighborhood V such that $\overline{V} \subseteq U$) and strong Choquet.

The Büchi topology

Theorem (Choquet)

A nonempty, second countable (countable basis) topological space is Polish if and only if it is T1 (singleton sets are closed), regular (for each open neighborhood U, there is a open neighborhood V such that $\overline{V} \subseteq U$) and strong Choquet.

The Büchi topology on a space $\Sigma^{\mathbb{N}}$ is :

- second countable (countable basis): A countable basis is constituted by the ω-regular sets.
- T1 (singleton sets are closed): The Büchi topology is finer than the usual Cantor topology,
- zero-dimensional: there is a basis of clopen sets (the ω-regular sets are closed under complements). This implies that the space $(\Sigma^{\mathbb{N}}, \tau_B)$ is regular: for each open neighborhood U, there is a open neighborhood V such that $\overline{V} \subseteq U$.

The Büchi topology is strong Choquet

In the spaces of the form $\Sigma^{\mathbb{N}}$, where Σ is a finite set with at least two elements, we consider a topology τ_Σ on $\Sigma^{\mathbb{N}}$, and a basis $\mathbb{B}_\Sigma$ for τ_Σ. We consider the following properties of the family $(\tau_\Sigma, \mathbb{B}_\Sigma)_\Sigma$:

(P1) $\mathbb{B}_\Sigma$ contains the usual basic clopen sets $N_w = w\Sigma^{\mathbb{N}}$,

(P2) $\mathbb{B}_\Sigma$ is closed under finite unions and intersections,

(P3) $\mathbb{B}_\Sigma$ is closed under projections, in the sense that if Γ is a finite set with at least two elements and $L \in \mathbb{B}_{\Sigma\times\Gamma}$, then $\pi_0[L] \in \mathbb{B}_\Sigma$,

(P4) for each $L \in \mathbb{B}_\Sigma$ there is a closed subset C of $\Sigma^{\mathbb{N}} \times \mathbb{P}_\infty$, where $\mathbb{P}_\infty = (0^\star \cdot 1)^{\mathbb{N}}$, (i.e. C is the intersection of a closed subset of the Cantor space $\Sigma^{\mathbb{N}} \times 2^{\mathbb{N}}$ with $\Sigma^{\mathbb{N}} \times \mathbb{P}_\infty$) which is in $\mathbb{B}_{\Sigma\times 2}$, and such that $L = \pi_0[C]$.

Theorem

Assume that the family $(\tau_\Sigma, \mathbb{B}_\Sigma)_\Sigma$ satisfies the properties (P1)-(P4). Then the topologies τ_Σ are strong Choquet.

Consequences

Let S be the set $\Sigma^{\mathbb{N}}$ with the Büchi topology.
Let Ult be the set of ultimately periodic words.

$$\mathrm{Ult} = \{uv^{\mathbb{N}} = uvvv \cdots \ : \ u, v \in \Sigma^*\}$$

Each ω-regular set contains an ultimately periodic word since each regular ω-language is of the form

$$L = \bigcup_{1 \leq j \leq n} U_j \cdot V_j^{\mathbb{N}}$$

for some regular finitary languages U_j and V_j.

Thus Ult is the set of isolated points in S and it is dense in S.
A set U is dense in S if and only if it contains Ult.

Then S is a Baire space because any intersection (even non-countable) of dense open sets is still dense.

Consequences

The disjoint union

$$S = P \uplus \mathrm{Ult}$$

is the Cantor-Bendixson decomposition, that is, P is perfect (closed without isolated point). Furthermore, P, as a Polish space is isomorphic to the Baire space $\mathbb{N}^{\mathbb{N}}$.

(We prove that every compact subset of (P, τ_B) has empty interior, which is sufficient since (P, τ_B) is a zero-dimensional Polish space)

Many other consequences follow from the rich theory of Polish spaces, for instance about the stratification of the Borel sets in a strict hierarchy of length ω_1.

The Büchi topology and the Cantor topology have the same Borel sets, but the level of a set in the two Borel hierarchies may be different.

Topologies on a space of trees

There is also a natural topology on the set T_Σ^ω.

Let t and s be two distinct infinite trees in T_Σ^ω. Then the distance between t and s is $\frac{1}{2^n}$ where n is the smallest integer such that $t(x) \neq s(x)$ for some word $x \in \{l, r\}^\star$ of length n.

The open sets are then in the form $T_0 \cdot T_\Sigma^\omega$ where T_0 is a set of finite labelled trees.

The set T_Σ^ω, equipped with this topology, is homeomorphic to the Cantor set 2^ω, hence also to the topological spaces Σ^ω, where Σ is a finite alphabet having at least two letters.

The Büchi topology

The notion of Büchi automaton has been extended to the case of a Büchi tree automaton reading infinite binary trees whose nodes are labelled by letters of a finite alphabet.

Muller tree automata are stronger and accept the whole class of regular tree languages, those definable in monadic second order of two successors S2S.

The Büchi and the Muller topologies are not Polish

Theorem

Let Σ be a finite alphabet having at least two letters.

1. *The Büchi topology on T_Σ^ω is strong Choquet, but it is not regular (and hence not zero-dimensional) and not metrizable.*
2. *The Muller topology on T_Σ^ω is zero-dimensional, regular and metrizable, but it is not strong Choquet.*

In particular, the Büchi topology and the Muller topology on T_Σ^ω are not Polish.

The Büchi topology is not metrizable

Theorem

Let Σ be a finite set with at least two elements. Then the Büchi topology on T_Σ^ω is not metrizable and thus not Polish.

In a metrizable topological space, every closed set is a countable intersection of open sets.

The set $\mathcal{L}$ of infinite trees in T_Σ^ω, where $\Sigma = \{0, 1\}$, having at least one path in the ω-language $\mathcal{R} = (0^\star \cdot 1)^\mathbb{N}$ is $\mathbf{\Sigma}_1^1$-complete for the usual topology, and it is open for the Büchi topology (it is accepted by a Büchi tree automaton).

Its complement $\mathcal{L}^-$ is the set of trees in T_Σ^ω having all their paths in $\{0, 1\}^\mathbb{N} \setminus (0^\star \cdot 1)^\mathbb{N}$; it is $\mathbf{\Pi}_1^1$-complete for the usual topology and closed for the Büchi topology.

The Büchi topology is not metrizable

Every tree language accepted by a Büchi tree automaton is a $\mathbf{\Sigma}^1_1$-set (for the usual Cantor topology). Moreover every open set for the Büchi topology is a countable union of basic open sets, and thus a $\mathbf{\Sigma}^1_1$-set for the usual topology (the class $\mathbf{\Sigma}^1_1$ is closed under countable union).

Assume now that $\mathcal{L}^-$ is a countable intersection of open sets for the Büchi topology. Then it is a countable intersection of $\mathbf{\Sigma}^1_1$-sets for the usual topology. But the class $\mathbf{\Sigma}^1_1$ in a Polish space is closed under countable intersections.

Thus $\mathcal{L}^-$ would be also a $\mathbf{\Sigma}^1_1$-set for the usual topology. But $\mathcal{L}^-$ is $\mathbf{\Pi}^1_1$-complete and thus in $\mathbf{\Pi}^1_1 \setminus \mathbf{\Sigma}^1_1$, $\rightarrow$ a contradiction.

References

O. Carton, O. Finkel, and D. Lecomte. Polishness of Some Topologies Related to Automata. Proceedings of CSL 2017.

O. Carton, O. Finkel, and D. Lecomte. Polishness of Some Topologies Related to Automata (Extended version). 2017. Preprint, available from ArXiv:1710.04002.

S. Hoffmann and L. Staiger. Subword metrics for infinite words. Proceedings of CIAA 2015.

S. Hoffmann, S. Schwarz, and L. Staiger. Shift-invariant topologies for the Cantor space X^ω. Theoretical Computer Science, 679:145-161, 2017.

References

A. S. Kechris. Classical descriptive set theory. Springer-Verlag, New York, 1995.

Y. N. Moschovakis. Descriptive set theory, volume 155 of Mathematical Surveys and Monograph. American Mathematical Society, Providence, RI, second edition, 2009.

D. Perrin and J.-E. Pin. Infinite words, automata, semigroups, logic and games, volume 141 of Pure and Applied Mathematics. Elsevier, 2004.

S. Schwarz and L. Staiger. Topologies refining the Cantor topology on X^ω. Proceedings of TCS 2010.

L. Staiger. ω-languages. In Handbook of formal languages, Vol. 3, pages 339–387. Springer, Berlin, 1997.

Open questions

- Can we have an explicit description of a complete distance inducing the Büchi topology on $\Sigma^{\mathbb{N}}$?
- Our results lead to applications of the polishness of the Büchi topology and of the other three topologies on a space of words, using the many results of the theory of Polish spaces in Descriptive Set Theory.

THANK YOU !

The Wadge Hierarchy on the four Polish spaces

- the alphabetic topology: $wA^{\mathbb{N}}$ for some word $w \in \Sigma^*$ and some alphabet $A \subseteq \Sigma$
- the strictly alphabetic topology: $wA^{\mathbb{N}} \setminus \bigcup_{B \subsetneq A} wB^{\mathbb{N}}$ for some word $w \in \Sigma^*$ and some alphabet $A \subseteq \Sigma$
- the automatic topology: all closed (for the Cantor topology) ω-regular sets.
- the Büchi topology: all ω-regular sets.

Theorem

The Wadge Hierarchy of each of these Polish spaces is equal to the Wadge Hierarchy of the Baire space.