ssdnm
środowiskowe
studia doktoranckie
z nauk matematycznych

Filip Mazowiecki

Uniwersytet Warszawski

Continuous reductions of regular languages

Praca semestralna nr 1

(semestr zimowy 2011/12)

Opiekun pracy: Filip Murlak

KAPITAŁ LUDZKI
NARODOWA STRATEGIA SPÓJNOŚCI

UNIA EUROPEJSKA
EUROPEJSKI
FUNDUSZ SPOŁECZNY

Continuous reductions of regular languages

Filip Mazowiecki

University of Warsaw

Abstract. By embedding into the space of infinite words one can define a metric on the space of finite words. We show that under weak assumptions only two essentially different metrics can be obtained this way. For both of them we describe hierarchies of regular languages induced by the existence of continuous, uniformly continuous and Lipschitz reductions, thus creating an analog of Wagner's result on ω-regular languages.

1 Introduction

In [10] Wagner described a hierarchy of ω-regular languages (see [9] for ω-regular languages), which corresponds to the existence of continuous reductions between languages, i.e., continuous functions $f : A^\omega \to B^\omega$ such that $f^{-1}(M) = L$ for $L \subseteq A^\omega$, $M \subseteq B^\omega$. We would like to develop analogous theory for regular languages of finite words. While the space of infinite words has a natural metric definition (that of the Cantor space), the space of finite words does not have a canonical metric. We consider metrics defined by embeddings into the space of infinite words that preserve prefix convergence. This last condition guarantees that the natural connection between finite and infinite words is not lost in the process of embedding. The space of finite words with a given metric can be considered a uniform space. We show that under mild assumptions embeddings give only two non-isomorphic uniform structures. In both cases the space is not compact, so the notions of continuity and uniform continuity diverge hierarchies for continuous and uniformly continuous reductions need to be described separately.

The first embedding gives a discrete metric and the hierarchies of all languages are simple. For the second embedding we define canonical automata, similar to those defined by Wagner in [10], and show that every regular language is equivalent to one of the languages recognized by them. This is stated in the main theorems of this paper – Theorem 6 and Theorem 7. In this case we describe only the hierarchy for regular languages and show an example of a non-regular language that is not equivalent to any of the regular languages with respect to any reduction. In the end we examine the hierarchy of Borel languages in Σ^ω and Σ^* induced by Lipschitz reductions.

The basic tool we use is a game characterization of the existence of reductions. For every kind of reduction – continuous, uniformly continuous and Lipschitz – we define an appropriate game. We also use basic theorems about the Cantor space from [6].

2 Embeddings

Throughout the paper Σ and Γ stand for finite alphabets, each containing letters 0 and 1. For any injective function $f : X \to Y$ if Y is a metric space one can define a metric d_f on X as $d_f(x_1, x_2) = d(f(x_1), f(x_2))$.

Definition 1. *We say that f guarantees singleton equivalence if for every $x, y \in \Sigma^*$ there exists a continuous function $g : f(\Sigma^*) \to f(\Sigma^*)$ such that $g^{-1}(\{f(x)\}) = \{f(y)\}$.*

The proposition below describes how the induced topology looks.

Proposition 1. *If f guarantees singleton equivalence then all points in Σ^* are isolated or none of them is isolated.*

Proof. Suppose that $x \in \Sigma^*$ is an isolated point and that $y \in \Sigma^*$ is a not isolated point. Then the set $\{f(x)\}$ is open and the set $\{f(y)\}$ is not open in $f(\Sigma^*)$. Then there is no continuous function $g : f(\Sigma^*) \to f(\Sigma^*)$ such that $g^{-1}(\{f(x)\}) = \{f(y)\}$. □

Definition 2. *For $x, x_n \in \Gamma^\omega$ we write $x_n \to_{pref} x$ if $\forall_k \exists_{n_0} \forall_{n>n_0}\ x_n|_k = x|_k$. We say that the embedding $f : \Sigma^* \to \Gamma^\omega$ preserves prefix convergence if $f(x_n) \to_{pref} x \implies f(x_n) \to x$.*

Notice that if f preserves prefix convergence then by König's Lemma $f(x_n) \to_{pref} x \iff f(x_n) \to x$. We are now ready to define the embeddings of Σ^*.

Definition 3. *Let $\Sigma \subseteq \Gamma$. We call f an embedding of Σ^* in Γ^ω if:*

1. *f is an injection;*
2. *f guarantees singleton equivalence; and*
3. *f preserves prefix convergence.*

Notice that the last condition implies that the space $f(\Sigma^*)$ is dense in the space Σ^ω. From now on we will write Σ^* instead of $f(\Sigma^*)$ if f is clear from the context. By Proposition 1 there are only two types of embeddings: the ones where all points are isolated; and the ones where no points are isolated. Below we define examples for both types.

Definition 4. *Let $0, 1 \in \Sigma$. We define $\Sigma_\$ = \Sigma \cup \{\$\}$, where $\$ \notin \Sigma$.*

1. *We define the* discrete embedding $f_1 : \Sigma^* \to \Sigma_\$^\omega$ *as $f_1(x) = x\$^\omega$;*
2. *We define the* rational embedding $f_2 : \Sigma^* \to \Sigma^\omega$ *as $f_2(x) = x10^\omega$.*

Proposition 2. *Functions f_1, f_2 are embeddings of Σ^* in the sense of Definition 3. Moreover $f_1(\Sigma^*)$ has all points isolated and $f_2(\Sigma^*)$ has no isolated point.*

Proof. We start with the function f_1. It is obvious that f_1 is an injection. Assume that we have $x \in \Sigma^\omega$ and a sequence x_n such that $x_n \to_{pref} x$. By definition $\forall_k \exists_{n_0} \forall_{n>n_0}\ x_n|_k = x|_k$. This implies that words x and $f_1(x_n)$ do not differ on first k letters so $d(x_n, x) \leq 2^{-k-1}$. Because we can choose any k then $f_1(x_n) \to x$. To prove that f_1 preserves prefix convergence all we need to show is that Σ^* has all points isolated. Take $u \in \Sigma^*$ and let n be the length of u. Then the ball centered at $f_1(u)$ of radius 2^{-n-2} contains only words of the form $f_1(u)\$\Gamma^\omega$. In $f_1(\Sigma^*)$ there is only one point of this form – $f_1(u)$.

Consider now f_2. It is an injection because for every $u \in \Sigma^*$ if we take the image $f(u)$ then we get u again if we remove 10^ω. Notice that if the function would be $x \to x0^\omega$ then it would not be an injection. For example for every n the image of 0^n would be 0^ω. The proof that f_2 preserves prefix convergence is similar as in the case of f_1. We stated that preserving prefix convergence implies that $f_2(\Sigma^*)$ is dense in Σ^ω and by definition we know that $f_2(\Sigma^*) \subseteq \Sigma^\omega$. This proves that in $f_2(\Sigma^\omega)$ no point is isolated. The fact that f_2 guarantees singleton equivalence will be shown in the next chapter after introducing some new definitions (Corollary 1). □

Now we would like to show that each embedding is uniformly isomorphic to one of these two examples. We consider (Σ^*, d_f) a uniform space. We will show that for any two embeddings f_1, f_2 of the same type there exists a uniformly continuous isomorphism between (Σ^*, d_{f_1}) and (Σ^*, d_{f_2}) (see [1] about uniform spaces). The case of discrete metrics is easy.

Theorem 1. *Let f_1, f_2 be Σ^* embeddings such that all points are isolated. Then (Σ^*, d_{f_1}) and (Σ^*, d_{f_2}) uniform spaces are isomorphic.*

Proof. We show that the function $g : \Sigma^* \to \Sigma^*$ defined as $g(f_1(u)) = g(f_2(u))$ is a uniformly continuous isomorphism. One should think about it as the identity function $g(u) = u$. This is of course a bijection so we just need to show that g and g^{-1} are uniformly continuous. Because f_1 and f_2 are arbitrary then all we need to show is that g is uniformly continuous. Suppose to the contrary that $\exists_{k<\omega} \forall_{n<\omega} \exists_{x_n, y_n \in \Sigma^*}\ d_{f_1}(x_n, y_n) < 2^{-n} \implies d_{f_2}(x_n, y_n) \geq 2^{-k}$. Because Γ^ω is compact we can assume that in the metric d_{f_1} the sequences (x_n) and (y_n) converge to a point $z \in \Gamma^\omega$. In (Σ^*, d_{f_1}) all points are isolated so $z \notin \Sigma^*$ and we can assume that elements of sequences (x_n) and (y_n) are different. By König's Lemma we can choose such subsequences (x_{n_k}) and (y_{n_k}) and words $x, y \in \Gamma^\omega$ that $x_{n_k} \to_{pref} x$ and $y_{n_k} \to_{pref} y$. For simplicity we assume that $(x_n) = (x_{n_k})$ and $(y_n) = (y_{n_k})$. But f_1 guarantees prefix convergence so $x = y = z$. Since f_2 also guarantees prefix convergence it is a contradiction to $d_{f_2}(x_n, y_n) \geq 2^{-k}$. □

In the case of rational metric we need a generalized Sierpiński's Theorem from [2] (third proof).

Theorem 2. *(Sierpiński) Let X and Y be T_0 spaces without isolated points and each possessing a countable basis consisting of clopen sets. Let A and B be countable dense subsets of X and Y respectively. Then there is a bijection*

$f : A \to B$ which is a homeomorphism from A onto B. If in addition Y is compact, then f extends uniquely to a relative homeomorphism f^* from X into Y . If X is also compact then f is a homeomorphism of X onto Y.

Theorem 3. *Let f_1, f_2 be embeddings of Σ^* such that no points are isolated. Then the uniform spaces (Σ^*, d_{f_1}) and (Σ^*, d_{f_2}) are isomorphic.*

Proof. The proof is an easy consequence of Sierpiński's Theorem. Let $X = Y = \Sigma^\omega, A = f_1(\Sigma^*)$ and $B = f_2(\Sigma^*)$. By Theorem 2 we get a homeomorphism $f : f_1(\Sigma^*) \to f_2(\Sigma^*)$ that extends to a homeomorphism $f^* : \Sigma^\omega \to \Sigma^\omega$. Since Σ^ω is compact f^* and f^{*-1} are uniformly continuous. Thus f and f^{-1} are also uniformly continuous, so f is the desired isomorphism. □

As a consequence of Theorems 1 and 3 we get that the hierarchies for continuous and uniformly continuous reductions of all languages are either like the hierarchies for the discrete metric or for the rational metric.

3 Games and reductions

We recall the definition of *Wadge reductions*. For two finite alphabets Σ, Γ and two languages $L \subseteq \Sigma^\omega, M \subseteq \Gamma^\omega$, we write $L \leq_W M$ iff there exists a continuous function $f : \Sigma^\omega \to \Gamma^\omega$ such that $f^{-1}(M) = L$. The function f is called a *Wadge reduction (continuous reduction)* between languages L and M. We call L and M Wadge equivalent if $L \leq_W M$ and $M \leq_W L$ and denote it $L \equiv_W M$. A strict inequality means that there exists a reduction only in one direction. If there is no reduction we denote it $L \perp_W M$.

We call f a c-Lipschitz reduction if it is a reduction and it meets Lipschitz condition with a constant c, i.e., for $x, y \in \Sigma^\omega$, $d(f(x), f(y)) \leq c \cdot d(x, y)$. In the case of $c = 2^i$ we write $L \leq_i M$.

Let us show some basic facts about continuous functions on the space Σ^ω.

Proposition 3. *For finite alphabets Σ, Γ and a function $f : \Sigma^\omega \to \Gamma^\omega$*

1. *if f is continuous then it is also uniformly continuous;*
2. *f is uniformly continuous iff $\forall_{k<\omega}\exists_{l<\omega}$ first k letters of $f(x)$ are determined by first l letters of x, i.e., $x|_l = y|_l$ implies $f(x)|_k = f(y)|_k$ for all y;*
3. *f is a 2^i-Lipschitz function iff $\forall_{k<\omega}$ first k letters of $f(x)$ are determined by first $k + i$ letters of x, i.e., $x|_{k+i} = y|_{k+i}$ implies $f(x)|_k = f(y)|_k$.*

Proof. 1. This is the Heine-Cantor Theorem;
2. We write the definition that f is uniformly continuous:

$$\forall_{k>0}\exists_{l>0}\forall_{x,y\in\Sigma^\omega}\ d(x, y) \leq 2^{-l} \implies d(f(x), f(y)) \leq 2^{-k}.$$

   From the definition of the metric on Σ^ω the equivalent condition is obvious;
3. Take $x, y \in \Sigma^\omega$. Let $d(x, y) \leq 2^{-k-i}$. The Lipschitz condition implies that $d(f(x), f(y)) \leq 2^{-k}$. Again from the definition of the metric on Σ^ω the equivalent condition is obvious.

□

We now recall the definition of the *Wadge Game*. For two finite alphabets Σ, Γ and two languages $A \subseteq \Sigma^\omega, B \subseteq \Gamma^\omega$ the game $W(A,B)$ is played by two players – Spoiler and Duplicator. Spoiler uses the alphabet Σ and Duplicator the alphabet Γ. In each turn the players play letters from their alphabets. Spoiler starts. Duplicator has an option to pass but he is obliged to play infinitely often. As a result we obtain two infinite words $x \in \Sigma^\omega$ and $y \in \Gamma^\omega$. Duplicator wins the game iff $x \in A \iff y \in B$. If Duplicator is allowed to pass only i times then we call this game 2^i-Lipschitz and denote it $W_i(A,B)$.

Using Proposition 3 one can prove the following characterizations.

Lemma 1. *(Duparc) For two languages $A \subseteq \Sigma^\omega, B \subseteq \Gamma^\omega$ Duplicator has a winning strategy in $W(A,B)$ iff $A \leq_W B$.*

Proof. $\Longleftarrow$ Let f be the reduction between A and B. By the Proposition 3 we know that for every k there is such an l that first l letters of x determine first k letters of $f(x)$. So Duplicator's strategy is as follows: If Duplicator has to play the k-th letter then he passes until Spoiler plays l letters and plays the letter determined by f. Thus if in the end Spoiler creates an infinite word $x \in \Sigma^\omega$, Duplicator creates the word $f(x)$. By the definition of the reduction this is a winning strategy.

$\Longrightarrow$ We define the reduction $f : \Sigma^\omega \to \Gamma^\omega$. For $x \in \Sigma^\omega$ let $f(x)$ be the word that Duplicator created using his winning strategy if Spoiler created x. By the definition of the game $W(A,B)$ when Spoiler plays letters then Duplicator eventually has to play a letter, because if he passes infinitely many times he loses the game. This implies that f is uniformly continuous. By Duplicator's winning conditions f is the desired reduction. ◻

Lemma 2. *For two languages $A \subseteq \Sigma^\omega, B \subseteq \Gamma^\omega$ Duplicator has a winning strategy in $W_i(A,B)$ iff $A \leq_i B$.*

Proof. The proof is similar to the proof of Lemma 1. This time we need to use facts about Lipschitz functions from the Proposition 3. ◻

Now we would like to define more types of games which correspond to reductions in the spaces Σ^* and Γ^*. Players play letters in turns like in the case of the infinite words spaces. Sometimes they can also choose an extra letter $\$ \notin \Sigma$ (for discrete metric). In the end players get words from the spaces $\Sigma^\omega_\$, \Gamma^\omega_\$$, which do not have to be from Σ^* or Γ^*.

Definition 5. *For $A \subseteq \Sigma^*, B \subseteq \Gamma^*$ we define three types of games depending on the winning conditions:*

1. *$W^*(A,B)$ game. Duplicator wins this game if for the final words $x \in \Sigma^\omega_\$, y \in \Gamma^\omega_\$$*

$$x \in A \iff y \in B \quad or \quad x \notin \Sigma^*.$$

   Duplicator also wins if he passes forever (creating a finite word), provided that $x \notin \Sigma^*$.

2. $W_u^*(A,B)$ *game. It is similar to* $W^*(A,B)$ *game except that Duplicator loses if he passes forever.*
3. $W_i^*(A,B)$ *game. Similar to* W_u^* *except that Duplicator also loses if he passes more then* i *times (altogether).*

We defined reductions in the case of infinite words in Section 3. For finite words this definition is similar. For two languages $A \subseteq \Sigma^*, B \subseteq \Gamma^*$ we call $f : \Sigma^* \to \Gamma^*$ a reduction if $f^{-1}(B) = A$. This reduction is continuous, uniformly continuous or 2^i-Lipschitz if the corresponding function is respectively continuous, uniformly continuous or 2^i-Lipschitz. Continuous functions do not have to be uniformly continuous so we will use the notation $A \leq_u B$ for uniformly continuous reductions.

Example 1. Let $0, 1 \in \Sigma$ and let Σ^* have the metric defined with the discrete embedding. We define the function $f : \Sigma^* \to \Sigma^*$ as

$$f(x) = \begin{cases} 0^n & \text{for } x = 0^{2n}, \\ 1^n & \text{for } x = 0^{2n-1}, \\ 01 & \text{otherwise.} \end{cases}$$

This function is continuous but not uniformly continuous. It turns out that in these spaces continuous functions do not have to be uniformly continuous. It is possible because the embedding of a countable set Σ^*, whose closure contains Σ^ω, is not a compact space.

Lemma 3. *For two languages* $A \subseteq \Sigma^*, B \subseteq \Gamma^*$*:*

1. *Duplicator has a winning strategy in* $W^*(A,B)$ *iff* $A \leq_W B$*;*
2. *Duplicator has a winning strategy in* $W_u^*(A,B)$ *iff* $A \leq_u B$*;*
3. *Duplicator has a winning strategy in* $W_i^*(A,B)$ *iff* $A \leq_i B$*.*

Proof. The proof is similar to the proofs of Lemmas 1, 2. □

Corollary 1. *The* f_2 *embedding from Proposition* 2 *guarantees singleton equivalence.*

Proof. Set two languages A, B over a finite alphabet Σ, such that both contain only one word. We identify words from Σ^* with words from $f_2(\Sigma^*)$. We want to show Duplicator's winning strategy in the game $W^*(A,B)$. As long as Spoiler is playing letters creating the word from his language Duplicator responses playing letters creating the word from his language. If Spoiler plays such a letter that he cannot create his word anymore then Duplicator plays analogously. The embedding f_2 is defined in such a way that Duplicator can always finish his word creating a word from Σ^*. He just has to play letters from the sequence 10^ω. Thus this is a winning strategy. □

4 The discrete embedding

In the case of the discrete embedding the metric is discrete. That is why the hierarchy of languages with respect to continuous reductions is very simple.

Theorem 4. *For the discrete embedding the hierarchy for continuous reductions is described by the figure:*

$$\begin{array}{ccc} \Sigma^* & & \\ & \searrow & \\ & & L \\ & \nearrow & \\ \emptyset & & \end{array}$$

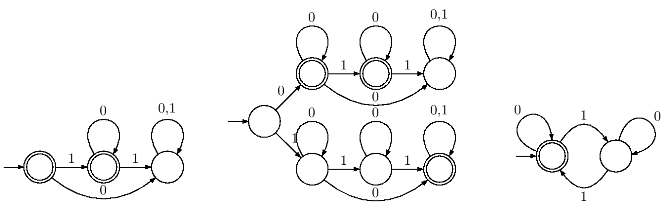
where $\rightarrow$ denotes $<_W$ and L is any language different from $\emptyset$ or Σ^*.

Proof. The empty language and the language of all words are always on the bottom of the hierarchy for any kind of reductions. It is because the preimage of the empty set is always the empty set and the preimage of the set of values is the whole domain. All other languages are equivalent because the metric is discrete, which makes every function continuous. □

We now give the key definition for the hierarchy induced by uniformly continuous reductions.

Definition 6. *For $k \in \mathbb{N}$ we say that a language $L \subseteq \Sigma^*$ is at most k-level if for every word $x \in \Sigma^*$ if $|x| \geq k$ then either $x\Sigma^* \subseteq L$ or $x\Sigma^* \subseteq L^{\complement}$. Language L is k-level if it is at most k-level and not at most $(k-1)$-level. Languages for which such a finite number does not exist are ω-level.*

For example $\emptyset$ and Σ^* are 0-level languages. With these classes of languages we can describe the hierarchy with respect to uniformly continuous reductions.



(a) C_2^* (b) E_2^* (c) E_ω^*

Fig. 1. Sample automata

Theorem 5. *For the discrete embedding the hierarchy for uniformly continuous reductions is shown below:*

$$\begin{array}{ccc} \Sigma^* & & \\ & \searrow & \\ & & L_f \rightarrow L_\omega \\ & \nearrow & \\ \emptyset & & \end{array}$$

where $\rightarrow$ *denotes* $<_u$*,* L_f *are all finite level languages and* L_ω *are* ω*-level languages.*

Proof. Set two languages $A, B \subseteq \Sigma^*$ both different from $\emptyset$ and Σ^*. Suppose language A is finite level. We show Duplicator's winning strategy in the game $W_u^*(A, B)$. Let k be the level of the language A. Duplicator strategy is to pass first k rounds. After k rounds Spoiler creates a finite word of length k. By definition of A words with a set prefix of the length k either all are in A or all are not in A. Then Duplicator can create a word that belongs to B iff Spoiler's word belongs to A. This shows that all finite level languages are equivalent and that $L_f \leq_u L_\omega$. We need to prove that $L_\omega \not\leq_u L_f$. Let A be an ω-level language and B a finite level language. We show Spoiler's winning strategy in the game $W(A, B)$. By König's Lemma there exists an infinite word $x \in \Sigma^\omega$ such that for any $n < \omega$ there exist finite words $u, v \in \Sigma^*$ such that $x|_n u \in A$ and $x|_n v \notin A$. Let k be the level of language B. Spoiler strategy is to play letters from the word x. If Duplicator wants to win then he cannot pass infinitely many times in a row. Thus after some rounds he has to play k letters creating a finite word. By definition words with a set prefix of length k are all in B or all are not in B. Then Spoiler has an easy winning strategy. □

5 The rational embedding

Throughout this section we assume for simplicity that $\Sigma = \{0, 1\}$. We are going to work mainly with regular languages i.e., languages recognized by finite automata. We assume that the reader is familiar with the notion of regular languages and finite automata (see e.g. [5]). A finite automaton can also be used to recognize images of languages under the rational embedding. A word of the form $x10^\omega$ is accepted by the automaton if the automaton after reading x is in an accepting state. If the word is not of the form $x10^\omega$ it is rejected.

We begin with the definition of canonical automata corresponding the levels of the hierarchies.

Definition 7. *For every natural number* $n > 0$ *we define* C_n^* *and* D_n^* *as the automaton with* $n+1$ *states* $\{0, .., n\}$*, and the transition function:*

$$i \xrightarrow{0} i \text{ for } i > 0, \qquad j \xrightarrow{1} j+1 \text{ for } j < n, \qquad 0 \xrightarrow{0} n, \qquad n \xrightarrow{1} n.$$

The initial state is 0. The accepting states are $0, 1, 3, 5, \ldots$ *in* C_n^* *and* $2, 4, 6, \ldots$ *in* D_n^**.*

Automaton E_n^* *has* $2n+3$ *states. From the initial state after reading 0 the automaton moves to the initial state of the copy of* C_n^* *and after reading 1 it moves to the copy of* D_n^**. The initial state is not accepting.*

Finally we define E_ω^* *as an automaton with two states, 0 and 1. This automaton after reading 0 stays in its current state and after reading 1 it moves to the other state. The initial and accepting state is 0.*

These automata correspond to automata from Wagner's hierarchy (see [7], [10]). To distinguish one from another we use the * symbol. The construction of E^*_ω automaton is similar to C_ω from Wagner's hierarchy. But the letter E is more appropriate because a dual incomparable language does not exist. Note that the automaton C^*_1 recognizes Σ^* and D^*_1 recognizes $\emptyset$. Some more example automata are on Figure 1.

In the proposition below we show that these automata form a strict hierarchy.

Proposition 4. *For any $0 < n < m < \omega$, the following holds:*

$$\begin{array}{ccccccc} C^*_n & & & & C^*_m & & \\ & \searrow & & \nearrow & & \searrow & \\ & & E^*_n & & & & E^*_\omega \\ & \nearrow & & \searrow & & \nearrow & \\ D^*_n & & & & D^*_m & & \end{array}$$

where $\rightarrow$ denotes $<_W$.

Proof. The inequalities $C^*_n, D^*_n \leq_W E^*_n \leq_W C_m, D_m$ are obvious because automata on the left side are contained in automata on the right side. Thus we can create a winning strategy for Duplicator in appropriate games. We prove that the inequalities are strict by showing $C^*_n \perp_W D^*_n$. We show that Spoiler has a winning strategy in the game $W^*(C^*_n, D^*_n)$. Spoiler's first move is to play 1. Then he plays 0 until Duplicator makes a move. We assume that Duplicator plays 1 (the case when Duplicator first letter is 0 is easier). Then Spoiler plays 0. Then his strategy is to copy Duplicator's moves. If Duplicator played 1 in the last state then Spoiler plays 10^ω to the end. If Duplicator passes then Spoiler plays 0. Thus Spoiler creates a word of the form $u10^\omega$. By the construction of C^*_n and D^*_n Spoiler wins the game because he changed a state iff Duplicator changed a state.

It remains to show that $C^*_n <_W E^*_\omega$. To prove the inequality $C^*_n \leq_W E^*_\omega$ we consider the game $W^*(C^*_n, E_\omega)$ and show that Duplicator has a winning strategy. The automaton E^*_ω is constructed in such a way that Duplicator can always change a state from accepting to a not accepting and reversely. Spoiler on the C^*_n automaton can change states like this only $n-1$ times. Thus Duplicator has a winning strategy. Because this inequality holds for all n then it is a strict inequality. □

The last thing to show is that these languages are sufficient to describe the hierarchy of regular languages. Abusing the notation of reduction we will write that automata are continuously or uniformly continuously equivalent, meaning that the languages recognized by those automata are equivalent. Definitions $8, 9, 10$ and Proposition 5 are needed only for the proof of Theorem 6, but they are not needed to understand Theorem 6 or the rest of this paper.

Definition 8. *Let $\mathcal{A} = (Q, \Sigma, \delta, q_0, F)$ be a deterministic automaton. We define a function $st_n : Q \to P(Q)$ as $st_n(q) = \{p \in Q : \exists_{x\in\Sigma^*} \delta(q, 10^n x) = p\}$.*

We define the function $st : Q \to P(Q)$ as $st(q) = \bigcap_n st_n(q)$. We also define a directed graph $G_\mathcal{A} = (V_\mathcal{A}, E_\mathcal{A})$ as $V_\mathcal{A} = Q$ and $(p, q) \in E_\mathcal{A}$ iff $q \in st(p)$.

Definition 9. *Let $\mathcal{A}$ be a deterministic automaton. In $G_{\mathcal{A}}$ we define an automaton path as regular path with an additional condition that two vertices (states) are connected iff one of them is accepting and the other one is not accepting. We allow repeating the states in a path. We call an automaton path the main path if it is the longest automaton path in $G_{\mathcal{A}}$. If such a path does not exist then by König's Lemma there exists an infinite path. Then we say that this infinite path is the main path.*

Definition 10. *Let Reg be a family of deterministic automata and Can be the family of canonical automata. We define the function Type : Reg → Kan by:*

$$Type(\mathcal{A}) = \begin{cases} E^*_\omega & \text{if in } G_{\mathcal{A}} \text{ the main path is infinite,} \\ C^*_n & \text{if all main paths start from an accepting state and are } n-1 \text{ long,} \\ D^*_n & \text{if all main paths start from a not accepting state and are } n-1 \text{ long,} \\ E^*_n & \text{if there are main paths starting both from an accepting state} \\ & \text{and a not accepting state and they are } n-1 \text{ long.} \end{cases}$$

Proposition 5. *Let $\mathcal{A}$ be a deterministic automata with the alphabet Σ and such that $Type(\mathcal{A}) = E^*_n$. Suppose there is such an infinite word $x \in \Sigma^\omega$ that for all $k < \omega$ the automata $\mathcal{A}$ after reading $x|_k$ is in a state from which the beginning of main paths starting from both accepting and not accepting state are available. Then $x \notin \Sigma^*$.*

Proof. Suppose $x \in \Sigma^*$. Then $x = u10^\omega$ for some finite word u and $x \in L(\mathcal{A})$ or $x \notin L(\mathcal{A})$. Because $Type(\mathcal{A}) = E^*_n$ then one could extend the main path in $G_{\mathcal{A}}$ – a contradiction. □

Theorem 6. *For any automaton there exists a continuously equivalent canonical automaton.*

Proof. We prove that $Type(\mathcal{A}) \equiv_W \mathcal{A}$.

(1) $Type(\mathcal{A}) \leq_W \mathcal{A}$. First suppose $Type(\mathcal{A}) \neq E^*_\omega$. We show by induction (over n) Duplicator's strategy in the game $W^*(Type(\mathcal{A}), \mathcal{A})$. For C^*_1, D^*_1 it is obvious. In C^*_n, D^*_n Spoiler is on the begging of his main path. In E^*_n Spoiler after his first move has to choose the beginning of one of his main paths. Duplicator can choose his main path such that the beginning of this path is going to be an accepting state iff Spoiler's beginning is an accepting state. Thus the strategy in case of E^*_n can be further copied from the strategies for C^*_n, D^*_n. As long as Spoiler plays next letters from the word 10^ω Duplicator can do the same. If they finish like this then Duplicator wins the game. If Spoiler chooses eventually to play 1 then his automaton changes the state. Duplicator can move to the next state on his main path and by induction he has a winning strategy there.

Now suppose $Type(\mathcal{A}) = E^*_\omega$. Then Duplicator has an infinite main path. Thus if Spoiler changes the state from an accepting to a not accepting then Duplicator can do the same and then play letters from the word 10^ω. If Spoiler changes finitely many times the states like this then this is a winning strategy. Otherwise Spoiler creates a word not from Σ^* so by definition he loses the game.

(2) $\mathcal{A} \leq_W Type(\mathcal{A})$. First suppose $Type(\mathcal{A}) \neq E^*_\omega$. We show by induction (over the length of the main path) Duplicator's strategy in the game $W^*(\mathcal{A}, Type(\mathcal{A}))$. The case when the main paths are 0 long are obvious – it corresponds to $\emptyset$ and Σ^*. In other cases the strategy is similar as in (1). Duplicator passes until Spoiler finds himself on the beginning of a main path or all automaton paths he can reach are shorter from the main path. If non of these two cases happen then by Proposition 5 Spoiler creates a word not in Σ^* so he loses the game. In the second case we just use the induction assumption and we are done. In the first case Duplicator can move to the beginning of his main path which starts from the beginning state iff Spoiler's main path starts from the main path. Then until Spoiler plays letters from the word 10^ω Duplicator can do the same. If they finish like this then Duplicator wins the game. If Spoiler eventually plays 1 then his path will be shorter and by induction Duplicator has a winning strategy.

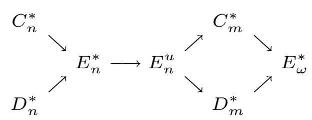
The case if $Type(\mathcal{A}) = E^*_\omega$ is similar as in (1). □

Now we would like to define the automata that are important for the hierarchy with respect to uniformly continuous reductions.

Definition 11. *For every natural $n > 0$ we define an automaton E^u_n. When it is in its initial state then after reading 1 it remains in it, and after reading 0 it moves to a copy of E^*_n. The initial state is accepting.*

In the proposition below we show that these new automata form new levels in the hierarchy.

Proposition 6. *For any $0 < n < m < \omega$, the following holds:*



where $\rightarrow$ denotes $<_u$.

Proof. The reductions in the proof of Proposition 4 were actually uniformly continuous. So all we have to prove is that $E^u_n \not\leq_u E^*_n$. We show Spoiler's winning strategy in the game $W^*_u(E^u_n, E^*_n)$. Spoiler's strategy is to play the letter 1 until Duplicator makes a move. If Duplicator passes forever then Spoiler wins the game. If he eventually chooses a letter then he will be in a copy of C^*_n or D^*_n and then Spoiler of course has a winning strategy. □

The interesting thing here is that because this space is not compact the class E^*_n splits into two classes. In the space Σ^ω the E_n automaton was constructed in such a way that after reading a finite number of letters it was in a copy of D_n or C_n. In the case of Σ^*, automata can read words out of the space. It turns out that automata from Proposition 6 form the whole hierarchy.

Theorem 7. *Every automaton is uniformly continuously equivalent to a canonical automaton $C^*_n, D^*_n, E^*_n, E^*_\omega$ or E^u_n.*

Proof. The games W^* and W^*_u differ only by the rule whether Duplicator can win passing infinitely many times if Spoiler's word is out of the space. Notice that in the proof of Theorem 6 we needed that only when we used Proposition 5. Thus all we need is to consider the case when $Type(\mathcal{A}) = E^*_n$. There are two cases – there is a k such that after reading any word at least k long the automaton $\mathcal{A}$ moves to states from which its main paths are contained in the main paths of C^*_n or D^*_n; or there exists a set M of infinite words such that after reading its prefixes the automaton can reach the main paths from both C^*_n and D^*_n. In the first case by proof of Theorem 6 it is clear that $\mathcal{A} \equiv_u E^*_n$. In the second case by Proposition 5 the words from M are not from Σ^*. Then we show that $\mathcal{A} \equiv_u E^u_n$. Duplicator has easy winning strategies in the games $W^*_u(\mathcal{A}, E^u_n)$ and $W^*_u(E^u_n, \mathcal{A})$. In both cases Duplicator can play letters from one of his infinite word from M until Spoiler has main paths contained in C^*_n or D^*_n. Then Duplicator has a winning strategy. If Spoiler never plays like this then both players create infinite words not from Σ^*. Then by definition Duplicator wins. ☐

In the case of the discrete embedding we described the hierarchy of all languages. For the rational embedding we described only the hierarchy for the regular languages and we are now going to show that it does not cover all languages. For simplicity we assume that $\Sigma = \{0, 1, 2\}$.

Definition 12. *We define language T as $\bigcup_{n<\omega} 2^n L(C^*_n)$, where $2^n = 22...2$.*

This language is not regular. In the proposition below we show that for continuous reductions it is not equivalent to any regular language (and thus for uniformly continuous reductions neither).

Proposition 7. *It holds that $L(C^*_n) \leq_W T <_W L(E^*_\omega)$ for all $n \in \mathbb{N}$.*

Proof. Consider the game $W^*(L(C^*_n), T)$. We show that Duplicator has a winning strategy. Duplicator plays n times letter 2. Then he stops playing 2 and he is in a copy of the language $L(C^*_n)$. So Duplicator obviously has a winning strategy. ☐

In the end we would like to show where other languages can be in the hierarchy. We first show that the language E^*_ω is on the top of the hierarchy for uniformly continuous reductions (and thus for continuous reductions too).

Proposition 8. *For any language $A \subseteq \Sigma^*, A \leq_u E^*_\omega$.*

Proof. Consider the game $W^*_u(A, L(E^*_\omega))$. We show that Duplicator has a winning strategy. The space Σ^* is countable so we can enumerate all words to a sequence (w_n). Suppose Spoiler chooses a letter. Then he has a finite word u. Duplicator seeks the first word from (w_n) such that u is its prefix. Call it w_m. If $w_m \in A$ then Duplicator plays the letter 1 to have an even numbers of 1 and at least two 1. If $w_m \notin A$ then Duplicator plays 1 so many times to have an odd number of 1. Then while Spoiler plays letters from the word w_m then Duplicator plays 0. If Spoiler plays a letter that his word is no longer a prefix of w_m then Duplicator looks again for a word from (w_n) and plays analogously. This is a winning strategy because if Spoiler's word is from Σ^* then Duplicator eventually has to find it in (w_n) and then he has a winning strategy. If not then Duplicator did not pass infinitely many times in a row so he wins the game. $\square$

Now we show that new levels in the hierarchy can only be above the language T. This means that such levels lay between T and E^*_ω.

Theorem 8. *For any language A if $A \leq_u T$ then A is uniformly continuously equivalent to one of the languages recognized by the canonical automata.*

Proof. It suffices to show that: if $L(C^*_n) <_u A$ or $L(D^*_n) <_u A$ then $L(E^*_n) \leq_u A$; if $L(E^*_n) <_u A$ then $L(E^u_n) \leq_u A$; if $L(E^u_n) <_u A$ then $L(C^*_n) \leq_u A$ or $L(D^*_n) \leq_u A$; and if $C^*_n \leq_u A$ then $A \leq_u T$. It is all easy to prove in terms of games. $\square$

6 Lipschitz reductions

In this section we consider only 1-Lipschitz reductions. To explain why, we consider an example with languages from Definition 6 and Σ^* with a discrete embedding. Take three languages $A = 1\Sigma^*, B = 11\Sigma^*, C = 111\Sigma^*$. They are respectively 1, 2 and 3-level. Consider 2-Lipschitz reductions. It is easy to see that languages A, B are equivalent and languages B, C are also equivalent. But languages A and C are not equivalent so being 2-Lipschitz reducible is not an equivalence relation. It is also easy to create analogous examples in the case of the rational embedding or the space of infinite words.

We start from 1-Lipschitz reductions on the space of infinite words and recall two basic theorems. We call a language A *self-dual* if $A \equiv_W A^{\complement}$.

Theorem 9. *(Wadge's lemma) Let $A, B \subseteq$ be Borel languages. Then $A \leq_0 B$ or $B^{\complement} \leq_0 A$.*

This theorem can be found for example in [6] (Theorem 21.14). It is formulated for Wadge reductions but it is proved for a game where Duplicator cannot pass. According to Lemma 2 this is our case.

Theorem 10. *(Martin, Wadge) A A be a Borel language. Then:*

$$A \text{ is not self-dual} \iff \exists_{x \in \Sigma^\omega} \forall_{n<\omega} (x|_n)^{-1} A \equiv_W A.$$

The proof of this theorem can be found in [3] (Theorem 13).

Proposition 9. *Let $A, B \subseteq \Sigma^\omega$ be Borel languages. Then $A <_W B \implies A <_0 B$.*

Proof. (1) Let $A, B \subseteq \Sigma^\omega$ be Borel languages. Then $A \perp_W B \implies B \equiv_W A^{\complement}$. By Wadge's Lemma $A \not\leq_W B \implies B^{\complement} \leq_W A \implies B \leq_W A^{\complement}$ and $B \not\leq_W A \implies A^{\complement} \leq_W B$. So $B \equiv_W A^{\complement}$.

(2) Suppose $A <_W B$ and $A \not<_0 B$. Then by Wadge's Lemma $B^{\complement} <_0 A$. But by (1) we have $A <_W B^{\complement}$ – a contradiction. $\square$

Proposition 10. *Let $A \subseteq \Sigma^\omega$ be a Borel language such that $A \bot_W A^{\complement}$. Then for any Borel language $A \equiv_W B \implies A \equiv_0 B$.*

Proof. (1) Let $A, B \subseteq \Sigma^\omega$ be Borel languages. Then $A <_W B \implies A <_W B^{\complement}$. By Wadge's Lemma $B \not\leq_W A \implies A^{\complement} \leq_W B \implies A \leq_W B^{\complement}$. Suppose the inequality is not strict. Then $B^{\complement} \leq_W A \leq_W B$. Thus $B \equiv_W B^{\complement} \equiv_W A$. That is a contradiction to $A <_W B$.

(2) Suppose $A \equiv_W B$ and $A \not\leq_0 B$. By Wadge's Lemma $B^{\complement} \leq_W A$. Since $A \equiv_W B$ and by (1) we have $B \bot_W B^{\complement}$. Thus $A \bot_W B^{\complement}$ – a contradiction. □

We are now ready to present the main definition and theorem in this section.

Definition 13. *Let Bor be the family of Borel languages. We define seq : $Bor \to \omega \cup \{\omega\}$, as the function associating to a language A the length of the longest finite word x such that $\forall_{n \leq |x|} (x|_n)^{-1} A \equiv_W A$. If there exists an infinite word x with this property then $seq(A) = \omega$.*

Let $A \subseteq \Sigma^\omega$ be a Borel language. If A is not self-dual then by Theorem 10 $seq(A) = \omega$. If A is self-dual then by the same theorem and König's Lemma $seq(A) < \omega$.

Theorem 11. *Let $A \subseteq \Sigma^\omega$ be a Borel language such that $A \equiv_W A^{\complement}$. Then for any Borel language B if $A \equiv_W B$ then $(A \leq_0 B \iff seq(A) \leq seq(B))$.*

Proof. Consider the game $W_0(A, B)$ where Duplicator cannot pass. Suppose $seq(A) \leq seq(B) = k$. We show Duplicator's winning strategy. Let x be the word of length k such that $\forall_{n<\omega} (x|_n)^{-1} A \equiv_W A$. Duplicator first k move are letters from x. Let y be the word that Spoiler receives after $k + 1$ moves. By definition $(y)^{-1} A <_W A$. Then $(y)^{-1} A <_W (x)^{-1} B$. By Proposition 9 we have $(y)^{-1} A <_0 (x)^{-1} B$ so Duplicator has a winning strategy. If $seq(A) > seq(B)$ then one can create an analogous winning strategy for Spoiler. □

We now have the full characterization of the hierarchy for Borel languages with respect to 1-Lipschitz reductions. By Proposition 10 and Theorem 11 we get the corollary below.

Corollary 2. *Let $A \subseteq \Sigma^\omega$ be a Borel language. Then if A is not self-dual then the class of languages Wadge equivalent to A is the same as* 1*-Lipschitz equivalent. If A is self-dual then the class of languages Wadge equivalent splits into ω levels, according to the value of seq.*

It is easy to prove analogous results for languages of finite words with a discrete or rational embedding. In the case of the discrete embedding the class of finite level languages splits into k-level languages for every natural k. For example in the case of the rational embedding the E_n^* classes split into ω levels.

7 Conclusions

We described the hierarchy of regular languages for continuous and uniformly continuous reductions for two metrics. For the discrete metric they turned out not to be very rich. In the case of the rational metric for continuous reductions we obtained a hierarchy similar to Wagner's hierarchy up to level ω. For uniformly continuous reductions as a result of the space not being compact we obtained new interesting levels in that hierarchy. In the end we considered the hierarchy for Lipschitz reductions for both Σ^ω and Σ^* spaces. They turned out to be a simple extension of the hierarchy for uniformly continuous reductions.

There are some open problems left. It would be interesting to see how the hierarchy for all languages looks in the case of the rational metric. The intuition is that for continuous reductions it looks like the Wadge hierarchy up to the language C_ω. As for the uniformly continuous reductions the hierarchy could look the same but with an extra language below every not self dual class. In particular the position in the hierarchy of the language from Definition 12 suggests that on the limit levels there will be new classes, unlike for infinite words where on the limit levels are not self dual classes (see [4], Proposition 6).

In Section 2 we showed that the hierarchy of all languages is the same for all embeddings as in the case of the rational or discrete embedding. Nevertheless, this does not mean that the hierarchy of only regular languages is the same. It would be interesting to see if it is true also in this case, perhaps under some additional assumptions on embeddings.

References

1. Nicolas Bourbaki. Elements of mathematics. General topology. Chapters 1 - 4. *Springer*, 1998.
2. Abhijit Dasgupta. Countable metric spaces without isolated points. *Topology Atlas*, 2005.
3. Jacques Duparc. Wadge hierarchy and Veblen hierarchy: Part 1: Borel sets of finite rank. *Journal of Symbolic Logic*, 66(1):56–86, 2001.
4. Jacques Duparc. A hierarchy of deterministic context-free ω-languages. *Theoretical Computer Science*, 290(3):1253–1300, 2003.
5. John E. Hopcroft, Rajeev Motwani and Jeffrey D. Ullman. Introduction to Automata Theory, Languages and Computation. Pearson Addison-Wesley, 2007.
6. Alexander Kechris. Classical descriptive set theory. Springer-Verlag, New York, 1995.
7. Dominique Perrin and Jean-Éric Pin. Infinite Words. Automata, Semigroups, Logic and Games. Elsevier, 2004.
8. Jean-Éric Pin and Pascal Weil. Uniformities on free semigroups. International Journal of Algebra and Computation, 9, (1999), 431-453.
9. Wolfgang Thomas. Languages, Automata and Logic. Technical Report 9607, Institut fr Informatik und Praktische Mathematik, Christian-Albrechts-Universitt zu Kiel, Germany, May 1996.
10. Klaus Wagner. On ω-regular sets. *Information and Control*, 43(2):123–177, 1979.