Degrees of Streams

Jörg Endrullis   Dimitri Hendriks   Jan Willem Klop

Vrije Universiteit Amsterdam

Challenges in Combinatorics on Words

Fields Institute, Toronto
25th of April 2013

Comparing Streams

Goal

Measure the complexity of streams in terms of their infinite pattern.

Measure should be invariant under

- insertion/removal of finitely many elements
- change of alphabet

Comparing Streams

Goal

Measure the complexity of streams in terms of their infinite pattern.

Measure should be invariant under

- insertion/removal of finitely many elements
- change of alphabet

Shortcomings of existing complexity measures:

Comparing Streams

Goal

Measure the complexity of streams in terms of their infinite pattern.

Measure should be invariant under

- insertion/removal of finitely many elements
- change of alphabet

Shortcomings of existing complexity measures:

- Recursion theoretic degrees of unsolvability
  Comparison of streams via transformability by Turing machines.

Comparing Streams

Goal

Measure the complexity of streams in terms of their infinite pattern.

Measure should be invariant under

- insertion/removal of finitely many elements
- change of alphabet

Shortcomings of existing complexity measures:

- Recursion theoretic degrees of unsolvability

All computable streams are identified.

Comparing Streams

Goal

Measure the complexity of streams in terms of their infinite pattern.

Measure should be invariant under

- insertion/removal of finitely many elements
- change of alphabet

Shortcomings of existing complexity measures:

- Recursion theoretic degrees of unsolvability

  All computable streams are identified.

- Kolmogorov complexity

  Size of the shortest program computing the stream.

Comparing Streams

Goal

Measure the complexity of streams in terms of their infinite pattern.

Measure should be invariant under

- insertion/removal of finitely many elements
- change of alphabet

Shortcomings of existing complexity measures:

- Recursion theoretic degrees of unsolvability

  All computable streams are identified.

- Kolmogorov complexity

  Can be increased arbitrarily by finite insertions.

Comparing Streams

Goal

Measure the complexity of streams in terms of their infinite pattern.

Measure should be invariant under

- insertion/removal of finitely many elements
- change of alphabet

Shortcomings of existing complexity measures:

- Recursion theoretic degrees of unsolvability

  All computable streams are identified.

- Kolmogorov complexity

  Can be increased arbitrarily by finite insertions.

- Subword complexity

  $\xi_\sigma : \mathbb{N} \to \mathbb{N}$ where $\xi_\sigma(n)$ number of subwords of length n in σ.

Comparing Streams

Goal

Measure the complexity of streams in terms of their infinite pattern.

Measure should be invariant under

- insertion/removal of finitely many elements
- change of alphabet

Shortcomings of existing complexity measures:

- Recursion theoretic degrees of unsolvability

  All computable streams are identified.

- Kolmogorov complexity

  Can be increased arbitrarily by finite insertions.

- Subword complexity

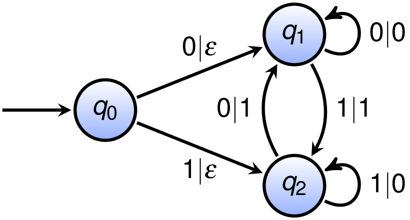
$u = 0\ 1\ 0\ 0\ 0\ 1\ 1\ \ldots$ w contains u

$w = 0\ 2\ 1\ 2\ 2\ 0\ 2\ 2\ 2\ 2\ 0\ \ldots$ but w has trivial complexity

Finite State Transducers

We propose: comparison via finite state transducers (FSTs).

Example: FST computing the difference of consecutive elements

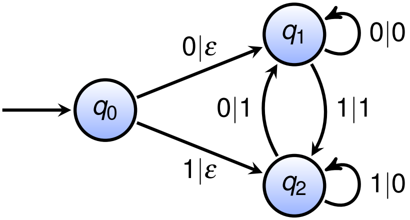


input letter | output word along the edges

Finite State Transducers

We propose: comparison via finite state transducers (FSTs).

Example: FST computing the difference of consecutive elements



input letter | output word along the edges

Transduces Thue-Morse sequence to period doubling sequence:

$$0\,1\,1\,0\,1\,0\,0\,1\ldots$$
$$\to\ 1\,0\,1\,1\,1\,0\,1\ldots$$

Degrees of Streams

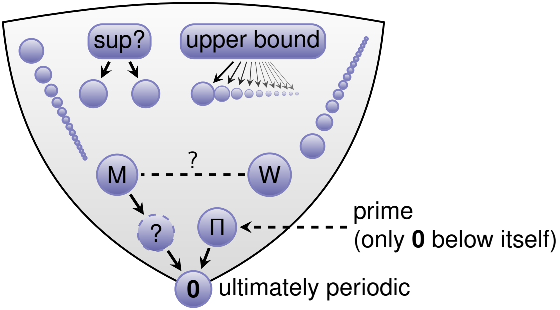
Principle: M is at least as complex as N if it can be transformed to N

$$M \rhd N \iff \text{there exists an FST transforming } M \text{ into } N$$

Degrees of Streams

Principle: M is at least as complex as N if it can be transformed to N

$$M \triangleright N \iff \text{there exists an FST transforming } M \text{ into } N$$



Partial order of degrees induced by $\triangleright$.

(degree is class of streams that can be transformed into each other)

Initial Observations

Theorem

Every degree is countable.

Initial Observations

Theorem

Every degree is countable. There are uncountably many degrees.

Initial Observations

Theorem

Every degree is countable. There are uncountably many degrees.

Theorem

Every degree has only a countable number of degrees below itself.

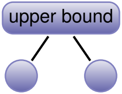
Initial Observations

Theorem

Every degree is countable. There are uncountably many degrees.

Theorem

Every degree has only a countable number of degrees below itself.



Theorem

A set of degrees has an upper bound $\iff$ the set is countable.

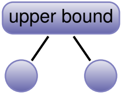
Initial Observations

Theorem

Every degree is countable. There are uncountably many degrees.

Theorem

Every degree has only a countable number of degrees below itself.



Theorem

A set of degrees has an upper bound $\iff$ the set is countable.

$$\mathrm{zip}(w_0, \mathrm{zip}(w_1, \mathrm{zip}(w_2, \ldots))),$$

$w_0(0)$	$w_1(0)$	$w_0(1)$	$w_2(0)$	$w_0(2)$	$w_1(1)$	$w_0(3)$	$w_3(0)$	$w_0(4)$	$w_1(2)$	$w_0(5)$	$w_2(1)$	$\cdots$

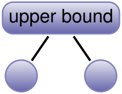
Initial Observations

Theorem

Every degree is countable. There are uncountably many degrees.

Theorem

Every degree has only a countable number of degrees below itself.



Theorem

A set of degrees has an upper bound $\iff$ the set is countable.

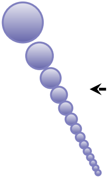
$$\text{zip}(w_0, \text{zip}(w_1, \text{zip}(w_2, \ldots))),$$

$w_0(0)$	$w_1(0)$	$w_0(1)$	$w_2(0)$	$w_0(2)$	$w_1(1)$	$w_0(3)$	$w_3(0)$	$w_0(4)$	$w_1(2)$	$w_0(5)$	$w_2(1)$	$\cdots$

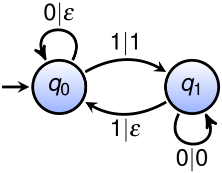
Theorem

There are no maximal degrees.

An Infinite Descending Chain



descending sequence of degrees

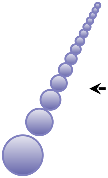


Theorem

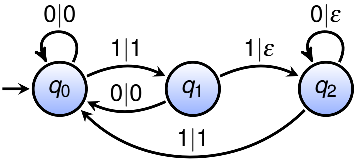
The following is an infinite descending sequence:

$$
\begin{aligned}
& D_0 = 10^{2^0} 10^{2^1} 10^{2^2} 10^{2^3} 10^{2^4} 10^{2^5} 10^{2^6} \dots \\
\mathrel{\underset{\neq}{\triangleright}}\ & D_1 = 10^{2^0} 10^{2^2} 10^{2^4} 10^{2^6} 10^{2^8} 10^{2^{10}} 10^{2^{12}} \dots \\
\mathrel{\underset{\neq}{\triangleright}}\ & D_2 = 10^{2^0} 10^{2^4} 10^{2^8} 10^{2^{12}} 10^{2^{16}} 10^{2^{20}} 10^{2^{24}} \dots \\
\mathrel{\underset{\neq}{\triangleright}}\ & \dots
\end{aligned}
$$

An Infinite Ascending Chain



ascending sequence of degrees

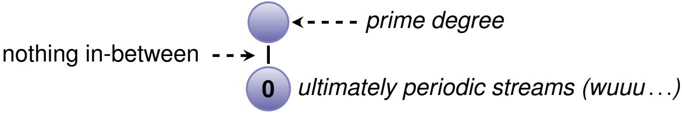


Theorem

The following is an infinite ascending sequence:

$$\begin{array}{l} \vdots \\ \underset{\neq}{\triangleright}\ A_3 = \color{red}{1}(10)^3\,\color{red}{1}(100)^3\,\color{red}{1}(10000)^3\,\color{red}{1}(100000000)^3\ldots \\ \underset{\neq}{\triangleright}\ A_2 = \color{red}{1}(10)^2\,\color{red}{1}(100)^2\,\color{red}{1}(10000)^2\,\color{red}{1}(100000000)^2\ldots \\ \underset{\neq}{\triangleright}\ A_1 = \color{red}{1}10\,\color{red}{1}100\,\color{red}{1}10000\,\color{red}{1}100000000\ldots \\ \underset{\neq}{\triangleright}\ A_0 = \color{red}{111111}\ldots \end{array}$$

Prime Degrees

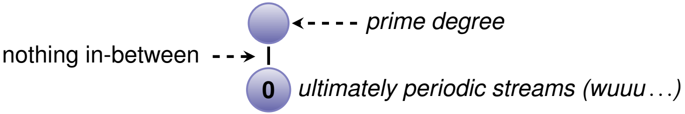


Definition

A degree $M \neq \mathbf{0}$ is prime if there is no N between M and $\mathbf{0}$:

$$\neg\exists N.\; M \mathrel{\underset{\neq}{\trianglerighteq}} N \mathrel{\underset{\neq}{\trianglerighteq}} \mathbf{0}$$

Prime Degrees



Definition

A degree $M \neq \mathbf{0}$ is prime if there is no N between M and $\mathbf{0}$:

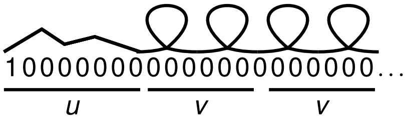
$$\neg\exists N.\ M \underset{\neq}{\triangleright} N \underset{\neq}{\triangleright} \mathbf{0}$$

Theorem

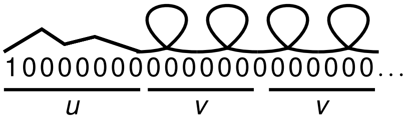
The degree of the following stream is prime:

$$\Pi = 10\ 100\ 1000\ 10000\ 100000\ 1\ldots$$
$$= 1\,0^1\ 1\,0^2\ 1\,0^3\ 1\,0^4\ 1\,0^5\ 1\,0^6\ 1\ldots$$

A Prime: $\Pi = 110100100010000100000 1\ldots$

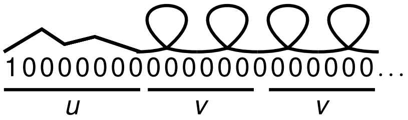


A Prime: $\Pi = 11010010001000001000001\ldots$



Let Z be the least common multiple of lengths of 0-loops in the FST.

A Prime: $\Pi = 110100100010000100000 1\ldots$



Let Z be the least common multiple of lengths of 0-loops in the FST.

Lemma

For all $q \in Q$, $n > |Q|$, there exist $u, v \in \Gamma^*$ s.t. for all $i \in \mathbb{N}$:

$$\delta(q, 10^{n+i \cdot Z}) = \delta(q, 10^n) \qquad \delta = \text{state transition function}$$

$$\lambda(q, 10^{n+i \cdot Z}) = u\,v^i \qquad \lambda = \text{output function}$$

Proof.

Analogous to the pumping lemma for regular languages. □

A Prime: $\Pi = 11010010001000010000001\ldots$

Lemma

Every transduct of Π is of the form

$$w \cdot \prod_{i=0}^{\infty} w_i \quad \text{where} \quad w_i = \prod_{j=0}^{n-1} u_j \cdot v_j^i$$

for some $n \in \mathbb{N}$ and finite words w, u_j, v_j.

A Prime: $\Pi = 11010010001000010000001\ldots$

Lemma

Every transduct of Π is of the form

$$w \cdot \prod_{i=0}^{\infty} w_i \quad \text{where} \quad w_i = \prod_{j=0}^{n-1} u_j \cdot v_j^i$$

for some $n \in \mathbb{N}$ and finite words w, u_j, v_j.

Proof.

By the pigeonhole principle we find blocks 10^k and 10^ℓ in Π s.t.:

- $|Q| < k < \ell$
- $k \equiv \ell \bmod Z$
- automaton enters 10^k and 10^ℓ with the same state q

A Prime: $\Pi = 11010010001000010000001\ldots$

Lemma

Every transduct of Π is of the form

$$w \cdot \prod_{i=0}^{\infty} w_i \quad \text{where} \quad w_i = \prod_{j=0}^{n-1} u_j \cdot v_j^i$$

for some $n \in \mathbb{N}$ and finite words w, u_j, v_j.

Proof.

By the pigeonhole principle we find blocks 10^k and 10^ℓ in Π s.t.:

- $|Q| < k < \ell$
- $k \equiv \ell \mod Z$
- automaton enters 10^k and 10^ℓ with the same state q

Define $n = \ell - k$.

A Prime: $\Pi = 110100100010000100000 1\ldots$

Lemma

Every transduct of Π is of the form

$$w \cdot \prod_{i=0}^{\infty} w_i \quad \text{where} \quad w_i = \prod_{j=0}^{n-1} u_j \cdot v_j^i$$

for some $n \in \mathbb{N}$ and finite words w, u_j, v_j.

Proof.

By the pigeonhole principle we find blocks $1\,0^k$ and $1\,0^\ell$ in Π s.t.:

- $|Q| < k < \ell$
- $k \equiv \ell \bmod Z$
- automaton enters $1\,0^k$ and $1\,0^\ell$ with the same state q

Define $n = \ell - k$. Then $Z \mid n$ and

- automaton also enters $1\,0^{k+1}$ and $1\,0^{\ell+1}$ in the same state q'
- $k+1 \equiv \ell+1 \bmod Z, \ldots$

A Prime: $\Pi = 110100100010000100000 1\ldots$

Lemma

Every transduct of Π is of the form

$$w \cdot \prod_{i=0}^{\infty} w_i \quad \text{where} \quad w_i = \prod_{j=0}^{n-1} u_j \cdot v_j^i$$

for some $n \in \mathbb{N}$ and finite words w, u_j, v_j.

Proof.

By the pigeonhole principle we find blocks $1\,0^k$ and $1\,0^\ell$ in Π s.t.:

- $|Q| < k < \ell$
- $k \equiv \ell \bmod Z$
- automaton enters $1\,0^k$ and $1\,0^\ell$ with the same state q

Define $n = \ell - k$. Then $Z \mid n$ and

- automaton also enters $1\,0^{k+1}$ and $1\,0^{\ell+1}$ in the same state q'
- $k+1 \equiv \ell + 1 \bmod Z, \ldots$

For all $i \in \mathbb{N}$, the blocks $1\,0^{k+j+i\cdot n}$ are entered in the same state. □

A Prime: $\Pi = 110100100010000100000 1\ldots$

Theorem

The degree of $\Pi = 10\ 100\ 1000\ 10000\ 100000\ 1\ldots$ is prime.

A Prime: $\Pi = 11010010001000010000001\ldots$

Theorem

The degree of $\Pi = 10\ 100\ 1000\ 10000\ 100000\ 1\ldots$ is prime.

Proof.

We consider a transduct T of Π:

$$T = w \cdot \prod_{i=0}^{\omega} w_i \qquad\qquad w_i = \prod_{j=0}^{n} u_j \cdot v_j^i$$

A Prime: $\Pi = 11010010001000010000001\ldots$

Theorem

The degree of $\Pi = 10\,100\,1000\,10000\,100000\,1\ldots$ is prime.

Proof.

We consider a transduct T of Π:

$$T = w \cdot \prod_{i=0}^{\omega} w_i \qquad\qquad w_i = \prod_{j=0}^{n} u_j \cdot v_j^i$$

Removing 'ambiguous' factors, that is, factors $j \leq n$ for which:

- $v_j^\omega = u_{j+1} v_{j+1}^\omega$ (here addition $j+1$ is modulo n)

A Prime: $\Pi = 110100100010000100000 1\ldots$

Theorem

The degree of $\Pi = 10\ 100\ 1000\ 10000\ 100000\ 1\ldots$ is prime.

Proof.

We consider a transduct T of Π:

$$T = w \cdot \prod_{i=0}^{\omega} w_i \qquad\qquad w_i = \prod_{j=0}^{n} u_j \cdot v_j^i$$

Removing 'ambiguous' factors, that is, factors $j \leq n$ for which:

- $v_j^\omega = u_{j+1} v_{j+1}^\omega$ (here addition $j+1$ is modulo n)

If everything is ambiguous, then T is ultimately periodic.

A Prime: $\Pi = 1101001000100001000001\ldots$

Theorem

The degree of $\Pi = 10\ 100\ 1000\ 10000\ 100000\ 1\ldots$ is prime.

Proof.

We consider a transduct T of Π:

$$T = w \cdot \prod_{i=0}^{\omega} w_i \qquad w_i = \prod_{j=0}^{n} u_j \cdot v_j^i$$

Removing 'ambiguous' factors, that is, factors $j \leq n$ for which:

- $v_j^{\omega} = u_{j+1} v_{j+1}^{\omega}$ (here addition $j+1$ is modulo n)

If everything is ambiguous, then T is ultimately periodic. Otherwise we can choose the v_j, u_j s.t. no u_{j+1} is not a prefix of v_j^{ω}.

A Prime: $\Pi = 110100100010000100000 1\ldots$

Theorem

The degree of $\Pi = 10\,100\,1000\,10000\,100000\,1\ldots$ is prime.

Proof.

We consider a transduct T of Π:

$$T = w \cdot \prod_{i=0}^{\omega} w_i \qquad\qquad w_i = \prod_{j=0}^{n} u_j \cdot v_j^i$$

Removing 'ambiguous' factors, that is, factors $j \leq n$ for which:

- $v_j^\omega = u_{j+1} v_{j+1}^\omega$ (here addition $j+1$ is modulo n)

If everything is ambiguous, then T is ultimately periodic. Otherwise we can choose the v_j, u_j s.t. no u_{j+1} is not a prefix of v_j^ω.

An FST can detect all transitions

A Prime: $\Pi = 110100100010000100000 1\ldots$

Theorem

The degree of $\Pi = 10\,100\,1000\,10000\,100000\,1\ldots$ is prime.

Proof.

We consider a transduct T of Π:

$$T = w \cdot \prod_{i=0}^{\omega} w_i \qquad\qquad w_i = \prod_{j=0}^{n} u_j \cdot v_j^i$$

Removing 'ambiguous' factors, that is, factors $j \leq n$ for which:

- $v_j^\omega = u_{j+1} v_{j+1}^\omega$ (here addition $j+1$ is modulo n)

If everything is ambiguous, then T is ultimately periodic. Otherwise we can choose the v_j, u_j s.t. no u_{j+1} is not a prefix of v_j^ω.

An FST can detect all transitions

- from $u_j v_j^i$ to $u_{j+1} v_{j+1}^i$,

A Prime: $\Pi = 11010010001000010000001\ldots$

Theorem

The degree of $\Pi = 10\ 100\ 1000\ 10000\ 100000\ 1\ldots$ is prime.

Proof.

We consider a transduct T of Π:

$$T = w \cdot \prod_{i=0}^{\omega} w_i \qquad\qquad w_i = \prod_{j=0}^{n} u_j \cdot v_j^i$$

Removing 'ambiguous' factors, that is, factors $j \leq n$ for which:

- $v_j^\omega = u_{j+1} v_{j+1}^\omega$ (here addition $j+1$ is modulo n)

If everything is ambiguous, then T is ultimately periodic. Otherwise we can choose the v_j, u_j s.t. no u_{j+1} is not a prefix of v_j^ω.

An FST can detect all transitions

- from $u_j v_j^i$ to $u_{j+1} v_{j+1}^i$, and thus
- from w_i to w_{i+1}

A Prime: $\Pi = 11010010001000010000001\ldots$

Theorem

The degree of $\Pi = 10\,100\,1000\,10000\,100000\,1\ldots$ is prime.

Proof.

We consider a transduct T of Π:

$$T = w \cdot \prod_{i=0}^{\omega} w_i \qquad\qquad w_i = \prod_{j=0}^{n} u_j \cdot v_j^i$$

Removing 'ambiguous' factors, that is, factors $j \leq n$ for which:

- $v_j^\omega = u_{j+1} v_{j+1}^\omega$ (here addition $j+1$ is modulo n)

If everything is ambiguous, then T is ultimately periodic. Otherwise we can choose the v_j, u_j s.t. no u_{j+1} is not a prefix of v_j^ω.

An FST can detect all transitions

- from $u_j v_j^i$ to $u_{j+1} v_{j+1}^i$, and thus
- from w_i to w_{i+1}

The function $i \mapsto |w_i|$ is linear, so FST can transduce w_i to 10^i. $\square$

Infima and Suprema

Theorem

There exist degrees X, Y that have no supremum.

Theorem

There exist degrees X, Y that have no infimum.

Infima and Suprema

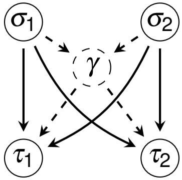
Theorem

There exist degrees X, Y that have no supremum.

Theorem

There exist degrees X, Y that have no infimum.

Idea: construct $\sigma_1, \sigma_2, \tau_1, \tau_2$ such that



and there exists no γ with the indicated properties.

Infima and Suprema

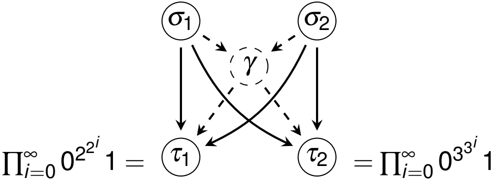
Theorem

There exist degrees X, Y that have no supremum.

Theorem

There exist degrees X, Y that have no infimum.

Idea: construct $\sigma_1, \sigma_2, \tau_1, \tau_2$ such that



and there exists no γ with the indicated properties.

Infima and Suprema

Theorem

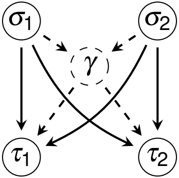
There exist degrees X, Y that have no supremum.

Theorem

There exist degrees X, Y that have no infimum.

Idea: construct $\sigma_1, \sigma_2, \tau_1, \tau_2$ such that

$$\prod_{i=0}^{\infty}(0^{2^{2^i}} 1 0^{3^{3^i}} 1) = \sigma_1 \qquad \sigma_2 = \prod_{i=0}^{\infty}(0^{3^{3^i}} 1 0^{2^{2^i}} 1)$$



$$\prod_{i=0}^{\infty} 0^{2^{2^i}} 1 = \tau_1 \qquad \tau_2 = \prod_{i=0}^{\infty} 0^{3^{3^i}} 1$$

and there exists no γ with the indicated properties.

The Subhierarchy of Computable Streams

It is also interesting to look at subhierarchies. For example

- computable streams
- morphic streams

are closed under finite state transduction.

The Subhierarchy of Computable Streams

It is also interesting to look at subhierarchies. For example

- computable streams
- morphic streams

are closed under finite state transduction.

Theorem

The subhierarchy of computable streams has a top degree.

The Subhierarchy of Computable Streams

It is also interesting to look at subhierarchies. For example

- computable streams
- morphic streams

are closed under finite state transduction.

Theorem

The subhierarchy of computable streams has a top degree.

Shuffling all computable streams does not work
(the resulting stream is not computable).

The Subhierarchy of Computable Streams

It is also interesting to look at subhierarchies. For example

- computable streams
- morphic streams

are closed under finite state transduction.

Theorem

The subhierarchy of computable streams has a top degree.

Shuffling all computable streams does not work
(the resulting stream is not computable).

Idea: for every Turing machine M define a stream

$$w(M) = x^{s(M,0)} o(M,0)\ x^{s(M,1)} o(M,1)\ x^{s(M,2)} o(M,2)\ \dots$$

The Subhierarchy of Computable Streams

It is also interesting to look at subhierarchies. For example

- computable streams
- morphic streams

are closed under finite state transduction.

Theorem

The subhierarchy of computable streams has a top degree.

Shuffling all computable streams does not work
(the resulting stream is not computable).

Idea: for every Turing machine M define a stream

$$w(M) = x^{s(M,0)} o(M,0)\ x^{s(M,1)} o(M,1)\ x^{s(M,2)} o(M,2)\ \ldots$$

where x is a fresh symbol and

- $o(M,n) =$ output of M on input n
- $s(M,n) =$ number of steps of M until termination on input n

The Subhierarchy of Computable Streams

It is also interesting to look at subhierarchies. For example

- computable streams
- morphic streams

are closed under finite state transduction.

Theorem

The subhierarchy of computable streams has a top degree.

Shuffling all computable streams does not work
(the resulting stream is not computable).

Idea: for every Turing machine M define a stream

$$w(M) = x^{s(M,0)} o(M,0)\ x^{s(M,1)} o(M,1)\ x^{s(M,2)} o(M,2)\ \ldots$$

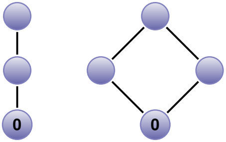
where x is a fresh symbol and

- $o(M,n) =$ output of M on input n
- $s(M,n) =$ number of steps of M until termination on input n

The shuffling of these streams is computable.

Open questions

- How to prove non-transducibility (e.g. for morphic streams)?
- Are Thue-Morse and Mephisto Walz transducible to each other?
- How many prime degrees are out there?
- Is Thue-Morse prime?
- Are there degrees forming the following structures?



- When does a set of degrees have a supremum?
- What is the structure of the subhierarchy of computable streams?
- What is the structure of the subhierarchy of morphic streams?
- …