Four Lectures On Normal Numbers - Talk 2

Pablo Ariel Heiber

March 21, 2013

Theorem. *If X is not normal then it is compressible.*

Proof. Fix X non-normal.

1. **Fix the blocks and positions with non-maximum entropy**

   Fix u_1 and positions $i_1^{(1)}, i_2^{(1)}, \ldots$ such that

   $$\lim_{n\to\infty} \text{blocc}(X[1..i_n^{(1)}], u)/i_n^{(1)} = f_u \neq b^{-|u|}.$$

   Let $u_2, \ldots, u_{b^{|u_1|}}$ be such that $\{u_i\} = \mathcal{D}^{|u_1|}$.
   For each $j = 1..b^{|u|} - 1$, let $i^{(j+1)}$ be a subsequence of $i^{(j)}$ such that

   $$\lim_{n\to\infty} \text{blocc}(X[1..i_n^{(j)}], u)/i_n^{(j)} = f_{u_j}.$$

   Let $i_j = i_j^{(b^{|u_1|})}$.

2. **Codify the compression scheme for those blocks**

   Let $\theta : (b^{|u_1|})^* \to b^*$ be such that:

   - $|\theta(v_1 \ldots v_k)| < -\log(\prod_{j=1}^k f_{v_j}) + 1$.
   - For each k, $\{\theta(v_1 \ldots v_k) : v_j \in \{0, \ldots, b-1\}^{|u_1|}\}$ is prefix free.

   For each k build a transducer T_k that looks like a trie tree of height $k-1$ that transduces each block of $k|u_1|$ digits into its output assigned by θ.

   $$\begin{aligned}
   \lim_{n\to\infty} \frac{|T_k[1..i_n]|}{i_n} &\leq \frac{\sum_{j=1}^{\lceil \frac{i_n}{k|u_1|} \rceil} |\theta(T_k[1+(j-1)k|u_1|..jk|u_1|])|}{i_n} \\
   &\leq \frac{\frac{i_n}{k|u_1|}(1 + \sum_{v_1,\ldots,v_k} -\log(\prod_{j=1}^k f_{v_j}) \prod_{j=1}^k f_{v_j})}{i_n} \\
   &\leq \frac{1 + k\sum_u (-\log f_u) f_u}{k|u_1|} \\
   &\leq \frac{1 + k|u_1|(1-\varepsilon)}{k|u_1|}
   \end{aligned}$$

3. **Group blocks to minimize rounding problems**

   Take k large enough such that the rounding does not matter (note that ε is independent from k).

4. **Profit**

$\square$

Theorem (too many people)**.** *If X is normal then it is not compressible.*

Proof. Fix X normal, $T = \langle \mathcal{Q}, \delta, o, q_0 \rangle$ with only reachable states and $\varepsilon > 0$ and show that for each sufficiently large n, $|T(X[1..n])| > (1-\varepsilon)^3 n$.

1. **Build a set of blocks with large contribution to the output**

$$a_u = \min\{|o^*(q,u)| : q \in \mathcal{Q}\}$$

$$\mathcal{S}_\ell = \{u : |u| = \ell \wedge a_u \geq (1-\varepsilon)\ell\}$$

2. **Show those are most of the blocks**

   Let t be the bound on the aebt1. For each pair of states q_1, q_2 and block v:

$$\begin{array}{rcl} |\{u : \delta^*(q_1,u) = q_2 \wedge o^*(q_1,u) = v\}| & \leq & t \\ |\{u : \exists q_1\ o^*(q_1,u) = v\}| & \leq & |\mathcal{Q}|^2 t \\ |\{u : a_u \leq (1-\varepsilon)\ell\} & \leq & |\mathcal{Q}|^2 t b^{(1-\varepsilon)\ell+2} \\ |\mathcal{S}_\ell| & \geq & b^\ell - |\mathcal{Q}|^2 t b^{(1-\varepsilon)\ell+2} \end{array}$$

   Fix ℓ such that

$$|\mathcal{S}_\ell| > b^\ell(1-\varepsilon)$$

3. **Consider only the output of those blocks**

   Let n_0 be large enough such that the frequency of each block of length ℓ is greater than $b^{-\ell}(1-\varepsilon)$. For each $n > n_0$.

$$\begin{array}{rcl} |T(X[1..n])| & \geq & \displaystyle\sum_{i=1}^{\lfloor n/\ell \rfloor} a_{X[1+(i-1)\ell..i\ell]} \\ & \geq & \displaystyle b^{-\ell}(1-\varepsilon)\frac{n}{\ell}\sum_{u\in\mathcal{S}_\ell} a_u \\ & \geq & \displaystyle b^{-\ell}(1-\varepsilon)\frac{n}{\ell}\sum_{u\in\mathcal{S}_\ell} \ell(1-\varepsilon) \\ & \geq & \displaystyle b^{-\ell}(1-\varepsilon)\frac{n}{\ell}(1-\varepsilon)b^\ell\ell(1-\varepsilon) \\ & \geq & (1-\varepsilon)^3 n \end{array}$$

4. **Profit**

$\square$

Theorem (Becher, Carton, H.)**.** *If X is normal then it is not compressible by k-var transducer.*

Proof. 1. **Build a set of blocks with large contribution to the output**

$$a_u = \min\{|o^*(q, \ldots, u)| : q \in \mathcal{Q}\}$$

$$\mathcal{S}_\ell = \{u : |u| = \ell \wedge a_u \geq (1 - \varepsilon)\ell\}$$

2. **Show those are most of the blocks**

   For each u of length ℓ, let v_u be a minimum length block such that $o^*(q, \ldots, u) = v_u$ ($|v_u| = a_u$). For each v_u there is a configuration of the machine $q, \ldots$ that justifies it.

   The $\ldots$ are the values of the k variables. Notice that changing the value of a variable's value from $x > k\ell$ to $y > k\ell$ does not change the behavior of the machine on the next ℓ steps (because the variable will always have non-zero value on each decision because it cannot decrease more than k at each of the ℓ steps). So, for each v_u consider a configuration to justify it with each variable having value less than $2k\ell$. After processing u from such configuration the configuration left has all variables less than $3k\ell$ for the same reasoning (each variable cannot grow more than k at each of the ℓ steps). As before, there are at most t blocks u that start from a given configuration, finish at the same configuration and give the same output v_u. Therefore, the size of $\{u : v_u = v\}$ for a single v is bounded by $(2k\ell)^k|\mathcal{Q}|(3k\ell)^k|\mathcal{Q}|t$ and therefore,

$$|\mathcal{S}_\ell| \geq b^\ell - (2k\ell)^k|\mathcal{Q}|(3k\ell)^k|\mathcal{Q}|tb^{(1-\varepsilon)\ell+2}.$$

   Then, we can proceed as in the previous proof.

   Configurations with a counter larger than ℓk have their behavior repeated with configurations with a smaller counter, so only consider this last one.

3. **Consider only the output of those blocks**
4. **Profit**

□

Lemma (adapted from Schnorr and Stimm)**.** *The set of sequences that go through every state of a reachable strongly connected component of an automata has positive measure.*

Proof. 1. **Fix a connected component and the path u that leads to it**

   Let $\{q_1, \ldots, q_m\}$ be a reachable strongly connected component and $\delta^*(q_0, u) = q_1$.

2. **Build accumulated paths u_i from each state to a given one**

   Let $u_{1,j} = \lambda$ and $u_{i+1,j}$ be a path from $\delta^*(q_{i+1}, u_{1,j} \ldots u_{i,j})$ to q_j.

3. **Consider the subset of $[u]$ that contains each $u_1u_2 \ldots u_n$ an infinite number of times**

   Each time $u_{1,j}u_{2,j} \ldots u_{n,j}$ occurs in the sequence, q_j will be revisited.

4. **Profit**

□

Theorem (Becher, Carton, H.)**.** *If X is normal then it is not compressible by non-deterministic transducer.*

Proof. 1. **Build a set of blocks with large contribution to the output**

$$a_u = \min\{\min\{|v| : v \in o^*(q, u)\} : q \in \mathcal{Q}'\}$$

   where $\mathcal{Q}'$ are the states visited infinitely often in a computation of $T(X)$.

2. **Show those are most of the blocks**

   Follow as before, note that there cannot be more than t ways to go from state q_1 to state q_2 with the same output and not violate aebt1 because there is a set of positive measure extensions that is accepted starting from q_2.

3. **Consider only the output of those blocks**

4. **Profit**

□

Theorem (Becher, Carton, H. on an idea of Boasson)**.** *There is a non-deterministic k-ary stack transducer that compresses a normal sequence.*

Proof. Proof steps:

1. **Build a palindromic version of Champernowne**

   Let $X = 01\ 10\ 00011011\ 11011000\ \ldots$ be as Champernowne but adding the blocks of a given length twice, the second time being a mirror of the first.

2. **Show it is normal**

   Same reasoning as Champernowne.

3. **Build a compressor of palindromes**

   Two states, in one we push the current input onto the stack and output it. We may move non-deterministically to the other. In the second state we only remove everything from the stack as long as it matches the input and otherwise reject the input. When the stack is empty, go back to the initial state, outputting a separator.

   Notice that $u_1\#u_2\#u_3\ldots$ is the output only of $u_1(u_1^r)u_2(u_2^r)\ldots$ the transducer is one-to-one, and it compresses X up to almost $1/2$, using a third digit as output (which requires $\log_3 2 < 1/2$ inflation to go back to 2 digits in a simple way).

4. **Profit**

□

Theorem (Agafonov; Becher and H.)**.** *Finite-state selectors preserve normality.*

Proof. By the way of contradiction. Notice selectors will always select at least linearly many digits (Staiger) and then compose the selector with a transducer that compresses its output while maintaining the rest.

□

Theorem (Merkle and Reimann)**.** *Finite-state* 1*-var selectors do not preserve normality.*

Proof. Let $X = 01\ 00011011\ 000001010\ \ldots$ be as Champernowne but concatenating in each part all blocks of a given length in lexicographic order. Note that the number of zeros is always greater than the number of ones inside any part, and its the same after each part. Then, build a 1-var selector that counts that difference (if 0, add 1, otherwise, subtract 1) and select only when the the variable is 0. This will select the first digit of each part, which is always 0, so the output of the selector is $00000\ldots$.

□

Theorem (Becher, Carton and H.)**.** *Non-deterministic selectors do not preserve normality.*

Proof. Guess the next digit, select it if it is zero.

□

Theorem (Merkle and Reimann). *Selection to the left belonging to a linear language does not preserve normality.*

Proof. Use palindromic Champernowne and the language of palindromes, for which all zeros are selected.

□

Theorem (Becher, Carton and H.). *Selection to the right suffix belonging to a set of infinite sequences recognizable by non-deterministic automata preserves normality.*

Proof. Basically the same idea than from the left, although much more technical. Show that the selector takes at least linearly many and then compose. □

Theorem (Becher, Carton and H.). *Selection of digits in between two zeros does not preserve normality.*

Proof. Fix a normal in base 2 X. For binary digits c and d, let x_{cd} be the "probability" that 0 appears next on the output given that cd are the last two digits seen on the input.

Simple analysis gives the following recurrences:

$$\begin{aligned} x_{00} &= 1/2 + x_{01} \\ x_{01} &= x_{11}/2 \\ x_{10} &= x_{00}/2 + x_{01}/2 \\ x_{11} &= x_{10}/2 + x_{11}/2 \end{aligned}$$

By solving them we obtain that $x_{00} = 6/10$, which makes the frequency of 00 in the output different than 1/2, so the frequency of 00 is not 1/2 of the frequency of 0.

□