

Windable Heads & Recognizing NL with Constant Randomness

Mehmet Utkan Gezer^[0000-0002-5022-178X]

Boğaziçi University, Bebek İstanbul, Türkiye
utkan.gezer@boun.edu.tr

Abstract. Every language in NL has a k -head two-way nondeterministic finite automaton ($2\text{nfa}(k)$) recognizing it. It is known how to build a constant-space verifier algorithm from a $2\text{nfa}(k)$ for the same language with constant-randomness, but with error probability $k^2 - 1/2k^2$ that can not be reduced further by repetition. We have defined the unpleasant characteristic of the heads that causes the high error as the property of being “windable”. With a tweak on the previous verification algorithm, the error is improved to $k_W^2 - 1/2k_W^2$, where $k_W \leq k$ is the number of windable heads. Using this new algorithm, a subset of languages in NL that have a $2\text{nfa}(k)$ recognizer with $k_W \leq 1$ can be verified with arbitrarily reducible error using constant space and randomness.

Keywords: Interactive Proof Systems · Multi-head finite automata · Probabilistic finite automata.

1 Introduction

Probabilistic Turing machines (PTM) are otherwise deterministic Turing machines with randomness as a resource. They can be standalone recognizers of languages, or be verifiers for the proofs of memberships. In either scenario, a measurable error is incorporated into their decision due to randomness involved in their execution. This error can usually be reduced via repeated execution in the PTM’s control.

The language class verifiable by the constant-randomness two-way probabilistic finite automata (2pfa) is the class NL. Curiously, however, the error of these verifiers in recognizing languages of this class seems to be irreducible beyond a certain threshold [6].

In this paper, we introduce a characteristic for the languages in NL. Based on this characteristic, we lower the error threshold established in [6] for almost all languages in NL. Finally, we delineate a subset of NL which are verifiable by the constant-randomness 2pfa with arbitrarily low error.

The remaining of the paper is structured as follows: Sections 2 and 3 provides the necessary background as well as our terminology in the domain. A key property of the multi-head finite automata is identified in section 4. The characterization of languages in NL and our algorithm for verification achieving aforementioned results are described in section 5.

Following notation will be common throughout this paper:

- $\mathcal{L}(M)$ denotes the language recognized by the machine M .
- $\mathcal{L}(X) = \{ \mathcal{L}(M) \mid M \in X \}$ for a class of machines X .
- $S_{\setminus q}$ denotes the set S without its element q .
- σ_i denotes the i th element of the sequence σ .
- w^\times denotes the substring of w without its last character.
- $\sigma \circ \tau$ denotes the sequence σ concatenated with the element or sequence τ .

2 Finite automata with k heads

Finite automata are the Turing machines with read-only tape heads on a single tape. A finite automata with only one head is equivalent to a DFA (deterministic finite automaton) in terms of language recognition [2], hence recognizes a regular language. Finite automata with $k > 1$ heads can recognize more than just regular languages. Their formal definition may be given as follows:

Definition 1 (Multi-head nondeterministic finite automata). A $2nfa(k)$ is a 5-tuple, $M = (Q, \Sigma, \delta, q_0, q_f)$, where;

1. Q is the finite set of states,
2. Σ is the finite set of input symbols,
 - (a) $\triangleright, \triangleleft$ are the left and right end-markers for the input on the tape,
 - (b) $\Gamma = \Sigma \cup \{ \triangleright, \triangleleft \}$ is the tape alphabet,
3. $\delta: Q \times \Gamma^k \rightarrow \mathcal{P}(Q_{\setminus q_0} \times \Delta^k)$ is the transition function, where;
 - (a) $\Delta = \{ -1, 0, 1 \}$ is the set of head movements,
4. $q_0 \in Q$ is the unique initial state,
5. $q_f \in Q$ is the unique accepting state.

Machine M is said to execute on a string $w \in \Sigma^*$, when $\triangleright w \triangleleft$ is written onto M 's tape, all of its heads rewind to the cell with \triangleright , its state is reset to q_0 , and then it executes in steps by the rules of δ . At each step, inputs to δ are the state of M and the symbols read by respective heads of M .

When $|\delta| = 1$ with the only member $(q', (d_1, \dots, d_k)) \in Q_{\setminus q_0} \times \Delta^k$, the next state of M becomes q' , and M moves its i th head by d_i . Whenever $|\delta| > 1$, the execution branches, and each branch runs in parallel. A branch is said to reject w , if $|\delta| = 0$, or if all of its branches reject. A branch accepts w , if its state is at q_f , or if any one of its branches accepts. A branch may also do neither, in which case the branch is said to loop.

A string w is in $\mathcal{L}(M)$, if the root of M 's execution on w is an accepting branch. Otherwise, $w \notin \mathcal{L}(M)$, and the root of M 's execution is either a rejecting or a looping branch.

Restricting δ to not have transitions inbound to q_0 does not detriment the language recognition of a $2nfa(k)$ in terms of its language recognition: Any $2nfa(k)$ with such transitions can be converted into one without, by adding a new initial state q'_0 and setting $\delta(q'_0, \triangleright, \dots, \triangleright) = \{ (q_0, 0, \dots, 0) \}$.

Lemma 1. *The containment $\mathcal{L}(2nfa(k)) \subsetneq \mathcal{L}(2nfa(k+1))$ is proper [4, 5, 8, 7, 3].*

Lemma 2. *There is a way to obtain a $2nfa(2k)$ that is guaranteed to halt, from any given $2nfa(k)$.*

Proof. A k -headed automaton running on an input w of length n has n^k distinct configurations. Additional k heads can count up to $n^k = (nnn \dots n)_n$, and halt the machine with a rejection.

Lemma 3. *Every $2nfa(k)$ can be converted into an equivalent $2nfa(k)$ which does not move its heads beyond the end markers.*

This is done via trivial modifications on the transition function.

Definition 2 (Multi-head deterministic finite automata). *A $2dfa(k)$ is a $2nfa(k)$ that is restricted to satisfy $|\delta| \leq 1$, where δ is its transition function.*

Lemma 4. *Following are shown in [1]:*

$$\bigcup_{k=1}^{\infty} \mathcal{L}(2nfa(k)) = NL \quad (1)$$

$$\bigcup_{k=1}^{\infty} \mathcal{L}(2dfa(k)) = L \quad (2)$$

Definition 3 (Multi-head one-way finite automata). *A $1nfa(k)$ is a restricted $2nfa(k)$ that does not move its heads backwards on the tape. In its definition, $\Delta = \{0, 1\}$. A $1dfa(k)$ is similarly a restriction of $2dfa(k)$.*

Definition 4 (Multi-head probabilistic finite automata). *A $2pfa(k)$ M is a PTM defined similar to a $2nfa(k)$ with the following modifications on definition 1:*

1' $Q = Q_D \cup Q_P$, where Q_D and Q_P are disjoint.

3' Transition function δ is overloaded as follows:

$$- \delta: Q_D \times \Gamma^k \rightarrow \mathcal{P}(Q_{\setminus q_0} \times \Delta^k)$$

$$- \delta: Q_P \times \Gamma^k \times \{0, 1\} \rightarrow \mathcal{P}(Q_{\setminus q_0} \times \Delta^k)$$

The output of δ may at most have 1 element.

States Q_D are called deterministic, and Q_P probabilistic. Depending on the state of the machine, δ receives a third parameter, where a 0 or 1 is provided by a random bit-stream.

A string w is in $\mathcal{L}(M)$, if and only if M accepts w with a probability greater than $1/2$.

Due to the probabilistic nature of a given $2pfa(k)$ M , following three measures of error in the language recognition are inherent to it:

$$\varepsilon_{\text{fail-to-accept}}(M) = \Pr[M \text{ does not accept } w \mid w \in \mathcal{L}(M)] \quad (\text{False rejection})$$

$$\varepsilon_{\text{fail-to-reject}}(M) = \Pr[M \text{ does not reject } w \mid w \notin \mathcal{L}(M)] \quad (\text{Failure to reject})$$

$$\varepsilon_{\text{false-accept}}(M) = \Pr[M \text{ accepts } w \mid w \notin \mathcal{L}(M)] \quad (\text{False acceptance})$$

Note that when a $2\text{pfa}(k)$ M does not reject a string w , then it could have either accepted it, or wound up in an infinite loop. Consequently, $\varepsilon_{\text{fail-to-reject}} \geq \varepsilon_{\text{false-accept}}$ is always true. Based on this fact, the overall weak and strong errors of a probabilistic machine M is defined as follows:

$$\begin{aligned}\varepsilon_{\text{weak}}(M) &= \max(\varepsilon_{\text{fail-to-accept}}(M), \varepsilon_{\text{false-accept}}(M)) && \text{(Weak error)} \\ \varepsilon_{\text{strong}}(M) &= \max(\varepsilon_{\text{fail-to-accept}}(M), \varepsilon_{\text{fail-to-reject}}(M)) && \text{(Strong error)}\end{aligned}$$

Given a k and $\varepsilon < 1/2$, let

$$\mathcal{L}_{\text{weak},\varepsilon}(2\text{pfa}(k)) = \{ \mathcal{L}(M) \mid M \in 2\text{pfa}(k), \varepsilon_{\text{weak}}(M) \leq \varepsilon \}$$

be the class of languages recognized by a $2\text{pfa}(k)$ with a weak error at most ε . Class $\mathcal{L}_{\text{strong},\varepsilon}(2\text{pfa}(k))$ is defined similarly.

3 Interactive Proof Systems

An interactive proof system (IPS) models the verification process of proofs. Of the two components in an IPS, the *prover* produces the purported proof of membership for a given input string, while the *verifier* either accepts or rejects the string, alongside its proof. The catch is that the prover is assumed to advocate for the input string's membership without regards to truth, and the verifier is expected to be accurate in its decision, holding a healthy level of skepticism against the proof.

The verifier is any Turing machine with capabilities to interact with the prover via a shared communication cell. The prover can be seen as an infinite state transducer that has access to both an original copy of the input string and the communication cell. Prover never halts, and its output is to the communication cell.

Our focus will be on the one-way IPS, which restricts the interaction to be a monologue from the prover to the verifier. Since there is no influx of information to the prover, prover's output will be dependent on the input string only. Consequently, a one-way IPS can also be modeled as a verifier paired with a certificate function, $c: \Sigma^* \rightarrow \Lambda^\infty$, where Λ is the communication alphabet. A formal definition follows:

Definition 5 (One-way interactive proof systems). *An IP(restriction-list) is defined with a tuple of a verifier and a certificate function, $S = (V, c)$. The verifier V is a Turing machine of type specified by the restriction-list. The certificate function c outputs the claimed proof of membership $c(w) \in \Lambda^\infty$ for a given input string w .*

The verifier's access to the certificate is only in the forward direction. The qualifier "one-way", however, specifies that the interaction in the IPS is a monologue from the prover to the verifier, not the aforementioned fact, which is true for all IPS.

The language recognized by S can be denoted with $\mathcal{L}(S)$, as well as $\mathcal{L}(V)$. A string w is in $\mathcal{L}(S)$, if and only if the interaction results in an acceptance of w by V .

If the verifier of the IPS is probabilistic, its error becomes the error of the IPS. The notation $\mathcal{L}_{\text{weak},\varepsilon}(\text{IP}(\text{restriction-list}))$ and $\mathcal{L}_{\text{strong},\varepsilon}(\text{IP}(\text{restriction-list}))$ is also adopted.

Say and Yakaryılmaz proved that [6]:

$$\text{NL} \subseteq \mathcal{L}_{\text{weak},\varepsilon}(\text{IP}(2\text{pfa}(1), \text{constant-randomness})) \quad \text{for } \varepsilon > 0 \text{ arbitrarily small,} \quad (3)$$

$$\text{NL} \subseteq \mathcal{L}_{\text{strong},\varepsilon}(\text{IP}(2\text{pfa}(1), \text{constant-randomness})) \quad \text{for } \varepsilon = \frac{1}{2} - \frac{1}{2k^2}, k \rightarrow \infty. \quad (4)$$

For the latter proposition, the research proves that any language $L \in \text{NL}$ can be recognized by a one-way IPS $S \in \text{IP}(2\text{pfa}(1), \text{constant-randomness})$, which satisfies $\varepsilon_{\text{strong}}(S) \leq 1/2 - 1/2k$, and where k is the minimum number of heads among the $2\text{nfa}(k)$ recognizing L that also halts on every input. Existence of such a $2\text{nfa}(k)$ is guaranteed by lemmas 2 and 4.

This work improves on the findings of [6]. For their pertinence, an outline of the algorithms attaining the errors in eqs. (3) and (4) is provided in the following sections.

3.1 Reducing weak error arbitrarily using constant-randomness verifier

Given a language $L \in \text{NL}$ with a halting $2\text{nfa}(k)$ recognizer M , verifier $V_1 \in 2\text{pfa}(1)$ expects a certificate to report (i) the k symbols read, and (ii) the non-deterministic branch taken for each transition made by M on the course of accepting w . Such a report necessarily contains a lie, if $w \notin \mathcal{L}(M) = L$.

Verifier V_1 has an internal representation of M 's control. Then, the algorithm for the verifier is as follows:

1. Repeat m times:
 - (a) Move head left, until \triangleright is read.
 - (b) Reset M 's state in the internal representation, denoted q_m .
 - (c) Randomly choose a head of M by flipping $\lceil \log k \rceil$ coins.
 - (d) Repeat until q_m becomes the accepting state of M :
 - i. Read k symbols and the nondeterministic branch taken by M from the certificate.
 - ii. *Reject* if the reading from V_1 's head disagrees with the corresponding symbol on the certificate.
 - iii. Make the transition in the internal representation if it is valid, and move the chosen head as dictated by the nondeterministic branch. *Reject* otherwise.

2. *Accept.*

For the worst case errors, it is assumed that there is a lie for the certificate to tell about each one of the heads alone and in any single one of the transitions, which causes V_1 to fail to reject a string $w \notin L$. Similar lies are assumed to exist for the false acceptances. Following are then the (upper bounds of) errors for V_1 :

$$\varepsilon_{\text{fail-to-accept}}(V_1) = 0 \quad \varepsilon_{\text{fail-to-reject}}(V_1) \leq \frac{k-1}{k} \quad \varepsilon_{\text{false-accept}}(V_1) \leq \frac{1}{k^m}$$

A discrepancy between $\varepsilon_{\text{false-accept}}$ and $\varepsilon_{\text{fail-to-reject}}$ is observed, because an adversarial certificate may wind V_1 up in an infinite loop on its first round of m repetitions. This is possible despite M being a halting machine. The lie in the certificate can present an infinite and even changing input string from the perspective of the head being lied about.

Being wound up counts as a failure to reject, but does not yield a false acceptance. The resulting weak error is $\varepsilon_{\text{strong}} = k^{-m}$, which can be made arbitrarily small.

3.2 Bringing strong error below $1/2$ using constant-randomness verifier

Presented first in [6], verifier V'_1 with the following algorithm manages to achieve $\varepsilon_{\text{strong}}(V'_1) < 1/2$, outlined as follows:

1. Randomly *reject* with $k - 1/2k$ probability by flipping $\lceil \log k \rceil + 1$ coins.
2. Continue as V_1 .

This algorithm then has the following upper bounds for the errors:

$$\varepsilon_{\text{fail-to-accept}}(V'_1) = \frac{k-1}{2k} \quad \varepsilon_{\text{fail-to-reject}}(V'_1) \leq \frac{k^2-1}{2k^2} \quad \varepsilon_{\text{false-accept}}(V'_1) \leq \frac{k+1}{2k^{m+1}}$$

Since $\varepsilon_{\text{fail-to-reject}}(V'_1)$ is potentially greater than $\varepsilon_{\text{fail-to-accept}}(V'_1)$, the strong error is bounded by $k^2 - 1/2k^2$.

4 Windable heads

This section will introduce a property of the heads of a $2\text{nfa}(k)$. It leads to a characterization of the $2\text{nfa}(k)$ by the number of heads with this property. A subset rNL of the class NL will be defined, which will also be a subset of $\mathcal{L}_{\text{strong},\varepsilon}(\text{IP}(2\text{pfa}(1), \text{constant-randomness}))$ for $\varepsilon > 0$ approaching zero.

A head of a $2\text{nfa}(k)$ M is said to be *windable* if these three conditions hold:

- There is a cycle on the graph of M 's transition diagram, and a path from q_0 to a node on the cycle.
- The movements of the head-in-question add up to zero in a full round of that cycle.

- The readings of the head is consistent along the said path and cycle.

The definition of a head being windable completely disregards the readings of the other heads, hence the witness path and the cycle need not be a part of a realistic execution of the machine M .

We will define the windable heads formally to clarify its distinguishing points. Some preliminary definitions will be needed.

Definition 6 (Multi-step transition function).

$$\delta^t: Q \times (\Gamma^t)^k \rightarrow \mathcal{P}(Q_{\setminus q_0} \times (\Delta^t)^k)$$

is the t -step extension of the transition function δ of a $2nfa(k)$ M . It is defined recursively, as follows:

$$\delta^1 = \delta$$

$$\delta^t(q, g_1, \dots, g_k) = \left\{ (r, D_1 \circ d_1, \dots, D_k \circ d_k) \left| \begin{array}{l} (r, d_1, \dots, d_k) \in \delta(s, g_{1t}, \dots, g_{kt}) \\ (s, D_1, \dots, D_k) \in \delta^{t-1}(q, g_1^\times, \dots, g_k^\times) \end{array} \right. \right\}$$

The set $\delta^t(q, g_1, \dots, g_k)$ contains a $(k+1)$ -tuple for each nondeterministic computation to be performed by M , as it starts from the state q and reads g_i with its i th head. These tuples, each referred to as a *computation log*, consist of the state reached, and the movement histories of the k heads during that computation.

The constraint of a constant and persistent tape contents that is present in an execution of a $2nfa(k)$ is blurred in the definition for multi-step transition function. This closely resembles the verifier's perspective of the remaining heads that it does not verify in the previous section. There, however, the verifier's readings were consistent in itself. This slight will be accounted for with the next pair of definitions.

Definition 7 (Relative head position during i th transition). *The relative position of the head since the before the first movement in the movement history D of length t and while making the i th transition of that history is given by the function $\rho_D(i): \mathbb{N}_1^{\leq t} \rightarrow (-t, t)$ defined as*

$$\rho_D(i) = \text{sum}(D_{1:i-1}).$$

If D is a movement history from a computation that does not attempt to move the head out of tape's bounds, then $\rho_D(i)$ is the position of the head while making the i th transition, relative to the position where the head was at the beginning of that computation.

Definition 8 (1-head consistent δ^t). $\delta_1^t: Q \times (\Gamma^t)^k \rightarrow \mathcal{P}(Q_{\setminus q_0} \times (\Delta^t)^k)$ is the *i th-head consistent* subset of δ^t of a $2nfa(k)$ M . It filters out the first-head inconsistent computation logs by scrutinizing the purportedly read characters by examining the movement histories against the readings. The formal definition

assumes that M does not attempt to move its heads beyond the end markers, and is as follows:

$$\delta_1^t(q, g_1, \dots, g_k) = \{ (r, D_1, \dots, D_k) \in \delta^t(q, g_1, \dots, g_k) \mid \forall p \in (-t, t), \forall x, y \in \rho_{D_i}^{-1}(p) [g_{ix} = g_{iy}] \}$$

For each pair of transitions departing from the same tape cell, it is checked whether the same symbol is read while being performed. This check is needed to be done only for $p \in (-t, t)$, since in t steps, a head may at most travel t cells afar, and the last cell it can read from will then be the previous one. This is also consistent with the definition of ρ_D .

This last definition is the exact analogue of the verifiers' perspective in the algorithms proposed by [6]. It can be used directly in our next definition, that will lead us to a characterization of the $2\text{nfa}(k)$.

Definition 9 (Windable heads). *The i th head of a $2\text{nfa}(k)$ M is **windable** iff there exists;*

1. $g_1, \dots, g_k \in \Gamma^t$ and $g'_1, \dots, g'_k \in \Gamma^l$, for t and l positive,
2. $(q, D_1, \dots, D_k) \in \delta_i^t(q_0, g_1, \dots, g_k)$,
3. $(q, D_1 \circ D'_1, \dots, D_k \circ D'_k) \in \delta_i^{t+l}(q_0, g_1 \circ g'_1, \dots, g_k \circ g'_k)$ where $\text{sum}(D'_i) = 0$.

When these conditions hold, g_1, \dots, g_k can be viewed as the sequences of characters that can be fed to δ to bring M from q_0 to q , crucially without breaking consistency among the i th head's readings. This ensures reachability to state q . Then, the sequences g'_1, \dots, g'_k wind the i th head into a loop; bringing M back to state q and the first head back to where it started the loop, all while keeping the i th head's readings consistent. The readings from the other heads are allowed to be inconsistent, and their position can change with every such loop.

A head is **reliable** iff the head is not windable.

It is important to note that a winding is not based on a realistic execution of a $2\text{nfa}(k)$. A head of a $2\text{nfa}(k)$ M might be windable, even if it is guaranteed to halt on every input. This is because the property of being windable allows other heads to have *unrealistic*, inconsistent readings that may be never realized with any input string.

5 Recognizing some languages in NL with constant-randomness and reducible-error verifiers

Consider a language $L \in \text{NL}$ with a $2\text{nfa}(k)$ recognizer M that halts on every input. In designing the randomness-restricted $2\text{pfa}(1)$ verifier V_2 , following three cases will be considered:

All heads are reliable. In this case, V_1 suffices by itself to attain reducible error. Without any windable heads in the underlying $2\text{nfa}(k)$, each round of V_1 will terminate. The certificate can only make V_1 falsely accept, and the chances for that can be reduced arbitrarily by increasing m .

All heads are windable. In this case, unless the worst-case assumptions are alleviated, any verification algorithm using a simulation principle similar to V_1 will be wound up on the first round. The head with the minimum probability of getting chosen will be the weakest link of V_2 , thus the head the certificate will be lying about. The failure to reject rate is equal 1 minus that probability. This rate is the lowest when the probabilities are equal, and is then $k^{-1/k}$.

It is a mix. Let k_W, k_R denote the windable and reliable head counts, respectively. Thus $k_W + k_R = k$. The new verifier algorithm V_2 is similar to V_1 , but instead of choosing a head to simulate with equal probability, it will do a *biased branching*. With biased branching, V_2 favors the reliable heads over the windable heads while choosing a head to verify.

Let P_W, P_R denote the desired probability of choosing a windable and reliable head, respectively. Note that $P_W + P_R = 1$. The probabilities of choosing a head within types (windable or reliable) are kept equal. Denote the probability of choosing a particular windable head as $p_W = P_W/k_W$, and similarly $p_R = P_R/k_R$. Assume P_W, P_R are finitely representable in binary, and with b digits after the decimal point. Then, the algorithm of V_2 is the same as V_1 , with the only difference at step 1c:

- 1c.' Randomly choose a head of M by biased branching:
- Instead of flipping $\lceil \log k \rceil$ coins, flip $b + \lceil \log(\max(k_W, k_R)) \rceil$ coins. Let z_1, z_2, \dots, z_b be the outcomes of the first b coins.
 - If $\sum_{i=1}^b 2^{-i} z_i < P_W$, choose one of the windable heads depending on the outcomes of the next $\lceil \log k_W \rceil$ coins. Otherwise, similarly choose a reliable head via $\lceil \log k_R \rceil$ coins.

For an input string $w \in L$. Verifier V_2 is still perfectly accurate. Certificate may provide any route that leads M to acceptance. Repeating this for m -many times, V_2 will accept after m rounds of validation.

For an input string $w \notin L$. To keep V_2 from rejecting, the certificate will need to lie about at least one of the heads. Switching the head to lie about in between rounds cannot be of any benefit to the certificate on its mission, since the rounds are identical both from V_2 's and the certificate's points of view. Hence, it is reasonable to assume that the certificate repeats itself in each round, and simplify our analysis.

The worst-case assumption is that the certificate can lie about a single (arbitrary) head alone and deceive V_2 in the worst means possible, depending on the head it chooses:

- If it chooses the head being lied about, V_2 detects the lie rather than being deceived.
- Otherwise, if a windable head was chosen, V_2 loops indefinitely.
- Otherwise (i.e. a reliable head was chosen), V_2 runs for another round or accepts w .

The head which the certificate fixes to lie about is either a windable head or a reliable one. Given a V_2 algorithm with its parameter P_W set, let $F_W(P_R)$ be the probability of V_2 failing to reject against a certificate that lies about any one windable head. Let $F_R(P_R)$ similarly be the probability for the reliable counterpart.

The most evil certificate would lie about the head that yields a higher error. Thus, the worst-case failure to reject probability is given by

$$F(P_R) = \max(F_W(P_R), F_R(P_R)).$$

Individually, $F_W(P_R)$ and $F_R(P_R)$ are calculated using the following formulae:

$$\begin{aligned} F_W(P_R) &= \sum_{i=0}^{m-1} P_R^i (P_W - p_W) + P_R^m \\ &= (1 - P_R^m) \cdot \left(1 - \frac{1}{k_W}\right) + P_R^m \\ &= 1 - \frac{1 - P_R^m}{k_W} \end{aligned}$$

$$\begin{aligned} F_R(P_R) &= \sum_{i=0}^{m-1} (P_R - p_R)^i P_W + (P_R - p_R)^m \\ &= \frac{1 - (P_R - p_R)^m}{1 - (P_R - p_R)} \cdot P_W + (P_R - p_R)^m \\ &= \frac{P_W}{P_W + p_R} + \left(1 - \frac{P_W}{P_W + p_R}\right) (P_R - p_R)^m \end{aligned}$$

The objective is to find the optimum P_R , denoted P_R^* , minimizing the error $F(P_R)$. We note that $F(1)$ is 1. Hence, $P_R^* < 1$.

Constant m may be chosen arbitrarily large. For $P_R < 1$, and m very large, approximations of F_W and F_R are, respectively, given as

$$F_W^*(P_R) = 1 - \frac{1}{k_W} \quad F_R^*(P_R) = \frac{P_W}{P_W + p_R}.$$

Error F_W^* is a constant between 0 and 1. For $0 \leq P_R \leq 1$, error F_R^* decreases from 1 to 0, and in a strictly monotonous fashion:

$$\frac{dF_R^*}{dP_R} = \frac{-p_R - P_W/k_W}{(P_W + p_R)^2} < 0$$

These indicate that $F_W^*(P_R)$ and $F_R^*(P_R)$ are equal for a unique $P_R = P_R^*$. The optimality of P_R^* will be proved shortly. It is easy to verify that

$$P_R^* = \frac{k_R}{k - 1}. \quad (5)$$

Using P_R^* we can define F^* as the following partial function:

$$F^*(P_R) = \begin{cases} F_R^*(P_R) & \text{for } P_R \leq P_R^* \\ F_W^*(P_R) & \text{for } P_R \geq P_R^* \end{cases}$$

Since F_R^* is a decreasing function, $F(P_R) > F(P_R^*)$ for any $P_R < P_R^*$. The approximation F_W^* is a constant function. Function F_W , however, is actually an increasing one. Therefore, given m large, probability P_R^* approximates the optimum for V_2 choosing a reliable head among the k heads of the M , while verifying for the language $\mathcal{L}(M) \in \text{NL}$. Consequently the optimum error for V_2 is

$$F(P_R^*) = 1 - \frac{1}{k_W}. \quad (6)$$

This points to some important facts.

Theorem 1. *The minimum error for V_2 depends only on the number of windable heads of the $2\text{nfa}(k)$ M recognizing $L \in \text{NL}$.*

Definition 10 (Reducible strong error subset of NL). *For $\varepsilon > 0$ approaching zero, the reducible strong error subset of NL is defined as*

$$r\text{NL} = \text{NL} \cap \mathcal{L}_{\text{strong},\varepsilon}(\text{IP}(2\text{pfa}(1), \text{constant-randomness})).$$

Theorem 2. *For $k_W \leq 1$ and k_W arbitrary,*

$$\mathcal{L}(2\text{nfa}(k_W + k_R)) \subseteq r\text{NL}.$$

Equations (5) and (6), and their consequent theorems 1 and 2, constitute the main results of this study.

Similar to how V_1' was obtained, the algorithm for V_2' is as follows:

1. Randomly *reject* with $k_W - 1/2k_W$ probability by flipping $\lceil \log k_W \rceil + 1$ coins.
2. Continue as V_2 .

The strong error of V_2' is then given by $\varepsilon_{\text{strong}}(V_2') \leq 1/2 - 1/2k_W$.

5.1 Example languages from rNL and potential outsiders

Let w_a denote the amount of symbols a in a string w .

Following two are some example languages with $2\text{nfa}(k_W + k_R)$ recognizers, where $k_W = 0$:

$$\begin{aligned} A_1 &= \{ a^n b^n c^n d^n \mid n \geq 0 \} \\ A_2 &= \{ w \in \{ a, b, c \} \mid w_a = w_b = w_c \} \end{aligned}$$

An example language with a $k_W \leq 1$ recognizer is the following:

$$A_3 = \{ a_1 a_2 \cdots a_n \# a_1^+ a_2^+ \cdots a_n^+ \mid n \geq 0 \}$$

Lastly, an example language that might be outside rNL is follows:

$$A_4 = \{ w \in \{ a, b, c \} \mid w_a \cdot w_b = w_c \}$$

6 Open Questions

It is curious to us whether $\mathcal{L}(2\text{nfa}(k_W + k_R))$ coincides with any known class of languages for $k_W = 0$ or 1 , or $k_W \leq 1$. The minimum number of windable heads required for a language in NL to be recognized by a halting $2\text{nfa}(k)$, could establish a complexity class. Conversely, one might be able to discover yet another infinite hierarchy of languages based on the number of windable heads. For some $c > 0$ and $k'_W = k_W + c$, this hierarchy might be of the form

$$\mathcal{L}(2\text{nfa}(k = k_W + k_R)) \subsetneq \mathcal{L}(2\text{nfa}(k' = k'_W + k'_R))$$

for $k = k'$, $k_R = k'_R$, or without any further restriction.

References

1. Hartmanis, J.: On non-determinacy in simple computing devices. *Acta Informatica* **1**(4), 336–344 (1972). <https://doi.org/10.1007/BF00289513>, <http://link.springer.com/10.1007/BF00289513>
2. Holzer, M., Kutrib, M., Malcher, A.: Complexity of multi-head finite automata: Origins and directions. *Theoretical Computer Science* **412**(1-2), 83–96 (Jan 2011). <https://doi.org/10.1016/j.tcs.2010.08.024>, <https://linkinghub.elsevier.com/retrieve/pii/S0304397510004597>
3. Ibarra, O.H.: On two-way multihead automata. *Journal of Computer and System Sciences* **7**(1), 28–36 (Feb 1973). [https://doi.org/10.1016/S0022-0000\(73\)80048-0](https://doi.org/10.1016/S0022-0000(73)80048-0), <https://linkinghub.elsevier.com/retrieve/pii/S0022000073800480>
4. Monien, B.: Transformational methods and their application to complexity problems. *Acta Informatica* **6**(1), 95–108 (Mar 1976). <https://doi.org/10.1007/BF00263746>, <http://link.springer.com/10.1007/BF00263746>
5. Monien, B.: Two-way multihead automata over a one-letter alphabet. *RAIRO. Informatique théorique* **14**(1), 67–82 (1980). <https://doi.org/10.1051/ita/1980140100671>, <http://www.rairo-ita.org/10.1051/ita/1980140100671>
6. Say, C., Yakaryilmaz, A.: Finite state verifiers with constant randomness. *Logical Methods in Computer Science* **10**(3) (Aug 2014). [https://doi.org/10.2168/LMCS-10\(3:6\)2014](https://doi.org/10.2168/LMCS-10(3:6)2014), <http://arxiv.org/abs/1102.2719>, arXiv: 1102.2719
7. Seiferas, J.I.: Relating refined space complexity classes. *Journal of Computer and System Sciences* **14**(1), 100–129 (Feb 1977). [https://doi.org/10.1016/S0022-0000\(77\)80042-1](https://doi.org/10.1016/S0022-0000(77)80042-1), <https://linkinghub.elsevier.com/retrieve/pii/S0022000077800421>
8. Sudborough, I.H.: Some remarks on multihead automata. *RAIRO. Informatique théorique* **11**(3), 181–195 (1977). <https://doi.org/10.1051/ita/1977110301811>, <http://www.rairo-ita.org/10.1051/ita/1977110301811>