

# COMPLETE AND INCOMPLETE RANDOMIZED NP PROBLEMS

Yuri Gurevich<sup>1</sup>

Electrical Engineering and Computer Science  
University of Michigan, Ann Arbor, MI 48109-2122

## 0 Introduction

A randomized decision problem is a decision problem together with a probability function on the instances. Leonid Levin [Lev1] generalized the NP completeness theory to the case of properly defined randomized NP (shortly, RNP) problems and proved the completeness of a randomized version of the bounded tiling problem with respect to (appropriately generalized) Ptime reductions. Levin's proof naturally splits into two parts; a randomized version of the bounded halting problem is proved complete and then reduced to Randomized Tiling. David Johnson [Jo] provided some intuition behind Levin's definitions and proofs, and challenged readers to find additional natural complete RNP problems.

A randomized version of the bounded Post Correspondence Problem and some other Ptime complete RNP problems are presented in Section 3 of this paper. A natural complete (though not Ptime complete) RNP problem (Randomized Graph Coloring) was recently found by Venkatesan Ramarathnam and his advisor Leonid Levin [Lev2]. We do not know any other announcements of complete RNP problems.

A partial explanation of the difficulty in finding natural Ptime complete RNP problems is given in Section 4. Assuming  $\text{NEXPTIME} \neq \text{DEXPTIME}$ , we prove there the incompleteness of any so-called flat RNP problem with respect to Ptime or even expected Ptime reductions. The natural randomizations of usual NP complete problems very often are flat. For example, consider any RNP graph problem where the conditional probability  $\mu\{G \mid G \text{ has } n \text{ vertices}\}$  is determined

by some edge probability function  $e(n)$ ; if there is a positive  $c < 2$  such that  $n^{-2+c} < e(n) < 1 - n^{-2+c}$  for sufficiently big  $n$  then the problem is flat.

It is important to stress that flat RNP problem are not necessarily easy on average. Randomized Graph Coloring mentioned above and some versions of Randomized Halting and Randomized Tiling are flat but complete for RNP with respect to randomizing (coin-flipping) Ptime reductions introduced and advocated by Levin [Lev2]. If such a problem is decidable in expected Ptime then every RNP problem is decided in expected Ptime by some coin-flipping algorithm. One possible implication of the incompleteness theorem is an increased role of randomizing algorithms in the field of RNP problems.

The sections of this paper not mentioned above are as follows. Section 1 contains basic definitions of the RNP theory, a simplified proof of the Ptime completeness of Randomized Halting is given in Section 2, and in Section 5 we prove the completeness of any RNP companion of any NEXPTIME complete decision problem for the class of sparse RNP problems with respect to expected Ptime reductions.

## 1 Randomized NP Problems

**Definition 1.1.** A randomized decision problem is a pair  $(D, \mu)$  where  $D$  is a decision problem and  $\mu$  is a probability function on the instances of  $D$ .

It is supposed that, in principle, the set of instances of any decision problem  $D$  is the set  $A^*$  of strings in some ordered alphabet  $A$ . The set of yes-instances of  $D$  is the language  $L(D)$  of  $D$ .  $A$ -strings themselves are ordered too: first by length and then lexicographically. The successor of a string  $x$  is

---

<sup>1</sup> Partially supported by NSF grant DCR 85-03275

denoted  $x^+$ . The letter  $A$  is reserved to denote ordered alphabets. The binary alphabet  $\{0,1\}$  is taken to be ordered in the natural way.

A probability function  $\mu$  on a nonempty finite or infinite countable set  $X$  assigns probabilities to elements of  $X$ ;  $\mu$  is positive if every value of  $\mu$  is positive. If  $\mu$  is a probability function on some  $A^*$  then  $\mu^*(x) = \sum_{y < x} \mu y$  is the corresponding probability distribution.

**Convention.** On any nonempty finite set, the uniform probability function is standard. On positive (resp. nonnegative) integers, the standard probability of a number  $n$  is proportional to  $n^{-2}$  (resp.  $(n+1)^{-2}$ ). On any  $A^*$ , the standard probability of a string  $w$  of length  $n$  is proportional to  $(n+1)^{-2} * |A|^n$ . Choosing randomly means choosing with respect to the standard function.

**Definition 1.2.** Let  $\mu$  be a probability function on some  $A^*$ .

(a) A function  $T$  from  $A^*$  to nonnegative reals is polynomial on average with respect to  $\mu$  if there is a positive integer  $k$  such that the expectation  $\sum \mu x * [(Tx)^{1/k} / |x|]$  converges.

(b) A function  $f$  from  $A^*$  to strings in some alphabet is Eptime (for 'expected Ptime') computable with respect to  $\mu$  if some Turing machine computes  $f$  within time which is polynomial on average wrt  $\mu$ . A decision problem over  $A^*$  is Eptime decidable wrt  $\mu$  if its characteristic function is Eptime computable wrt  $\mu$ .

For justification of Definition 1.2(a), see [Lev1], [Jo] or the full version of this paper. Strictly speaking, the sum should be restricted to nonempty strings. Notice that

$$\sum [\mu x * (Tx)^{1/k} * |x|^{-1}] < \infty \text{ if } \sum [\mu x * Tx * |x|^{-k}] < \infty,$$

but the converse is not necessarily true.

**Definition 1.3.** Let  $\mu_1, \mu_2$  be probability functions on some  $A_1^*, A_2^*$  respectively, and  $f$  be a function from  $A_1^*$  to  $A_2^*$ .

(a)  $\mu_2$  dominates (or P-dominates)  $\mu_1$  if  $A_1 = A_2$  and there is a polynomial  $p$  such that  $\mu_2 x * p(|x|) \geq \mu_1 x$ , and  $\mu_2$  EP-dominates  $\mu_1$  if  $A_1 = A_2$  and there is a function  $p$  polynomial on average wrt  $\mu_1$  such that  $\mu_2 x * p(|x|) \geq \mu_1 x$ .

(b)  $f$  transforms  $\mu_1$  into  $\mu_2$  if  $\mu_2 y = \sum_{fx=y} \mu_1 x$ .

(c)  $f$  reduces (or P-reduces)  $\mu_1$  to  $\mu_2$  if some  $\mu$  dominates  $\mu_1$  and is  $f$ -transformed into  $\mu_2$ .  $f$  EP-reduces  $\mu_1$  to  $\mu_2$  if some  $\mu$  EP-dominates  $\mu_1$  and is  $f$ -transformed into  $\mu_2$ .

(d)  $f$  Ptime (resp. Eptime) reduces a randomized decision problem  $(D_1, \mu_1)$  to a randomized decision problem  $(D_2, \mu_2)$  if it reduces  $D_1$  to  $D_2$ , P-reduces (resp. EP-reduces)  $\mu_1$  to  $\mu_2$ , and is Ptime computable (resp. Eptime computable wrt  $\mu_1$ ).

**Lemma 1.1.** (a)  $f$  P-reduces (resp. EP-reduces)  $\mu_1$  to  $\mu_2$  if  $\mu_2$  P-dominates (resp. EP-dominates) the result of the  $f$ -transformation of  $\mu_1$ .

(b) Ptime (resp. Eptime) reductions are closed under composition.

**Definition 1.4.** A function  $f$  from some  $A^*$  to reals is Ptime computable if there exists a Ptime Turing machine which, given an  $A$ -string  $x$  and the unary notation for a natural number  $k$ , computes the binary notation for an integer  $i$  with  $|fx - i/2^k| < 1/2^k$ .

**Definition 1.5.** A randomized decision problem  $(D, \mu)$  is RNP (for 'Randomized NP') if  $D$  is NP and the probability distribution  $\mu^*$  is Ptime computable.

The responsibility for Definition 1.4 lies entirely with us; cf. [Ko]. Levin had in mind probability functions with rational values [Lev2].

**Lemma 1.2.** Every RNP problem Ptime reduces to some RNP problem over the binary alphabet.

**Proof.** If the alphabet  $A$  of the given problem contains at least two letters then the desired reduction assigns the  $n$ -th binary string to the  $n$ -th  $A$ -string. Q.E.D.

**Lemma 1.3.** On binary strings, every probability function  $\mu$  with Ptime computable  $\mu^*$  is dominated by a positive probability function  $\mu_1$  with Ptime computable  $\mu_1^*$  such that every value of  $\mu_1$  is a (finite) binary fraction.

**Proof.** Let  $x$  and  $y$  be binary strings and  $dx = 1/2^{2|x|}$ . Without loss of generality, every  $\mu^*x < 1$ . Since  $\mu^*$  is Ptime computable, some Ptime computable function  $N$  assigns a binary fraction  $Nx = 0.y$  to each binary string  $x$  in such a way that  $|y| \leq 2|x|$  and  $|\mu^*x - Nx| < dx$ . Let  $e$  be the empty string. Define

$$4\mu_1 x = Nx^+ - Nx + 2dx \text{ if } x \neq e, \text{ and } 4\mu_1 e = Ne^+ + 1.$$

$$\text{Then } 4\mu_1 x = Nx^+ - Nx + 2dx > (\mu^*x^+ - dx) - (\mu^*x + dx) + 2dx = \mu x \text{ if } x \neq e, \text{ and } 4\mu_1 e > \mu^*e^+ - de^+ + 1 > \mu e,$$

$$4\mu_1^*x = 1 + Nx + \sum_{e < y < x} 2dy \text{ if } x \neq e, \text{ and}$$

$$\lim_{|x| \rightarrow \infty} 4\mu_1^*x = 1 + 1 + 2\sum_{n \geq 1} 1/2^{-n} = 4. \quad \text{Q.E.D.}$$

## 2 Randomized Halting

**Randomized Halting Problem RH(M)** for a non-deterministic Turing machine M with binary input alphabet is defined as follows.

**Instance:** A binary string  $w01^n$  with  $n > |w|$ .

**Question:** Is there a halting computation of M on w with at most n steps ?

**Probability:**  $(6/\pi^2) * n^{-3} * 2^{-k}$  where  $k = |w|$ .  
(Randomly choose n; randomly choose  $k < n$ ; randomly choose a string w of length k.)

**Theorem 2.1.** For every RNP problem  $(D, \mu)$  there is an NTM M with binary input alphabet such that  $(D, \mu)$  Ptime reduces to RH(M).

**Proof.** By virtue of Lemmas 1.2 and 1.3, we may suppose that the alphabet of D is  $\{0,1\}$  and every value of  $\mu$  is a binary fraction. Since D is NP, there exist an NTM, called the D-machine below, and a polynomial p such that a binary string x is in L(D) if and only if the D-machine has a halting

computation on x if and only if the D-machine has a halting computation on x with at most  $p(|x|)$  steps.

Let  $x'$  be the shortest binary string with  $\mu^*x < 0.x'1 \leq \mu^*(x^+)$ . Then  $0.x'1 - 2^{-|x'|} \leq \mu^*x \leq \mu^*(x^+) < 0.x'1 + 2^{-|x'|}$ , and therefore  $2*2^{-|x'|} > \mu x$ . Set  $x'' = [ \text{if } 2^{-|x'|} > \mu x \text{ then } 0x, \text{ else } 1x' ]$ . Then  $2*2^{-|x''|} > \mu x$ .

Given a binary bit b followed by a string w, the desired NTM M does the following:

- (1) If  $b = 0$  then  
if  $2^{-|w|} \leq \mu w$  then loop else set  $x = w$  and go to (4).
- (2) Find the unique x with  $\mu^*x < 0.w1 \leq \mu^*(x^+)$ .

(3) If  $2^{-|x|} > \mu x$  or  $x' \neq w$  then loop.

(4) Simulate the D-machine on x.

There exist a polynomial q such that M has a halting computation on  $x''$  with at most  $q(|x|)$  steps if and only if M has a halting computation on  $x'$  if and only if x is in L(D). The desired reduction is  $fx = x''01^{q(|x|)}$ . The probability function of RH(M) dominates the result of f-transformation of  $\mu$ :

$$(\pi^2/3) * |fx|^3 * \text{Prob}(fx) > 2*2^{-|x''|} > \mu x. \quad \text{Q.E.D.}$$

### Randomized Halting Problem:

**Instance:** An NTM M with binary input alphabet and an instance  $w01^n$  of RH(M).

**Question:** Is there a halting computation of M on w with at most n steps?

**Probability:** Choose M with respect to your favorite positive Ptime computable probability distribution, and then choose an instance of RH(M) as above.

### Corollaries.

- (1) Randomized Halting is Ptime complete for RNP.
- (2) There is an NTM M with binary input alphabet such that RH(M) is Ptime complete for RNP.

## 3 Randomized Post Correspondence Problem et al.

The bounded PCP is NP complete [CHS]. We define

### Randomized Post Correspondence Problem (RPCP):

**Instance:** A nonempty list  $(u_1, v_1), \dots, (u_k, v_k)$  of pairs of binary strings, and the unary notation for a positive integer n.

**Question:** Are there  $m \leq n$  and a function  $F: \{0, \dots, m\} \implies \{1, \dots, k\}$  with  $u_{F0} \dots u_{Fm} = v_{F0} \dots v_{Fm}$  ?

**Probability:** Randomly choose k and n; then randomly (and independently) choose 2k binary strings.

**Theorem 3.1.** RPCP is Ptime complete for RNP.

**Corollary.** The following RANDOMIZED PALINDROME PROBLEM is Ptime complete for RNP :

**Instance:** A context-free grammar with productions  
 $S \rightarrow u_1 S v_1 \mid \dots \mid u_k S v_k \mid e$ ,  
 and the unary notation for a positive integer  $n$ .  
 ( $u_i, v_i$  are binary strings;  $e$  is the empty string.)

**Question:** Is it possible to derive a nonempty palindrome in at most  $n$  steps ?

**Probability:** Randomly choose  $k$  and  $n$ ; then randomly (and independently) choose  $2k$  binary strings.

**Theorem 3.2.** The following problems are Ptime complete for RNP.

(a) A RANDOMIZED VERSION OF SATISFIABILITY IN FINITE ARITHMETIC (Gurevich and Shelah):

**Instance:** A formula  $F(P)$  in the first-order language of arithmetic augmented with the unary predicate symbol  $P$ , the unary notation for a positive integer  $n$ , a positive integer  $k \leq n$ , and a unary relation  $R$  on  $\{0, \dots, k-1\}$ .

**Question:** Is there an extension  $R'$  of  $R$  to  $\{0, \dots, n-1\}$  such that  $F(R')$  holds in the arithmetic modulo  $n$  ?

**Probability:** Choose  $F(P)$  with respect to your favorite positive Ptime computable probability distribution, randomly choose  $n$ , randomly choose  $k$ , randomly choose  $R$ .

(b) A RANDOMIZED VERSION OF SATISFIABILITY OF FIRST-ORDER FORMULAS:

**Instance:** A first-order sentence  $F(P)$ , the unary notation for a positive integer  $n$ , a positive integer  $k \leq n$ , and a unary relation  $R$  on  $\{0, \dots, k-1\}$ .

**Question:** Is there a model for  $F(P)$  on  $\{0, \dots, n-1\}$  with the interpretation of  $P$  extending  $R$  ?

**Probability:** Similar to that in (a).

**Remark.** Utilizing the undecidability proofs for different fragments of first-order logic, one can put severe restrictions on the sentence  $F(P)$  in (b).

## 4 Incompleteness

**Definition 4.1.** A probability function  $\mu$  on strings in some alphabet is flat if there exists a positive real  $c$  such that  $\mu x \leq 2^{-|x|^c}$ , i.e.  $-\log_2 \mu x \geq |x|^c$ , for all sufficiently long  $x$ . A randomized decision problem  $(D, \mu)$  is flat if  $\mu$  is.

The intuition behind the definition is that all values of a flat probability function are small; none of them juts out.

In this section the term "exponential" and its relatives are used in the broader sense. A function  $T$  from some  $A^*$  to non-negative reals is exponential if there is a polynomial  $p$  with  $T(x) \leq 2^{p(|x|)}$ . A decision problem  $D$  is DEXPtime (resp. NEXPtime) decidable if there is an exponential-time deterministic (resp. nondeterministic) Turing machine that recognizes  $L(D)$ .

**Theorem 4.1.** Let  $(D, \mu)$  be a flat randomized decision problem such that  $D$  is DEXPtime (e.g. NP). If  $(D, \mu)$  is EPTIME hard for RNP then NEXPtime = DEXPtime.

**Proof.** We assume that  $(D, \mu)$  is EPTIME hard for RNP and show that an arbitrary NEXPtime decision problem  $D_0$  is DEXPtime decidable. Without loss of generality, the alphabet of  $D_0$  is  $\{0,1\}$ .

Fix a polynomial  $p$  such that some  $2^p$ -time-bounded NTM recognizes  $L(D_0)$ . Let  $x$  range over nonempty binary strings,  $n = |x|$  and  $x'$  be the string  $x0$  followed by  $2^{p(n)}$  occurrences of 1. Consider a randomized decision problem  $(D_1, \mu_1)$  where  $L(D_1) = \{x' : x \text{ belong to } L(D_0) \text{ and } \mu_1 x' = (6/\pi^2) * n^{-2} * 2^{-n}\}$ .

$(D_1, \mu_1)$  is RNP; hence there is an EPTIME reduction  $f$  of  $(D_1, \mu_1)$  to  $(D, \mu)$ . This gives the following decision algorithm for nonempty instances  $x$  of  $D_0$ : compute  $y = fx'$ , and then check whether  $y$  belongs to  $L(D)$ . We prove the decision algorithm to be EXPTIME.

Firstly, we show that  $y$  is computable from  $x$  in time exponential in  $n$ . It suffices to show that  $y$  is computable from  $x'$  in time exponential in  $n$ . Since  $f$  is EPTIME computable wrt  $\mu_1$ , one can compute  $y$  from  $x'$  within time  $T(x')$  polynomial on average wrt  $\mu_1$ . We have

$$\sum [\mu_1 x' * (T x')^{1/i} * |x'|^{-1}] \leq a < \infty \text{ for some } i \text{ and } a.$$

$$\text{Then } [\mu_1 x' * (T x')^{1/i} * |x'|^{-1}] \leq a,$$

and  $T(x') \leq [(a\pi^2/6) * n^2 * 2^n * (n + 1 + 2^{p(n)})]^i$ .

Secondly, we show that the question whether  $y$  belongs to  $L(D)$  is decidable in time exponential in  $n$ . Let  $m = |y|$ . Since  $D$  is DEXptime, it suffices to show that  $m$  is bounded by a polynomial of  $n$ . Since  $\mu$  is flat, it suffices to show that  $-\log_2 \mu y$  is bounded by a polynomial of  $n$ .

Since  $f$  EP-reduces  $\mu_1$  to  $\mu$ , there is a probability distribution  $\mu_2$  which EP-dominates  $\mu_1$  and is  $f$ -transformed into  $\mu$ . Hence,  $\mu y \geq \mu_2 x'$  and there is a function  $p(x')$ , polynomial on average wrt  $\mu_1$ , such that  $\mu_2 x' * p(x') \geq \mu_1 x'$ . Thus,  $\mu y \geq (px')^{-1} * \mu_1 x'$  and  $-\log_2 \mu y \leq \log_2(px') - \log_2 \mu_1 x'$ . Since  $-\log_2 \mu_1 x'$  is bounded by a polynomial of  $n$ , it remains to prove that  $\log_2(px')$  is bounded by a polynomial of  $n$ .

But  $p$  is polynomial on average wrt  $\mu_1$ ; hence there are  $j$  and  $b$  such that  $\sum (px')^{1/j} * |x'|^{-1} * \mu_1 x' = b < \infty$ . Then  $(px')^{1/j} * |x'|^{-1} * \mu_1 x' < b$  and  $(1/j) * \log_2(px')$  is bounded by  $\log_2 b + \log_2 |x'| - \log_2 \mu_1 x'$  which is bounded by a polynomial of  $n$ . Q.E.D.

**Examples.** (1) Let  $M$  be an NTM  $M$  with binary input alphabet and  $i$  be a positive integer. The restriction of the randomized halting problem for  $M$  to inputs  $w01^n$  with  $|w| \geq n^{1/i}$  is flat.

(2) Let  $i$  be a positive integer. The restriction of Randomized Tiling Problem [Lev1] to inputs where the length of the given portion of the first row is at least the  $i$ -th root of the row length, is flat.

(3) Let  $\mu$  be any probability function on graphs such that the conditional probability  $\mu\{G \mid G \text{ has } n \text{ vertices}\}$  is determined by some edge probability function  $e(n)$ :

$$\mu\{G \mid G \text{ has } n \text{ vertices}\} = [e(n)]^m * [1 - e(n)]^{n(n-1)/2 - m}$$

where  $m$  is the number of edges of  $G$ . If there is a positive real  $c < 2$  with  $n^{-2+c} < e(n) < 1 - n^{-2+c}$  for each sufficiently big  $n$  then  $\mu$  is flat.

**Corollary.** Let  $\mu$  be any probability function on graphs such that the conditional probability  $\mu\{G \mid G \text{ has } n \text{ vertices}\}$  is determined by some edge probability function  $e(n)$ . The  $\mu$ -randomization of Hamiltonian Path Problem is not Eptime complete for RNP unless NEXptime = DEXptime.

Flat RNP problems are not necessarily easy on average. Contrast the first two examples with the following easy modifications of Ptime completeness results discussed earlier.

**Theorem 4.2.** (a) There is an NTM  $M$  with binary input alphabet such that the restriction of  $RH(M)$  to inputs  $w01^n$  with  $n = |w|+1$  is complete for RNP with respect to RPtime (randomizing Ptime) reductions.

(b) The restriction of Randomized Tiling Problem to inputs where the given portion of the first row is the whole row, is RPtime complete for RNP.

## 5 Sparse problems

**Definition 5.1.** A probability function  $\mu$  on strings in some alphabet is sparse if there is a polynomial  $p$  with  $|\{x: \mu x > 0 \text{ and } |x| = n\}| < p(n)$ . A randomized decision problem  $(D, \mu)$  is sparse if  $\mu$  is sparse.

If the definition is tightened by requiring  $D$  to be sparse as well, Theorem 5.1 below will remain true.

In this section, the term 'exponential' and its relatives are used in the narrow sense. A function  $T$  from some  $A^*$  to nonnegative reals is exponential if there is a constant  $c$  with  $T(x) \leq c^{|x|}$ . A function  $f$  is EXptime computable if some exponential-time Turing machine computes  $f$ . A decision problem  $D$  is NEXptime if some exponential-time NTM recognizes  $L(D)$ . A NEXptime decision problem  $D$  EXptime reduces [Lew] to a NEXptime decision problem  $E$  if there exist a constant  $c$  and an exponential-time computable function  $f$  such that  $x$  belongs to  $L(D)$  if and only if  $fx$  belongs to  $L(E)$ , and  $|fx| \leq c|x|$ .

**Definition 5.2.** Let  $D$  be a NEXptime decision problem over strings in an  $m$ -letter alphabet. An exponential function  $gn = c^n$  with an integer  $c > 1$  is a guard for  $D$  if some  $g$ -time-bounded NTM accepts  $L(D)$ . The RNP companion of  $D$  wrt a guard  $g$  is the randomized decision problem  $(E, \mu)$  such that

$$L(E) = \{w01^{g|w|}: w \text{ belongs to } L(D)\} \text{ and}$$

$$\mu(w01^{g|w|}) = (6/\pi^2) * (|w| + 1)^{-2} * m^{-|w|}.$$

The companion is a sparse RNP problem. The requirement that  $c$  is integer can be relaxed of course.

**Lemma 5.1.** (a) Suppose that  $(D_1, \mu_1)$  is an RNP companion of some NEXPTIME decision problem, and  $F$  is a function from some  $A^*$  to  $\{y: \mu_1 y \neq 0\}$ . Then  $F$  P-reduces any probability function  $\mu$  on  $A^*$  to  $\mu_1$ .

(b) If a NEXPTIME decision problem  $D_1$  EXPTIME reduces to a NEXPTIME decision problem  $D_2$  then any RNP companion of  $D_1$  PTIME reduces to any RNP companion of  $D_2$ .

**Proof.** (a) There is a polynomial  $p$  such that if  $\mu_1 y \neq 0$  then  $p(|y|) * \mu_1 y > 1$ .

(b) Let  $c$  and  $f$  witness the EXPTIME reducibility, and  $(E_1, \mu_1)$  be an RNP companion of  $D_1$ . Let  $x$  be an instance of  $D_1$ ,  $x'$  be the corresponding instance of  $E_1$ ,  $y = fx$ , and  $y'$  be the corresponding instance of  $E_2$ . The function  $F(x') = y'$  reduces  $E_1$  to  $E_2$ . By (a),  $F$  reduces  $\mu_1$  to  $\mu_2$ .  $F(x')$  is computable in time exponential in  $|x| + |y|$ . Since  $|y| \leq c|x|$ ,  $F(x')$  is computable in time exponential in  $|x|$ , hence in time polynomial in  $|x'|$ . Q.E.D.

**Theorem 5.1.** Let  $E$  be any decision problem EXPTIME complete for NEXPTIME. Any RNP companion of  $E$  is EPTIME complete for the class of sparse RNP problems.

**Proof.** For every NTM  $M$  with binary input alphabet, define the exponential halting problem  $EH(M)$  for  $M$  as follows:

**Instance:** A binary string  $w$ .

**Question:** Is there a halting computation of  $M$  on  $w$  with at most  $2^{|w|}$  steps?

It suffices to prove that for every sparse RNP problem  $(D, \mu_0)$  there is some NTM  $M_1$  with binary input alphabet such that  $(D, \mu_0)$  EPTIME reduces to an RNP companion of  $EH(M_1)$ . For, by Lemma 5.1, this companion of  $EH(M_1)$  EPTIME reduces to the designated companion  $E_0$  of  $E$ , and therefore  $(D, \mu_0)$  EPTIME reduces to  $E_0$ .

Without loss of generality, instances of  $D$  are binary strings. By the proof of Lemma 1.3, there is a probability function  $\mu$  with PTIME computable  $\mu^*$  such that values of  $\mu$  are binary fractions and  $4\mu x > \mu_0 x$ . The rest of the proof is similar to that of Theorem 2.1. Let the  $D$ -machine,  $x''$  and  $M$  be as in the proof of Theorem 2.1. Given an input  $bu0v$ , where  $b$  is a binary bit and  $v$  is a string of 1's, the desired machine  $M_1$

simulates  $M$  on  $bu$ . There is a polynomial  $q$  such that if  $x$  is a binary string of some length  $n$  and  $v$  is a string of 1's of length  $\log_2 q(n)$  then  $x$  belongs to  $L(D)$  if and only if there is a halting computation of  $M_1$  on  $x''0v$  if and only if there is a halting computation of  $M_1$  on  $x''0v$  of length at most  $q(n)$ .

Let  $F(x) = x''01^i01^j$  where  $i = \log_2 q(|x|)$  and  $\log_2 j = |x''| + 1 + i$ . We show that  $F$  EPTIME reduces  $(D, \mu_0)$  to the RNP companion  $(D_1, \mu_1)$  of  $EH(M_1)$  wrt guard  $gn = 2^n$ .  $Fx$  belongs to  $L(D_1)$  if and only if there is a halting computation of  $M_1$  on  $x''01^i$  of length at most  $j$  if and only if there is a halting computation of  $M_1$  on  $x''01^i$  of length at most  $q|x|$  if and only if  $x$  belongs to  $L(D)$ . Thus,  $F$  reduces  $D$  to  $D_1$ . By Lemma 5.1(a),  $F$  reduces  $\mu_0$  to  $\mu_1$ .

To show that  $F$  is EPTIME computable, it suffices to check that the function  $q(|x|) * 2^{|x''|}$  is polynomial on average wrt  $\mu_0$ .

$$\sum [\mu_0 x * [q(|x|) * 2^{|x''|}]^{1/k} * |x|^{-1}] < \infty \quad \text{if}$$

(by the remark after Definition 1.2)

$$\sum [\mu_0 x * q(|x|) * 2^{|x''|} * |x|^{-k}] < \infty \quad \text{if}$$

$$\sum \{[\mu_0 x * q(|x|) * 2^{|x''|} * |x|^{-k}]: \mu_0 x > 0\} < \infty \quad \text{if}$$

(since  $4\mu x > \mu_0 x$ )

$$\sum \{[\mu x * q(|x|) * 2^{|x''|} * |x|^{-k}]: \mu_0 x > 0\} < \infty \quad \text{if}$$

(since  $2 * 2^{|x''|} > \mu x$ )

$$\sum \{[q(|x|) * |x|^{-k}]: \mu_0 x > 0\} < \infty.$$

But the last inequality is true for sufficiently large  $k$  depending on  $q$  and the polynomial witnessing that  $\mu_0$  is sparse. Q.E.D.

Some NEXPTIME complete problems can be found in [KV] and [Lew].

**Acknowledgements.** Discussions with Andreas Blass, David Harel, Saharon Shelah, Quentin Stout, Martin Tompa and especially Leonid Levin were very enjoyable and very helpful. A preliminary version of this paper appeared in lecture notes for my theory class of Winter 1987; the reaction of the students - Murtaza Doctor, Shantanu Dutt, David Eliezer, David McCauley, Philip MacKenzie, Brian Murray, and Durriya Netterwala - was very helpful as well.

## References

- CHS Robert L. Constable, Harry B. Hunt, III, and Sartaj K. Sahni, On the Computational Complexity of Scheme Equivalences, Report 74-201, 1974, Dept. of Computer Science, Cornell University, Ithaca, NY.
- Jo David S. Johnson, The NP-Completeness Column, Journal of Algorithms 5 (1984), 284-299.
- Ko Ker-I Ko, On the Definition of Some Complexity Classes of Real Numbers, Math. Systems Theory 16 (1983), 95-109.
- KV Phokion G. Kolaitis and Moshe Y. Vardi, The Decision Problem for the Probabilities of Higher-Order Properties, STOC 1987.
- Lev1 Leonid Levin, Average Case Complete Problems, SIAM Journal of Computing, 1986.
- Lev2 Leonid Levin, Private communication, July 1987.
- Lew Harry R. Lewis, Complexity Results for Classes of Quantificational Formulas, J. Computer and System Sciences 21 (1980), 317-353.