Precise Average Case Complexity

Rüdiger Reischuk and Christian Schindelhauer

Technische Hochschule Darmstadt*

Abstract. A new definition is given for the average growth of a function $f : \Sigma^* \to \mathbb{N}$ with respect to a probability measure μ on Σ^*. This allows us to define meaningful average case distributional complexity classes for arbitrary time bounds (previously, one could only distinguish between polynomial and superpolynomial growth). It is shown that basically only the ranking of the inputs by decreasing probabilities are of importance.
To compare the average and worst case complexity of problems we study average case complexity classes defined by a time bound and a bound on the complexity of possible distributions. Here, the complexity is measured by the time to compute the rank functions of the distributions. We obtain tight and optimal separation results between these average case classes. Also the worst case classes can be embedded into this hierarchy. They are shown to be identical to average case classes with respect to distributions of exponential complexity.
These ideas are finally applied to study the average case complexity of problems in $\mathcal{NP}$. A reduction between distributional problems is defined for this new approach. We study the average case complexity class $\mathcal{A}v\mathcal{P}$ consisting of those problems that can be solved by DTMs on the average in polynomial time for all distributions with efficiently computable rank function. Fast algorithms are known for some $\mathcal{NP}$–complete problems under very simple distributions. For langugages in $\mathcal{NP}$ we consider the maximal allowable complexity of distributions such that the problem can still be solved efficiently by a DTM, at least on the average. As an example we can show that either the satisfiability problem remains hard, even for simple distributions, or $\mathcal{NP}$ is contained in $\mathcal{A}v\mathcal{P}$, that means every problem in $\mathcal{NP}$ can be solved efficiently on the average for arbitrary not too complex distributions.

1 Introduction and Overview

Levin observed that a sound definition of average case complexity and complexity classes is not at all obvious ([Levi86]). The classical notion of average-case time complexity of a machine M with respect to given probability distributions μ_n on inputs x of length n takes the expectation

$$\boldsymbol{Time}_M^{\mu}(\boldsymbol{n}) \;:=\; \sum_{|x|=n} \mu_n(x) \cdot time_M(x) \;,$$

* Institut für Theoretische Informatik, Alexanderstraße 10, 6100 Darmstadt, Germany
email: reischuk/schindel @iti.informatik.th-darmstadt.de

where $\text{time}_M(x)$ denotes the running time of M on x and $\mu := \mu_1, \mu_2, \ldots$. The machine M is **μ–average T–time bounded** (in the expected sense) for a resource bound $T : \mathbb{N} \to \mathbb{N}$, if $Time_M^\mu \leq T$, that means for all n

$$\sum_{|x|=n} \mu_n(x) \cdot \frac{time_M(x)}{T(|x|)} \;\leq\; 1 \,.$$

The problem with this definition is that polynomial time simulations of polynomial average time machines can result in superpolynomial average time complexity. It was resolved by Levin by applying the inverse of T to the fraction, thus requiring

$$\sum_{|x|=n} \mu_n(x) \cdot \frac{T^{-1}(\text{time}_M(x))}{|x|} \;\leq\; 1 \,.$$

This definition does not take into account that the weights of different input length may be very unequal. Thus one considers only distributions μ defined over the whole set of inputs and requires

$$\sum_{x} \mu(x) \cdot \frac{T^{-1}(time_M(x))}{|x|} \;\leq\; 1 \,.$$

M is then called (Levin)**–μ–average T–time bounded**. For a discussion of this approach see the detailed exposition in [Gure91].

Still there remains an unpleasant property, the influence of the functional growth of $\mu(x)$ on the time bound T. If, for example, one takes the "standard" *uniform probability distribution,* which assigns probability $\mu_{\text{uniform}}(x) := 6/\pi^2 \cdot |x|^{-2} \cdot 2^{-|x|}$ to a string $x \in \{0,1\}^*$ a machine using n^2 steps on every input of length n would already be average $O(n^{1+\epsilon})$–time bounded for arbitrary $\epsilon > 0$. This problem can be resolved to a certain extent (see [Gure91]), but not completely.

Our first contribution to the average case analysis will be a new definition of average T–time bounded, which gets rid of this problem. It will allow us to differenciate between bounds T_1 and T_2 for any $T_1 \leq o(T_2)$. The idea is to bound the complexity of a machine not only with respect to the probability distribution μ, but with respect to all monotone transformations of μ. At first glance, it seems that this complicates the analysis even more. But we will show that this larger set of conditions is equivalent to a very simple property of the distribution μ, which does not involve probabilities explicitly anymore. The only thing that matters is the ranking of the inputs by μ, that is the sequence of inputs ordered by decreasing probabilities.

In practice, one often does not know the values of the distribution exactly, but for each pair of inputs at least one can decide which input is more likely. This way, the whole analysis is greatly simplified. Each ranking of the input space describes a whole equivalence class of distributions, and we get rid of the influence of the asymptotic growth of the probability measure.

A **distributional problem** is a pair (L, μ), consisting of a language $L \subseteq \{0,1\}^*$ and a probability distribution μ on $\{0,1\}^*$. We define *distributional complexity classes*

$\boldsymbol{DistDTime(T)}$ containing all pairs (L, μ), for which there exists a DTM accepting L that is μ–average T–time bounded in this generalized sense.

Given a language $L \in DTime(T)$ and a DTM M for L, it is easy to see that by cycling through all inputs of length n one can find an x, on which M spends the maximal time for inputs of length n. If a probability measure μ gives all its weight for inputs of length n to this x then the average time of M (in the expected sense) with respect to this μ equals the worst case complexity. $\mu(x)$ can be computed in time $O(2^{|x|} \cdot T(|x|))$. Using this idea, Miltersen has shown that allowing exponential time overhead a measure μ can be constructed that is *malign* for all expected T–time bounded machines ([Milt91]). That means their expected time complexity with respect to this μ is no more than a constant factor smaller than their worst case complexity.

On the other hand, restricting an average case analysis to some simple distributions may yield results with little practical value. The satisfiability problem, for example, has been shown quickly solvable for certain symmetric distributions, but the input space generated this way seems to be of not interest for applications in AI (see for example the discussion in [MiSeLe92]). These observation motivate to consider *average case complexity classes* $\boldsymbol{AvDTime(T, C)}$ consisting of all languages L that can be recognized in μ–average time T for distributions μ of complexity at most C, for certain bounds C. That way, average case complexity classes are directly comparable to the standard worst case classes, because both contain only languages.

In this paper a notation different from the one in previous research on average case complexity will be used, because we feel that this new one is more appropriate and natural. There should be a clear distinction between distributional classes, where distributions appear explicitly, and average case classes the elements of which are languages in the usual sense. From a complexity theoretic point of view one is more interested in the second kind of classes.

The complexity of a distribution is taken w.r.t. its **rankability**, that is the effort to compute the rank of an input x. Previous approaches have bounded the complexity of distributions using the notion of **computable** and **sampleable**. A distribution is POL-computable if the sum of all weights of inputs lexicographically lower than x can be computed in polynomial time w.r.t. the length of x and the binary expansion ([Levi86],[Gure91]). A distribution μ is POL-samplable, if there exists a randomized algorithm that outputs the string x with probability $\mu(x)$ in polynomial time w.r.t. $|x|$ ([BCGL92]). These concepts are not directly comparable to rankability, a discussion of their relation will be given in the full version of this paper.

As an analog to the worst case class $\mathcal{P}$ the average case complexity class

$$\mathcal{A}v\mathcal{P} \;:=\; AvDTime(\text{POL}, \text{POL–rankable})$$

seems to be the most natural candidate. Problems in this class are efficiently solvable in practice, because for all not too complex distributions their average time complexity is bounded by a polynomial. Our second main contribution is a tight separation and inclusion results for average case complexity classes within $\mathcal{A}v\mathcal{P}$.

Finally, we consider reductions between distributional problems and relations between nondeterministic and average case complexity classes. Of particular interest

are distributional problems (L, μ) such that $L \in \mathcal{NP}$ and the complexity of μ is polynomially bounded (again, we consider here the rankability). In Levin's model this class has been called distributional $\mathcal{NP}$ or randomized $\mathcal{NP}$, but both notions are somehow misleading (distributional $\mathcal{NP}$ should better be used for the class $DistNTime(\mathrm{POL})$).

For a meaningful reduction between distributional problems one needs an additional property called *domination* (see [Levi86] or [Gure91]). In our model this becomes a simple condition on the transformation of the ranks between two probability distributions. Similar to the previous models one can show that the bounded halting problem for NTM together with a natural ranking is complete for distributional problems taken from $\mathcal{NP}$.

Finally, we discuss the relation between $\mathcal{NP}$ and $\mathcal{AVP}$. To analyse the average case behaviour of problems in $\mathcal{NP}$ we propose to classify them w.r.t. the largest amount of rankability one can allow such that the average time complexity stays polynomial. We call this the *nose* of a problem.

Some $\mathcal{NP}$–complete problems are known that can be solved very fast on the average for simple distributions. Examples are 3-colourability of graphs or Hamiltonian circuits (for a discussion and references see [John84] and [Gure91]). If a problem is complete in the sense above simple distributions, which might yield a polynomial time complexity on the average, probably do not exist. Otherwise, by a result of Ben-David and Luby (see [Gure91]) deterministic and nondeterministic exponential time would be identical. Our last result shows that satisfiability has no nose, that means it will require superpolynomial time for almost all distributions, unless $\mathcal{NP} \subseteq \mathcal{AVP}$.

Most of the proofs have to be omitted in this short report. For a complete version see [ReSc92], some of the results can already be bound in [Schi91].

2 Notations

Let $\mathcal{N}$ denote the identity function on the natural numbers $\mathbb{N}$. A complexity bound is a function $T : \mathbb{N} \to \mathbb{N}$. All complexity bounds in this paper are assumed to be monotone increasing and time-constructible. The following sets of complexity bounds will be of special interest: $\mathrm{POL} := \bigcup_{k \in \mathbb{N}} O(\mathcal{N}^k)$, $\mathrm{EXL} := \exp \Theta(\mathcal{N})$ and $\mathrm{EEXL} := \exp\exp \Theta(\mathcal{N})$. For a complexity bound T, which does not necessarily have to be injective, we define the inverse T^{-1} by

$$T^{-1}(m) \;:=\; \min\{n \mid T(n) \geq m\} \; .$$

Let $M_1, M_2, \ldots$ be an enumeration of all deterministic Turing machines (in some cases we also consider nondeterministic machines). We may assume that all machines have only 2 work tapes, implying that one can use a universal machine with only a constant factor slowdown.

When talking about an ordering of binary strings, $x \leq y$ we refer to the lexicographical ordering. We consider probability measures (density functions) $\mu : \Sigma^* \to [0, 1]$ over the input space. μ has to satisfy $\sum_x \mu(x) \leq 1$. $\mathrm{bin} : \mathbb{N} \to \{0,1\}^*$ denotes the standard correspondence between binary strings and natural numbers.

3 Refinement of Levin's Average Case Measure

In the introduction we have already discussed the problem to measure precisely the average complexity of a time bound T with respect to a probability distribution μ. Levin's solution essentially can only distinguish between polynomial and superpolynomial growth.

Definition 1 *The pair (f, μ) consisting of a function $f : \Sigma^* \to \mathbb{N}$ and a distribution μ belongs to the class $\boldsymbol{l Av}(\mathbf{POL})$ with respect to a distribution μ iff for some number k*

$$\sum_x \mu(x) \frac{f(x)^{1/k}}{|x|} < \infty \ .$$

The problem with the standard uniform distribution mentioned above can somehow be diminished, by giving $\{0,1\}^n$ a total weight proportional to $n^{-1} \cdot \log^{-2} n$ or even less, instead of n^{-2}. Still, it can never be resolved completely. Below, we will present a precise average case measure. The idea is to consider simultaneously all distributions $\tilde{\mu}$ that yield the same ordering of inputs by decreasing probabilities as μ, that means if $\mu(10001) < \mu(11)$ then $\tilde{\mu}(10001) \leq \tilde{\mu}(11)$. Thus, only the ranking of the inputs by decreasing weights matters.

Definition 2 $\quad \mathbf{rank}_\mu(x) \ := \ |\{z \in \Sigma^* \mid \mu(z) \geq \mu(x)\}| \ .$

μ–average bounded by T will then defined to be $\tilde{\mu}$–average bounded by T in the sense above for all such $\tilde{\mu}$. The set of such $\tilde{\mu}$ can be generated by *monotone transformations* of μ.

Definition 3 *A real-valued monotone function $m : [0,1] \to [0,1]$ is called a* **monotone transformation** *of the distribution μ if $\sum_x m(\mu(x)) \leq 1$.*
The set $\mathbf{Av}(\boldsymbol{T})$ *contains all pairs (f, μ) consisting of a function $f : \Sigma^* \to \mathbb{N}$ and a distribution μ such that for all monotone transformations m of μ*

$$\sum_x m(\mu(x)) \frac{T^{-1}(f(x))}{|x|} \leq 1 \ .$$

Because of the universal quantifier over all monotone transformations the above definition for $\mathrm{Av}(T)$ is even more complicated than the one given by Levin. But there exists an equivalent, very simple characterization of $\mathrm{Av}(T)$. Consider the special case of *threshold functions* $\mathtt{thr}_l : [0,1] \to [0,1]$ as monotone transformations, where for $l = \mathrm{rank}_\mu(x)$ we define $\mathtt{thr}_l(z) := 1/l$ if $z \geq \mu(x)$ and 0 else.

Lemma 1

$$(f, \mu) \in Av(T) \qquad \Longleftrightarrow \qquad \forall l \quad \sum_x \mathtt{thr}_l(\mu(x)) \ \frac{T^{-1}(f(x))}{|x|} \ \leq \ 1 \ .$$

As an immediate consequence of this lemma we obtain the following fundamental result, which shows that an average bound can be computed without considering all possible transformations.

Proposition 1

$$(f,\mu) \in Av(T) \qquad \Longleftrightarrow \qquad \forall l \quad \sum_{\mathrm{rank}_\mu(x)\leq l} \frac{T^{-1}(f(x))}{|x|} \;\leq\; l \;.$$

In the following we will only use this characterization to verify membership in $\mathrm{Av}(T)$. This generalization keeps polynomial bounds, which are increased by a polynomial, in that class. Each rank function represents a whole set of distributions, namely those which are equivalent with respect to the definition of the sets $\mathrm{Av}(T)$.

We are now ready to define the following **distributional complexity classes**.

Definition 4

$$\boldsymbol{DistDTime(T)} := \{(L,\mu) |\, \exists \; DTM\; M \; with\; L(M) = L,\; (time_M, \mu) \in Av(T)\} \;,$$
$$\boldsymbol{Dist\mathcal{P}} := \bigcup_{T\in \mathrm{POL}} DistDTime(T) \;.$$

Note that in this setting already for simple distributions like the uniform one there can only be an exponential difference between the distributional and the worst case complexity of a problem. That means, if $(L, \mathrm{uniform}) \in Dist\mathcal{P}$ then $L \in DTime(\mathrm{EXL})$. For Levin's model separation results between polynomial distributional complexity classes are given in [WaBe92], but the separating languages are in classes higher than exponential time. The technical trick to achieve this is to let the uniform probabilities for inputs of length n converge very fast to 0 with n. Such a separation result does not seem to yield much insight into average case complexity.

4 Hierarchies of Average Case Complexity Classes

Since for this new average case measure all essential information of a distribution is the rank function we will identify both in the following. Thus (L,μ) and (L, rank_μ) denote the same distributional problem. The straightforward way to restrict distributions is a time limit for computing the rank.

Definition 5 *Let* $\boldsymbol{T}$**–rankable** *be the set of all distributions* μ *for which there exists a DTM* M *that on input* x *computes* $bin(rank_\mu(x))$ *in time* $T(|x|)$.

In order to compare the worst case and the average case complexity of problems we consider the distributional complexity of languages L with respect to a set C of distributions and define

Definition 6

$$\boldsymbol{AvDTime(T,C)} := \{L \mid \forall \mu \in C\ (L,\mu) \in DistDTime(T)\}\ ,$$
$$\boldsymbol{Av\mathcal{P}} := AvDTime(POL, POL\text{–}rankable)\ .$$

Let us first show that for complex distributions there is no difference between the average and the worst case complexity. We will construct a rank function that for any DTM M with $L(M) \notin DTime(T)$ gives small ranks ρ to inputs with long computations, thus (L,ρ) does not belong to $DistDTime(T)$. For this purpose, the following $\mathcal{NP}$–complete language is helpful.

Definition 7

$$\mathbf{H_T} := \{(w,1^i) \mid M_i \textit{ is a DTM and } \exists\, x \leq w \textit{ with } time_{M_i}(x) > T(|x|)\}\ .$$

Let $\mathbf{h(T)}\ \geq\ \Omega(T^2)$ *be a time bound such that* $H_T \in DTime(h(T))$.

Obviously, $h(T)$ is of order at most $2^n \cdot T(n)$ (remember that all bounds were assumed to be monotone).

Theorem 1 *For all* $T \geq \mathcal{N}$ *and for all* $\delta > 1$ *holds*

$$AvDTime(T, h(T)\text{–}rankable)\ \subseteq\ DTime(T(\delta\mathcal{N})).$$

Since the rank functions of these distributions can be computed in time $T \cdot \mathrm{EXL}$, compared to the worst case, no machine works significantly faster on the average with respect to this set of distributions.

Corollary 1 *For all* $T \in POL$*:*

$$AvDTime(T, (T \cdot EXL)\text{–}rankable) = DTime(T)\ .$$

Miltersen has shown that there exists a distribution μ malign for $DTime(\mathcal{N}^k)$, which can be computed in polynomial time with an $\Sigma_2^{\mathcal{P}}$–oracle ([Milt91]). The proof of this theorem yields that for the more general situation we consider here already an $\mathcal{NP}$–oracle suffices.

The equality above cannot be generalized to arbitrary large time bounds T. This has technical reasons when taking the inverse of a large bound (the price one has to pay for the closure under polynomial growth). Indeed, we can show

Theorem 2 *For* $T \geq EEXL$ *and the set of all distributions* U *holds:*

$$DTime(T) \subset AvDTime(T, U)\ .$$

For average case complexity classes with a fixed bound on the rankability of the distributions we can establish a tight hierarchy, comparable to the situation in the worst case.

Theorem 3 *For time bounds* $T_1, T_2, V \geq (1+\omega(1)) \cdot \mathcal{N}$ *with* $T_1 \leq o(T_2)$ *holds:*

$$AvDTime(T_1, V\text{–}rankable) \subset AvDTime(T_2, V\text{–}rankable) .$$

Proof Sketch: First we show that even under the simplest distributions, the uniform ones, not all problems with worst case time bound T_2 can be solved in time T_1 on the average, that means

$$DTime(T_2) \setminus AvDTime(T_1, \{\text{uniform}\}) \neq \emptyset .$$

The idea is to diagonalizes slowly enough over the sequence of DTM $M_1, M_2, \ldots$ such that either an input can be found, on which M_i differs from the diagonal language L to be constructed, or M_i spends too much time on sufficiently many inputs. These inputs will have enough weight to contradict that L is accepted by M_i in average time T_1.

Now observe that if $V_1 \leq V_2$ then V_1–rankable $\subseteq V_2$–rankable. Therefore,

$$DTime(T) \subseteq AvDTime(T, V_1\text{–rankable}) \subseteq AvDTime(T, \{\text{uniform}\})$$

∎

Furthermore, we can show an optimal separation of these averge case classes with respect to the complexity of the distributions.

Theorem 4 *For* $\delta > 1$, $\mathcal{N} \leq V_2 \leq o(V_1)$ *and* $V_1(\delta\mathcal{N}) \leq O(T)$ *holds:*

$$AvDTime(T, V_1\text{–}rankable) \subset AvDTime(T, V_2\text{–}rankable) .$$

The proof of this separation result with a fixed time bound is a rather complicated diagonal construction. We construct a distribution that is V_1– but not V_2–rankable with the property that some long inputs are assigned small ranks.

The left side of fig. 1 shows a pictorial description of the hierarchies implied by the last two theorems. Each point in the diagram represents a complexity class $AvDTime(T, V\text{–rankable})$ defined by the two complexity bounds T and V.

5 Reductions and Completeness for Average Case Complexity Classes

A meaningful reduction between distributional problems (L_1, ρ_1) and (L_2, ρ_2) has to relate the distributions μ_i, resp. rank functions ρ_i in order to guarantee that a good average case behaviour of one problem is transferred to the other. For this prupose, Levin introduced the notion of *dominance*. Considering the ranking, this property can be expressed by a simple condition if the reduction is injective. This is not a real restriction for reductions between standard $\mathcal{NP}$–complete problems. For technical reasons we assume in the following that all distributions μ have the property that all ranks are unique, that means the corresponding rank function $\rho = \text{rank}_\mu$ is injective. By a slight perturbation of the probabilities, this can always be achieved.

Definition 8 *An injective function $f : \Sigma^* \to \Sigma^*$ is a* **distributional reduction** *from the distributional problem (L_1, ρ_1) to the distributional problem (L_2, ρ_2) if the following conditions hold:*

1. *f is a polynomial time reduction from L_1 to L_2 in the classical sense, that means f can be computed in deterministic polynomial time and $x \in L_1 \Leftrightarrow f(x) \in L_2$.*
2. Domination: *There exist constants $c_0, c_1 > 0$ such that for all $x \in \Sigma^*$*
$$\rho_2(f(x)) \ \leq \ c_0 \ |x|^{c_1} \ \rho_1(x) \ .$$

In order to analyse the average case complexity of problems in $\mathcal{NP}$ we first consider distributional problems.

Definition 9

$$\boldsymbol{\mathcal{NP}}^{\mathrm{dist}} \ := \ \mathcal{NP} \times \mathit{POL\text{–}rankable} \ = \ \{(L, \mu) | L \in \mathcal{NP} \text{ and } \mu \in \mathit{POL\text{–}rankable}\} \ .$$

A distributional problem (L, ρ) is $\boldsymbol{\mathcal{NP}}$**–distributional complete** *if $(L, \rho) \in \mathcal{NP}^{\mathrm{dist}}$ and if for all distributional problems in $\mathcal{NP}^{\mathrm{dist}}$ there exists a distributional reduction to (L, ρ).*
A language L is $\boldsymbol{\mathcal{NP}}$**–average complete** *if $L \in \mathcal{NP}$ and for all $L' \in \mathcal{NP}$ and $\rho' \in$ POL–rankable there exists a distribution (ranking) $\rho \in$ POL–rankable such that (L', ρ') has a distributional reduction to (L, ρ).*

Lemma 2 *If $(L_1, \rho_1) \in \mathcal{NP}^{\mathrm{dist}}$, $(L_2, \rho_2) \in Dist\mathcal{P}$, and (L_1, ρ_1) has a distributional reduction to (L_2, ρ_2) then $(L_1, \rho_1) \in Dist\mathcal{P}$.*
If (L, ρ) is $\mathcal{NP}$–distributional complete then L is $\mathcal{NP}$–average complete.

Theorem 5 *If an $\mathcal{NP}$–average complete language belongs to $Av\mathcal{P}$ then $\mathcal{NP} \subseteq Av\mathcal{P}$.*

Definition 10 *The bounded halting problem* NBH *for NTM is the language*

$$\mathtt{NBH} \ := \ \{(x01^t0^i) \mid \mathit{time}_{M_i}(x) \leq t\} \ .$$

Let $cod(n)$ be a self-delimiting binary encoding of the natural number n of length $O(\log n)$ that can be computed in time $O(n)$. Define a distribution for NBH *by*

$$\mathit{rank}_{\mathtt{NBH}}(w) \ := \ \begin{cases} \mathit{bin}^{-1}(x \ \mathit{cod}(t) \ \mathit{cod}(i)) & \text{if } w = x01^t0^i, \\ \infty & \text{else.} \end{cases}$$

Observe that $\mathrm{rank}_{\mathtt{NBH}}$, resp. any distribution with this rank function, is linear rankable. The following reduction uses Levin's idea in case of computable distributions (see [Levi86] and [Gure91]).

Theorem 6 $(\mathtt{NBH}, rank_{\mathtt{NBH}})$ *is $\mathcal{NP}$–distributional complete.*

Proof: Let $(L,\mu) \in \mathcal{NP}^{\text{dist}}$ with rank function r and M_i be a NTM that accepts $x \in L$ in time $q(|x|)$, where q is a polynomial. For inputs not in L the machine M_i does not halt on any computation. Let M_j be a NTM that on input by, where b is a single bit and $y \in \Sigma^*$, does the following:
If $b = 0$ then find a string x such that $r(x) = \text{bin}(y)$, else set $x := y$.
Simulate M_i on input x. There exists a polynomial p such that M_j halts on input x in time $p(|x|)$ iff M_j accepts. We define a reduction f by

$$f(x) := \begin{cases} 1x01^{p(|x|)}0^j & \text{if } r(x) \geq \text{bin}^{-1}(x), \\ 0\text{bin}(r(x))01^{p(|x|)}0^j & \text{if } r(x) < \text{bin}^{-1}(x). \end{cases}$$

The reduction property is obvious for f. Domination is achieved because a string is coded by its rank in case the rank is smaller than its binary length. ∎

Corollary 2 $\mathtt{NBH}$ *is $\mathcal{NP}$–average complete.*

Let us consider a standard (worst-case) reduction f that is injective, invertible in polynomial time and honest, that means $|f^{-1}(y)| \leq R(|y|)$ for some polynomial R. Then the reduction can be translated into one for the average case.

Theorem 7 *Let f be an injective, polynomial time invertible and honest reduction from L_1 to L_2. If L_1 is $\mathcal{NP}$–average complete then the same holds for L_2.*

Proof: To get a distributional reduction from a distributional problem (L_1, ρ_1) to L_2 define the rank function for L_2 by $\rho_2(y) := \rho_1(f^{-1}(y))$. Thus the dominance property is trivially fulfilled and because of the invertibility and honesty of f the complexity of ρ_2 is polynomially bounded if this holds for ρ_1. ∎

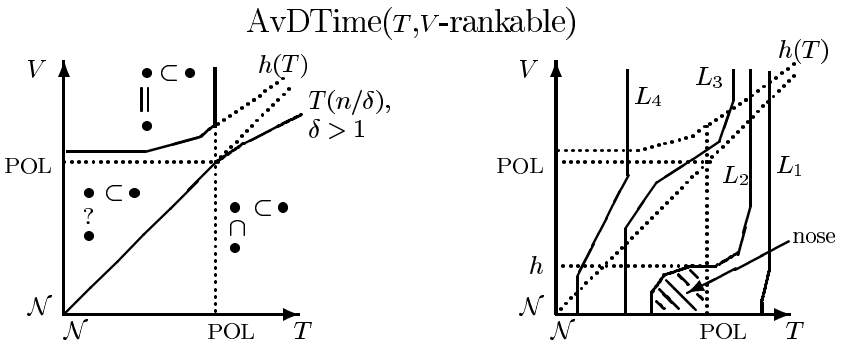

Fig. 1. On the left: hierarchies between average case complexity classes; on the right: different average case behaviour of languages L_i

6 The Average Case Complexity of Problems in $\mathcal{NP}$

An interesting question is whether $\mathcal{NP}$ problems for certain bounds V can be solved efficiently on average, at least for V–rankable distributions. Therefore, for $L \in \mathcal{NP}$ we look at the set of all pairs $(T, V) \in (\mathrm{POL}, \mathrm{POL})$ such that $L \in AvDTime(T, V\text{–rankable})$ and call this the **nose** of L.

An $\mathcal{NP}$-problem that has a nontrivial nose can be considered feasible on average for most practical applications. The height of a nose is defined by the supremum V of all pairs (T, V) contained in the nose. If L has a nose of height h we know that there is a polynomial average time algorithm for L unless the inputs are supplied by an adversary that needs at least $O(h(|x|))$ steps (in the worst case) to compute the rank of x.

The right side of figure 1 visualizes the possible average case behaviour of languages in $\mathcal{NP}$. The right halfspace of a line corresponding to a language L_i contains all pairs (T, V) such that $L_i \in AvDTime(T, V\text{–rankable})$. L_1 is an example of a language that cannot be solved efficiently, even on the average with respect to very simple distributions. L_2 represents a feasible problem for all distributions that can be computed within the complexity bound h, but not for more complex distributions. If $\mathcal{NP}$ contains such a language with $h \in \mathrm{POL}$ then $\mathcal{NP} \not\subseteq \mathcal{A}\mathrm{v}\mathcal{P}$. On the other hand, L_3 can be solved efficiently for all polynomially rankable distributions and if such a language were $\mathcal{NP}$–average complete then $\mathcal{NP} \subseteq \mathcal{A}\mathrm{v}\mathcal{P}$. If L_4 were $\mathcal{NP}$–complete then $\mathcal{NP}$ would collapse to $\mathcal{P}$ since, as we have shown, for ranking bounds above $h(T)$ the average case complexity equals the worst case complexity.

Using the completeness of the bounded halting problem and invertible distributional reductions, standard $\mathcal{NP}$–complete problems can be shown to be $\mathcal{NP}$–average complete for a linear rankable distribution. As an example, we can prove this for the satisfiability problem SAT.

Theorem 8 *There exists a ranking ρ of linear complexity such that the distributional problem $(\mathtt{SAT}, \rho)$ is $\mathcal{NP}$–distributional complete. Thus,* SAT *is $\mathcal{NP}$–average complete.*

Therefore SAT has no nontrivial nose unless $\mathcal{NP} \subseteq \mathcal{A}\mathrm{v}\mathcal{P}$. An explicit distribution that turns SAT into a hard distributional problem can efficiently be computed from the ranking ρ, for example as $\mu(x) := c/(\rho(x) \log^2 \rho(x))$. It seems that with respect to Levin's notion of computability no such distribution is known that can be computed in polynomial time.

7 Conclusions

We have shown that the average case time complexity of an algorithm can be estimated as precisely as in the worst case. Ranking the input space and measuring

the complexity of a distribution with respect to its rankability has been proved to an appropriate and natural concept. Classical results like tight hierachies can be obtained this way, both for the time complexity and the complexity of the distributions. Based on these notions, starting with distributional complexity classes we have presented meaningful definitons of average case complexity classes the elements of which are languages in the standard sense. They are directly comparable to worst case classes.

Definitions for reductions and completeness have been given for distributional and average case classes. This way, one overcomes problems with flat distributions in Levin's approach as observed by Gurevich [Gure91]. In a natural way, standard $\mathcal{NP}$–completeness translates into a completeness for average case analysis. In contrast to computability, for many $\mathcal{NP}$–problems a hard polynomial time bounded rank function can be constructed, and in contrast to sampleability this function can be used to construct computable distributions efficiently. We have shown this for the basic complete problem `SAT`. The maximal complexity of distributions such that a problem can be solved in average polynomial time – the height of the nose – has been proposed as a measure for the average case complexity of problems above $\mathcal{P}$. It may be possible to prove the existence of problems with nontrivial noses by using one-way-functions as reductions (compare [VeLe88] and [ImLe90]).

These ideas can also be applied to other cases like the analysis of average space complexity.

References

[BCGL92] S. Ben-David, B. Chor, O. Goldreich, M. Luby, *On the Theory of Average Case Complexity*, J. CSS 44, 1992, 193-219; see also Proc. 21. STOC, 1989, 204-261.

[Gure91] Y. Gurevich *Average Case Completeness*, J. CSS 42, 1991, 346-398.

[ImLe90] R. Impagliazzo, L. Levin, *No Better Ways to Generate Hard $\mathcal{NP}$ Instances than Picking Uniformly at Random*, Proc. 31. FoCS, 1990, 812-821.

[John84] D. Johnson, *The $\mathcal{NP}$–Completeness Column*, J. of Algorithms 5, 1984, 284-299.

[Levi86] L. Levin, *Average Case Complete Problems*, SIAM J. Computing 15, 1986, 285-286.

[Milt91] P. Miltersen, *The Complexity of Malign Ensembles,* Proc. 6. Structure in Complexity Theory, 1991, 164-171.

[MiSeLe92] D. Mitchell, B. Selman, H. Levesque *Hard and Easy Distributions of SAT Problems*, Proc. 10. Nat. Conf. on Artificial Intelligence, 1992, 459-465.

[ReSc92] R. Reischuk, Chr. Schindelhauer, *Precise Average Case Complexity Measures,* Technical Report, Technische Hochschule Darmstadt, 1992.

[Schi91] Chr. Schindelhauer, *Neue Average Case Komplexitätsklassen,* Diplomarbeit, Technische Hochschule Darmstadt, 1991.

[VeLe88] R. Venkatesan, L. Levin, *Random Instances of Graph Coloring Problems are Hard*, Proc. 20. SToC, 1988, 217-222.

[WaBe92] J. Wang, J. Belanger, *On Average $\mathcal{P}$ vs. Average $\mathcal{NP}$*, Proc. 7. Struc. Compl., 1992.

This article was processed using the LaTeX macro package with LLNCS style