

# Chaitin's halting probability and the compression of strings using oracles

George Barmalias

Joint work with Andrew Lewis



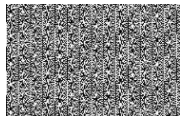
**ISCAS**

Institute of Software  
Chinese Academy of Sciences

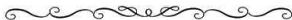
Barcelona, July 2011

# Question

Random data lack internal structure and patterns



... so they are **incompressible**.



If a **computer is given access** to external information

this may affect its ability to **compress data**.

*Given an oracle  $A$ , **how many** oracles can compress data **at most as well as  $A$** ?*

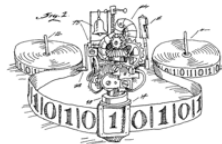
# Kolmogorov 1960s

The complexity of a binary string is the length of its shortest description.



Descriptions should be given in an algorithmic way:

If  $M$  is a Turing machine and  $M(\sigma) = \tau$ , then  $\sigma$  is an  $M$ -description of  $\tau$ .



# Kolmogorov complexity of strings

Let  $|M|$  be the **size** of the machine  $M$ .

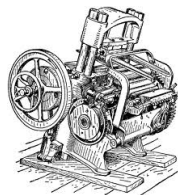
Let  $K_M(\sigma)$  be the complexity of  $\sigma$  w.r.t.  $M$ .

*The **complexity of  $\sigma$**  is the least sum  $|M| + K_M(\sigma)$  where  $M$  ranges over all machines.*

Let  $K(\sigma)$  denote the complexity of  $\sigma$ .



A string is **c-compressible** if it has a description that is shorter than its length by at least  $c$  bits.



# Remark

Chaitin and Levin observed in the 70s:

Given a string, one can recover information  
from the bits of the string **but also from its length.**

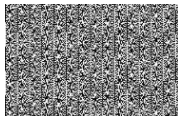


Obtain a more faithful complexity measure by restricting to . . .

*Prefix-free machines cannot extract information from  
the length of a string.*

$$K(\sigma) \leq |\sigma| + K(|\sigma|) \quad \text{and} \quad K(n) \leq 2 \log n.$$

# Algorithmic randomness



A stream  $X$  is **random** if there is a constant  $c$  such that  $K(X \upharpoonright_n) \geq n - c$  for all  $n$ .

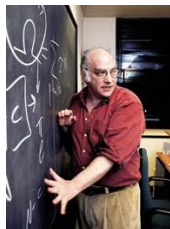
This notion of randomness is **robust**:

- ▶ **Coincides with other approaches** (betting strategies, statistics)
- ▶ Random reals form a set of **measure 1**
- ▶ Meets **laws of large numbers**, normality etc.
- ▶ **Relativizes** giving randomness of various strengths

# Chaitin's halting probability

In 1975 Chaitin considered the halting probability of a universal prefix-free machine.

$$\Omega = \sum_{U(\sigma)\downarrow} 2^{-|\sigma|}$$



- ▶  $\Omega$  is **random**
- ▶  $\Omega$  has the same information as the Halting problem

Halting probabilities of universal machines were characterized as the

*... random left c.e. reals*

by Solovay, Kučera, Slaman, Calude, Khoussainov, Hertling, Wang.

# Walace's universality probability

$P_U$  is the probability that if  $X$  is random then  $\sigma \mapsto U(X \upharpoonright_n * \sigma)$  is universal for all  $n$ .



Barnali and Dowe showed that

- ▶  $P_U$  is random relative to  $\mathbf{0}^{(3)}$
- ▶ The Turing degree of  $P_U$  depends on the choice of  $U$

Universality probabilities of universal prefix-free machines were characterized as the

*... 4-random right c.e. relative to  $\mathbf{0}^{(3)}$  reals*

Finally ...  $P_U + \Omega_V^{\mathbf{0}^{(3)}} = 1$ .

## Back to the question

Given an oracle  $A$ , *how many* oracles can compress data *at most as well as  $A$* ?

or, more formally...

What is the *cardinality* of  $\mathcal{C}_A = \{X \mid \exists c \forall \sigma K^A(\sigma) \leq K^X(\sigma) + c\}$ ?

or even...

Given an oracle  $A$ , what is the *cardinality* of  $\mathcal{C}_A = \{X \mid X \leq_{LK} A\}$ ?

# More than 10 years ago...

Ambos-Spies and Kučera asked this in 1999 for  $A = \emptyset$ .



*How many low for  $K$  sets are there?*

**Motivation:** there are non-computable low for  $K$  sets.



Nies answered this in 2004 by showing that this is a subclass of  $\Delta_2^0$ .

# About 5 years ago...

Barnali, Lewis and Soskova showed in 2006 that

$A = \emptyset'$  then it *is uncountable*.

This was quickly extended to

If  $A$  “*not very close to computable*” (*not generalized low<sub>2</sub>*) then it *is uncountable*.



J. Miller exhibited a ‘large’ class of oracles  $A$  for which  $C_A$  *is countable*.



# Weakly low for $K$

He used a **generalization** of the low for  $K$  sets.

$A$  is **weakly low for  $K$**  if infinitely many programs achieve the same compression with or without  $A$ .

... if it is *infinitely often low for  $K$*

... if  $K(\sigma) \leq K^A(\sigma) + c$  for some constant  $c$  and infinitely many  $\sigma$ .



**J. Miller** also showed that if  $A$  is weakly low for  $K$  then  $C_A$  is **countable**.

# A reasonable guess

J. Miller showed that  $A$  is weakly low for  $K$  iff  $\Omega$  is random relative to  $A$ .



... the weakly low for  $K$  sets form a large class.



Conjecture (J. Miller)

$\mathcal{C}_A$  is countable if and only if  $A$  is low for  $\Omega$ .

# In the effective world



In 2007 I showed that for  $\Delta_2^0$  sets,  $A$  is **low for  $K$**  iff  $C_A$  is countable.

## In the $\Delta_2^0$ world

*If an oracle can compress **better than some oracle** ...*

*... then it can compress **better than uncountably many oracles**.*



The perfect set I exhibited was  $\Pi_1^0$ ...

and later (with M. Baartse) **lacking low for  $K$  members**.

## Conjecture (J. Miller)

*If  $A$  is not low for  $\Omega$  then  $\mathcal{C}_A$  contains a perfect set .*

### Tools:

- ▶ **Measure-permitting approximation** argument translating the **power of the oracle** into **compression power**.
- ▶ Using **compressions of  $\Omega$**  for achieving **uniform compression on all programs**.

**Recall:**  $A$  can **compress  $\Omega$**  iff  $\lim_{\sigma} (K(\sigma) - K^A(\sigma)) = \infty$ .

# Recycling lost measure and $\Omega$

## Plan:

- ~> Before taking the risk of losing some measure, **transform it into  $\Omega$ -form** and **compress** it.
- ~> If you lose it, you **lose a compressed form of it**.
- ~> Transform **lost measure** into **better guesses** in next cycles.



## Problem

*A-computable constructions produce **A-computable parameters**.*

*We want  **$\Omega$  products**; **not  $\Omega^A$** .*

# Answer

Simulate **computable procedures** within the **oracle construction**.



**Pre-cooked** computable procedures work in a **program managed by  $A$**   
... producing **versions of  $\Omega$** , which are then **processed by  $A$** .

## Theorem (Barnmpalias and Lewis)

*$\mathcal{C}_A$  is countable if and only if  $\Omega$  is  $A$ -random.*

$A$  can **compress more than uncountable collection** of oracles iff it can **compress segments of  $\Omega$** .

# References

- ▶ Barmpalias/Lewis, Chaitin's halting probability and the compression of strings using oracles (Proc. Royal Soc.)
- ▶ Barmpalias/Dowe, Universality probability of a prefix-free machine
- ▶ J. Miller, The  $K$ -degrees, low for  $K$  degrees, and weakly low for  $K$  sets (NDJFL)
- ▶ Nies, Computability and Randomness, Oxford Press.

Webpage: <http://www.barmpalias.net>