

## Bounded Arithmetic $S_2$ – part I

Recall that our ultimate goal is to come up with a theory  $T$  so that given a relation  $P(x, y)$  for which  $T \vdash \forall x \exists y P(x, y)$  one can find an efficient algorithm which on input  $x$  computes  $y$  so that  $P(x, y)$ .

The first reasonable candidate theory was  $I\Delta_0$  with  $\Delta_0$ -definable relation  $P(x, y)$ , since by Parikh's theorem

$$I\Delta_0 \vdash \forall x \exists y P(x, y) \implies I\Delta_0 \vdash \forall x \exists y \leq t(x) P(x, y),$$

with  $t(x)$  - an arithmetical term, i.e. a polynomial.

As we have discussed, the algorithm for computing  $y$  from  $x$  just loops through all numbers in increasing order until one is found satisfying  $P(x, y)$  and  $t(x)$  serves as an upper bound on the length of the cycle.

Of course, since  $x$  is represented as a string of length  $\approx \log(x)$  and there are  $t(x)$  numbers to check, the total runtime of the algorithm is  $\geq t(x) \approx 2^{C \log(x)}$  for a suitable constant  $C$ , since  $t(x)$  is a polynomial.

Notice that, even if one comes up with a substantial reduction of the search space, in principle even verifying whether  $P(x, y)$  holds for a  $\Delta_0$ -definable relation  $P(x, y)$  still takes time exponential in lengths of  $x, y$ .

Can we do better?

**Exercise 1.** \* Show that, if one is allowed to use nested (co)non-determinism, the task of computing  $y$  given  $x$  as above can be done in time linear in the length of  $x$ .

Formally, the above task can be solved by an algorithm from the functional analog of the *linear-time hierarchy*.

**Fact 2.** Any total relation  $P(x, y)$  computable by a function from the linear-time hierarchy is definable by a suitable  $\Delta_0$ -formula  $\varphi(x, y)$ . Moreover, for any such  $P(x, y)$  and  $\varphi(x, y)$   $I\Delta_0$  proves  $\forall x \exists y \varphi(x, y)$ .

In particular,  $I\Delta_0$  cannot be the theory we seek, unless we first derive new groundbreaking results. For the exact same reason, we cannot show that  $I\Delta_0$  does not give the witnessing we want.

Before we proceed further, we first need to inspect our main question once again.

Recall that we have started with  $PA$ , which is obviously too strong for any reasonable witnessing, and then moved to a weaker theory  $I\Delta_0$  which again turned too strong. A natural choice is to move to an even weaker theory. Notice, however, that all this time we actually had the weakest theory of them all right before us - the *empty theory*  $E$ .

If one believes that some efficient witnessing is achievable and is dependent on the strength of  $T$ , then there is no better choice than to let  $T$  be  $E$ . One then needs to inspect what happens if  $E \vdash \forall x \exists y P(x, y)$  and whether we do get an algorithm of any kind.

As we will later see, there is a fundamental witnessing result for  $E$  which is also known as Herbrand's theorem. We will prove it later. At this point, however, we need to refine the original question we started with.

It is true that the weaker the theory, the stronger the proofs, and proofs in  $E$  are as strong as possible. However, such theories are actually too weak, in the sense that for any reasonable  $P(x, y)$   $E$  would probably not be able to prove  $\forall x \exists y P(x, y)$ . Witnessing results, which are of course interesting on their own, are primarily used as a tool for further investigations in the complexity theory. And so we want our  $T$  to be weak to derive efficient witnessing, but at the same time, we want it to be strong enough to actually prove  $\forall x \exists y P(x, y)$  for non-trivial  $P(x, y)$ .

And so we refine our original question and get the following.

**Question 3.** Is there a theory  $T$  so that for a total relation  $P(x, y)$  there is an efficient algorithm computing  $y$  from the given  $x$  satisfying  $P(x, y)$  if and only if  $P(x, y)$  is definable by a formula  $\varphi(x, y)$  so that  $T \vdash \forall x \exists y \varphi(x, y)$ ?

We have already seen that  $I\Delta_0$  can give only nested (co)non-deterministic linear-time witnessing and proving whether or not it is equivalent to efficient (i.e. poly-time deterministic) witnessing is beyond the current state of knowledge.

However, we can show that  $I\Delta_0$  fails to be the  $T$  we seek by investigating the other implication of 3.

**Exercise 4.** In the following, we assume  $P(x, y)$  is already represented as a suitable  $\Delta_0$ -formula. Show that for all of the below relations  $I\Delta_0 \not\vdash \forall x \exists y P(x, y)$ . Furthermore, argue that there is an efficient algorithm which on input  $x$  outputs  $y$  satisfying  $P(x, y)$ .

- $x$  represents two binary  $n$ -dimensional vectors  $v, w$  and  $y$  represents their *outer product*  $v * w$ , i.e. a matrix of shape  $(n, n)$  such that  $(v * w)_{i,j} = v_i w_j$ .
- $x$  represent a formula  $\psi(x)$  together with a term  $t$  and  $y$  represent  $\psi(t)$ .
- For a fixed efficient super-linear-time algorithm  $A$ ,  $x$  represents an input to  $A$  and  $y$  represents finished accepting/rejecting computation of  $A$  on  $x$ .
- \* For a fixed *secure pseudorandom number generator*  $G$  with super-linear stretch,  $x$  represents an input to  $G$  and  $y$  represents the output of  $G$  on  $x$ .

And so it is not enough just to weaken  $I\Delta_0$ , we first need to bypass the above problems. This will be done by expanding the language  $L_{PA}$  and adding suitable axioms for newly introduced symbols.