A SIMPLE WAY TO REDUCE FACTORIZATION PROBLEMS TO SAT

DAVIDE MARAN

ABSTRACT. As Cook-Levin theorem showed, every NP problem can be reduced to SAT in polynomial time. In this paper I show a simpler and more efficient method to reduce some factorization problems to the satisfiability of a boolean formula.

1. Introduction

Definition 1.1. SAT is the problem which consist in determining whether or not a boolean formula $S$ consisting only of OR AND NOT and TRUE/FALSE variable is satisfiable, which means if there exist a truth assignment which satisfies $S$.

As all the logical operators like $\iff$ and $\implies$ can be substituted with a combination of the three logical operators named above, we can also consider the formulas which contain $\iff$ and $\implies$ as instances of SAT. In computational complexity the importance of the SAT is given from this theorem, which can be considered the most important in this field.

Theorem 1.2. (Cook-Levin) Every NP problem can be reduced to SAT in polynomial time.

It is important to clarify that if $NP = co-NP$ is false the complement of an NP problem could not be in NP. This fact can lead to counterintuitive result, for example in NP it is much more difficult to determine whether a number is composite than to determine if the same number is prime. However some other results can solve this issue. Call $SAT^c$ the problem of saying if a boolean formula is NOT satisfiable.

Corollary 1.3. every co-NP problem can be reduced to $SAT^c$ in polynomial time.

2010 Mathematics Subject Classification. Primary 03D15; Secondary 68Q15.
Proposition 1.4. Call \( P^{SAT} \) the set of decision problems which can be decided in polynomial time from a Turing machine with oracle \( SAT \). Then \( P^{SAT} = NP \cup co-NP \)

Proof. Since every problem in \( NP \) can be reduced to \( SAT \) and every problem in \( co-NP \) can be reduced to \( SAT^c \), the union of these two sets contains only problems that can be solved in polynomial time with the help of a one step decision box which provides both the affirmative and the negative response to \( SAT \). This is, by definition, a \( SAT \) oracle. On the other hand, if a decision problem is not in \( NP \cup co-NP \), it cannot be reduced in polynomial time to \( SAT \) or \( SAT^c \). This means that it cannot be solved in polynomial time by a Turing machine with oracle \( SAT \). \( \square \)

In force of this result we can agree that \( P^{SAT} \) is the class of decision problems whose positive or negative solution can be verified in polynomial time (it contains in particular both the problems of deciding whether a number is prime and whether a number is composite). The proof of theorem 1.2 consists of building, for every problem \( p \in NP \), the non deterministic Turing machine which solves that problem and assigning to every cell of the accepting computational path, in every step of the computation, a boolean variable. Now, since there are precise condition for a computational path to be valid (for example a cell must contain one symbol in every step and only one cell can change its content in every step) it is possible to impose some condition to the variables just defined which turn out to form a boolean expression. Unfortunately, this process is extremely expensive for any problem \( p \) for two reasons: first is very difficult to build the Turing machine which solve a given problem, second, and most important, the theorem requires to initialize an incredibly large (although non exponential) number of boolean variables. I will show a much simpler way to turn factorization problems to SAT instances.

2. The problems which I face

Definition 2.1. Call \( \text{prime}(n) \) the problem of defining whether an integer number \( n \) is composite.

In year 2002 this problem has been proven to belong to \( P \), so I will deal with a more general instance of the problem.
Definition 2.2. Call exprime \((v[n])\), where the vector \(v[n]\) contains in every position either "1", "0" or "-", the problem of deciding whether exists a prime number which has the numbers of the vector \(v\) as the digits in its binary expression.

It is difficult to understand this definition without an adequate example

Example 2.3. The problem \(Exprime(1, 0, 1, -, 1, -, 0, 0, -, 1)\) consists of deciding whether exists a prime number whose binary form is \(101a1b00c1\), where \(a, b, c\) are unknown numbers in the set \(\{0, 1\}\).

It is simple to note that the same polynomial algorithm used to solve \(prime(n)\) cannot be used in this case since it could take an exponential time to examine all the possible numbers. However is is also obvious that this problem belongs to co-NP since it is extremely easy to check the correctness of a solution of its complement problem (which can be called \(Expcomposite(n)\)) once it is found. We can simply ask the non-deterministic Turing machine to "guess" the values of the free variables in the input number and then solve the problem as \(composite(n)\). We can generalize again this problem many time. So, it is useful to define a set \(F\) of generalizations of \(prime(n)\) and \(composite(n)\).

Definition 2.4. Let \(F\) be the set of problems \(p(\rho) \in F\) such that \(p\) can be expressed as "there exist a number \(\rho\) whose binary representation is \(\rho_1\rho_2...\) which can be factorized as \(\gamma_1\gamma_2... \ast \theta_1\theta_2...\) such \(\rho_n, \theta_n, \gamma_n\) respect \(C_1, C_2...\)"

where \(\forall n \rho_n\) can be either binary digits fixed and free 1/0 variables. \(C_1, C_2, ...\) are conditions about the digits of these three numbers which can be verified in polynomial time. Note that every decision problem \(p\) of \(F\) recives as input a vector which represent a number with some unknown digits, but in fact, as all the symbols of the variables can be codified in a binary number, \(p\) recives a sequence of bit which form a binary number that has nothing to do with the number that is \(represents\). Some of there problems in this class are quite useless:

Example 2.5. The problem of deciding whether a given number \(n\) can be factorized into a product of \(p \ast q\), where the last 3 digits of \(n\) are exactly the same last 3 digits of \(q\), belongs to \(F\).

some others are "famous"
Example 2.6. Let $\text{Factoring}(N)[L,U]$ be the problem of determing whether exist or not a natural number $\in [L,U]$ which divides $N$. This problem belongs to $F$

Remark 2.7. As we have defined the class $F$ as set of problems which recives only a number as input, the two numbers $L,U$ must be a-priori chosen. If we would prefer to add these couple to the input, we would had to define another class $F'$ where some condition can be given as input.

As S. Arora ans B. Barak proved (2007) this problem is NP-complete, so all the NP problems can be reduced to factoring and, in particular, all the $F$ problams can be reduced to Factoring.

3. Main idea

The idea of the solution is quite simple, altough we will see that is does not work without an important adjustment. Every problem of $F$ is characterized by

1) One natural number $N$ recived as input with, in case, some unknown digits
2) A set of conditions which are chosen a priori;

If $N$ can be factorized as $A \ast B$ where the digits of $A, B, N$ respect all the conditions, then the problem accepts $N$, if $N$ cannot be facorized or the factorization does not respect the conditions, the problem rejects $N$.

The heart of all these problem is a multiplication.

There are many algorithms to multiplicate two numbers. We are going to use one of them called standard multiplication. Altough it is one of the worst in terms of time complexity, it serves to our purpose since we are searching for a simple way to reduce factorization problems to SAT and not to multiplicate numbers in the easiest way. Standard multiplication between binary is, for some reasons, naturally prepared to be transformed into SAT. Essentially we are going to substitute the digits of the numbers involved in standard multiplication with boolean variables and their arithmetic relations with boolean formulas. Before that, we have to present a very simple result about the factorizability.

Lemma 3.1. A number $r$ with $n$ binary cifres is composite if and only if it is the product of a number with at most $n-1$ digits with a number with at most $(n+1)/2$ digits.
Proof. The if-clause is banal. If a number $r$ is non prime it can always be written as the product of two numbers, the biggest of them cannot be greater than $r/2$, so it cannot have more than $n-1$ digits in base 2. On the other hand, the lower cannot be higher than $\sqrt{r}$, otherwise the product would result bigger than $r$ and, of course, $\sqrt{r}$ has at most $(n+1)/2$ digits. This proves the only-if clause. □

4. Resolution

So, every 3-digit binary composite number $ABC$ can be written as $ab * c$. We are trying to use the algorithm of standard multiplication to determine if these two numbers exist. The algorithm for standard multiplication which is taught in schools consists of multiply the multiplicand by each digit of the multiplier and then add up all the properly shifted results. It requires memorization of the multiplication table for single digits, we can express it with this schema

$$A \ B \ C =$$
$$a \ b \ *$$
$$c \ d =$$
$$e \ f +$$
$$g \ h \ 0$$

where the number $ef$ is equal to $ab * d$ and $gh0 = ab * c0$. Of course it must result that $ef + gh0 = ABC$ in order to let $ABC$ being factorizable.

1) $d = 0 \implies (1 - e) + (1 - f) > 0$
2) $d = 1 \implies (e = a), (f = b)$
3) $c = 0 \implies (1 - h) + (1 - g) > 0$
4) $c = 1 \implies (g = a), (b = h)$
5) $f = C$
6) $B = e + f \pmod{2}$
7) $e + g + h < 11_2$
8) $\max(g, \min(e, f)) = a$

$A, B, C, a, b, c, d, e, f, g, h$ are digits that can assume the values 1, 0. We are going to treat them as boolean variables ($1 = true, 2 = false$) and to translate the previous facts into boolean expressions.

1) $\neg d \implies (\neg e \land \neg f)$
2) \( d \Rightarrow (e \iff a \land f \iff b) \)
3) \( \neg c \Rightarrow (\neg h \land \neg g) \)
4) \( c \Rightarrow (g \iff a \land h \iff b) \)
5) \( f \iff C \)
6) \( ((e \land h) \lor (\neg e \land \neg f)) \iff \neg B \)
7) \( \neg (e \land g \land h) \)
8) \( (g \lor (e \land h)) \iff A \)

Where the first four equations express the connection between \( a, b, c, d \) and \( e, f, g, h \), the seventh express that the result cannot have more than 3 digits and the remaining equation express how to find the values of \( A, B, C \). It is important to note that the fact that we could have a ”carry” as a result of the sum of \( e \) and \( h \) has obliged us to add the seventh expression and to make the eight much more complicated. Now, we have reduced the problem of knowing if a general three digit number is composite to the satisfability of the conjunction of eight boolean expressions in eleven variables. This means, for example, that if we wanted to know if there exist one composite number with 1 as unit digit, it would be sufficient to substitute C with ”false” and try to solve the SAT formula. This reduction can seem pretty idiotic since there is literally no person in the world who would be interested in using this method to determine if number less than 8 is composite. However we should compute how the numbers of variables and expression grows with augmenting of the number of digits in the input number.

**Example 4.1.** So here it is how the algorithm works for an 8-digit number:

\[
\begin{align*}
A & B & C & D & E & F & G & H = \\
& a & b & c & d & e & f & g * \\
& h & i & l & m = \\
& a_1 & b_1 & c_1 & d_1 & e_1 & f_1 & g_1 + \\
& a_2 & b_2 & c_2 & d_2 & e_2 & f_2 & g_2 0 + \\
& b_3 & c_3 & d_3 & e_3 & f_3 & g_3 0 0 + \\
& c_4 & d_4 & e_4 & f_4 & g_4 0 0 0 =
\end{align*}
\]
We have to express the relation between these digits with a boolean formula in the corresponding variables. In this case we have needed 44 variables to represent all the digits of the problem. This is not a problem because it can be easily proved that the number of variables is a quadratic function in the length of the input.

**Proposition 4.2.** let \( N(x) \) be the function which expresses the number of boolean variables assigned with respect to the length \( x \) of the input. Then \( N(x) \in O(x^2) \)

**Proof.** We prove that \( \forall x, N(x) < x^2 + 3x \). All the assigned variables are either missing digits of the input number or digits of the multiplicand/multiplier or digits obtained as product of a digit of the multiplicand with a digit of the multiplier. Since \( x \) is the number of digits of the input number, the total number variables corresponding to digits of the multiplicator and the multiplier is strictly less than \( 3x \). All the other variables are the product of a digit of the multiplicand with a digit of the multiplier, so each one corresponds to one and only one pair \((y, z)\) where \( z \) is a digit of the multiplicand and \( y \) is a digit of the multiplier. Since there are less than \( x^2 \) pairs such, the total number of variables of this type is less than \( x^2 \). This mean that the total number of boolean variables is strictly less than \( x^2 + 3x \)

\[ \square \]

Unfortunately, the boolean expression which comes out is much larger. Coming back to the Example 4.1, let us try to translate the relation between the variables \( d_1, e_2, f_3, g_4, E \) into a boolean expression. We know that \( E = 1(true) \) if and only if there is an odd number of "ones" between \( d_1, e_2, f_3, g_4 \). This obliges us to write an expression like

\[ (d_1 \land \neg(e_2 \lor f_3 \lor g_4)) \lor (e_2 \land \neg(d_1 \lor f_3 \lor g_4))... \]

and, in fact, list all the subsets of \( d_1, e_2, f_3, g_4 \). This turns even worse if we try to consider the carry from the sum of the digits of the precedent column! And the worst thing of all is that the number of subsets increase exponentially with the dimension of the set. So, since we are trying to sum \( (x+1)/2 \) digits, the total expression is longer than the disjunction of \( 2^{(x+1)/2} \) simple expressions. This means that we cannot use the standard multiplication to make a polynomial reduction from the problems \( \in F \) to SAT.
5. Alternative solution

In order to make more efficient this transformation it is necessary to insert an apparently stupid trick. As the precedent issue has been caused by the necessity to sum multiple binary digits at the same time, we have to make an algorithm that avoids this operation. This can be done quite easily by adding the numbers “step by step” making use of partial sums and carries.

Example 5.1. This is the adjusted version for an 8-digit number:

\[
\begin{array}{cccccccc}
A & B & C & D & E & F & G & H = \\
A & B & C & D & E & F & G & H \ast \\
h & i & l & m = \\
\begin{array}{cccccccc}
a_1 & b_1 & c_1 & d_1 & e_1 & f_1 & g_1 & + \\
a_2 & b_2 & c_2 & d_2 & e_2 & f_2 & g_2 & = \\
\mid r_1.1 & r_1.2 & r_1.3 & r_1.4 & r_1.5 & r_1.6 & r_1.7 & 0 \\
s_{1.1} & s_{1.2} & s_{1.3} & s_{1.4} & s_{1.5} & s_{1.6} & s_{1.7} & s_{1.8} + \\
b_3 & c_3 & d_3 & e_3 & f_3 & g_3 & 0 & 0 = \\
\mid r_2.1 & r_2.2 & r_2.3 & r_2.4 & r_2.5 & r_2.6 & r_2.7 & 0 \\
s_{2.1} & s_{2.2} & s_{2.3} & s_{2.4} & s_{2.5} & s_{2.6} & s_{2.7} & s_{2.8} + \\
c_4 & d_4 & e_4 & f_4 & g_4 & 0 & 0 & 0 = \\
\mid r_3.1 & r_3.2 & r_3.3 & r_3.4 & r_3.5 & r_3.6 & r_3.7 & 0 \\
\end{array}
\end{array}
\]

In the last example, every row put between “|” is said auxiliary and contains all the carries of the sum of the two precedent rows. Any auxiliary row is made of variables whose only purpose is to make the final boolean formula easier to understand. Every row of \( s \), which we can call sumline, is the sum of the precedent two non auxiliary rows. So, for example:

\[s_{1.1}s_{1.2}s_{1.3}s_{1.4}s_{1.5}s_{1.6}s_{1.7}s_{1.8} = a_1b_1c_1d_1e_1f_1g_1 + a_2b_2c_2d_2e_2f_2g_20\]

And, of course:

\[ABCDEFGH = s_{2.1}s_{2.2}s_{2.3}s_{2.4}s_{2.5}s_{2.6}s_{2.7}s_{2.8} + c_4d_4e_4f_4g_4000\]

this force us to add some boolean formulas to express the fact that every row of \( s \)-es is the sum of the previous rows and every row of \( r \)-es contains the carries from the previous sum. The rest of the formulas, which
contain the way to obtain the other rows from the initial product can remain unchanged. The boolean formulas which expressed the fact that $ABCDEFGH$ is the sum of all the rows except for the first three must be, of course, deleted. In this case we have only to sum two numbers at a time.

**Example 5.2.** Sum of two binary numbers using boolean variables.

\[
\begin{array}{cccccccc}
\text{a} & \text{b} & \text{c} & \text{d} & \text{e} & \text{f} \\
\text{g} & \text{h} & \text{i} & \text{l} & \text{m} & \text{n} \\
\hline
r_1 & r_2 & r_3 & r_4 & r_5 & r_6 \\
\end{array}
\]

\[
s_1 \quad s_2 \quad s_3 \quad s_4 \quad s_5 \quad s_6
\]

$s_6$ is the sum of $f$ and $n$, so $s_6 \iff [(f \lor n) \land \neg(f \land n)]$. for the next variables the formula becomes more difficult because of the existence of the carry from the previous sum. So,

\[
r_4 \iff [(e \land m) \lor (e \land r_5) \lor (m \land r_5)]
\]

\[
s_4 \iff [(l \land d \land r_5) \lor (m \land \neg(e \lor r_4)) \lor (e \land \neg(m \lor r_4)) \lor (r_4 \land \neg(e \lor m))]
\]

This last formula, which seem to be very difficult, is in fact the disjunction of all the subset of \{l, d, r_4\} with an odd number of true variables. This is because $s_4$ is true (=1) if and only if there is an odd number of true variables in that set. For $s_2, s_3$ we can find a formula completely identical to the one just fund except for the name of the boolean variables and for $s_5$ a simpler version of the same formula (because the variable $r_6$ has value 0). For $s_1$ we find

\[
[s_1 \iff (a \lor g \lor (b \land h)) \land \neg((a \lor g) \land b \land h)]
\]

which contains also the fact that the carry may not exceed the number of cells.

So, any instance which expresses

\[
a_1a_2a_3\ldots a_n + b_1b_2b_3\ldots b_n = s_1s_2s_3\ldots s_n
\]

can be translated into a boolean formula containing no more than $42n$ symbols (which can be either $\land, \lor, \neg$, $\iff, \implies, \lbrace, \rbrace$ or variables). Since it is a conjuncion of some clauses none of which can be longer than the previously found

\[
s_4 \iff [(l \land d \land r_5) \lor (m \land \neg(e \lor r_4)) \lor (e \land \neg(m \lor r_4)) \lor (r_4 \land \neg(e \lor m))]
\]
Thanks to this result we can derive an approximation of the length of the total boolean formula obtained by the reduction algorithm shown in Example 5.1.

**Theorem 5.3.** The problem Expcomposite can be reduced to an instance of SAT with input length \( \in O(n^2) \) where \( n \) is the input length of Expcomposite().

*Proof.* Let \( ABCD... \) be the input of Expcomposite(\( ABCD... \)). Let \( Z : \{z_1, z_2, z_3, ..\} \in \{A, B, C, D, ...\} \) the set of fixed digits of \( ABCD... \). We are going to reduce this problem to \( \neg z_1 \land \neg z_2 \land \neg z_3, ... \land SAT' \), where every variable \( z_1, z_2, z_3, .. \) is denied if his corresponding digit in the input has value 0, and \( SAT' \) is the boolean expression defined in the next steps. Now, the variables of \( SAT' \) for every input of excomposite() can be visualized in this schema (where all the letters represent boolean variables):

\[
A B C D... = \\
\begin{align*}
& a b c d...* \\
& h i ... = \\
& a_1 b_1 c_1 ...+ \\
& a_2 b_2 c_2 ...0 = \\
& \mid r_{1.1} r_{1.2} r_{1.3}... \\
& s_{1.1} s_{1.2} s_{1.3}...+ \\
& b_3 c_3 d_3 e_3...0 0 = \\
& \mid r_{2.1} r_{2.2} r_{2.3}... \\
& s_{2.1} s_{2.2} s_{2.3}...+ \\
& c_4 d_4 e_4 ...0 0 0 = \\
& ... 
\end{align*}
\]

Call *prodlines* all the rows that are neither a sumline nor an auxiliary one except for the first three. Note that the first row can be considered a *sumline*, as it is the sum of the last two rows. As I showed in the previous examples, we are going to build an instance \( SAT' \) that is the conjunction of some condition which link the value of every variable to the variables of the previous rows, such that at the end all the variables are connected with \( A, B, C, D... \) and so with the variables of the input number which we are factorizing. \( SAT' \) is the conjunction of three types of expressions: the formulas which determine the value of a variable in one sumline from
variables in the previous rows (as \( s_{1,1} \iff (a_2 \lor r_{1,2}) \land \neg (a_2 \land r_{1,2}) \)), the
ones which determine the value of a variable in a prodline from variables
in the previous rows (as \( a_1 \iff (a \land h) \)) and the one which determine
the variable of an auxiliary line from the previous rows (as \( r_{1,1} \iff 
[(a_1 \land b_2) \lor (a_1 \land r_{1,2}) \lor (b_2 \land r_{1,2})] \)).

We have just seen that an expression of the first type cannot contain
more than 42 symbols (as it is involved in a sum of only two numbers)
and that any expression of the third type is shorter than 21 symbols. We
can also tell for sure that all the expression of the second type are not
longer that 7 symbols, as they have always the form \( v \iff (w \land u) \).
Of course every variable of a prodline corresponds to the product of two
digits of the two initial numbers, so it has value true (=1) if and only if
the two variables corresponding to the digits that we are multiplicityng
are both true. We have now found that there is a fixed limit (42 or or 21 or
7) to the length of every conjunctive clause of this instance of SAT'. Now
since every in clause (except for the first \( n \) which determine the first
prodline) appears one and only one variable which had never appared
before, the total number of the clauses grows linearly with the number
of variables. Since the number of variables grows linearly with the square
of the imput and all the clauses have a fixed limit of length, the length
of the final instance SAT' is \( \in O(n^2) \), where \( n \) is the length of the input
number. This means that the total length of \( \neg z_1 \land z_2 \land \neg z_3, \ldots \land SAT' \) is
also \( \in O(n^2) \).

We can generalize this result to all the problems in class \( F \), if the
set \( C_1, C_2, \ldots \) can be translated into an instance of SAT in the boolean
variables associated to the digits of the input number, the multiplier
and the multiplicator whose length is, at most, quadratic with respect to
the input. We could call \( FP \) this subset of \( F \)

**Corollary 5.4.** Every problem \( f \in FP \) can be reduced to an instance of
SAT with with input length \( \in O(n^2) \) where \( n \) is the input length of \( f \)

**Proof.** From the definition of \( F \) we know \( f \) requires to find a factorization
of a number \( n \) which satisfies some conditions. Let \( S \) be the formula of
SAT obtained by the reduction of Excomposite\( (n) \). Let \( A, B, C \ldots \) be the
boolean variables associated to the digits of the input number, \( a, b, c \ldots \)
be the boolean variables associated to the digits of the multiplicand and
$a', b', c'$ be the boolean variables digits of the multiplicator. By definition, all the conditions $C_1, C_2, \ldots$ imposed by $f$ can be expressed with boolean formulas in the variables $A, B, C, \ldots a, b, c, \ldots a', b', c'$ with quadratic length with respect to the input. Calling $J$ their conjunction, since $J$ is a conjunction of a finite set of boolean formulas with quadratic length, $J$ too has quadratic length with respect to $n$. Now, imposing that the same variables have been indicated with the same name in $S$ and $J$, $f$ can be reduced to the satisfability of the formula $S \land J$. $S \land J$ has, obviously, length $\in O(n^2)$.

Fortunately, there are lots of proposition about the digits of a binary number that can be reduced to a SAT formula with quadratic length or less!

**Corollary 5.5.** Factoring$(N)[L,U]$ can be translated into a SAT instance with length $\in O(n^2)$

**Proof.** It suffices to prove that this problem $\in FP$. Factoring imposes one conditions: one of the two numbers in the factorization of $N$ must be $L < x < U$. Let $a_1a_2a_3\ldots$ be the digits of the multiplicand ($A$), $l_1l_2l_3\ldots$ be the digits of $L$ and $u_1u_2u_3\ldots$ be the digits of $U$. We have to reduce the fact that $A > L$ to a boolean formula $S$ with quadratic length or less, then the instance of SAT which expresses the condition just will certainly result as long as $4S$. Call $A_n, L_n$ the expressions obtained from $A$ and $L$ respectively by cancelling the first $n$ digits. This means that $A = A_0$. Of course, $A > L$ if $a_1 > l_1$ and $A < L$ if $a_1 < l_1$. If $a_1 = l_1$ we can simply repeat the process on $a_2$ and $l_2$. This means that $A > L \iff (a_1 > l_1) \lor (\neg(a_1 < l_1) \land (A_1 > L_1))$. This fact can be generalized as $A_n > L_n \iff (a_{n+1} > l_{n+1}) \lor (\neg(a_{n+1} < l_{n+1}) \land (A_{n+1} > L_{n+1}))$, of course if $n$ equals the number of digits of $A$ and $L$, $A_n > L_n$ is false. Now, considering the boolean variables associated to the digits of $A$ and $L$, $a_n > l_n$ corresponds to $a_n \land \neg l_n$ and $a_n < l_n$ corresponds to $l_n \land \neg a_n$. This means that we can translate the previous recursive arithmetical formula to a recursive boolean formula:

$$A > L \iff (a_1 \land \neg l_1) \lor (\neg(l_1 \land \neg a_1) \land (A_1 > L_1))$$

$$A_n > L_n \iff (a_{n+1} \land \neg l_{n+1}) \lor (\neg(l_{n+1} \land \neg a_{n+1}) \land (A_{n+1} > L_{n+1}))$$
As the total length of this formula, which express the relation of ”>” between binary numbers, grows linearly with the number of digits of $L$ and $A$, Factoring $\in FP$.

5.1. **Final observation.**

**Proposition 5.6.** If $P = NP$ and SAT is solvable in time $\in O(n^a)$, any problem $f \in F$ is solvable in time $\in O(n^{2a})$.

**Proof.** Since $f$ can be reduced to an instance of SAT with length $l \in O(n^2)$, if we were able to solve SAT in time $\in O(l^a)$ for an input of length $a$, we could also solve $f$ in time $\in O(n^{2a})$.

**References**

[BC] D. Bovet, P. Crescenzi, *Introduction to the theory of complexity*, Prentice Hall, 1993.

[AB] S. Arora, B. Barak, *Computational complexity, a modern approach*, Princeton, 2007.

[AKS] M. Agrawal, N. Kayal, N. Saxena, *PRIMES in P*, Indian institute of technology Kanpur, 2002.

*E-mail address: davide.maran@mail.polimi.it*