Description of the project Feasibility, Logic and Randomness (FEALORA)

This project was proposed as an Advanced Grant of the European Research Council. The panel recommended it for funding and preparation of the grant agreement has been initiated. It is scheduled to start January 1, 2014 and last five years.

The principal investigator will be me (Pavel Pudlák). There will be four permanent positions for co-investigators partly funded by the grant: Pavel Hrubeš, Emil Jeřábek, Michal Koucký and Neil Thapen.

There will be funds for one post-doc position during the whole period and two PhD-student positions. These positions will be filled on a competitive basis and start January 1, 2014. The Mathematical Institute will publish an official call on its web-page. I recommend, however, that you contact me directly as soon as you start considering to apply. Also do not hesitate to contact me, or the project manager Beata Kubis (kubisb@math.cas.cz) if you have any questions.

The **post-doc positions** are intended for young researchers with a background in computational complexity and logic. The typical duration should be 1 or 2 years.

The **student positions** are for talented students who want to do their PhD at Charles University in Prague (with me as the supervisor) working on a topic related to the project. The support will be given for up to 4 years.

The topic of the project is *proof complexity* with *feasible incompleteness* as a central theme. Feasible incompleteness, roughly speaking, refers to the study of the incompleteness phenomenon in systems with bounded resources. (For more detail, see below or [P13, Section 6.4].) My motivation for focusing on this topic is twofold:

1. It is an approach to explaining why problems in computational complexity are so difficult. It may also lead to proofs of independence (at least unprovability in some relatively weak theories or proof systems).

2. It is a source of problems in proof complexity as well as in computational complexity. In fact, many results in proof complexity can be viewed as confirming certain conjectures about feasible incompleteness.

Since the topic of feasible incompleteness is rather special, we do not expect that all students, postdocs and visitors will work on it. One should view feasible incompleteness as a an approach to problems in computational complexity, not as the ultimate goal. So if participating researchers make progress in solving fundamental problems in computational complexity theory or proof complexity using different means, it will be absolutely fine.

The Scientific Content of the Project

In this project we want to study theoretical aspects of computational complexity. Computational complexity is one of the central areas of theoretical computer science with deep and difficult open problems. Since its origin in the 1960s computational complexity has made tremendous progress. Many important concepts have been introduced, many difficult problems have been solved and a number of new proof methods have been developed. In spite of this the main open problems, such as **P** vs. **NP**, remain as widely open as they were before.

Therefore it is necessary to study the reasons why these problems cannot be solved using the current methods. The aim of this project is to study the limitations of the methods systematically and develop a theory that will explain why the problems are difficult. Our approach is based on studying the incompleteness phenomenon in the context of polynomial time computability, which we call *feasible incompleteness.* The basic idea is that important conjectures in complexity theory, including **P** not equal to **NP**, are equivalent, or follow from plausible assumptions about unprovability, or lower bounds on the lengths of proofs of certain statements. The study of feasible incompleteness will be combined with the study of *pseudorandomness*, which is a concept that is connected with a number of important problems in computational complexity and proof complexity. Most of the technical work will concern proving special cases of feasible incompleteness conjectures and their relativized versions in order to provide evidence for their truth.

A brief overview of the research area

1. Results about limitations of methods. One of the first results showing limitations of methods in computational complexity was the application of the concept of relativization [BGS75]. It was shown that for most of the open questions about pairs of complexity classes one can find oracles such that the question relativized to oracles is decided in both ways ways (equal and unequal). This shows that simple methods, such as straightforward application of diagonalization, cannot be used to solve these problems. Another early result about limitations of methods is due to Razborov [R89]. He showed that his method of approximation, which had been successfully used to prove exponential lower bounds on monotone circuits and bounded depth circuits with modular gates, cannot work for general circuits. Razborov's result on the limited applicability of the approximation method can be stated in plain words as the fact that one cannot base a lower bound on circuit complexity on the concept of progress towards the computed function. For Boolean formulas, it possible to use "measures of progress", but still we do not have more than cubic lower bounds. One result that explains the difficulties of using this approach for formula size complexity is [HKPJ10]. It shows that many complexity measures are convex and for those one can only prove quadratic lower bounds. The most important of these results, because of its wide applicability, is the result of Razborov and Rudich [RR97] that shows that one cannot prove superpolynomial lower bounds on circuit complexity using *natural proofs. Natural proofs* are a precisely defined concept that captures the typical form of lower bound proofs where one first defines a property that implies high complexity and then shows that the given Boolean function has the property. Their result uses an unproven conjecture about the existence of strong pseudorandom generators. More recently, the method of relativization was extended to proofs that use approximations by low degree polynomials [AW08]. The concept is called *algebraic relativization*, or *algebrization.* Further results were proved using self-reducibility of some problems. Eg., in [PP10] we proved that a the existence of slightly subexponential algorithms for circuit satisfiability implies the existence of much faster algorithms for this problem. Therefore it is unlikely that there there are subexponential algorithms of this type.

2. Feasible incompleteness. This is the idea that the incompleteness phenomenon, as we know it from first-order logic, manifests itself also on the level of (polynomially) bounded proofs and computations. It is not a completely new idea (e.g., already in the 1950s Georg Kreisel mentioned a feasible version of Godel's theorem), but very little has been published and essentially no systematic research has been done. The first result in this direction was proved by H. Friedman [F79] (and independently by the PI [P86], since [F79] has never been published). It is a lower bound on the lengths of proofs of finite consistencies – statements asserting that a theory is consistent up to proofs of a given length. This result is interesting, but it concerns only proofs in a theory T of the consistency statements about T. A much more interesting question is what happens if we consider provability in a theory S of the consistency statements in another theory T where S is weak and T is strong. Then it is natural to conjecture that the proofs must have exponential length. Using the fact that the consistency statements are universal in a certain class of formulas, Krajíček and the PI showed that conjectures of this kind are equivalent to several natural statements about computational complexity [KP89]. These conjectures are stronger than the central open problems in complexity theory, such as **P** *not equal to* **NP**, hence if proved, they would settle fundamental problems. The reason for lack of publications

about feasible incompleteness is that the problems one needs to solve are apparently at least as hard as those in computational complexity theory. Therefore the PI proposed to develop *a system of conjectures that would explain,* rather than prove, the phenomenon of feasible incompleteness and thus would also give justification for commonly accepted conjectures about the basic complexity classes. In a short paper [P06] the PI presented more conjectures and proposed the term *the Feasible Incompleteness Thesis* for the system of these conjectures. The Feasible Incompleteness Thesis is treated in more detail in the forthcoming book [P13]. However, this project is just at the beginning.

An example of a feasible conjecture is: *For every recursively axiomatized theory T there, there exists a total polynomial search problem P such that T is unable to prove that P is total for any formalization of the problem by a low complexity formula*[1]. An equivalent statement purely in terms of computational complexity is: *For every total polynomial search problem P, there exists a total polynomial search problem S that is not polynomially reducible to P.* While the second statement does not give us much clue why it should be true, the first one has an explanation: because of incompleteness of r. e. theories.

3. Proof complexity. Proof complexity is a research field that studies the question of how difficult it is *to prove* a theorem (in contrast to computational complexity which studies the question of how difficult it is *to compute* a function value). Proof complexity can be divided, like computational complexity, into uniform and nonuniform parts. The part of proof complexity that studies nonuniform concepts deals with the lengths of proofs in various proof systems for (classical) propositional calculus. The part of proof complexity that studies uniform concepts deals with weak arithmetical theories. A generic proof system for propositional calculus can be viewed as a nondeterministic Turing machine accepting the **coNP**-complete set of tautologies. A proof, then, is an accepting computation; its length is the length of this computation. We can also view it as a certain framework for deterministic algorithms. For example, the Davis-Putnam procedure corresponds to tree-like resolution proofs. Thus by proving exponential lower bound on tree-like resolution proofs, it has been shown that any deterministic algorithm that is based on Davis-Putnam procedure must run in exponential time on some inputs. Two general methods for proving lower bounds on the lengths of proofs are known. The first is the *random restriction method* that is known in circuit complexity. Its application in proof complexity is always technically more involved. The second method is the *feasible interpolation method*. Both methods can only be applied to relatively weak proof systems. In Boolean circuit complexity there is another important method: the method of approximation. This method has not been applied in proof complexity yet. It is a challenge to find a version of this method suitable for proof complexity. The weak first-order theories are called bounded arithmetic. Each such theory is usually associated with a complexity class. There are several ways how to formalize this association. One of these is as follows. Given a complexity class C, we pick a suitable set of formulas F that define sets in C. Then we define a theory using some finite set of basic axioms plus the schema of induction postulated for formulas from F. Further, these theories are associated with propositional proof systems, which can be viewed as their nonuniform counterparts. The association with propositional proof systems enables us: 1. to construct propositional proofs more efficiently by first finding a proof in the theory and then translating it into the propositional calculus, 2. to prove independence of sentences by proving superpolynomial lower bounds on the proofs in the associated proof system. The main reason why these theories are studied is that they are on the border where one can prove independence of statements about complexity classes. While proving unprovability of such statements in Peano Arithmetic is a hopelessly difficult problem, for some weak theories in bounded arithmetic, this is possible. One result of this kind is due to Razborov [R95]. He proved that one cannot prove superpolynomial lower bounds on circuit complexity in a certain theory. In terms of complexity classes, the theory is unable to prove that **NP** *is not a subset of* **P/poly**. On the other hand, one can prove for a stronger (and more natural) theory that it is consistent with **NP** *not equal to* **coNP** [folklore].

4. Randomness, pseudorandomness and lower-bound methods. The concept of pseudorandomness will play one of the key roles in our project. Pseudorandom generators were introduced in order to simulate sources of genuine random bits that are needed in probabilistic algorithms. This concept is very important in

1 Low complexity means Sigma^b_1.

theoretical cryptography because of is close connection with one-way functions. Furthermore, it is also important in proof complexity. There is a growing body of evidence that the concept is connected with circuit lower bounds and thus connected with fundamental problems about separating complexity classes. Already Claude Shannon observed that a random Boolean function has exponential circuit-size complexity. So, in principle, it suffices to imitate random functions by pseudorandom ones in order to get explicit boolean functions with exponential circuit complexity. There are much more concrete connections. In particular, derandomization of certain randomized algorithms implies lower bounds on circuit complexity and, vice versa, sufficiently large lower bounds on Boolean circuit complexity imply the existence of pseudorandom generators [NW94,IW97,KI04]. Pseudorandomness is also important for the aforementioned result of Razborov and Rudich – their theorem uses a conjecture about the existence of strong pseudorandom generators. Pseudorandom generators are further used in an important conjecture of Krajíček and Razborov about hard tautologies [R03,K11]. According to their conjecture, the existence of pseudorandom generators implies superpolynomial lower bounds on the lengths of proofs of tautologies in Frege systems, which is a big open problem in proof complexity. This conjecture is important because, in spite of being very strong, it is not known to imply **NP** *not equal to* **coNP.** Furthermore, it has been proved for a certain class of proof systems (instead of Frege system) [Pi11].

5. Pseudorandomness in number theory. Pseudorandomness plays an important role also in number theory. In number theory considerations about random behavior of some sequences have been used to justify conjectures in an informal way and to explain why some algorithms are fast. There are now results that consider pseudorandomness as a formal concept. In particular, a certain concept of pseudorandomness plays an important role in the famous result of Green and Tao about arithmetic progressions of primes [GT08]. Subsequently, their idea was analyzed in complexity theory and interesting connection were found [RTTV08]. Another example is the *Mobius Randomness Principle,* a conjecture proposed informally by Peter Sarnak. It can be formalized by saying that the sequence of values of the Mobius function is pseudorandom in a well-defined sense. This means that it is not possible to distinguish the sequence from a random sequence using polynomial time computable tests. (Since the values of the Mobius function are 0, 1, and -1, and we are comparing it with a random sequence of 1s and -1s, we have to ignore the 0s.) The simplest special case of the conjecture is equivalent to the Prime Number Theorem. Very recently Ben Green proved the conjecture for **AC0** tests [G13]. These connections are important for our project. We will use models of arithmetic to study this kind of pseudorandomness.

Objectives of research

1. Develop a system of conjectures in order to justify commonly conjectured inequalities between complexity classes. The idea of explaining the difficulty of the fundamental problems in complexity theory attracted a lot of attention in the past few years. The term *barriers in computational complexity* has become fashionable, but the results obtained so far only give more precise delimitation of the barriers without explaining why there are any barriers. We want to take this challenge seriously and develop a global theory of the open problems in complexity theory. We hope that this theory will also suggest which directions of research should be pursued in order to make progress in solving the fundamental problems.

2. Develop better understanding of feasible incompleteness. Working towards objective 1, we will focus on feasible incompleteness. The reason is that we believe that the nature of the fundamental problems in complexity theory is logical. The results obtained and methods developed in proof complexity enable us to analyze and test the conjectures. To this end we will prove special cases and relativized versions of the conjectures. In this way, the conjectures will guide us in which direction we should pursue the research in proof complexity.

3. Find connections between pseudorandomness and feasible incompleteness. Although pseudorandomness and feasible incompleteness seem to be of fundamentally different nature, some connections are already known and there are good reasons to expect that more will be found. One reason is that in several cases conjectures about pseudorandom generators have been used to give conditional solutions to problems in proof complexity. Another reason is the connection of derandomization with circuit lower bounds. A concrete

example of a conjecture connecting these two concepts that has already been stated is the aforementioned Krajíček-Razborov's conjecture.

A sample of specific problems

1. Find a natural feasible incompleteness conjecture that implies all feasible conjectures considered so far.
2. Find connection of feasible incompleteness conjectures with standard conjectures about pseudorandom generators and extend the system to include conjectures that imply them.
3. Find connections with proof-complexity generators of Krajíček and Razborov [R03,K11].
4. Explain pseudorandomness using a system of models with a probability measure or using Boolean valued models [K11]. A preliminary result in this direction is in [P12].
5. Find a nontrivial example of a trade-off between the complexity of an algorithm and strength of a theory in which it is provable that the algorithm solves a given problem.
6. Prove a lower bound on a theory or a proof system in which it is possible to prove some pseudorandom properties of particular structures. Except for simple kind of circuits, properties of pseudorandom generators have not been formally proved. We believe that there are some fundamental reasons why it is so.
7. Prove separation for relativized total search problems associated with the Bounded Arithmetic Hierarchy. One of the conjectures is that the hierarchy of total search problems is unbounded and this is because the theories with which they are associated have increasing strength. If we prove that total search problems associated with Bounded Arithmetic Hierarchy have increasing strength relative to an oracle, it will be evidence for the truth of this conjecture.
8. Construct a hierarchy of total search problems indexed by ordinal less than some transfinite ordinal greater than omega. Countable ordinals have been successfully used in proof theory as a parameter that determines the strength of theories. This approach failed for weak theories studied in proof complexity. It seems, however, conceivable that it may work for concepts such as total search problems because they correspond to fast growing functions in strong theories and those can be classified by ordinals.
9. Find a concept in proof complexity that corresponds to natural proofs of Razborov-Rudich or find an argument that there is none. It has been shown that many concepts of computational complexity have their counterparts in proof complexity. The concept of natural proofs in proof complexity is elusive so far.
10. Formalize pseudorandom generators in weak theories. This is an important problem because proving lower bounds on circuit complexity is tightly connected with constructions of pseudorandom generators.
11. Find a natural theory in which one can formalize current lower bound techniques, but which is unable to prove lower bounds for general circuits. A significant progress towards this goal has been made by Razborov [R95], but the theory defined in his paper is not natural.
12. Prove lower bounds on the complexity of proofs of some properties of combinatorial structures, such as Ramsey graphs and expanders. A related problem is to prove lower bounds on the complexity of proving properties of some pseudorandom generators.

References

[AW08] S. Aaronson and A. Wigderson: Algebrization: A New Barrier in Complexity Theory. Proceedings of STOC 08, 731--740, (2008)
[AK10] E. Allender, M. Koucký. Amplifying Lower Bounds by Means of Self-Reducibility. Journal of ACM, 57(3), 1--36 (2010)
[BGS75] T.P. Baker, J. Gill, R. Solovay: Relativizations of the P =? NP Question. SIAM Journal on Computing, 4(4), 431--442 (1975)
[F79] H. Friedman: On the consistency, completeness and correctness. Unpublished typescript, (1979)
[G13] B. Green: On (not) computing the Möbius functions using bounded depth circuits. Combinatorics,

Probability and Computing, to appear
[GT08] B. Green, T. Tao: The primes contain arbitrarily long arithmetic progressions,. Annals of Math. 167 (2008), 481--547
[HKPJ10] P. Hrubeš, A. Kulikov, P. Pudlák, S. Jukna: On convex complexity measures, Theoretical Computer Science 411, 1842--1854 (2010)
[I95] R. Impagliazzo: A personal view of average-case complexity. In: 10th Annual Structure in Complexity Theory Conference (SCT'95), 134--147 (1995)
[IW97] Impagliazzo, R., Wigderson, A.: P=BPP unless E has Subexponential Circuits: Derandomizing the XOR Lemma. Proceedings of the 29th STOC, 220--229 (1997)
[KI04] V. Kabanets, R. Impagliazzo: Derandomizing polynomial identity tests means proving circuit lower bounds. Computational Complexity, 13(1-2), 1-46 (2004)
[K95] J. Krajíček: Bounded arithmetic, propositional logic, and complexity theory. Encyclopedia of Mathematics and Its Applications, Vol.60, Cambridge University Press (1995)
[K97] J. Krajíček: Interpolation theorems, lower bounds for proof systems, and independence results for bounded arithmetic. Journ. of Symbolic Logic 62(2), 457--486 (1997)
[K11a] J. Krajíček: On the proof complexity of the Nisan-Wigderson generator based on a hard NP \cap coNP function. J. of Mathematical Logic 1, 11--27 (2011)
[K11b] J. Krajíček: Forcing with random variables and proof complexity. London Mathematical Society Lecture Note Series, No.382, Cambridge University Press, (2011)
[KP89] J. Krajíček, P. Pudlák: Propositional proof systems, the consistency of first order theories and the complexity of computations. Journ. of Symbolic Logic 54(3), 1063--1079 (1989)
[KP90] J. Krajíček, P. Pudlák: Propositional provability and models of weak arithmetic. In: Proc. Computer Science Logic'89, Eds. Borger, Kleine-Buning, Richter, Springer-Verlag LNCS 440, 193—210 (1990)
[KNT11] L. Kolodziejczyk, P. Nguyen and N. Thapen: The provably total NP search problems of weak second order bounded arithmetic. Annals of Pure and Applied Logic 162(6), 419--446 (2011)
[KNP11] M. Koucký, P. Nimbhorkar, and P. Pudlák. Pseudorandom generators for group products. Proc. 43rd Annual ACM Symposium on Theory of Computing (STOC'11), 263--272 (2011)
[NW94] N. Nisan, A. Wigderson: Hardness vs. randomness. Journal of Computer Systems and Sciences 49(2), 149--167 (1994)
[PP10] R. Paturi, P. Pudlák: On the complexity of circuit satisfiability. Proc. 42nd Annual ACM Symposium on Theory of Computing (STOC'10), 241--249 (2010)
[Pi11] J. Pich: Nisan-Wigderson generators in proof systems with forms of interpolation. Mathematical Logic Quarterly 57(4), 379–383 (2011)
[P86] P. Pudlák: On the length of proofs of finitistic consistency statements in first order theories. In: Logic Colloquium 84. North Holland, 165--196 (1986)
[P06] P. Pudlák: G\"odel and computations. ACM SIGACT News 37(4), 13--21 (2006)
[P13] P. Pudlák: Logical Foundations of Mathematics and Computational Complexity, a gentle introduction. Springer-Verlag, 2013
[P12] P. Pudlák: Randomness, pseudorandomness and models of arithmetic. arXiv:1210.4692
[PT12] P. Pudlák, N. Thapen: Alternating minima and maxima, Nash equilibria and Bounded Arithmetic. Annals of Pure and Applied Logic 163 (2012), pp. 604-614.
[R89] A. Razborov: On the Method of Approximation. In Proc. of the 21st ACM Symposium STOC'89, 169--176 (1989)
[R95a] A. Razborov: Unprovability of lower bounds on the circuit size in certain fragments of bounded arithmetic. Izvestiya of the Russian Academy of Science, Mathematics 59(1), 201--224 (1995)
[R95b] A. Razborov, A.: Bounded arithmetic and lower bounds in Boolean complexity. In: Feasible Mathematics II, eds. P. Clote and J. Remmel, Birkhauser, 344--386 (1995)
[R03] A. Razborov: Pseudorandom generators hard for k-DNF resolution and polynomial resolution calculus. Preprint (2003)
[RR97] A. Razborov, S. Rudich: Natural Proofs. Journal of Computer and System. Sciences 55(1), 24--35 (1997)
[R05] O. Reingold: Undirected ST-connectivity in log-space. Proc. STOC'05, 376-385 (2005)
[RTTV] O. Reingold, L. Trevisan, M. Tulsiani, S. Vadhan: New Proofs of the Green-Tao-Ziegler Dense Model Theorem: An Exposition. ArXiv:0806.0381 (2008)