and longwavelengths, SAPHIR with six channels for atmospheric WV distribution and MADRAS operating in microwave region (89 and 157 GHz) for study of convective systems. The data from this mission are expected to provide better insights into the convective processes in the tropical regions. Table 4 summarizes the future space missions of direct relevance to our region.

With several advanced missions planned in the near future providing operational meteorological parameters, efforts will be focused towards assimilation of the space data in prediction models and to achieve improved forecasts in short, medium and long ranges.

1. Ramakrishnan, S. *et al.*, *Curr. Sci.*, 1999, **77**, 1038–1045.
2. Kelkar, R. R., in *Advances in Tropical Meteorology*, 1993, Tata McGraw Hill, p.179.
3. Arkin, P. A. and Ardanuy, P. E., *J. Climate*, 1989, **2**, 1229–1238.
4. Soden, B. J. and Bretherton, F. P., *J. Geophys. Res.*, 1997, **98**, 16669.
5. Velden, C. S. *et al.*, *Bull. Am. Meteorol. Soc.*, 1997, **78**, 173–179.
6. Menzel, W. P. and Purdom, J. F. W., *ibid*, 1994, **75**, 757–767.
7. Joshi, P. C., Simon, B. and Thapliyal, P. K., *Mausam*, 2001, **52**, 109–116.
8. Bhatia, R. C., Brij Bhushan and Rajeswara Rao, V., *Curr. Sci.*, 1999, **76**, 1448–1450.
9. Sikka, D. R. and Gadgil, S., *Mon. Weather Rev.*, 1980, **108**, 1840–1853.

ACKNOWLEDGEMENTS. The planning and execution of METSAT mission was a large team effort involving several scientists and engineers from the various ISRO centres. The focused and coordinated scientific efforts resulted in achieving a successful mission and the contributions of the team are gratefully acknowledged. We thank Dr K. Kasturirangan, Chairman, ISRO; Dr G. Madhavan Nair, Director, VSSC; Dr P. S. Goel, Director, ISAC and Dr A. K. S. Gopalan, Director, SAC for their encouragement and guidance.

Received 7 October 2002; revised accepted 23 October 2002

Fields Medals and Nevanlinna Prize: 2002

C. S. Rajan, Nitin Nitsure and Jaikumar Radhakrishnan

John Charles Fields, a Canadian mathematician who was Secretary of the ICM (International Congress of Mathematicans) held at Toronto in 1924, mooted in 1931 the idea of two medals to be awarded at successive ICMs (held once in four years) for outstanding achievements in mathematics, intended at the same time to be an encouragement for further achievement. When he passed away in 1932, in his will he had left his estate for the prizes. Fields apparently felt strongly about the lack of Nobel Prize in mathematics, which he tried to make up for with his limited means, and the prizes, though modest in monetary value, have held great prestige and emotional appeal to the mathematical community, similar to the Nobel Prizes.

Two to four medals are now awarded at each ICM, to mathematicians under the age of 40 (keeping in view the twin objectives). At the latest ICM held in Beijing in August this year, the medals were awarded to Laurent Lafforgue and Vladimir Voevodsky.

The IMU (International Mathematical Union) which operates the Fields Medal awards, and also organizes the ICMs, established in 1981, a medal and prize similar to the Fields Medal, for outstanding achievements in mathematical aspects of information science. The prize is named after the Finnish mathematician Rolf Nevanlinna who was President of the IMU, and is awarded along with the Fields Medals. This time the award went to Madhu Sudan.

We present here glimpses of the work of the three awardees: the notes are prepared by Rajan on Lafforgue, Nitsure on Voevodsky, and Radhakrishnan on Madhu Sudan.

Laurent Lafforgue

Laurent Lafforgue was awarded the Fields Medal for the proof of the Langlands conjectures for the case of the general linear group over function fields. The conjectures of Langlands represent a vast generalization of the classical quadratic reciprocity law. These conjectures envisage an intricate relationship between arithmetical objects on the one hand and analytical data on the other.

Reciprocity laws can be traced back to the observation of Fermat that a prime number p can be expressed as a sum of two integral squares, i.e. $p = x^2 + y^2$, where x and y are natural numbers, if and only if $p - 1$ is divisible by 4; a well-known identity which goes back to Brahmagupta then shows that a general natural number is a sum

C. S. Rajan and Nitin Nitsure are from the School of Mathematics, and Jaikumar Radhakrishnan is from the School of Technology and Computer Science, Tata Institute of Fundamental Research, Homi Bhabha Road, Mumbai 400 005, India.

of two squares if all its prime factors satisfy the condition. Fermat asserted also that primes p of the form $x^2 + 2y^2$ are precisely those that are congruent to 1 or 3 modulo 8 (namely, leave a reminder of 1 or 3 upon dividing by 8), and primes of the form $x^2 + 3y^2$ are precisely those that are congruent to 0 or 1 modulo 3.

Euler continued on the theme and investigated the condition for a prime p to divide an expression of the form $x^2 + ay^2$, where x and y are natural numbers with no common factors. Considering the expression modulo the prime p, we see that a necessary condition for this is that a has to be a square modulo p. The quadratic reciprocity conjecture essentially is a statement that the set of primes p for which a is a square modulo p are precisely those occurring in certain specific congruence classes modulo $4a$.

The law of quadratic reciprocity was finally proved by Gauss. As a step towards understanding possible higher reciprocity laws, Gauss gave many proofs of the quadratic reciprocity law during his lifetime. As an example of such a relationship, it may be mentioned that the cubic reciprocity law implies the fact, conjectured initially by Euler, that a prime p is of the form $x^2 + 27y^2$, if and only if 3 divides $p - 1$, and 2 is a cube modulo p.

It was realized by Gauss and Jacobi, that the theory of quadratic reciprocity laws and even the formulation of higher reciprocity laws, involve intimately the arithmetic of what are called algebraic numbers. These are numbers which are linear combinations, with rational coefficients, of the roots of a (fixed) rational polynomial, such as $a + b\sqrt{2}$, or $a_0 + a_1\omega + \ldots + a_{n-1}\omega^{n-1}$, where ω is an nth root of unity, a, b, $a_0, \ldots, a_{n-1}$ being rational numbers. An attempt of Kummer at Fermat's last theorem, which turned out to be erroneous in a curious way, led him to a deep study of the arithmetic of the number fields generated by roots of unity. In an effort to restore the unique factorization property of rational integers into prime numbers in a more general framework, Kummer in a bold and decisive step, defined 'ideal' numbers, and showed that they indeed have a unique factorization property. Over a period of nearly fifteen years Kummer formulated and proved the higher reciprocity laws over the cyclotomic fields. Subsequently, Kronecker, Dedekind, Weber and Hilbert amongst others, laid out the basics of the arithmetic of algebraic number fields needed for a study of the reciprocity laws.

Meanwhile, Galois had considered the group of symmetries (called the Galois group) of a rational polynomial as those permutations of the roots preserving the polynomial, and used it to study the structure of subfields of algebraic extensions of a given field. This incidentally, is the same work from which he deduced that general quintic equations cannot be solved by radicals. In the work of Takagi, a Japanese mathematician, an intrinsic description was obtained of the Galois group of the algebraic extensions of a given number field, in terms of the base number field, provided the Galois group of the extension field is commutative. The crowning moment came with the formulation and proof by Artin in the 1920s of the reciprocity law for such extension fields, based on earlier work done by Frobenius and Chebotarev.

In a vague sense, one of the goals of the reciprocity laws is to define intrinsically the collection of prime numbers which are expressible as the absolute value of the product of transforms of a given element in the number field, with respect to the elements of the Galois group. For example, the problem of describing the primes which can be written as a sum of two squares, can be reformulated in terms of the ring of Gaussian integers defined by all integral linear combinations of 1 and i, where i is a square root of -1. The primes p for which $p - 1$ is divisible by 4, are precisely those that can be written as a product $p = (x + iy)(x - iy)$, where x and y are rational integers.

However, the formulation and the proof of the reciprocity laws are neither easy nor direct. An important point here is that in any given arithmetical situation, there arises along with the Galois group, a collection of conjugacy classes in it (called Frobenius classes) indexed by the primes in the number field. Given such data, a way of dealing with it is to consider the group concretely as a subgroup of the group of invertible $n \times n$ matrices, say over the complex numbers, via representations of the group. The trace of a matrix being invariant under conjugation, then defines numerical data on the collection of conjugacy classes. These numerical data, denoted by $g_p(A, r)$, depend on the arithmetical data A, the representation r of the Galois group and the prime number p. The aim of the reciprocity laws is to understand this collection of conjugacy classes and the arithmetical data $\{g_p(A, r) \mid p \text{ a prime}\}$, in terms of data which are constructed out of the base field. We will call these data as *Galois* data.

Another way of constructing arithmetical data, is from the study of rational or integral solutions to systems X of polynomial equations defined over rationals; for example, the Fermat equation $x^n + y^n = z^n$, or a cubic equation $y^2 = x^3 + ax + b$, in two variables, for some rational numbers a, b (the so-called elliptic curves). Over the finite fields $\boldsymbol{F}_p$ of congruence classes modulo p (where p is any prime number), the number of such solutions gives arithmetically defined data (called *diophantine* data) depending on the system of equations X, and indexed by the set of primes p. It is a consequence of the deep work of Grothendieck (for which he won the Fields Medal in 1966), that diophantine data can be suitably interpreted as data of Galois kind.

In another stream, Dirichlet initiated in the 1830s a study of properties of numbers using analytical methods. Dirichlet proved that given two integers a and d without any common factors, there are infinitely many primes p such that $p - a$ is divisible by d. For proving this Dirichlet defined, inspired by a result of Euler concerning the infinitude of primes, analogues of the (Riemann) zeta

function, that are now called as Dirichlet L-functions. The zeta and L-functions were studied further by Riemann and Hecke respectively, and they were shown to have nice analytic properties. This theme was taken up by Artin who defined generalizations of Dirichlet L-functions, associated to Galois data (and depending on them). He conjectured that these functions also have nice analytic properties.

A third way of obtaining arithmetical data started with the work of Ramanujan, who considered the arithmetical properties of modular forms. Ramanujan considered the Fourier coefficients $\tau(n)$, indexed by natural numbers n, of the discriminant function $\Delta(z)$ defined on the collection of elliptic curves (the discriminant function measures when the polynomial $x^3 + ax + b$ fails to have distinct roots). Ramanujan conjectured that the $\tau(n)$ satisfy some natural multiplicative properties, one of them being that $\tau(mn) = \tau(m)\,\tau(n)$, provided m and n have no common prime factors. The conjectures of Ramanujan were proved and given a geometric and group-theoretic framework by the German mathematician Hecke, a contemporary and colleague of Artin. The study of modular forms continued in the work of Petersson, Maass, Rankin, Selberg and Weil. It is possible to define zeta-type functions in this setting, and also to show that these have nice analytical properties.

Langlands brought together the two strands of mathematical thought started by Artin and Hecke, and formulated the reciprocity laws in a new manner, generalizing the abelian reciprocity law of Artin. Langlands immensely generalized the work of Hecke and defined what are called as *automorphic* data, constructed analytically and group theoretically out of the base field and the group of invertible $n \times n$ matrices, containing as a special case the discriminant function considered by Ramanujan.

The reciprocity conjecture, as formulated by Langlands, is that *any Galois or diophantine data are automorphic*. More generally, any arithmetical data arising in a natural way are automorphic in nature. Conversely, it is possible to single out the Galois data from the automorphic data.

For $n = 1$, the conjecture of Langlands reduces to Artin's reciprocity law. Early confirmation of Langlands conjectures came from the representation theory of real reductive groups; for example, the group of invertible $n \times n$ real or complex matrices. Harish-Chandra had built up the edifice of representation theory for these groups. Langlands applied Harish-Chandra's theory to prove the appropriate reciprocity laws over the fields of real and complex numbers.

The second striking application of the conjectures of Langlands came with Deligne's proof of the Weil conjectures. The conjectures of Weil predict for diophantine data arising from smooth, projective X, a precise estimate for the growth of the number $n_{p^k}(X)$ of the solutions over the finite fields $\boldsymbol{F}_{p^k}$ consisting of p^k number of elements, as a function of k. It was shown before, also by Deligne, that the conjectures of Weil imply the estimate conjectured by Ramanujan for the size of $\tau(p)$, namely $|\tau(p)| \leq 2p^{11/2}$. Deligne's proof of Weil conjectures (for which he received the Fields Medal in 1978) relies, on the other hand, upon the ideas of Langlands and a method of Rankin providing a slightly weaker estimate for the size of $\tau(p)$ than that proposed by Ramanujan, coupled with the geometric machinery developed by Grothendieck.

A third confirmation for Langlands conjectures came with his own work proving conjectures of Artin in the special cases when the Galois group is solvable and the representation is of degree two. Together with the work of Eichler–Shimura, this is the starting point for the proof of Wiles of the Shimura–Taniyama–Weil conjecture (again a special case of the Langlands conjectures, but made before), which incidentally leads to a proof of Fermat's last theorem.

The work of Lafforgue concerns the Langlands conjectures over what are called *function fields*. Such a field is a finite extension of a field consisting of quotients $p(t)/q(t)$, where $p(t)$ and $q(t)$ are polynomials over a (fixed) finite field. Dedekind observed striking arithmetical similarities between the ring of integers and the ring of polynomials over a finite field; this accords the function fields the significance analogous to the algebraic number fields. This analogy of function fields with number fields was pursued by Artin, Hasse and Weil, and led to the formulation of the Weil conjectures for this case. An interesting feature of the function fields is that the reciprocity laws can be formulated geometrically, and in a stronger form, to say that there is a coincidence of Galois, diophantine and automorphic data.

A major breakthrough towards Langlands conjectures in the function field case was achieved by the Ukrainian mathematician Drinfeld, who settled the conjectures for the 2×2 general linear group over these fields. Drinfeld was awarded the Fields Medal in 1990 for this work. Again, the starting point of Drinfeld is to define analogues of the objects used by Eichler and Shimura, and to show that indeed the modular forms can be defined over such spaces. Drinfeld defined new objects called shtukas. The collection of shtukas is parametrized by a space analogous to a classical modular curve.

The work of Lafforgue generalizes Drinfeld's results to general linear groups of all ranks, over the function fields. He shows that in this set-up, the automorphic data are necessarily of Galois-type. His proof develops on the constructions introduced by Drinfeld. In generalizing Drinfeld's work to higher ranks, Lafforgue had to deal with the problems posed by the noncompleteness of the moduli spaces constructed by Drinfeld, and to construct geometric analogues of Hecke's theory to the completed spaces. This involves, by all accounts, formidable technical power and deep insight, which fetched him the coveted award this year.

Vladimir Voevodsky

Vladimir Voevodsky received the Fields Medal in recognition of his having developed new cohomology theories for algebraic varieties, thereby providing new insights into number theory and algebraic geometry.

Historically, the notion of cohomology comes from the subject of topology, which is sometimes called 'geometry of rubber sheets', as it studies those properties of a space that remain unchanged under contracting/stretching without tearing. Homology and cohomology theories (including 'generalized' cohomology theories) associate to any topological space some algebraic objects, such as abelian groups or vector spaces, with a multiplicative structure, and often some gradation or filtration. These algebraic objects associated to a space encode important data about the original space, but in a more usable form, which allows us to make algebraic calculations. Some examples of such theories are singular cohomology, de Rham cohomology, and topological K-theory (which is a generalized cohomology). In the last 60 years, homology and cohomology theories were developed for objects other than topological spaces. Modern mathematicians routinely use a large number of (co)homology theories, which variously apply to groups, Lie algebras, sheaves, etc. as basic computational tools. Topological K-theory was in fact inspired by the development of an algebraic analogue (called algebraic K-theory) which is a generalized cohomology theory that applies to various algebraic and algebro-geometric objects. Incidentally, algebraic K-theory is the idea of Alexander Grothendieck, and topological K-theory was developed by Michael Atiyah (and others), and both Grothendieck and Atiyah were awarded the Fields Medals in 1966 for their extensive contributions.

In contrast to topology, which studies the geometry of 'rubber sheets', algebraic geometry is the study of spaces which arise as loci of systems of polynomial equations in several variables. These spaces are known as algebraic varieties (or schemes), and are far more rigid than the spaces studied in topology. In modern algebraic geometry following Grothendieck, the polynomial equations may have coefficients in arbitrary fields, or even rings, which is what has led to a 'unification' of geometry and arithmetic, and has allowed geometric insights to be applied (with impressive results) to number theoretic problems.

It has been an important challenge to develop powerful homology and cohomology theories for algebraic varieties. When the field is that of complex numbers, the varieties come equipped with a topology, and one can apply the usual algebraic topology to deduce important geometric theorems. This strategy, described by Lefschetz as 'planting the harpoon of algebraic topology in the whale of algebraic geometry', was successful in the first half of the 20th century in the study of complex projective varieties. The famous conjectures of Weil gave rise to Grothendieck's programme of setting up certain cohomology theories for varieties over arbitrary fields (ℓ-adic cohomologies) which can do much of the job that singular cohomology does over complex numbers. This programme, pursued by Grothendieck and his students, was successful in finally proving the Weil conjectures around 1970, which is one of the towering achievements of 20th century mathematics; for his proof of the final Weil conjecture, Pierre Deligne (a student of Grothendieck) received the Fields Medal in 1978. The existence of a powerful cohomology theory for schemes over an arbitrary field, called motivic cohomology, which was to satisfy certain properties, was conjectured by Beilinson and Lichtenbaum in the late 1980s.

A visionary programme of Grothendieck calls for the creation of a theory of 'Motives', which will be the mother of all cohomology theories for algebraic varieties, in the sense that all other homology and cohomology theories will arise from it just by algebraic manipulations. In categorical language, the motivic theory will be the universal cohomology theory. The work of Voevodsky is a big stride towards fulfilling this vision of Grothendieck.

Progress in fulfilling Grothendieck's vision in its original form – that is, to set-up an abelian category of motives – has been held up because two fundamental conjectures of Grothendieck (which he proposed in 1968 at the International Colloquium on Algebraic Geometry held at the Tata Institute of Fundamental Research, Mumbai) are yet unproven. Grothendieck regarded proving these so-called 'standard conjectures' as one of the two most urgent tasks facing algebraic geometry (the other urgent task is to prove resolution of singularities in all characteristics). Meanwhile, the theory of derived and triangulated categories became well-developed as a means of addressing problems related to non-compact and singular spaces, and some mathematicians began trying to set-up a theory of motives (valid for non-compact, singular varieties as well) in a so-called *triangulated category*, rather than in an abelian category as Grothendieck had envisaged for non-singular projective varieties. (Note: Given an abelian category, say the category of all abelian groups, the new category which has chain complexes of abelian groups as objects, and homotopy classes of chain maps as morphisms, is a typical example of a triangulated category. If, moreover, we formally invert all chain maps which induce isomorphisms at the level of cohomologies of the complexes, then we get the so-called 'derived category' of the original abelian category, which is again a triangulated category.)

We now attempt to describe some highlights of the achievement of Voevodsky, without going into the technicalities – which are quite formidable! – to the extent possible.

Construction of motivic cohomology

One of the great successes of Voevodsky's construction of a triangulated category of motives is that he could extract from it the motivic cohomology envisaged by Beilinson and Lichtenbaum. Motivic cohomology was the dream of Beilinson and Lichtenbaum, who speculated that to any smooth variety over an arbitrary field k, there are naturally associated a family of groups denoted by $H^i_M(X, \mathbf{Z}(j))$, where i and j are integers – to be called as the *motivic cohomology groups* of X – which satisfy certain nice properties. The properties demanded of these groups $H^i_M(X, \mathbf{Z}(n))$ are of both geometric and arithmetical kind – they have to have good relations with both the Milnor K-groups and Chow groups!

The path towards Motivic cohomology taken by Voevodsky is to set-up a certain rigid tensor triangulated category for each field k, denoted by $DM(k)$ and called the *triangulated category of mixed motives over k*, together with an invertible object $\mathbf{Z}(1)$ of $DM(k)$ called the *Tate object*, and to define a functor

$$M : Sch/k \to DM(k),$$

where Sch/k is the category of schemes over k. This functor is Voevodsky's candidate for a universal cohomology theory, and by definition it associates to any scheme its triangulated motive (in the sense of Voevodsky). Having constructed this functor, Voevodsky *defines* motivic cohomology by the equation

$$H^i_M(X, \mathbf{Z}(j)) = Hom_{DM(k)}(M(X), \mathbf{Z}(j)[i]),$$

where $\mathbf{Z}(j)$ is the jth tensor power of $\mathbf{Z}(1)$, and shows that it has all the conjectural properties envisaged by Beilinson and Lichtenbaum. Moreover, he shows that some other theories such as algebraic singular homology $H_1^{alg}(X, \mathbf{Z})$ of Suslin, and Bloch's higher Chow groups $CH^i(X, j)$ can also be expressed in terms of the functor M, be bolstering the claims of M of being a universal cohomology.

It is expected that the triangulated category of motives constructed by Voevodsky will turn out to be the derived category of the abelian category of mixed motives of Beilinson. This requires in particular, the construction of a so-called t-structure on the triangulated category of motives, and even this has not been accomplished so far.

We end this brief description of Voevodsky's motivic cohomology by mentioning that there are two other independent constructions of triangulated categories of motives, besides that of Voevodsky, one by Levine and another by Hanamura, using different approaches. On the other hand, Madhav Nori has been developing a more direct construction of an abelian category of motives by following the model of singular homology of pairs!

The $\mathbf{A}^1$-homotopy theory

The concept of homotopy is one of the most important fundamental ideas of algebraic topology. Two maps f_0 and f_1 are said to be homotopic if one can be continuously deformed into the other by means of a one-parameter family of maps f_t, where $t \in [0, 1]$. The very simple idea behind $\mathbf{A}^1$-homotopy theory of schemes, to put it naively, is to try to mimic homotopy within algebraic geometry, by using the affine line $\mathbf{A}^1$ in place of the interval $[0, 1]$. Great difficulties had to be overcome to achieve this, which were accomplished by Voevodsky and Morel, thus completing a programme begun by Suslin. Also note that singular homology or cohomology is defined by means of singular simplices, which are maps from standard simplices Δ_n into a topological space. The idea of Suslin and Voevodsky was to replace Δ_n by the affine n-space $\mathbf{A}^n_k$, and to construct an algebro-geometric analogue of singular (co-)homology. Most algebraic varieties X do not admit enough morphisms from $\mathbf{A}^n_k$ into X, but maps can be successfully replaced by correspondences, thanks to the Dold–Thom theorem.

Homotopy theory is the technical heart of algebraic topology, with everything else flowing out of it (via classifying spaces, Eilenberg–Mac Lane spaces, Postnikov towers, Brown representability, spectra, K-theory, etc.). The creation of a theory very similar to algebraic topology but inside algebraic geometry, has now made it possible to repeat all kinds of topological constructions in geometry, and to transfer one's topological intuition to geometric domain. In the foundations of algebraic geometry provided by Grothendieck, classical geometry is unified with arithmetic. This is how the topological insights made available to geometry by the theory developed by Voevodsky, can have an impact on arithmetic.

We should emphasize that Voevodsky's construction of a triangulated category of motives is really fashioned out of his philosophy of having an algebraic homotopy theory as the master theory.

Proof of the Milnor conjecture

We now give two concrete achievements of Voevodsky, which come out of his general philosophy and constructions.

The first concrete achievement is the proof of the conjecture of Milnor which connects the Milnor K-theory of a field with its étale cohomology, and has deep arithmetical repercussions. It asserts that the quotient $K^{\text{Milnor}}_*(F)/2$ of the Milnor K-theory of a field F of characteristic $\neq 2$ is isomorphic to the étale cohomology $H^*_{et}(Spec(F), \mathbf{Z}/2)$.

Voevodsky's method of proving the Milnor conjecture is to convert the statement into the vanishing of certain motivic cohomology groups. He proves these vanishing statements by transferring from algebraic topology certain gadgets known as cohomology operations. Topologists have long used cohomology operations (such as Steenrod squares, etc.), and thanks to Voevodsky's

general constructions, these tools can now be applied to field theory!

Higher Chow groups and étale cohomology

Another concrete achievement of Voevodsky, which comes out of his general philosophy and constructions, is the result proved in collaboration with Suslin, which gives an isomorphism between higher Chow groups and étale cohomology.

As suggested by the above all too brief account of his work, Voevodsky has brought about a revolutionary new synthesis of topology with geometry and arithmetic, fulfilling old dreams. It is expected that many important, new results will be proved in the future using the tools and insights fashioned by Voevodsky.

Madhu Sudan

Madhu Sudan was awarded the Nevanlinna Prize for his contributions to probabilistically checkable proofs, non-approximability of optimization problems, and error-correcting codes. The early applications to probabilistically checkable proofs yielded a remarkable new characterization of the class of nondeterministic polynomial-time (NP) problems and sharpened our understanding of the non-approximability of hard optimization problems. The key ingredient in these works was the ingenious use of results from error-correcting codes to complexity theoretic problems. More recently, by developing efficient decoding algorithms for several algebraic and number-theoretic codes, Madhu Sudan has taken ideas in the converse directions: from complexity theory to the theory of error-correcting codes. In this article, we describe his work in these areas.

Proofs and computational complexity

Madhu Sudan's work on probabilistically checkable proofs and non-approximability of optimization problems is closely related to the so-called P = NP problem in complexity theory. We briefly present the background to this problem and then discuss Madhu Sudan's contribution.

One expects to use a computer to answer questions of the form 'Is 17 a perfect square?', 'Is M a symmetric matrix?', 'Is the graph G connected?', 'Is *abbabbabba* a palindrome?', etc. In complexity theory, such *computational problems* are usually modelled as sets; for example, the set of squares of integers, the set of symmetric matrices, the set of connected graphs and the set of palindromes. A *computational solution* to such a problem S is an algorithm A for answering membership questions of the form 'Is x in S?'. Let $t_A(n)$ denote the maximum time (the number of steps) algorithm A takes when x has length n. (From now on n will denote the length of the input.) The set S is *tractable* if there is an algorithm A that answers questions of the form 'Is x in S?' correctly, and whose running time, $t_A(n)$, is bounded by a polynomial in n.

The P = NP question arose from attempts to separate the tractable problems from the intractable, based on the time required to solve them. The *class* P is the collection of all tractable problems. Several natural problems fall in this class: the examples of sets given above, the set of non-singular matrices, the set of two-colourable graphs, and now, thanks to the work of Agrawal, Kayal and Saxena, the set of prime numbers! Various other problems, however, are not known to be tractable in the above sense: for example, graphs with Hamilton cycles, three-colourable graphs, Boolean expressions that hold for some choice of values for the variables. Yet, no formal justification for their apparent intractability is known. A large class of these problems, however, have the following property: they admit proofs of membership that are *verifiable* in polynomial-time. The notion of efficient verification of membership proofs is central to the original definition of the class NP as well as to the recent developments in the area of probabilistically checkable proofs. A set T is in the *class* NP if there is a polynomial-time computable predicate (that is, a function taking values in the set {true, false}) $N_T(x, y)$ so that $T = \{x : \exists y, N_T(x, y)\}$, where the length of y is bounded by a polynomial in n, the length of x. It will be helpful to view this formal definition of NP as a two-step interaction between a Prover and a Verifier.

Prover: Claims that x is in T, and supplies a proof y whose length is bounded by a fixed polynomial in n, the length of x.

Verifier: Computes in polynomial-time, the predicate $N(x, y)$, and accepts the claim if the predicate is true.

If x is indeed in T, then there must exist a y that causes the verifier to accept the claim; if x is not in T, the verifier should reject the claim for *every* y. Thus, one regards the string y as a *short proof* for the statement 'x is in T', and the predicate $N(x, y)$ as a polynomial-time verification of this proof. For example, consider the set S each of whose elements is a system of polynomial equations that have a solution in the set $\{0, 1\}$. So, the input x in this case is the system of equations. The *proof* is just a point $y \in \{0, 1\}^n$, where n is the number of variables. When the prover supplies a y, the verifier checks if y is indeed a solution, by evaluating each polynomial in the system explicitly. This is easily seen to be a polynomial-time computation. Thus, the set S is in NP. Every set in NP has such a proof system. Is it true that every set that has such a proof system is tractable? Intuitively, it appears that in general, finding proofs should be harder than just verifying a given proof. Showing that P ≠ NP will justify this intuition formally from a computational point of view.

Probabilistically checkable proofs: In 1985, Goldwasser, Micali and Rackoff, and independently, Babai considered generalizations of NP, where one allowed interaction between the prover and a *randomized* verifier, and also tolerated some error. Lund, Fortnow, Karloff and Nisan (1990) showed that this seemingly minor modification to the original definition of NP made these *interactive proof systems* surprisingly powerful. Building further on these ideas, Babai, Fortnow and Lund (1990), considered the model of probabilistically checkable proofs (defined earlier by Fortnow, Rompel and Sipser (1988) in a different setting), and showed that in this system, proofs of membership even for sets that have exponential complexity, can be verified in polynomial-time. Later developments built up on this result, so we will describe it in a little more detail.

Let NEXP be the exponential time analogue of NP. That is, NEXP is the class of sets T that can be defined by a polynomial-time-computable two-place predicate $N_T(x, y)$ as before, but the length of y is now allowed to be as much as $2^{p(n)}$ for some fixed polynomial $p(n)$, instead of just $p(n)$. The verifier is allowed to examine the proof for polynomial-time in the length of x and y, but since y is allowed to be long, the verifier's computation could take exponential time. Indeed, it can be proved that the class NEXP is strictly larger than NP, that is, there are sets in NEXP for which proofs of membership cannot be verified in polynomial-time.

With randomness and allowance for some small error, however, one can do much better. Babai, Fortnow and Lund (1991) showed that if the verifier is randomized and allowed to err with some small probability, then even these exponentially long membership proofs can be verified in polynomial-time. That is, every set in NEXP has a probabilistically checkable proof system of the following form:

Prover: Claims that x is in T, and supplies a proof y, whose length is bounded by $2^{p(n)}$.

Verifier: Picks a random binary string R whose length is bounded by polynomial in n, the length of x. Based on R she reads a polynomial number (bounded by $q(n)$, for some polynomial q) of bits from y. Then, using R, x and the bits read, y', she computes in polynomial-time a predicate $N_T(R, x, y')$, and accepts the claim if the predicate is true.

If x is in T, then there is a proof y that the verifier will accept with probability (taken over the choices of R) one. That is, every true statement about membership has a perfect proof – the proof system is complete. We compromise on soundness. If x is not in T, then for all y, the probability that the verifier ends up accepting the claim is bounded by 1/100 (this can be made arbitrarily small). Note that the predicate $N_T(x, y)$ in the definition of a set T in NEXP (see above) immediately yields a proof system where the proof y is exponentially long in the length of the candidate element x. The above probabilistic proof system also uses a long proof. The crucial difference is that now the verifier uses a polynomial number of random bits and performs only a polynomial-time (in the length of the input x, *not* y) computation. The class of languages where membership claims can be verified by using $O(r(n))$ random bits and by reading $O(q(n))$ bits of the proof, is denoted by PCP(r, q). For brevity, one writes PCP(poly, poly) for the class $\cup_{c>0}$ PCP(n^c, n^c). With this terminology, the result of Babai, Fortnow and Lund can be stated as follows.

Theorem 1 (*Babai, Fortnow and Lund 1991*): NEXP $\subseteq$ PCP(*poly, poly*).

It is immediate from the definition of NEXP that every problem in NEXP is contained in PCP(1, $2^{p(n)}$), for some polynomial p, and every language in NP is in PCP(1, $p'(n)$), for some polynomial p'. Using this analogy, Babai, Fortnow, Levin and Szegedy (1991) asked whether something similar to Theorem 1 could be proved for the class NP itself. Their work was refined further by Feige, Goldwasser, Lovász, Safra and Szegedy (1991), and Arora and Safra (1992) to give the following.

Theorem 2: NP $\subseteq$ PCP($\log n$, $(\log \log n)^c$), *for some* $c > 0$.

A key ingredient in these works, starting from the work of Babai, Fortnow and Lund, is the use of polynomials of small degree to encode of Boolean functions. This was inspired by the work in the area of self-testing and self-correcting programs. We will give an example of a typical result in this area. Let $\mathbb{F}_2$ be the field with two elements and let $\tilde{A}: \mathbb{F}_2^m \to \mathbb{F}_2$ be such that $\tilde{A}(x+y) = \tilde{A}(x) + \tilde{A}(y)$ for all $x, y \in \mathbb{F}_2^m$. Then, it is easy to see that $\tilde{A}$ is a linear function. Now, suppose we only know that the condition $\tilde{A}(x+y) = \tilde{A}(x) + \tilde{A}(y)$ holds for a fraction $1 - (\delta/2)$ of pairs $x, y \in \mathbb{F}_2^m$. Then, what can we say about $\tilde{A}$? Blum, Luby and Rubinfeld (1990) showed that if $\delta < \frac{1}{3}$, then there is linear function $A: \mathbb{F}_2^m \to \mathbb{F}_2$ that agrees with $\tilde{A}$ on a fraction δ of the places. Furthermore, one can efficiently reconstruct A from $\tilde{A}$. In his Ph D thesis (1992), Madhu Sudan made remarkable use of this theorem to construct a probabilistically checkable proof for every set in NP. Using linear functions over $\mathbb{F}_2$, he constructed membership proofs that could be verified in polynomial-time by reading only a constant number of bits of the proof.

Theorem 3: NP $\subseteq$ PCP(n^3, 1).

The proofs of membership in this construction are exponentially long (the classical proofs used to have poly-

nomial length), but the crucial point is that they can be verified by reading only a constant number of bits. Madhu Sudan's Ph D thesis also addressed questions of testing and correction of low-degree polynomials. By combining this work with that of Arora and Safra, and Theorem 3, Arora, Lund, Motwani, Sudan and Szegedy (1992) showed the following.

Theorem 4 (*the* PCP *theorem*)*:* NP = PCP(log n, 1).

Let us return to the set S, which we used as an example for the class NP. One natural, easily verifiable proof that a system of equations has a solution is the solution itself. Now, the PCP theorem says that there is another short proof, to verify which one needs to just read a constant number of randomly chosen bits from it. If the system of equations has a solution in $\{0, 1\}^n$, then for a suitable proof, the verifier will succeed with probability one. But, if there is no such solution, then with high probability the verification will fail. This result is important as an alternative characterization of the class NP, and a compelling demonstration of the role of randomness in computation. As such, it represents important progress in understanding the class NP. (It is easy to see that the class P is the same as PCP(1, log n); thus, the P = NP question is just the palindrome PCP(1, log n) = PCP(log n, 1)!) It also has deep implications for approximation algorithms for hard optimization problems. It is to this connection that it owes its 'celebrity status'.

PCP and approximation

There are many natural optimization problems for which one does not know efficient algorithms. In fact, for a large class of these problems, it is known that if even one of them can be solved exactly in polynomial-time, then P = NP. When faced with a hard optimization problem, researchers have tried to side-step the difficulty by considering approximation algorithms that produce solutions within a small factor of the optimum. There are problems where this factor can be made arbitrarily close to 1 (the running time of the algorithms increase when we demand it to be closer to 1). However, there are several others where no such solution is known. We describe two examples that illustrate the role of the PCP theorem in understanding approximation problems.

The travelling salesman problem: We are given a set of cities and distances between them, and are required to construct the shortest tour that visits all the cities and returns to the origin. No polynomial-time algorithm is known for this problem. In fact, it is known that if this problem can be solved exactly in polynomial-time, then P = NP. In the absence of efficient methods for obtaining exact solutions, researchers asked if one could obtain approximate solutions for this problem. Christofides (1976) showed that there is a polynomial-time algorithm that produces a tour whose length is at most 1.5 times the optimum.

The clique problem: We are given an undirected graph, and required to determine the maximum-sized clique in it; that is, a set of vertices where each pair is connected by an edge. Again, no polynomial-time algorithm is known for this problem, and one knows that if there exists one such, then P = NP. Unfortunately, here no efficient algorithm that produces a solution within a small factor of the optimum is known: the best polynomial-time algorithm known, due to Boppana and Halldórsson (1990), has an approximation factor (the ratio of the optimum solution and the solution produced by the algorithm) of $\Omega(n/(\log n)^2)$, where n is the number of vertices in the graph.

The decision versions of these two problems are 'equally hard': the question 'Does a graph have a clique of size t' can be transformed in polynomial-time to a question of the form 'Does a set of cities have a travelling salesman tour of length at most d' and vice-versa. That is, if one of them has a polynomial-time solution, then so does the other. Yet, their optimization versions appear to be very different.

The PCP theorem implies that there is a constant $\varepsilon > 0$ such that if there is a polynomial-time algorithm to produce a tour whose length is guaranteed to be at most $1 + \varepsilon$ times the optimum, then P = NP; that is, the factor 3/2 of Christofides' algorithm cannot be brought arbitrarily close to 1 if P $\neq$ NP. This is not an isolated problem where the PCP theorem imposes limits to approximability.

Papadimitriou and Yannakakis (1991) initiated the classification of approximation problems based on their approximability, and defined approximation-preserving reductions and a class (called MAXSNP) of problems. This class includes a large number of problems for which one can construct approximate solutions in polynomial-time that come within some constant factor of the optimum, but one does not know if they can be approximated for factors arbitrarily close to 1. The travelling salesman problem considered above is in MAXSNP. Papadimitriou and Yannakakis showed that if certain problems in this class (referred to as MAXSNP-complete problems) can be approximated in polynomial-time to within a factor arbitrarily close to 1, then all problems in MAXSNP can be approximated to within a factor arbitrarily close to 1.

Theorem 5 (Arora, Lund, Motwani, Sudan and Szegedy (1991): Unless P = NP*, for every* MAXSNP*-complete problem, there is an* $\varepsilon > 0$*, such that no polynomial-time algorithm can produce a solution with an error of at most* ε *times the optimum.*

For the clique problem described above, the implications of the PCP theorem are even more spectacular. It implies

that unless P = NP, there is a $\delta > 0$, such that no polynomial-time algorithm can produce a solution that is guaranteed to be at least $n^{-\delta}$ times the optimum, where n is the number of vertices in the graph.

Over the years, the PCP theorem has been sharpened by several researchers, including Madhu Sudan. As a result, several technical improvements have been made, resulting in stronger non-approximability results, sometimes even matching the performance of the best approximation algorithms known.

Error-correcting codes

Recently, Madhu Sudan has contributed new ideas and developed decoding algorithms for error-correcting codes. We now briefly touch upon this facet of his work.

A t-error-correcting code over the alphabet $\mathbb{F}_q$, the field of q elements, is a set C such that for any vector $\boldsymbol{r} \in \mathbb{F}_q^n$, there is at most one vector $\boldsymbol{c} \in C$ within Hamming distance (the distance between two vectors is the number of places where they differ) t of $\boldsymbol{r}$. Such a code is useful in transmitting information in the presence of errors. If vectors from C are used as messages, then one can tolerate up to t errors. That is, if at most t symbols in the message sent are changed, the intended message can be recovered uniquely from the received message.

Traditional error-correction algorithms were concerned with recovering the sent message uniquely from the message received, assuming that the number of errors is at most t. What if there were actually $t' > t$ errors? Now, in principle, there might be several vectors in C that are within distance t' of a single vector in $\mathbb{F}_q^n$. However, if t' is not much larger than t, the number of such code-words is small: it can, for example, be shown that if $t' < n - \sqrt{n(n-2t-1)}$, then the number of such code-words is bounded by a polynomial. One can ask if there is a polynomial-time algorithm for producing the list of all these code-words. This problem is known as the list-decoding problem. Although, combinatorial facts about the number of code-words above the error-correction threshold t were known, and the list-decoding problem has been studied since the 1950s, there was no efficient algorithm that admitted significantly more errors than the error-correcting threshold t.

Madhu Sudan (1997) gave the first such algorithm for the class of Reed–Solomon codes. This algorithm was built on the work of Ar, Lipton, Rubinfeld and Sudan (1992), and made elegant use of linear algebra and unique factorization in the ring $\mathbb{F}_q[X, Y]$. Later this algorithm was improved by Guruswami and Sudan (1999), who showed how to produce the list of code-words provided $t' < n - \sqrt{n(n-2t-1)}$. They also showed similar results for algebraic-geometric codes. Since then, list-decoding algorithms have been discovered for several other codes; these algorithms were placed in a fairly general setting of Ideal Error Correcting Codes by Sudan (2001).

Biographical sketches of the awardees

Laurent Lafforgue was born on 6 November 1966 in France. He studied at the Ecole Normale Supérieure. He entered the CNRS in 1990, at the Laboratoire de mathematiques d'Orsay and completed his thesis in 1994. From 2000, he is a permanent professor at the Institut des Hautes Études Scientifiques, Bures sur Yvette, France. He is also an Adjunct Professor in the School of Mathematics, Tata Institute of Fundamental Research, Mumbai.

Vladimir Voevodsky was born on 4 June 1966 in Russia. He did his B Sc in Mathematics from Moscow State University in 1989, and obtained his Ph D in mathematics from Harvard University 1992. He held visiting positions at the Institute for Advanced Study, Harvard University and the Max Planck Institute for Mathematics before joining the faculty of Northwestern University in 1996. From 2002, he is a permanent member of the faculty of the Institute for Advanced Study, Princeton, USA.

Madhu Sudan was born on 12 September 1966 in Chennai, India. He did B Tech in Computer Science at the Indian Institute of Technology, Delhi in 1987, and obtained his Ph D in Computer Science from the University of California, Berkeley in 1992. He was on the staff of the IBM Thomas J. Watson Research Center during 1992–97, and since 1997 has been Associate Professor in the Department of Electrical Engineering, Massachusetts Institute of Technology, Cambridge, USA. He is also an Adjunct Professor in the School of Technology and Computer Science, Tata Institute of Fundamental Research, Mumbai.

ACKNOWLEDGEMENT. N.N. thanks Kapil Paranjape, Najmuddin Fakhruddin and V. Srinivas for helpful comments.

Received 9 October 2002; accepted 11 October 2002