Mathematical Problems of Computer Science 38, 38–41, 2012.

Advanced Combinatorial Optimization

L. Aslanyan

Institute for Informatics and Automation Problems
of the National Academy of Sciences
Laboratory of Discrete modelling, analysis and recognition
hasmik@ipia.sci.am

The research aims at constructing efficient algorithms and approximations for hard, open and novel combinatorial optimization problems with emphasis on the practical applicability of results. Boolean minimization (hard), Isoperimetry and tomography (open), decentralized and stream analysis (novel) algorithms are considered. Combinatorial reduction is another target, being related with two sets of problem instances, together with a structure of reduction. This triplet is studied to recover the tractability and approximation boundaries. The concept of optimality that is not yet clear for complex systems with conflicting goals is addressed in terms of heuristics, stability and equilibria. These issues play a key-role to the emergence of a major research pole in the area of combinatorial optimization, highly responsive to the industry needs and tightly linked to different research areas with design, analysis and management of complex systems. Applied areas addressed include Population Genomics and Wireless Sensor Networks.

Research problem

For several years scientific research went through an important transformation process and research perspectives started supporting interests that were not only and exclusively the purely cognitive aspects, but also - the needs of society. Complexity Theory and Combinatorial Optimization manifest themselves as a new kind of relation between theoretical and applied science, and these disciplines meet needs of society through an interdisciplinary approach. Such an approach has to model hard and complex systems (**complex system** is composed of interconnected parts/agents that as a whole exhibit properties not obvious from the properties of the individual parts), and to face hard mathematical problems arising from this modeling. The organization of complex systems together with their possible applications in a variety of industrial sectors represents a great challenge to these disciplines.
Mathematical models for almost any real-world problem arising in various areas such as genomics, information technology (software and hardware including), communication networks, and (radioactive and other) waste management, give rise to combinatorial problems that can be characterized as **hard ones**, in the sense that any optimal solution process for large scale instances of them takes unacceptably long computation time. Most combinatorial optimization problems of great practical relevance are, indeed, computationally intractable in the above sense. In formal terms, they are classified as NP-hard optimization problems [1]. Despite many years of research it has not been formally proved that NP-hard problems

cannot be solved by means of algorithms running in polynomial time [1]. Nevertheless, there is very strong theoretical and practical evidence that such algorithms may not exist and we have to make use of approximation algorithms.

P/NP theory uses the problem reduction technique. An individual problem – SAT, TSP, 3DM, etc. include the whole set of individual instances [2]. Reduction is required to be polynomial and its degree is not a critical factor. This framework, valid in theory is to be further investigated from a practical point of view. We intend to construct the minimal complexity reductions, sub-problems of individual problems will be considered, and for-approximation reductions will be composed. This approach will transfer the solvable instances from one problem to the other, and transitively will compose the larger tractable subsets of problem instances. Approximation reductions will form the knowledge on approximability boundaries. Therefore, the practical counterpart of problem reductions theory will be set up and applied to complex systems tasks.

Indeed, a new conceptual framework is needed for dealing with optimality issues in complex systems. First of all, one has to take into account the presence of conflicting or joint interests of the agents involved in these systems, as well as the absence of any authority able to centralize information and to impose **optimal** (in fact, any kind of) decisions concerning the agents' behavior. Such aspects do induce major differences between classic optimization and the agent network context.

Besides the theory we addressed two basic applications with complex systems inside – one in Population Genomics (PG) and the second in Wireless Sensor Networks (WSN).

a) Wireless Sensor Networks (WSN) serves many potential application areas: from environment monitoring to health, and from defense models to social networks. WSN are made of large number of almost identical small sensor nodes that are integrated on base of distributed protocols and algorithms of data communication and data analysis.

WSN has no predefined infrastructure, i.e., nodes do not have any a priori knowledge about the rest of the network. Hence a WSN must be able to self-organize. Nodes cooperate for this self-organization and it is very basic to this proposal that this cooperation is in its nature local – neighborhood based. Current WSN models deal with this local analysis in a hidden mode, which doesn't help to achieve the best performance.

b) Population genetics is the study of allele frequency distribution and change under the influence of the four main evolutionary processes: natural selection, genetic drift, mutation and gene flow. Here the pure parsimony problem is to infer a maximally parsimonious collection of genetic donations that can combine to form a new population's diversity over portions of the chromosome [8]. We partially order a collection of genotypes so that we can represent the problem of inferring the least number of haplotypes in terms of substructures known as g-lattices. This representation allows proving that if the genotypes partition into chains with certain structure, and then the NP-Hard problem can be solved efficiently. Even without the specified structure, the decomposition shows how to separate the underlying integer programming model into smaller models.

Implementation

Analysis of problem reductions /hardness and approximation/.

A 'gadget' is a finite combinatorial optimisation problem, which translates constraints of one optimization problem into a set of constraints of a second optimization problem. Problem

---

[1] here is the link to the latest attempt in solving this problem http://michaelnielsen.org/polymath1/index.php?title =Deolalikar_P_vs_NP_paper)

hardness may mean the existence of a hard CPI (classes of problem instances) only. Is it large or not? May we determine this area? Additionally, in different problems there are known effectively solvable CPI. In IP such known area is the instances based on unimodular matrices. Therefor it is novel the structure, which translates constraints of one optimization problem into a set of constraints of a second optimization problem [7]. Polynomial reductions used in NP area are based on plenty of gadgets. Practically, computationally simpler gadgets provide us with more efficient reduction of problems or specific CPI. Two notions are specifically important. The notion of promising to find out different polynomial CPI in hard optimisation tasks, - use them in problem reductions, and enlarge in this way the areas of effectively solvable instances. This study required constructing new gadgets and special gadgets will be considered for reduction of approximate solutions.
Optimality criteria in complex systems.
The specific goal of this task is in deep analysis of local algorithms as the basis of all WSN tasks - to understand how well they serve the algorithmic needs of WSN: such as coverage, connectivity, security, etc. Start point is the basic research [6], which enabled discovering problems irresolvable by means of local algorithms. Then the NP complete tasks by local algorithms will be investigated. This series will show the way not to go in WSN design. Although hardness of NP complete problems is well recognized, at this point less attention is paid to NP hardness and irresolvable problems because of limitations to the classes of local algorithms. This is due to mix of use and complexity estimates of ordinary optimization tasks with ones obliged to use the local algorithms. In fact, the task will develop local criteria in terms of current state and behavior, and will model the global optimization criteria in these terms. WSN simulation will use the free, open source software project ns-3 http://www.nsnam.org, known as a discrete-event network simulator for Internet systems.
Applied level.
Population genomics example problem will be considered in this task. In short, the mathematical content is the following. Let a set of intervals in n-dimensional unit cube is given. It is necessary to construct a set of vertices so that in each interval there exist at least one pair of complementary (a vertex together with its negation inside the interval) vertices among the selected ones, and the selection size is to be minimal. This formulation is very close to the pruning problem, which is known as NP-hard. Due to huge sizes of genomic information it is necessary to find approximate algorithmic constructions. The technique to be used includes chain split, Boolean minimisation, and reduction to IP instances. This task considers a particular optimisation problem in approximation, considers reductions to find out more approximate solutions, and finds the same local concepts that may replace the concept of a global optimum. Testing and validation of the algorithm and software will be performed using publicity available data from Gene Expression Omnibus repository (http://www.ncbi.nlm.nih.gov/geo/).

Future work

Our goal was to advance in the research for new optimality paradigms, getting insight from the complex systems sciences and (to some extend) from socioeconomic sciences: economics, biology of populations, control theory and others. We strongly believe that in the near future, these thematic will be reshaping the research landscape in combinatorial optimization. They are equally expected to influence all the active research for new computing machine paradigms based upon properties of natural systems that are not exploited by conventional computers. We can, thus, be seen as an initiative to drastically renovate the research agenda in combinatorial optimization, by addressing open and novel problems arising from complex systems. In particular, our further research will enrich the tractability and approximation areas of combinatorial optimization by the basic set of paradigmatic problems SAT, IP, TSP, 3DM, etc. The concept of optimization for complex systems will be modeled in terms of local criteria stability and equilibria. The set of practical complex systems will be designed and solved algorithmically.

References

1. Garey, M. R. and Johnson, D. S. (1979). Computers and Intractibility: A Guide to the Theory of NP-Completeness. W. H. Freeman and Company, New York.

2. Crescenzi P., Kann V., Approximation on the web: A compendium of NP optimization problems, Randomization and Approximation Techniques in Computer Science, Lecture Notes in Computer Science, 1997, Volume 1269/1997, 111-118.

3. Papadimitriou C.H., Steiglitz K., Combinatorial optimisation: algorithms and complexity, Dover Publications Inc., 1998. New York, 498p.

4. Korte B., Jens V., Combinatorial optimization: Theory and algorithms (21 Algorithms and Combinatorics), 4th Ed., Springer, 2008, 627p.

5. Trevisan L., Sorkin G.B., Sudan M. and Williamson D.P., Gadgets, Approximation, and Linear Programming, SIAM Journal on Computing, Volume 29, Issue 6, 2074-2097(2000).

6. Zhuravlev Yu., Elected Research Works, Magister, Moscow, 1998, 420p.

7. Greenberg H., Hart W.E., Lancia G., Opportunities for combinatorial optimization in computational biology, INFORMS Journal on Computing 16, 221-231 (2004).

8. Arakelyan A., Boyajian A., Aslanyan L., Sahakyan H., Ivanova K., Mitov I., Growing support set systems in analysis of high-throughput gene expression data, Classification, forecasting, Data Mining conference, Varna, 22-24 June, 2010.