A propos du centre ou de la direction fonctionnelle

Contexte et atouts du poste

The goal of this PhD is to foster the next generation of large-scale parallel optimizers, by contributing to the design of advanced algorithms accounting for optimization challenges of parallel computing on the one hand, optimization is ubiquitous to countless modern engineering and scientific applications and problems and algorithms increasingly large-scale and heterogeneous in particular to deal with a huge amount of variables. The large-size and heterogeneous characteristics of nowadays complex applications lead to “big” optimization problems [1] raising new important and difficult scientific challenges for researchers and practitioners, that traditional approaches will hardly succeed when facing them. On the other hand, the advent of massively parallel and/or distributed computers including up to millions of processing cores is a challenging opportunity. In fact, harnessing such amount of computing resources raises several performance issues related to parallel scalability and heterogeneity. All these scalability issues constitute numerous difficult challenges to the optimization community and give rise to new great research directions. In this respect, the ambition of this PhD is to contribute to design and analysis of efficient optimization algorithms, in a large-scale setting, and to foster the next generation of modern optimization approaches in order to be able to scale both with the massive variable space and the massive parallel computing dimension. Motivation. More specifically, the project targets the design and analysis of novel gray-box optimization algorithms [2] both at the theoretical and applied levels. In fact, one can classify optimization problems and/or methodologies in several classes according the amount of knowledge that is available a priori before running the optimization process. At one extreme, in a black-box optimization scenario, the objective function to be minimized/maximized can be unknown, which is the case of optimization problem in multi-disciplinary engineering design where the cost of one solution is provided by some simulation process. In such a case, even a mathematical formulation of the problem might not be available. Evolutionary optimization techniques are method of choice as they do not require any information about the problem being solved and are generic enough to be applied to a wide range of problems. On the other extreme, in a white-box optimization scenario, a full information about the problem is available, e.g., in a continuous domain, the derivatives and the gradient are available. It should then be clear that the information available about the optimization problem have a substantial impact on the difficulty of its solving and the design of effective algorithms, the more information is available, the more the solving should be easier. In between these two extreme cases, one can seek for some information about the problem so that the (evolutionary) optimization process can be improved and accelerated [3]. This is precisely what we term gray-box optimization. For instance, knowing the interaction between variables can highly help the optimization process by guiding the evolution of solutions. This can be handled, for instance, by integrating techniques from graph theory to model the search regions of special interest. This can be performed either off-line, that is before the optimization process is started, or on-line typically using a machine-learning mechanism on the basis of the solutions explored so-far. Besides, knowing the components, in the decision space, that has the most important impact on solution quality, is another information that can be used for designing highly effective recombinant procedures and non-oblivious evolutionary search engines. Recently, such a methodology has shown its efficiency in tackling large-scale problems with hundreds of thousands of variables [4].

Mission confiée

General statement. The research topic of this project falls at the crossroads of the boarder field of optimization and parallel computing. On the one hand, optimization is ubiquitous to countless modern engineering and scientific applications and problems and algorithms increasingly large-scale and heterogeneous in particular to deal with a huge amount of variables. The large-size and heterogeneous characteristics of nowadays complex applications lead to “big” optimization problems [1] raising new important and difficult scientific challenges for researchers and practitioners, that traditional approaches will hardly succeed when facing them. On the other hand, the advent of massively parallel and/or distributed computers including up to millions of processing cores is a challenging opportunity. In fact, harnessing such amount of computing resources raises several performance issues related to parallel scalability and heterogeneity. All these scalability issues constitute numerous difficult challenges to the optimization community and give rise to new great research directions. In this respect, the ambition of this PhD is to contribute to design and analysis of efficient optimization algorithms, in a large-scale setting, and to foster the next generation of modern optimization approaches in order to be able to scale both with the massive variable space and the massive parallel computing dimension. Motivation. More specifically, the project targets the design and analysis of novel gray-box optimization algorithms [2] both at the theoretical and applied levels. In fact, one can classify optimization problems and/or methodologies in several classes according the amount of knowledge that is available a priori before running the optimization process. At one extreme, in a black-box optimization scenario, the objective function to be minimized/maximized can be unknown, which is the case of optimization problem in multi-disciplinary engineering design where the cost of one solution is provided by some simulation process. In such a case, even a mathematical formulation of the problem might not be available. Evolutionary optimization techniques are method of choice as they do not require any information about the problem being solved and are generic enough to be applied to a wide range of problems. On the other extreme, in a white-box optimization scenario, a full information about the problem is available, e.g., in a continuous domain, the derivatives and the gradient are available. It should then be clear that the information available about the optimization problem have a substantial impact on the difficulty of its solving and the design of effective algorithms, the more information is available, the more the solving should be easier. In between these two extreme cases, one can seek for some information about the problem so that the (evolutionary) optimization process can be improved and accelerated [3]. This is precisely what we term gray-box optimization. For instance, knowing the interaction between variables can highly help the optimization process by guiding the evolution of solutions. This can be handled, for instance, by integrating techniques from graph theory to model the search regions of special interest. This can be performed either off-line, that is before the optimization process is started, or on-line typically using a machine-learning mechanism on the basis of the solutions explored so-far. Besides, knowing the components, in the decision space, that has the most important impact on solution quality, is another information that can be used for designing highly effective recombinant procedures and non-oblivious evolutionary search engines. Recently, such a methodology has shown its efficiency in tackling large-scale problems with hundreds of thousands of variables [4].

Research Objectives. The purpose of this project is to push further the limits of existing solvers by tackling even larger problems, of different nature, and more importantly with variable degree of difficulty. Unique of the proposed project is to renovate the existing approaches to fit the massively parallel and heterogeneous nature of modern computing platforms. In fact, we argue that most of the knowledge used by gray-box optimization techniques is used in order to infer some decomposition of the decision space [3]. This decomposition approach has a decentralized nature which makes it very appealing for parallel computing. On the one hand, designing decomposition, at the very-first step of the development of a gray-box optimization algorithm, should be thought in parallel, not only to accommodate the fact that almost all optimization effort at the deployment step is done in parallel, not only to accommodate the fact that almost all optimization effort at the deployment step is done in parallel, not only to accommodate the fact that almost all optimization effort at the deployment step is done in parallel, not only to accommodate the fact that almost all optimization effort at the deployment step is done in parallel, not only to accommodate the fact that almost all optimization effort at the deployment step is done in parallel, but also to open new algorithmic and computational perspective that can be proved efficient in a parallel computing setting. On the other hand, using parallelism can allow to tackle increasingly complex problems and to design more powerful approaches [6]. To the best of our knowledge, investigating the synergies between gray-box optimization and parallel computing has not been so far explored by the community. As such, the conducted research will have four main orthogonal objectives: (i) design new evolutionary gray-box and parallel optimization approaches, (ii) improve the understanding of evolutionary gray-box optimization techniques and their theoretical foundations, (iii) analyze their effectiveness as a function of the benchmark problem at hand, (iii) provide evidence on the efficiency of the designed approaches when tackling large-scale problems using large-scale computing resources.

Informations générales

- Thème/Domaine : Optimisation, apprentissage et méthodes statistiques
- Ville : Villeneuve d’Ascq
- Centre Inria : CRI Lille – Nord Europe
- Date de prise de fonction souhaitée : 2019-10-01
- Durée de contrat : 3 ans
- Date limite pour postuler : 2019-04-22

Contacts

- Equipe Inria : BONUS
- Directeur de thèse : Derbel Bilel / bilel.derbel@inria.fr

A propos d’Inria

Inria, l’institut national de recherche dédié aux sciences du numérique, promeut l’excellence scientifique et le transfert pour avoir le plus grand impact. Il emploie 2400 personnes. Ses 200 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3000 scientifiques pour relever les défis des sciences informatiques et mathématiques, souvent à l’interface d’autres disciplines. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 160 start-up. L’institut s’efforce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l’économie.

L’essentiel pour réussir

Consignes pour postuler

CV + application letter + recommendation letters + references + school transcripts

Sécurité défense : Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement : Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.

Attention : Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.
Principales activités

Given the difficult challenges raised by the project, a possible research approach that we propose to investigate is as follows:

1. Analysis of the decentralized nature of existing gray-box optimization techniques
   - Elicit the decomposition components used in the design of state-of-the-art algorithms
   - Classify the decomposition components with respect to their dependency to the problem being tackled. We will consider three classes of combinatorial optimization problems: pseudo-boolean functions, permutation-based problems, graph-based problems.

2. Refine the already existing gray-box evolutionary techniques by incorporating techniques from other connected fields such as graph theory and machine learning
   - Learning the gray-box structure of black-box optimization problems and develop novel scalable parallel decomposition techniques
   - Adapt the existing gray-box evolutionary operators to operate with the solearned structure
   - Accommodate the core design of existing techniques to fit the massively parallel nature of a modern CPUs/GPUs compute facilities
   - Implement and deploy the so-developed heterogenous algorithms and study their parallel scalability.

3. Massively parallel gray-box decomposition
   - Develop fitness landscape analysis tools to better understand the landscape implied by search operators based on gray-box techniques
   - Design variable dependency structures and study their impact on search performance
   - Develop benchmark problems to both assess the performance of existing gray-box techniques, and to better understand their weaknesses and strengths
   - Adopt a systematic approach for the design of autonomous gray-box search solvers, i.e., selecting the most appropriate variants depending on instance features

5. Dissemination
   - Publications in international conferences and/or journal(s)
   - The developed programs should be made available for the expert community.

Compétences

Candidates with the following skills will be preferred:

- Fluent English, excellent communication skills, keen to team working
- Good background in (combinatorial) optimization and/or evolutionary algorithms
- Good background in parallel and distributed computing
- Good background in graph theory and machine learning
- Good experience in programming

Avantages

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours)
- Access to vocational training
- Social security coverage

Rémunération

1st and 2nd year: 1593.50€ Net monthly salary (after taxes)

3rd year: 1676.31€ Net monthly salary (after taxes)