Automatic Synthesis of Mathematical Programming Models for Business Processes

This research project is aimed at minimization of human-related effort and cost of modeling business processes by automating this task. A *business process* is a series of steps needed to reach a goal, e.g., manufacture a product or deliver goods. A model of the process is a formal specification that includes process variables, relationships between them and process outcome. A commonly employed representation is *Mathematical Programming* (MP) model that consists of a set of constraints reflecting relationships between the process variables and an objective function representing the process outcome. The types of constraints and objective function are not restricted in general, however for certain types like e.g., linear, efficient polynomial-time optimization algorithms exist, which makes use of such types of MP models reasonable.

A common practice is to build MP models manually. This is laborious task, since some of the process constraints may be difficult to identify or not expressible under adopted type of MP model (e.g., linear), requiring thus advanced modeling techniques. Moreover, human experts tend to build MP models simplified with respect to the modeled reality by skipping seemingly irrelevant relationships. This, however, may be deceptive since optimization of an oversimplified MP model may lead to solutions infeasible in practice. An MP model that includes infeasible solutions is considered incorrect and needs to be revised, possibly many times. Given a correct MP model, its use in optimization or simulation is fully automated by specialized software packages called *solvers*.

We *automate synthesis of MP models for real-world processes* using examples of process states that are relatively inexpensive to acquire by monitoring execution of the process. We decompose MP model synthesis problem into constraints synthesis and objective function synthesis problems, and tackle each one separately. We develop algorithms based on methods from *Machine Learning*, *Computational Intelligence* and *Operational Research*. Particularly, we employ *one-class classifiers* (e.g., Support Vector Data Description, POSC4.5), *Genetic Programming*, *local search heuristics* (e.g., Simulated Annealing and Tabu Search) and *Linear Programming* to handle constraint synthesis problem. For objective function synthesis we use *linear regression* using least-squares method, *non-linear regression* using Gauss-Newton algorithm, and *symbolic regression* using Genetic Programming. We assess properties of the developed algorithms and MP models that they produce theoretically and empirically in computational experiments. In the experiments, we employ synthetic benchmarks of controlled complexity that enable us to assess fidelity of the synthesized MP models to the actual (benchmark) ones by means of syntactic measures (e.g., *conformance of used terms*) and semantic measures (e.g., *Jaccard index*). We also employ data sets acquired from real-world processes, e.g., production of energy at a power plant and production of concrete mixture at a concrete plant to verify applicability of the algorithms in real-world scenarios. Particularly, we use the synthesized MP models in *process optimization* and *simulation* under changing working conditions.

We identified the problem of automating synthesis of MP models by analyzing research demands coming from business environment in discussions with business representatives. We found limited literature on this subject, and the found methods tackle only synthesis of MP models of Constraint Programming type, leaving no methods for Linear Programming and/or Non-Linear Programming types. Fortunately, the methods and tools that we developed or used in our previous research, particularly algorithms for Genetic Programming, classifiers and Linear Programming solvers may be adapted to the synthesis of MP model of these types. This allows us for smooth start in this new thread of research. We expect that the developed algorithms contribute to better availability of process modeling and optimization services, reduction of human-related effort and costs of these services. Automatically built MP models are to be used for explaining rules of conduct of processes (or phenomena), their simulation and optimization, improving so their efficiency. This should directly contribute to increase of income and competitiveness of organizations employing these processes and to increase of gross domestic product and wealth of society.