Complex Knowledge Resource Management in Model-Based Professional Reasoning

Sergey V. Kovalchuk, Anna V. Kalyuzhnaya, Klavdiya O. Bochenina.

ITMO University, Birzhevaya line, 4, Saint Petersburg, Russia, kovalchuk@itmo.ru.

Within the study we consider a problem of accessing, integration and usage of knowledge sources within a context of professional AI expert systems development. Such system should be able to reason within the defined problem domain to assess problem solving and consult the user within certain problem. By a knowledge resource we consider everything helping the system to reason and explain the results of reasoning (forward and backward reasoning). This includes a) structured, semi-structured, and unstructured data; b) formalized and non-formalized knowledge sources; c) models of various kind for model-based reasoning; d) functional reasoning units to support forward and backward reasoning. The proposed idea lies in between several areas: problem solving environment, eScience technologies, knowledge-based reasoning, data mining, and a bunch of directions within AI/machine learning (ML) (commonsense reasoning, meta-learning, explainable AI, etc.). All the available resources should be integrated and connected to a) identify and formalize the request; b) select an integrative structure of the available resources which are needed for reasoning; c) call and reason within each of the required resources; d) collect and reason within the integrative structure; e) infer the answer; f) explain elements of the answer by backward reasoning and connection of the backward reasoning path to domain-specific concepts.

Each of the mentioned components has own specific features and ways of integration:

- The data organized within data lakes includes observations and measurements as well as modelling and simulation results. This include both directly available and collectable (e.g., through web crawling) data. Data integration could be done through semantic interpretation and mapping, quality assessment, and domain-specific reasoning on the facts obtained from data (see, e.g. [1]).
- Formalized and non-formalized knowledge may be provided in various forms: ontologies, rules, functional dependencies, even textual description. As for data lake, knowledge lake may be considered as a weakly structured static, dynamic, or potentially available (e.g. acquirable trough machine learning) (see, e.g. [2]).
- Models including numerical, simulation, data-driven, etc. could be connected within composite models, which may be transferred into workflow structures ready for run with the available computational infrastructure [3].
- Reasoning units are ML/AI elements to support management of all available knowledge resources during forward and backward reasoning.

Our research is focused on building evolving structure of knowledge resources, organized in an automatic way to support the reasoning process for solving a problem. First, a general set of all available resources within the problem domain could be considered as a graphical structure providing dependencies between the resources within semantics, control and data dependencies, reasoning and causal dependencies. Each resource has own specific calling procedures with certain exploratory characteristics (precision, level of trust, request time, etc.). The graphical structure is used as a basic structure for uncertainty quantification and management; inference of particular application solution; scheduling and calling of resources. An important issue could be resolved by generalized resource structure is automatic or semi-

automatic explanation of the results using methods of explainable AI (XAI) integrated with backward analysis of uncertainty and semantic reasoning.

Second, according to the current formal problem definition a sub-graph could be identified within the basic structure. Next, through a series of resource calling, interpretation, and reasoning an answer is obtained which may be a) provided as a solution for the problem, and b) used for evolution of both elements and structure of the solution.

Within the presented study we are focused on the evolution of both basic structure of knowledge resources and inferred solutions (in a form of sub-graphs) evolving over time. Within the evolution both complex model [4] and integrative structure is evolving. The evolved structure is considered as an ensemble of possible solutions applied within a context of particular domain problem. A series of unified approaches and algorithms for automatic and semi-automatic building, identification, management, and application of resource structure is proposed using meta-learning approaches.

We consider a series of sample problems from various problem domains including modelling and simulation in healthcare [5], hydrometeorology [6], finances[7], and social networks analysis [8] for implementation of our approaches and increase automation, interpretability and quality of existing and developing solutions through the proposed approaches for knowledge resource management.

References

[1] P. A. Smirnov and S. V. Kovalchuk, “Linked-Data Integration for Workflow-Based Computational Experiments,” *Commun. Comput. Inf. Sci.*, vol. 468, pp. 175–183, 2014.

[2] P. A. Smirnov, S. V. Kovalchuk, and A. V. Boukhanovsky, “Knowledge-Based Support for Complex Systems Exploration in Distributed Problem Solving Environments,” *Commun. Comput. Inf. Sci.*, vol. 394, pp. 147–161, 2013.

[3] K. V. Knyazkov, S. V. Kovalchuk, T. N. Tchurov, S. V. Maryin, and A. V. Boukhanovsky, “CLAVIRE: e-Science infrastructure for data-driven computing,” *J. Comput. Sci.*, vol. 3, no. 6, pp. 504–510, Nov. 2012.

[4] S. V. Kovalchuk *et al.*, “A Conceptual Approach to Complex Model Management with Generalized Modelling Patterns and Evolutionary Identification,” *Complexity*, vol. 2018, pp. 1–15, Nov. 2018.

[5] S. V. Kovalchuk, A. A. Funkner, O. G. Metsker, and A. N. Yakovlev, “Simulation of patient flow in multiple healthcare units using process and data mining techniques for model identification,” *J. Biomed. Inform.*, vol. 82, pp. 128–142, Jun. 2018.

[6] S. V Kovalchuk, A. V Krikunov, K. V Knyazkov, and A. V Boukhanovsky, “Classification issues within ensemble-based simulation: application to surge floods forecasting,” *Stoch. Environ. Res. Risk Assess.*, vol. 31, no. 5, pp. 1183–1197, Jul. 2017.

[7] D. Vaganov, A. Funkner, S. Kovalchuk, V. Guleva, and K. Bochenina, “Forecasting Purchase Categories with Transition Graphs Using Financial and Social Data,” 2018, pp. 439–454.

[8] O. Severiukhina, K. Bochenina, S. Kesarev, and A. Boukhanovsky, “Parallel Data-Driven Modeling of Information Spread in Social Networks,” *Lect. Notes Comput. Sci.*, vol. 10860, pp. 247–259, 2018.