2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C)

Development of Intelligent Virtual Assistant for Software Testing Team

Iosif Itkin
Exactpro Systems
London, EC4N 7AE, UK
iosif.itkin@exactprosystems.com

Andrey Novikov
Higher School of Economics
Moscow, 101000, Russia
agnovikoff@gmail.com

Rostislav Yavorskiy
Surgut State University and
Exactpro Systems
Surgut, 628403, Russia
rostislav.yavorski@exactprosystems.com

***Abstract*—This is a vision paper on incorporating of embodied virtual agents into everyday operations of software testing team. An important property of intelligent virtual agents is their capability to acquire information from their environment as well as from available data bases and information services. Research challenges and issues tied up with development of intelligent virtual assistant for software testing team are discussed.**

***Index Terms*—software testing, virtual assistant, artificial intelligence**

I. Introduction

Recent advances in microelectronics and artificial intelligence are turning smart machines and software applications into first class employees. Quite soon work teams will routinely unite humans, robots and artificial agents to work collaboratively on complex tasks, which require multiplex communication, effective coordination of efforts, and mutual trust built on long term history of relationships. Virtual assistants are actively applied in health care [1], education [2], [3], [4], elder care [5], laboratory automation [6], team management [7], etc. Following this trend we discuss here research challenges and issues tied up with development of virtual assistant for a software testing team.

II. Data Driven Service Workflow

A. General Data Analysis Process

Different data analysis services have similar structure, which is presented on fig. 1, see [8].

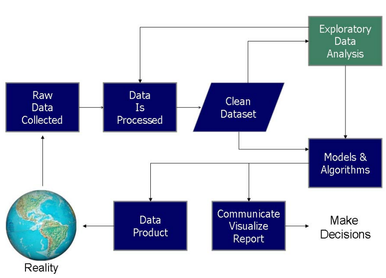

Fig. 1. Data science process [8]

Data driven software testing is not an exclusion. We start with a real world application to be tested, collect raw data, clean it up, process and explore, apply models and algorithms, and then communicate reports to the decision maker. As the systems to be tested scale up, so do tools and algorithms for the data analysis, which collect tons of information about the system behavior in different environments, contexts and configurations. Then, all the collected knowledge is simplified to be presented in a visual form of a short table or a standard graphical chart. Interface of the analytic services with a human is becoming a bottleneck, which limits the capacity of the whole process, see fig. 2. That bottleneck is another reason behind the demand for new types of interfaces.

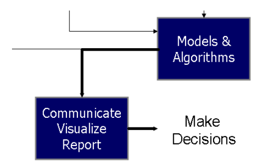

Fig. 2. Fragment of scheme 1: transferring results from machine to a human is the communication bottleneck, which limits the capacity of the whole service

Embodied conversational agents seem to be quite promising in resolving that kind of limitation due to their ability to acquire information from their environment as well as from available data bases and information services, see [9].

B. Testing Distributed Financial Applications

Big data analysis in software testing has its specifics, especially when testing distributed financial applications, see scheme on Fig. 3.

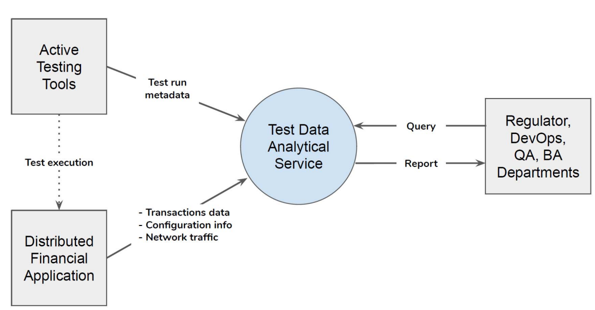

Fig. 3. Software Testing Data Analysis Service

All difficulties typical for testing of high load distributed big scale systems are amplified by numerous regulations from

978-1-7281-3925-8/19/$31.00 ©2019 IEEE
DOI 10.1109/QRS-C.2019.00036

local and international financial authorities. This is discussed in details in [10], [11].

III. Virtual Assistants in Testing Team

A. Motivating Example

Our study was inspired by papers [12] and [13], which report on experience of using robot teams in inspecting littoral environments for military, environmental, and disaster-response applications. Field report [13] describes two deployments of heterogeneous unmanned marine vehicle teams at the 2011 Great Eastern Japan Earthquake response and recovery by the Center for Robot-Assisted Search and Rescue (USA) in collaboration with the International Rescue System Institute (Japan) for port clearing and victim recovery missions using sonar and video. The team was able to cover 80,000 m^2 in 6 hours, finding submerged wreckage and pollutants in areas previously marked clear by divers. Paper [12] shows that in the exploration missions the cooperation between the vehicles extends their capabilities beyond the capabilities of a single vehicle.

Although these examples are not directly connected to software testing, one can see many similarities. Indeed, the testing process is about exploring the space of the system states and traces with the goal to find hidden "wreckage and pollutants". Involving robots and artificial agents into this kind of work looks quite a natural idea to consider. See e. g. recent report by Accenture Labs, which presents an intelligence testing advisor [14].

B. Virtual Assistant Architecture

In order to build an intelligent virtual assistant one needs to combine in one project different technologies and tools from different areas:

- **User input understanding:** audio speech recognition, natural language understanding, user face recognition, sentiment analysis and opinion mining, etc.
- **Intelligent reasoning:** context understanding, dialogue management, social reasoning, domain specific knowledge, user model, etc.
- **Output generation:** sentence construction, gaze, posture and gesture generation, body animation control, audio speech generation, etc.

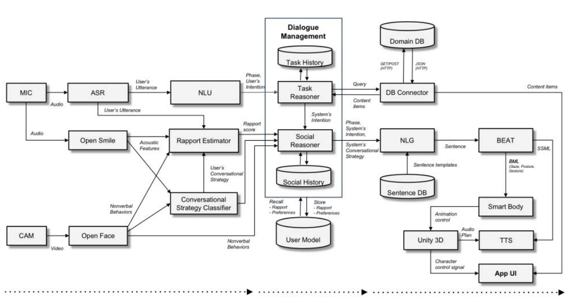

Fig. 4. Architecture of SARA: the Socially Aware Robot Assistant [15]

See for example architecture of SARA [15]: the Socially Aware Robot Assistant on fig. 4.

The good news is one does not have to develop all of these modules from the scratch. Fig. 5 below shows components of SARA architecture, which cover the domain specifics.

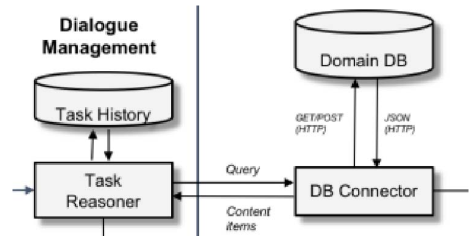

Fig. 5. Fragment of SARA architecture [15], which covers the domain specifics

So a reasonable approach is to take existing of-the-shelf modules related to natural language processing and conversation control and extend them with a systematic knowledge base management system able to keep everything related to software testing skills and the domain expertise.

C. Software Testing Domain Knowledge

Quite many sources could be used to provide a virtual assistant with software testing specific knowledge and skills.

First, countless test plans, defect reports and information from bug tracking system form a meaningful corpus for machine learning algorithms, see e.g. [16].

Second, existing learning materials and certification programs such as those produced by ISTQB [17] may provide a good input for algorithms of information extraction from textual definitions through deep syntactic and semantic analysis, see e.g. [18].

Third, in order to represent and to associate semantics to a large volume of test information a Reference Ontology on Software Testing (ROoST) have been developed in [19]. The ontology defines shared vocabulary to facilitate communication, domain specific knowledge representation, storage and search. Fragment of ROoST's testing process and activities ontology is presented on fig. 6.

Similar projects are going on in other related fields. See [20] for Common Domain Model developed by the International Swaps and Derivatives Association, which provides a standard digital representation of events and actions that occur during the life of a derivatives trade, expressed in a machine-readable format. That could be valuable when using virtual assistants in testing the corresponding financial software.

For data collection one may use Parceval framework to deal with technical complexities related to data acquisition, see [21]. Another related project for for advanced data analysis in software development is Q-Rapids, see e. g. [22].

D. Sample Application: System Logs Analysis

Error logs are one of the high-volume data types that are generally analyzed in QA processes. Log database of a

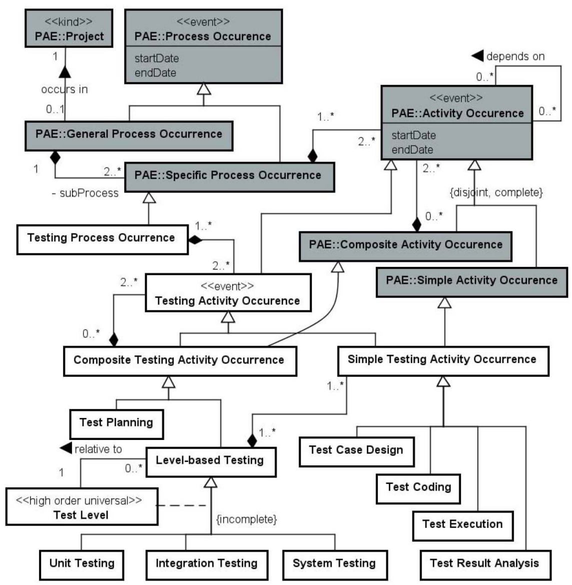

Fig. 6. ROoST's testing process and activities sub-ontology [19]

production system can normally contain up to a million records a day, which makes automation a necessity. Some of the typical questions that a test engineer needs to provide answers to, include:

- What new errors appeared today?
- Are there any traces of recent bugs we fixed? Especially, related to issue we saw yesterday?
- When was the last major change in logs structure and characteristics?
- What applications logged abnormal number of errors today?
- What kind of anomaly was detected in recent logs?

With any of these questions, using virtual assistants features would much facilitate day to day work of a test engineer. Behind the answers, there should be an ML-based bug classification system, such as one described in our previous work [23], to which the virtual assistant would be an interface. Besides, this virtual assistant should also feature context-awareness and understanding questions in natural language, which makes a positive difference to traditional analytic and visualization tools. Being part of a hybrid team, a virtual assistant of this kind would streamline normal work and provide a shorter learning curve for new engineers.

E. Deployment and Team Integration Issues

It is rather clear that all sides of team work will be affected when a robot or a virtual agent is added as a new member to a software testing team. Here are some of the issues:

- **Task distribution** will heavily depend on the intellectual and social abilities of virtual assistants, see [24].
- **Work structure** will change due to new communication structure and role differentiation (see research on multimodal human-computer interaction [25], models of conversations [26], and situation awareness [27]).
- **Team characteristics** will be affected too due to rearranged power distribution, new levels of member diversity and the resulting team climate (see research on trust in hybrid teams [28] and culture related specifics of virtual agents [29]).
- **Team processes** seem to be affected a lot (see papers on human-robot collaboration [30], [31]).
- **Team interventions** will have to include specific training for humans and calibration or update procedures for virtual agent, see e.g. paper [32], which discusses the role of explanation in coordination in human-agent teams, and also [33].
- Measurements of **team performance, team viability, team member satisfaction** will have to take into account presence of artificial team members (see research on building close and harmonious relationship with virtual agents [34], and hybrid human-robot team efficiency [35]).

IV. Conclusion

Software systems become more and more complex and life critical. In order to keep their quality up on the necessary level we build software to test software. Smart testing software should incorporate all state of the art algorithms of complex data analysis, but also it should implement modern communication interfaces to transfer all findings about the tested system behavior from machine to human. Recent advances in machine learning and artificial intelligence provide us with a new type of interface, artificial virtual assistant, which may improve software testing process as well.

References

[1] K. Denecke, M. Tshanz, T. L. Dorner, and R. May, "Intelligent conversational agents in healthcare: Hype or hope?" *Studies in health technology and informatics*, vol. 259, pp. 77–84, 2019.

[2] Z. Lv and X. Li, "Virtual reality assistant technology for learning primary geography," in *International Conference on Web-Based Learning*. Springer, 2015, pp. 31–40.

[3] A. K. Goel and L. Polepeddi, "Jill watson: A virtual teaching assistant for online education," Georgia Institute of Technology, Tech. Rep., 2016.

[4] M. J. Callaghan, G. Bengloan, J. Ferrer, L. Cherel, M. A. El Mostadi, A. G. Eguíluz, and N. McShane, "Voice driven virtual assistant tutor in virtual reality for electronic engineering remote laboratories," in *International Conference on Remote Engineering and Virtual Instrumentation*. Springer, 2018, pp. 570–580.

[5] A. König, A. Malhotra, J. Hoey, and L. E. Francis, "Designing personalized prompts for a virtual assistant to support elderly care home residents," in *Proceedings of the 10th EAI International Conference on Pervasive Computing Technologies for Healthcare*. ICST (Institute for Computer Sciences, Social-Informatics and , 2016, pp. 278–282.

[6] J. Austerjost, M. Porr, N. Riedel, D. Geier, T. Becker, T. Scheper, D. Marquard, P. Lindner, and S. Beutel, "Introducing a virtual assistant to the lab: A voice user interface for the intuitive control of laboratory instruments," *SLAS TECHNOLOGY: Translating Life Sciences Innovation*, vol. 23, no. 5, pp. 476–482, 2018.

[7] M. Harbers and M. A. Neerincx, "Value sensitive design of a virtual assistant for workload harmonization in teams," *Cognition, Technology & Work*, vol. 19, no. 2-3, pp. 329–343, 2017.

[8] C. O'Neil and R. Schutt, *Doing data science: Straight talk from the frontline*. " O'Reilly Media, Inc.", 2013.

[9] J. Cassell, J. Sullivan, E. Churchill, and S. Prevost, *Embodied conversational agents*. MIT press, 2000.

[10] A. Alexeenko, I. Itkin, A. Averina, D. Sharov, and P. Protsenko, "Compatibility testing of clients' protocol connectivity to exchange and broker systems," in *2013 Tools & Methods of Program Analysis*. IEEE, 2013, pp. 63–67.

[11] A. Averina, I. Itkin, and N. Antonov, "Special features of testing tools applicable for use in trading systems production," in *2013 Tools & Methods of Program Analysis*. IEEE, 2013, pp. 39–43.

[12] M. Lindemuth, R. Murphy, E. Steimle, W. Armitage, K. Dreger, T. Elliot, M. Hall, D. Kalyadin, J. Kramer, M. Palankar *et al.*, "Sea robot-assisted inspection," *IEEE robotics & automation magazine*, vol. 18, no. 2, pp. 96–107, 2011.

[13] R. R. Murphy, K. L. Dreger, S. Newsome, J. Rodocker, B. Slaughter, R. Smith, E. Steimle, T. Kimura, K. Makabe, K. Kon *et al.*, "Marine heterogeneous multirobot systems at the great eastern japan tsunami recovery," *Journal of Field Robotics*, vol. 29, no. 5, pp. 819–831, 2012.

[14] Accenture, "Precision testing," Accenture Labs, 2017.

[15] Y. Matsuyama, A. Bhardwaj, R. Zhao, O. Romeo, S. Akoju, and J. Cassell, "Socially-aware animated intelligent personal assistant agent," in *Proceedings of the 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, 2016, pp. 224–227.

[16] A. Gromova, "Using cluster analysis for characteristics detection in software defect reports," in *International Conference on Analysis of Images, Social Networks and Texts*. Springer, 2017, pp. 152–163.

[17] D. Graham, E. Van Veenendaal, and I. Evans, *Foundations of software testing: ISTQB certification*. Cengage Learning EMEA, 2008.

[18] C. D. Bovi, L. Telesca, and R. Navigli, "Large-scale information extraction from textual definitions through deep syntactic and semantic analysis," *Transactions of the Association for Computational Linguistics*, vol. 3, pp. 529–543, 2015.

[19] E. Souza, R. Falbo, and N. L. Vijaykumar, "Using ontology patterns for building a reference software testing ontology," in *2013 17th IEEE International Enterprise Distributed Object Computing Conference Workshops*. IEEE, 2013, pp. 21–30.

[20] C. D. Clack, "Design discussion on the isda common domain model," *Journal of Digital Banking*, vol. 3, no. 2, pp. 165–187, 2018.

[21] S. Dueñas, V. Cosentino, G. Robles, and J. M. Gonzalez-Barahona, "Perceval: Software project data at your will," in *Proceedings of the 40th International Conference on Software Engineering: Companion Proceeedings*. ACM, 2018, pp. 1–4.

[22] L. López, S. Martínez-Fernández, C. Gómez, M. Choraś, R. Kozik, L. Guzmán, A. M. Vollmer, X. Franch, and A. Jedlitschka, "Q-rapids tool prototype: supporting decision-makers in managing quality in rapid software development," in *International Conference on Advanced Information Systems Engineering*. Springer, 2018, pp. 200–208.

[23] I. Itkin, A. Gromova, A. Sitnikov, D. Legchikov, E. Tsymbalov, R. Yavorskiy, A. Novikov, and K. Rudakov, "User-assisted log analysis for quality control of distributed fintech applications," in *The First IEEE International Conference On Artificial Intelligence Testing, April, 4 - 9, 2019, San Francisco East Bay, California, USA*. IEEE, 2019, pp. 11–17.

[24] F. Grimaldo, M. Lozano, and F. Barber, "Coordination and sociability for intelligent virtual agents," in *International Workshop on Coordination, Organizations, Institutions, and Norms in Agent Systems*. Springer, 2007, pp. 58–70.

[25] D. Traum, S. C. Marsella, J. Gratch, J. Lee, and A. Hartholt, "Multi-party, multi-issue, multi-strategy negotiation for multi-modal virtual agents," in *International Workshop on Intelligent Virtual Agents*. Springer, 2008, pp. 117–130.

[26] A. Sagae, J. R. Hobbs, S. Wertheim, M. H. Agar, E. Ho, and W. L. Johnson, "Efficient cultural models of verbal behavior for communicative agents," in *International Conference on Intelligent Virtual Agents*. Springer, 2012, pp. 523–525.

[27] S. Ranathunga and S. Cranefield, "Improving situation awareness in intelligent virtual agents," in *International AAMAS Workshop on Cognitive Agents for Virtual Environments*. Springer, 2012, pp. 134–148.

[28] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. J. De Visser, and R. Parasuraman, "A meta-analysis of factors affecting trust in human-robot interaction," *Human factors*, vol. 53, no. 5, pp. 517–527, 2011.

[29] B. Endrass, M. Rehm, and E. André, "Culture-specific communication management for virtual agents," in *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*. International Foundation for Autonomous Agents and Multiagent Systems, 2009, pp. 281–287.

[30] B. Chandrasekaran and J. M. Conrad, "Human-robot collaboration: A survey," in *SoutheastCon 2015*. IEEE, 2015, pp. 1–8.

[31] T. Bagosi, K. V. Hindriks, and M. A. Neerincx, "Ontological reasoning for human-robot teaming in search and rescue missions," in *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*. IEEE Press, 2016, pp. 595–596.

[32] M. Harbers, J. M. Bradshaw, M. Johnson, P. Feltovich, K. van den Bosch, and J.-J. Meyer, "Explanation in human-agent teamwork," in *Coordination, Organizations, Institutions, and Norms in Agent System VII*, S. Cranefield, M. B. van Riemsdijk, J. Vázquez-Salceda, and P. Noriega, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 21–37.

[33] T. Bickmore and D. Schulman, "A virtual laboratory for studying long-term relationships between humans and virtual agents," in *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*. International Foundation for Autonomous Agents and Multiagent Systems, 2009, pp. 297–304.

[34] L. Huang, L.-P. Morency, and J. Gratch, "Virtual rapport 2.0," in *International Workshop on Intelligent Virtual Agents*. Springer, 2011, pp. 68–79.

[35] M. C. Gombolay, R. A. Gutierrez, S. G. Clarke, G. F. Sturla, and J. A. Shah, "Decision-making authority, team efficiency and human worker satisfaction in mixed human–robot teams," *Autonomous Robots*, vol. 39, no. 3, pp. 293–312, 2015.