Shuyue LI, Jiaqi GUO, Yan GAO, Jianguang LOU, Dejian YANG, Yan XIAO, Yadong ZHOU, Ting LIU, 2022. How to manage a task-oriented virtual assistant software project: an experience report. *Frontiers of Information Technology & Electronic Engineering*, 23(5):749-762. https://doi.org/10.1631/FITEE.2100467

How to manage a task-oriented virtual assistant software project: an experience report

Front Inform Technol Electron Eng

Key words: Experience report; Software project management; Virtual assistant; Machine learning

Corresponding author: Ting LIU
E-mail: tingliu@mail.xjtu.edu.cn
ORCID: https://orcid.org/0000-0002-7600-0934

Motivation

1. Task-oriented virtual assistants (VAs) are software systems that provide users with a natural language interface to complete domain-specific tasks, such as booking flights and ordering foods. With new technological advances in machine learning (ML), we are witnessing an increasing number of commercial VAs being developed.

2. It is, however, particularly challenging to manage a VA software project, as observed in our hands-on experience.

3. The development management approaches and experiences have been hardly studied or shared.

Front Inform Technol Electron Eng

Main idea

To bridge the knowledge gap, in this paper, we share our experience and the lessons that we have learned in managing a task-oriented VA software project at Microsoft, named XTalk.

Front Inform Technol Electron Eng

Method

1. We describe five main concrete software engineering problems that we have encountered in our project, including the problems in requirement management, development management, and quality management.

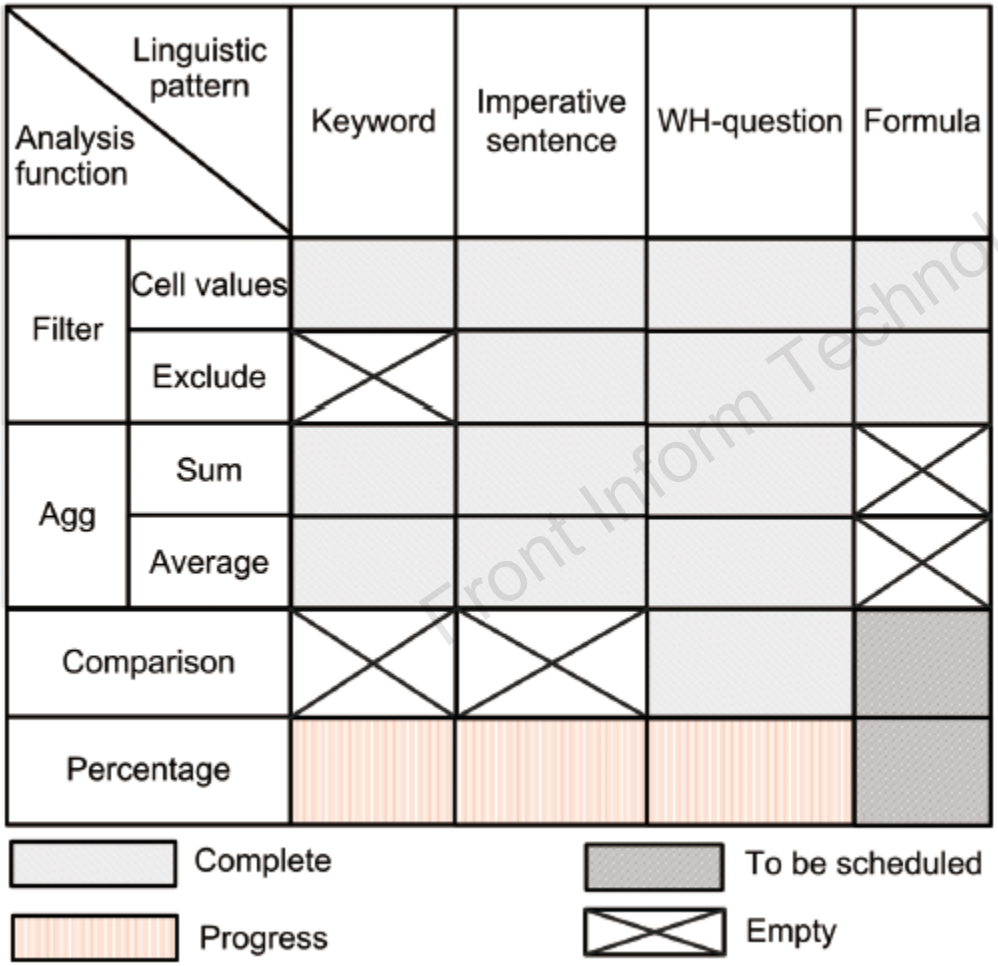
2. We share seven practices adopted in the XTalk project to tackle or mitigate the problems.

3. We summarize three lessons learned from the development of XTalk and present four open challenges.

Front Inform Technol Electron Eng

Major results

1. Software requirement management



The requirement management tool we proposed

Front Inform Technol Electron Eng

Major results (Cont’d)

2. Software development management

Problem 1 The emergent properties of ML, e.g., data-driven and statistical nature, hinder efficient development.

Problem 2 Efficient collection of data for VA projects is under the constraints of domain knowledge and data privacy regulations.

Practice 3 VAs are developed using a hybrid of deep neural network (DNN) and rule-based methods.

Practice 4 The generated NL questions are paraphrased automatically with synthesized oracles.

Front Inform Technol Electron Eng

Major results (Cont'd)

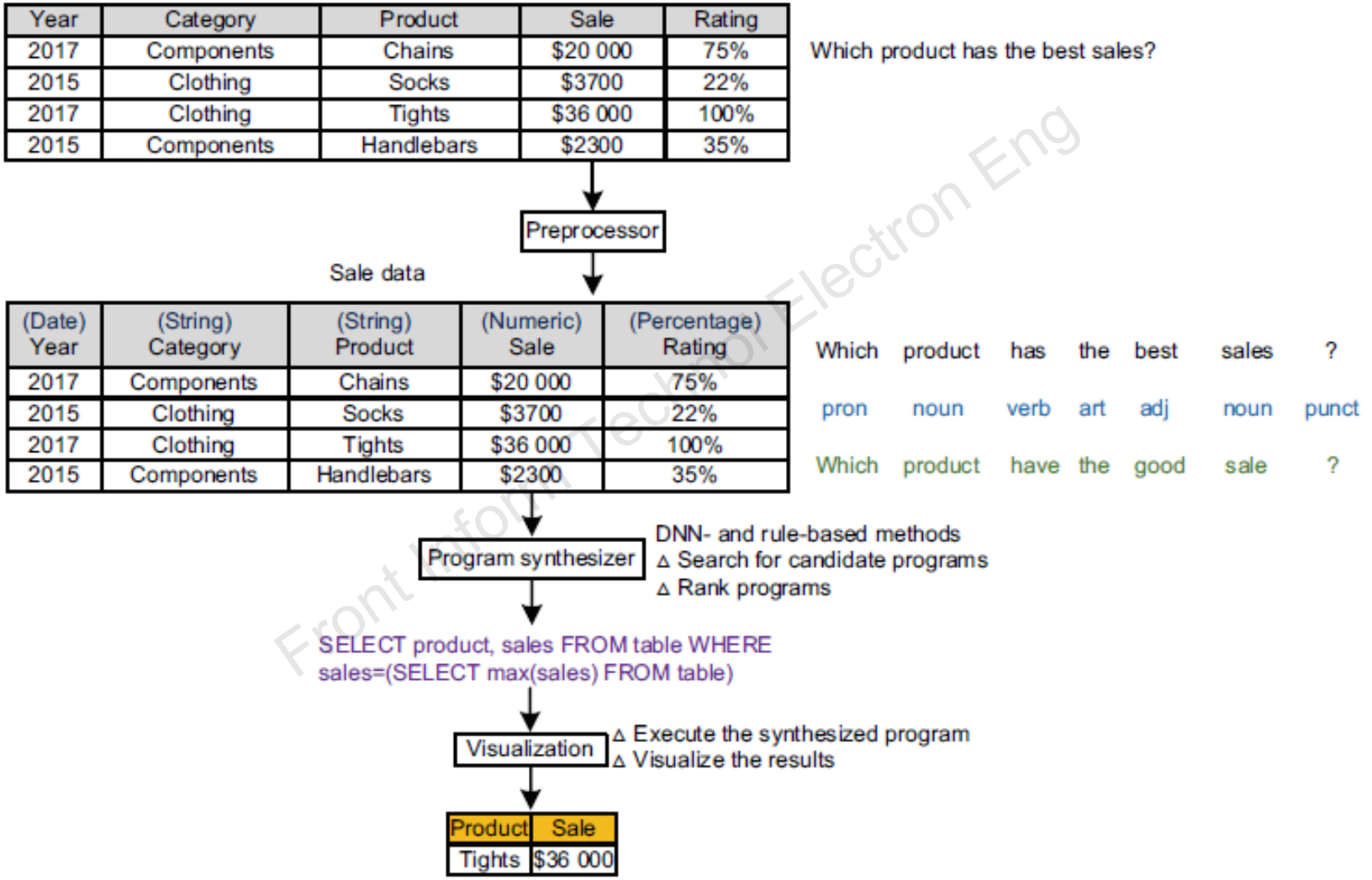


Fig. 2 Overall architecture of XTalk

Front Inform Technol Electron Eng

Major results (Cont'd)

3. Software quality management

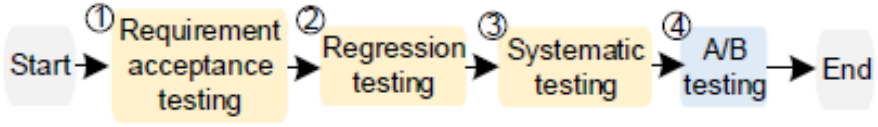


Fig. 4 Testing workflow of XTalk

Testing activities ①–③ are conducted offline, while activity ④ is conducted online

XTalk bug report

NL question: sales in 2010
Table: product sales

Correct answer:

Sales	Year
$1800	2010
$3600	2010

Returned answer:

Sales
$2010
$2010

Test cases:

1. Sales for 2010
2. 2010 sales
3. Sales of 2010
4. Sales for year 2010

Priority	High
Category	Program synthesizer error
Assigned to	XXX
Created date	2019-11-18
Status	Resolved
Product version	1.0.1

Fig. 5 An example bug report of XTalk

In XTalk 1.0.1, for natural language (NL) question "sales in 2010," given the table "product sales," the expected result is filtering the sales in the year 2010, while XTalk wrongly filtered sales that were equal to 2010; it was caused by the program synthesizer module

Front Inform Technol Electron Eng

Major results (Cont’d)

Lessons that we have learned:

1. Making full use of the available data will tremendously facilitate solving the problems occurring while managing a VA software project.
2. We find that the scope of natural language could be organized through its domain-specific tasks and linguistic patterns.
3. To enjoy the benefit of advanced DNN techniques and make the development process manageable, we argue that it is necessary to find a good balance between DNN- and rule-based methods.

Front Inform Technol Electron Eng

Conclusions

1. In this study, we described five critical problems in managing a task-oriented VA software project and shared seven practices adopted in our project XTalk, which is a VA for data analytics. practices
2. We summarized three lessons learned at managing XTalk and four open challenges that require more research efforts.

Front Inform Technol Electron Eng

Shuyue LI received her BE degree from Xi'an Jiaotong University, Xi'an, China, in 2018. She is currently a PhD candidate in the Faculty of Electronic and Information Engineering, Xi'an Jiaotong University. Her research interests include software testing and machine learning.

Jiaqi GUO is currently a PhD candidate in the Faculty of Electronic and Information Engineering, Xi'an Jiaotong University. His research interests include program synthesis, semantic parsing, and deep learning.

Yan GAO is now a researcher at Microsoft Research Asia. Her research interests are primarily on natural language understanding and its applications in business intelligence and natural language interface for database.

Jianguang LOU is a senior principal researcher at Microsoft Research Asia. He obtained his MS degree from Zhejiang University in 2000 and his PhD degree from Institute of Automation, Chinese Academy of Sciences in 2003. His research areas include multimedia systems, data mining for software systems, machine learning, and data mining for business intelligence.

Front Inform Technol Electron Eng

Dejian YANG is a researcher SDE at Microsoft Research Asia. His research interests are primarily on data platforms and analytics.

Yan XIAO is currently a researcher SDE at Microsoft Research Asia. His research interests are primarily on natural language understanding.

Yadong ZHOU is now an associate professor with the Faculty of Electronic and Information Engineering, Xi'an Jiaotong University. His research interests include online social network and cyber-physical system security.

Ting LIU is now a professor in the Faculty of Electronic and Information Engineering, Xi'an Jiaotong University. He obtained his PhD degree from Xi'an Jiaotong University in 2010. His research interests include security and reliability in computer network, cyber-physical system, and software system.

Front Inform Technol Electron Eng