A Bayesian Based Machine Learning Application to Task Analysis

B

Shu-Chiang Lin
Purdue University, USA

Mark R. Lehto
Purdue University, USA

INTRODUCTION

Many task analysis techniques and methods have been developed over the past decades, but identifying and decomposing a user's task into small task components remains a difficult, impractically time-consuming, and expensive process that involves extensive manual effort (Sheridan, 1997; Liu, 1997; Gramopadhye and Thaker, 1999; Annett and Stanton, 2000; Bridger, 2003; Stammers and Shephard, 2005; Hollnagel, 2006; Luczak et al., 2006; Morgeson et al., 2006). A practical need exists for developing automated task analysis techniques to help practitioners perform task analysis efficiently and effectively (Lin, 2007). This chapter summarizes a Bayesian methodology for task analysis tool to help identify and predict the agents' subtasks from the call center's naturalistic decision making's environment.

BACKGROUND

Numerous computer-based task analysis techniques have been developed over the years (Gael, 1988; Kirwan and Ainsworth, 1992; Wickens and Hollands, 2000; Hollnagel, 2003; Stephanidis and Jacko, 2003; Diaper and Stanton, 2004; Wilson and Corlett, 2005; Salvendy, 2006; Lehto and Buck, 2008). These approaches are similar in many ways to methods of knowledge acquisition commonly used during the development of expert systems (Vicente, 1999; Schraagen et al., 2000; Elm et al., 2003; Shadbolt and Burton, 2005). Several taxonomies exist to classify knowledge elicitation approaches. For example, Lehto et al. (1992) organize knowledge elicitation methods (including 140 computer-based tools), identified in an extensive review of 478 articles, into three categories: manual methods, interactive or semi-automated methods, and automated or machine learning methods. Manual methods such as protocol analysis or knowledge organization are especially useful as an initial approach because they can be used to effectively retrieve structure and formalize knowledge components, resulting in a knowledge base that is accurate and complete (Fujihara, et al., 1997). Studies such as Trafton et al. (2000) have shown this technique can capture the essence of qualitative mental models used in complex visualization and other tasks. The drawbacks of this technique are similar to those of classic task analysis techniques in that they involve extensive manual effort and may interfere with the expert's ability to perform the task. Semi-automated methods generally utilize computer programs to simplify applications of the manual methods of knowledge acquisition. The neural network model is one of the methods in common use today, especially when learning and recognition of patterns are essential (Bhagat, 2005). A neural network can self-update its processes to provide better estimates and results with further training. However, one arguable disadvantage is that this approach may require considerable computational power should the problem be somewhat complex (Dewdney, 1997).

Automated methods or machine learning based methods primarily focus on learning from recorded data rather than through direct acquisition of knowledge from human experts. Many variations of commonly used machine learning algorithms can be found in the literature. In general, the latter approach learns from examples-guided deductive/inductive processes to infer rules applicable to other similar situations (Shalin, et al., 1988; Jagielska et al., 1999; Wong & Wang, 2003; Alpaydın, 2004; Huang et al., 2006; Bishop, 2007).

MAIN FOCUS

The Bayesian framework provides a potentially more applicable method of task analysis compared to competing approaches such as neural networks, natural language processing methods, or linguistic models. Two Bayesian methods are often proposed: naïve Bayes and fuzzy Bayes. Over the decades, studies such as those of Bookstein, (1985), Evans and Karwowski (1987), Lehto and Sorock (1996), Chatterjee (1998), Yamamoto and Sagisaka (1999), Zhu and Lehto (1999), Qiu and Agogino (2001), Hatakeyama et al. (2003), Zhou and Huang (2003), Leman and Lehto (2003), Wellman et al. (2004), and Bolstad (2004) have shown that statistical machine learning within the framework of fuzzy Bayes can be more efficient when the assumptions of independence are violated. McCarthy (2002) found that fuzzy Bayes gave the highest success rate for print defect classification compared to ID3, C4.5, and individual keyword comparison algorithms. Noorinaeini and Lehto (2007) compare the accuracy of three Singular Value Decomposition (SVD) based Bayesian/Regression models and conclude that all three models are capable of learning from human experts to accurately categorize cause-of-injury codes from injury narrative.

Case studies have contributed to both theoretical and empirical research in the naturalistic decision making environment (Zsambok, 1997; Klein, 1998; Todd & Gigerenzer, 2001; Hutton et al, 2003). The following discussion presents a brief case study illustrating the application of a Bayesian method to task analysis. This particular study here focuses on describing what takes place in a call center, when the customer calls to report various problems and the knowledge agent helps troubleshoot remotely. In this example, the conversation between agent and customer was recorded and manipulated to form a knowledge database as input to the Bayesian based machine learning tool.

Model Development

Figure 1 illustrates important elements of the dialog between a call center knowledge agent and customer. The arrows indicate data flow. The dialog between the customer and the knowledge agent can be recorded using several methods. For example, if the customer uses e-mail, these conversations are directly available in written form. The knowledge agent's troubleshooting processes similarly could be recorded in video streams, data screens, time-stamp streams of keystrokes, mouse-clicks, data streamed to the agent's monitor, or various forms of data entry used by agents. These data streams can be synchronized with a time-stamp as input for the Bayesian based machine learning tool.

Figure 1, Model of Bayesian based machine learning tool for task analysis[1]
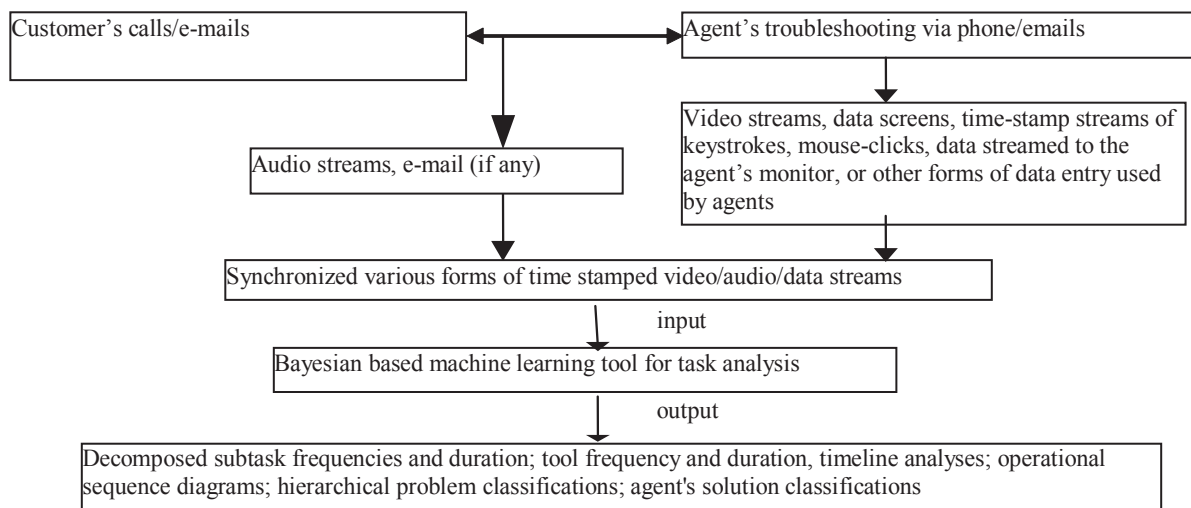
Related Content

Distributed Data Mining
Grigorios Tsoumakas (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 709-715).*
www.irma-international.org/chapter/distributed-data-mining/10898

Using Prior Knowledge in Data Mining
Francesca A. Lisi (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 2019-2023).*
www.irma-international.org/chapter/using-prior-knowledge-data-mining/11096

Data Quality in Data Warehouses
William E. Winkler (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 550-555).*
www.irma-international.org/chapter/data-quality-data-warehouses/10874

Automatic Genre-Specific Text Classification
Xiaoyan Yu, Manas Tungare, Weiguo Fan, Manuel Pérez-Quiñones, Edward A. Fox, William Cameronand Lillian Cassel (2009). *Encyclopedia of Data Warehousing and Mining, Second Edition (pp. 120-127).*
www.irma-international.org/chapter/automatic-genre-specific-text-classification/10808

Unleashing the Potential of Every Child: The Transformative Role of Artificial Intelligence in Personalized Learning
Natalia Riapina (2024). *Embracing Cutting-Edge Technology in Modern Educational Settings (pp. 19-47).*
www.irma-international.org/chapter/unleashing-the-potential-of-every-child/336189