Mastering Situation Awareness: The Next Frontier?

Dieter Gawlick
Oracle
500 Oracle Parkway
Redwood City, CA 94065
dieter.gawlick@oracle.com

ABSTRACT

John Boyd noticed in the 1960's the importance of Situation Awareness for military operations and introduced the notion of the OODA loop (Observe, Orient, Decide, and Act) [1]. There are many applications that have to deal with Situation Awareness: CRM, HCM, SCM, patient care, power grid management, and cloud services management; the list seems to be endless. Situation awareness requires applications to integrate the management of data, knowledge, processes, and other services such as social networking and middleware. These applications require high personalization as well as rapid and continuous evolution. Additionally, a wide variety of operational and functional requirements must be met, including real time execution and strong temporal support with provenance. Handcrafting these applications is a significant task that requires exhaustive resources for development and maintenance. Due to the resources and time involved in their development, these applications typically fall way short of the desired functionality, operational characteristics, and ease and speed of evolution.

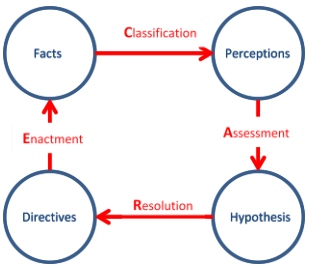

Figure 2: The KIDS CARE Loop

We – Eric S. Chan, Adel Ghoneimy, Zhen Hua Liu, and the author (all at Oracle) – have developed a model to support Situational Awareness; we call it **KIDS** – Knowledge Intensive Data-processing System [2].

The basic ideas are as follows: KIDS manages data, knowledge and processes. KIDS assumes that there are four categories of data (Facts, Perceptions, Hypotheses, and Directives) and four categories of knowledge (Classification, Assessment, Resolution, and Enactment). Facts are what we measure – facts are quantitative. Facts are classified to derive perceptions – perceptions are compact qualitative interpretations of facts designed to be easily useable by the human brain. Perceptions are assessed to derive one or more hypotheses, hypotheses are used to derive directives. The enactment of the directives will create new facts; the loop starts again until the problem on hand has been solved.

The four categories of data require state of the art data management. This starts with the support of structured and unstructured data and with extensibility for domain specific data. Temporal support with provenance and selective long term retention is a must as is real time processing of queries (rules, models). The management of data has to be done with rich declarative interfaces. OLTP as well as real-time analytics have to be handled concurrently and with great efficiency.

While the use of tacit knowledge is always the last resort, the use of formal knowledge is preferred. For formal representations of classification knowledge the preferred choices include Support Vector Machines, Naïve Bayesian Networks, and Neural Networks; for assessments the preferred choices include Bayesian Belief Networks and Least-Squares Optimization; for resolutions the preferred choices include Influence Diagrams and Prognosis of Remaining Useful Life; and for enactments the preferred choices include control structures encoded in scripts, plans, schedules, BPEL workflows, and business processes in BPMN.

The execution of the various categories of knowledge represents a process, which we call the CARE loop; the existing process management technologies can be leveraged. An additional loop supports the evolution of knowledge based on tacit knowledge and/or improved machine interpretation of the data. All data, knowledge, and processes are stored and managed using temporal databases to enable unambiguous provenance. KIDS poses many technical challenges; real processing of knowledge seems to be one of the major challenges. CEP (Complex Event Processing) is certainly a good start; however, the model, the language, the provenance support, and other components need major enhancements.

A challenge as demanding as situation awareness will ultimately only succeed with a solid mathematical foundation. The best fit seems to be the Theory of Categories [3]. The KIDS model is currently used in Oracle to develop cloud management services as well as customer support; it is a domain where the IT experts, including the authors of KIDS, are also the domain experts. The expected benefits of the CARE loop processing are well visible even though the KIDS software is not yet available. We also know that KIDS is a very good model for the management of the bug databases such as JIRA. A project at the Medical Center of the University of Utah in 2010 was instrumental for the evolution of KIDS [4]; this project showed especially the importance of the classification of facts into perceptions.

REFERENCES

[1] http://en.wikipedia.org/wiki/OODA_loop

[2] E.S. Chan, D. Gawlick, A. Ghoneimy, and Z.H. Liu, "Situation Aware Computing for Big Data," SemBIoT 2014

[3] D.I. Spivak: "Category Theory for the Scientists," ISBN-13: 978-0262028134

[4] D. Guerra, U. Gawlick, P. Bizarro, D. Gawlick: "An Integrated Data Management Approach to Manage Health Care Data," BTW 2011, pp 596-605.

7th Biennial Conference on Innovative Data Systems Research (CIDR '15) January 4-7, 2015, Asilomar, California, USA.