Research Statement

Eric P. Kasten

In a broad stroke, an overarching theme of my research addresses bridging the gap between real-world sensor input and *in silico* interpretation and perception. My studies often integrate different technologies in an interdisciplinary fashion. For instance, a project that addresses environmental monitoring may require innovations in sensor technology, distributed computing, and informatics to enable the collection, search, and analysis of large volumes of data to better understand how the natural world is changing and how it will affect our lives. Similarly, adaptive and autonomic software may need to sense both the physical and virtual world to successfully reconfigure while maintaining performance objectives, address system failures, or to adapt to changing user requirements in real-time. Such systems often need to address the challenges presented by extremely large data sets, or big data, to extract important information, make decisions and adapt in a timely fashion.

My research strives to improve our capacity for sensing and interpreting the world around us using computation as an instrument that can extend our senses and comprehension. My research is interdisciplinary in that the broader goals of my research are often grounded by the goals of other disciplines. My research typically combines three elements to achieve timely distillation of raw data into usable information: distributed computing, informatics, and machine learning. The goal is to enable understanding, discovery and decision making by humans, or adaptivity in autonomic software systems.

Research Background

During pursuit of my M.S., I focused on computer networks and distributed computing. In this early work, I helped investigate efficient algorithms for distributing data among a set of computers connected by an ATM network [1]. Later, I applied my knowledge of distributed systems to the design and implementation of a data acquisition system [2,3] to support nuclear physics research.

My doctoral research would build on my early studies and experience designing and supporting distributed software systems. I focused on the design and implementation of adaptive and autonomic software, that can manage and protect itself with only high-level human guidance [4,5]. My investigations produced the key elements of my Ph.D. dissertation [6]. My dissertation describes my studies on the integration of three major components found in the design of autonomic software: *adaptive mechanisms* [7–9] that enable software to be recomposed during execution; *state maintenance* [10] that orchestrates adaptation to avoid data loss and ensure continued operation; and *decision making* [11] that enables software to adapt autonomously.

One focus of my research was on autonomous computation in data streaming applications that benefit from dynamic recomposition and reconfiguration during execution. Such applications include: data acquisition, sensor data analysis, and software for streaming communication or multimedia data. For instance, an application that streams acoustic data over a wireless network to a mobile receiver needs to avoid packet loss while conserving bandwidth. In mobile applications of this type, automated, run-time exchange of error control filters can avoid packet loss and improve audio quality without interruption of service. A notable product of my investigation of autonomous decision making in software was the design and implementation of the MESO [11,12] perceptual memory system. Perceptual memory, refers to a type of long-term memory for remembering external stimulus patterns [13], that can offer a useful model for an important component of decision making in context-aware, adaptive software. The ability to remember complex, high-dimensional patterns that occur as a product of interaction between application users and the environment, and to quickly recall associated actions, enables timely, autonomous system response.

Software that can learn how to interact with complex pervasive computing and ubiquitous [14] environments while accounting for user preferences are becoming increasingly important. For instance, to explore the use of MESO to support learning in adaptive software, I conducted a case study involving adaptive error control [12]. Specifically, MESO was used to implement the decision maker in a data streaming network application, that adapts to changes in packet loss rate as a user roams about a wireless cell. This application used MESO to "remember" user preferences for balancing packet loss with bandwidth consumption. The decision maker gains this knowledge through *imitative learning*. A user shows the application how to adapt to a rising loss rate by selecting an error correction setting with greater redundancy. If the new setting reduces the loss rate to an acceptable level, the user reinforces the new configuration (e.g., by pressing a particular key), and the application uses MESO to associate the sensed environment and the selected configuration. Later, when operating autonomously, the decision maker senses current environmental conditions and queries MESO for a system configuration that most likely addresses current conditions. Then, the decision maker emulates the user's actions and adapts, changing the configuration to match that returned from MESO. This case study demonstrated that through human-computer interaction an application can "learn" to significantly reduce loss rate while conserving bandwidth under good channel conditions. In addition, working with other members of the Software Engineering and Network System (SENS) Laboratory, enabled integration of my work with that of others and produced a number of related publications in areas such as wearable computers [15], intrusion detection [16] and others [17–19].

Current Research

After completing my Ph.D. I sought to explore interdisciplinary postdoctoral studies at the Remote Environmental Assessment Laboratory (REAL) and the Global Observatory for Ecosystem Services (GOES). In collaboration with ecologists and other scientists I extended my research to ecological sensor applications that target the advancement of remote environmental monitoring. Thus began my research on automated processing of sensor data streams for ecological and environmental applications. During my tenure in these laboratories, I have worked on custom sensor hardware and software systems, and played a key role in the design and implementation of a web-based system that enables the collection, cataloging and search of acoustic recordings acquired using environmental sensors deployed in natural settings.

During this project I have studied run-time recomposable infrastructure and automated processing and analysis of acoustic data streamed from sensor platforms deployed in natural areas [20, 21], such as forest or wetlands. This project had two key goals. The first was the near real-time extraction of acoustic events that can be used for the classification and detection of species of birds. The second was the automated summarization of recordings to enable organization and searching of sensor data repositories. The collection of acoustic sensor data can produce a large time series of complex data that must be organized and searchable to facilitate further research. My investigations helped developed the numeric and machine learning techniques that enable search and analysis of these repositories [22, 23]. During these projects, I have enjoyed being part of a collaboration with the University of Notre Dame's Environmental Research Center and with the University of Wisconsin's Trout Lake Station. At these sites, I have gained insight and field experience on site selection and sensor deployment.

In addition to my work with terrestrial sensors and data analysis, I have also begun research on the analysis of very high resolution satellite imagery. The goal of this project is to support reliable, remote monitoring of the carbon stores sequestered by forests and individual trees. This represents a significant remote sensing and image understanding problem that requires advancement in classification and object recognition techniques in a spatial context.

Research Agenda

In the future, I intend to pursue research that extends our ability to process data streams and big data to support automated analysis, visualization and distillation of raw data into useful knowledge. Second, I intend to continue to study evolutionary computation and analytical methods that can be used to further our understanding of ecological and biological processes. For instance, numerical and machine learning methods can be used to mine remotely sensed data to provide insight on ecological processes in remote regions that are difficult or costly to visit. Third, I intend to continue my pursuit of collaborations that enable the integration of computer science and software engineering methods with other disciplines. Finally, I wish to continue my investigations on adaptive and autonomic computing with an eye toward better understanding autonomous decision making and human-computer interaction in software to better understand how software and biological systems can adapt through interaction with their environment.

References Cited

[1] C. Huang, E. P. Kasten, and P. K. McKinley, "Design and implementation of multicast operations for ATM-based high performance computing," in *Proceedings of Supercomputing'94*, (Washington, D.C., USA), pp. 164–173, November 1994.

[2] R. Fox, E. Kasten, K. Orji, C. Maurice, and J. Venema, "Real-time results without real-time systems," in *Proceedings of the 13th Real Time Computer Applications in Nuclear, Particle and Plasma Physics Conference (RT2003)*, (Montreal, Quebec, Canada), May 2003.

[3] R. Fox, E. Kasten, K. Orji, C. Maurice, and J. Venema, "Real-time results without real-time systems," *IEEE Transactions on Nuclear Science*, 2004.

[4] J. O. Kephart and D. M. Chess, "The vision of autonomic computing," *IEEE Computer*, pp. 41–50, January 2003.

[5] P. K. McKinley, S. M. Sadjadi, E. P. Kasten, and B. H. Cheng, "Composing adaptive software," *IEEE Computer*, vol. 37, pp. 56–64, July 2004. Featured article.

[6] E. P. Kasten, *An Integrated Approach to Autonomous Computation in Data Streaming Applications*. PhD thesis, Department of Computer Science and Engineering, Michigan State University, East Lansing, Michigan, USA, 2007.

[7] E. P. Kasten, P. K. McKinley, S. M. Sadjadi, and R. Stirewalt, "Separating introspection and intercession to support metamorphic distributed systems," in *Proceedings of the 22nd IEEE International Conference on Distributed Computing Systems ICDCS'02*, (Vienna, Austria), July 2002.

[8] P. K. McKinley, E. P. Kasten, S. M. Sadjadi, and Z. Zhou, "Realizing multi-dimensional software adaptation," in *Proceedings of the ACM Workshop on Self-Healing, Adaptive and self-MANaged Systems (SHAMAN), held in conjunction with the 16th Annual ACM International Conference on Supercomputing*, (New York, New York, USA), June 2002.

[9] S. Sadjadi, P. K. McKinley, and E. P. Kasten, "Metasockets: Run-time support for adaptive communication services," in *Addendum to the proceedings of the International Symposium on Distributed Object and Applications (DOA)*, (Irvine, California, USA), pp. 42–45, November 2002.

[10] E. P. Kasten and P. K. McKinley, "Perimorph: Run-time composition and state management for adaptive systems," in *Proceedings of the 24th IEEE International Conference on Distributed Computing Systems ICDCS'04*, (Tokyo, Japan), March 2004.

[11] E. P. Kasten and P. K. McKinley, "MESO: Supporting online decision making in autonomic computing systems," *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, vol. 19, pp. 485–499, April 2007. Featured article.

[12] E. P. Kasten and P. K. McKinley, "MESO: Perceptual memory to support online learning in adaptive software," in *Proceedings 3rd International Conference on Development and Learning (ICDL'04)*, (La Jolla, California, USA), October 2004.

[13] J. M. Fuster, *Memory in the Cerebral Cortex: An Empirical Approach to Neural Networks in the Human and Nonhuman Primate*. Cambridge, Massachusetts, USA: The MIT Press, 1995.

[14] M. Weiser, "The computer for the twenty-first century," *Scientific American*, pp. 94–110, 1991.

[15] P. K. McKinley, S. M. Sadjadi, E. P. Kasten, and R. Kalaskar, "Programming languages support for adaptive wearable computing," in *Proceedings of the International Symposium on Wearable Computers (ISWC02)*, (Seattle, Washington, USA), October 2002.

[16] P. K. McKinley, S. M. Sadjadi, and E. P. Kasten, "An adaptive software approach to intrusion detection and response," in *Proceedings of the 10th International Conference on Telecommunication Systems, Modeling and Analysis (ICTSM10)*, (Monterey, California,USA), October 2002.

[17] S. Sadjadi, P. K. McKinley, and E. P. Kasten, "Metasockets: Run-time support for adaptive communication services," Tech. Rep. MSU-CSE-02-22, Department of Computer Science, Michigan State University, East Lansing, Michigan, USA, July 2002.

[18] S. M. Sadjadi, P. K. McKinely, and E. P. Kasten, "Architecture and operation of an adaptable communication substrate," in *Proceedings of the 9th International Workshop on Future Trends of Distributed Computing Systems (FTDCS '03)*, (San Juan, Puerto Rico), May 2003.

[19] S. M. Sadjadi, P. K. McKinley, E. P. Kasten, and Z. Zhou, "Metasockets: Design and operation of run-time reconfigurable communication services," *Software: Practice and Experience (SP&E). Special Issue: Experiences with Auto-adaptive and Reconfigurable Systems.*, vol. 36, pp. 1157–1178, 2006.

[20] E. P. Kasten, P. K. McKinley, and S. H. Gage, "Automated ensemble extraction and analysis of acoustic data streams," in *Proceedings of the 1st International Workshop on Distributed Event Processing, Systems and Applications (DEPSA), held in conjunction with the 27th IEEE International Conference on Distributed Computing Systems (ICDCS)*, (Toronto, Ontario, Canada), June 2007.

[21] E. P. Kasten, P. K. McKinley, and S. H. Gage, "Ensemble extraction for classification and detection of bird species," *Ecological Informatics*, vol. 5, pp. 153–166, 2010.

[22] W. Joo, S. H. Gage, and E. P. Kasten, "Analysis and interpretation of variability in soundscapes along an urban-rural gradient," *Landsacpe and Urban Planning*, vol. in press, p. 18 pages, 2011. doi:10.1016/j.landurbplan.2011.08.001.

[23] E. P. Kasten, S. H. Gage, and J. Fox, "The REAL Digital Acoustics Library: An acoustic archive for studying the natural soundscape," 2011. in preparation.