Verification and Validation of Intelligent Systems: Five Years of AAAI Workshops

Daniel E. O'Leary*
3660 Trousdale Parkway, University of Southern California, Los Angeles, California 90089-1421

I. INTRODUCTION

One of the longest sequences of ongoing workshops at the AAAI (American Association for Artificial Intelligence) meeting has been the "Workshop on Verification, Validation and Testing of Intelligent Systems" ("V&V workshop" or "workshop"). These workshops have been actively attended by a V&V researchers, tool developers, and practitioners. They have played an important role in facilitating communication between these groups and in the dissemination of major new results and systems.

The first five workshops on V&V occurred from 1988–1992. Over that time period there were about 125 articles accumulated in the yearly workshop proceedings. However, those proceedings have been limited primarily to those in attendance at those workshops (e.g., Bellman,[1] Culbert,[2] O'Leary,[3] Miller,[4] and Preece[5]). This edited collection presents 11 of the papers presented at these meetings.

Although this set of articles focuses on those AAAI workshops from 1988–1992 there are other meetings that have focused on V&V issues. The International Joint Conference on Artificial Intelligence (IJCAI) has had workshops on V&V at each meeting since 1989. First, the 1989 AAAI V&V workshop was a joint AAAI and IJCAI workshop in Detroit. Second, there were workshops on V&V organized at both the 1991 Australia and 1993 Chambery IJCAI meetings. Further, the European Conference on AI (ECAI) has had a number of workshops on verification and validation. In addition, EUROVAV recently began yearly organized meetings on V&V. Finally, in 1992 (at Houston, NASA) and 1993 (at the IEEE Conference on Artificial Intelligence Applications in Or-

* oleary@rcf.usc.edu
Tel. No.: 213-740-4856

INTERNATIONAL JOURNAL OF INTELLIGENT SYSTEMS, VOL. 9, 653–657 (1994)
 CCC 0884-8173/94/080653-05

lando) there were meetings of the Winter Workshop on Verification and Validation.

The remainder of this article is organized as follows. Sec. II provides a brief background, Sec. III summarizes the lines of research discussed characterizing the articles in this collection, and Sec. IV briefly discusses future meetings and article collections on V&V.

II. BACKGROUND: VERIFICATION AND VALIDATION

There has been an explosion of activity in the areas of V&V over the past 5 to 10 years. This interest has derived, in part, from the need to test the large number of expert systems that have been developed since the mid-1980s. It also has derived from the increasing role that intelligent systems are taking in critical situations, such as those in aerospace and medicine.

The importance of V&V has been recognized in a number of ways. Large projects on V&V have been funded in Europe, Japan, and the US, through organizations such as ESPRIT, DARPA, and NASA. In addition, there have been some collections of articles on V&V in the literature. Culbert,[6] Ayel and Laurent,[7] and Preece and Suen[8] have developed edited collections of original articles. Gupta[9] generated an edited collection of over 50 previously published articles.

The focus of the research articles presented in this collection and those presented at the V&V workshop is aimed at intelligent systems; thus, they address issues that exploit or concern unique V&V issues in intelligent systems. However, in many cases the results generalize to other kinds of systems.

While there is controversy over how to define the terms *verification* and *validation*, there is general consensus that validation refers to the process of building the right system, while verification refers to the process of building the system right (e.g., O'Keefe et al.[10]). (A recent survey of the literature on V&V is provided by O'Keefe and O'Leary[11]).

III. LINES OF RESEARCH IN THIS COLLECTION

This collection investigates some of the primary lines of research in V&V. The articles presented in this collection fall into five basic categories exploring those lines of research:

(1) Automated Tools to Verify Knowledge Bases (Becker et al.; Preece and Shinghal; Zlatareva)
(2) Design Languages that Support V&V (French and Hamilton; Highland and Kornman; Lee et al.)
(3) Mathematical and Set-based Models, Facilitating V&V (Gold and Plant; Laita et al.)
(4) Theory of V&V (O'Leary)
(5) Summary of International Projects (Meseguer and Plaza; Terano)

A. Automated Tools

One of the primary areas of research in V&V has been the generation of automated tools to perform V&V. For example, Chang et al.[12] provide a history of automated tools over the time period 1979–1989. In that analysis they discuss the contributions of 12 articles.

The current collection provides three excellent examples of automated verification and validation processes. Becker et al. ("Automated Test Generation and Evaluation for Real-time Expert Systems") discuss a software development tool, the Activation Framework (AF), which can be used to support the development of distributed real-time AI applications. Generating test data must take into account the basic nature of the system being tested. As a result, pure random generation of characteristics is unlikely to provide any insight since the resulting portfolio of characteristics may not be feasible or similar to any real world situation.

Preece and Shinghal ("Foundation and Application of Knowledge-Based Verification") provide formal definitions of four types of anomalies (redundancy, ambivalence, circularity, and deficiency) addressed by automated systems. In addition, Preece and Shinghal also provide some empirical evidence of the system developed by Preece (Cover) on its ability to find anomalies in five different expert systems. Those times range from around 6 minutes to over 3 hours, while running on a Sun 4/300 workstation.

Zlatareva ("A Framework for Verification, Validation, and Refinement of Knowledge Bases: The VVR System") discusses another automated system, DIVER. The system seems to be more efficient than Cover, finding the anomalies in one system in 10 minutes, while Cover took 3.5 hours. Both parallel and nonparallel versions of the verification procedures are discussed. The parallel version led to only slight improvements in the necessary time to solve process all the rules.

B. Design Languages that Support V&V

One of the more recent lines of V&V research seems to be the development of programming environments designed to facilitate V&V processes. This collection has three different, yet interrelated, articles focusing on the design and development of such environments.

French and Hamilton, in their article "A Comprehensive Framework for Knowledge-Based Verification and Validation," outline the TOP (Terms, Operators, and Productions) system. TOP designs C-Language Production System (CLIPS) rules. The principle goal of TOP is to integrate generally accepted software engineering practice into the design of expert systems, by providing a language and methodology. TOP does that in three ways. First, it encourages the definition of knowledge units. Second, TOP supports the definition of expectations about the behavior of the knowledge (e.g., sequencing). Third, TOP separates domain knowledge and control knowledge for verification purposes.

Lee et al. ("A Tool for Task-Based Knowledge and Specification Acquisition") discuss their system, TAME (Task-based knowledge Acquisition Methodology for Expert systems). TAME is a hypertext-based knowledge acquisition tool designed to support specification elicitation and refinement. TAME provides feedback to the user about incomplete refinements and duplications.

Highland and Kornman ("A Design Language and the Use of Cleanroom Methodology for Knowledge-Based System Development") discuss a design language and associated verification techniques. The approach makes extensive use of English language specifications in a process of stepwise refinement. The article illustrates the use of the design language in the development of a prototype flight replanner.

C. Mathematical and Set-Based Models, Facilitating V&V

One approach to V&V is to develop mathematical representations of knowledge bases, which are then verified by exploiting the mathematical structure. Laita et al. ("A Formal Model for Knowledge-Based Systems Verification") developed logico–algebraic method for the development of a formal representation of a knowledge base. The article provides a theoretical construct for the investigation of V&V issues.

In the article "Towards the Formal Specification of an OPS5 Production System Architecture," Gold and Plant present a formal specification of the OPS5 production systems framework. They use "Z" notation, a language based on typed set theory.

D. V&V Theory

There have been limited efforts to develop a theory of verification and validation. However, this collection includes one article with that goal. O'Leary ("Artifacts: Toward a Theory of Verification and Validation") presents a theory for verification and validation based on the notion of "artifacts." The use of artifacts in AI and other disciplines, such as archeology, is discussed. Artifacts are found to provide a language and structure that is useful in V&V. The theory suggests that V&V methods attach to each artifact and partially ordered set of artifacts. The choice of the artifacts then guides the choice of the V&V methods. The notion of artifacts expands the focus of V&V beyond classic knowledge-base verification, and suggests that a wide range of artifacts be considered as part of the V&V process.

E. International Projects

Although each of the five AAAI workshops have been in the US, there has been an effort to internationalize the workshop. At least two of the workshops have focused on international developments in V&V. Accordingly, two of the articles summarize large-scale projects in Europe and Japan, respectively.

The article "The VALID Project: Goals, Development, and Results," by

Meseguer and Plaza derives from one such discussion. Meseguer and Plaza summarized V&V methods and tools developed under the ESPRIT project. Ultimately, the project has resulted in a V&V tool kit that consists of seven different tools ranging from anomaly detectors to interactive knowledge-based inspectors.

In "The JIPDEC Checklist-Based Guideline for Expert System Evaluation," Terano summarizes work on a project sponsored by JIPDEC. Although V&V is a primary concern of this article, Terano is also concerned with other factors such as productivity, security, usability, and other issues. He summarizes some case studies of firms that have used the methodology in the process of expert system evaluation.

IV. THE FUTURE

While this collection summarizes selected articles over the years 1988–1992, there is ongoing activity in the V&V community. During 1993 there were V&V workshops sponsored at both AAAI (organized by A. Preece[12]) and IJCAI in Chambery (organized by M. Ayel and J.-P. Laurent). There will be a V&V workshop at the 1994 AAAI meeting (being organized by R. Plant).

I hope that in 5 years another collection of articles is published, covering AAAI V&V workshops for the years 1993–1998.

Putting together this collection of articles required the efforts of many. However, particular thanks go to: all the referees; Ken Ford for his efforts; and Ron Yager for allowing us the space of two issues of the journal for this collection.

References

1. K. Bellman, Ed., *Proceedings of the 1989 AAAI/IJCAI Workshop on Knowledge-Based System Verification, Validation and Testing,* Detroit, MI, 1989.
2. C. Culbert, Ed., *Proceedings of the 1990 AAAI Workshop on Knowledge-Based System Verification, Validation and Testing,* Boston, MA, 1990.
3. D. O'Leary, Ed., *Proceedings of the 1991 AAAI Workshop on Knowledge-Based System Verification, Validation and Testing,* Anaheim, CA, 1991.
4. L. Miller, Ed., *Proceedings of the 1992 AAAI/IJCAI Workshop on Knowledge-Based System Verification, Validation and Testing,* San Jose, CA, 1992.
5. A. Preece, Ed., *Proceedings of the 1993 AAAI Workshop on Knowledge-Based System Verification, Validation and Testing,* Anaheim, CA, 1991.
6. C. Culbert, Ed., *Expert Syst. Appl.,* **1**(3) (1990).
7. M. Ayel and J.-P. Laurent, Eds., *Verification, Validation and Test of Knowledge-Based Systems,* Wiley, Chichester, 1991.
8. A. Preece and C. Suen, Eds., *Int. J. Expert Syst.: Res. Appl.* **6**(2), (3) (1993).
9. U. Gupta, Ed., *Validating and Verifying Knowledge-Based Systems,* IEEE Press, Los Alamitos, CA, 1990.
10. R. O'Keefe, O. Balci, and E. Smith, "Validating expert system performance," *IEEE Expert,* Winter, 81–87 (1987).
11. R. O'Keefe and D. O'Leary, "Expert system verification and validation," *Artif. Intell. Rev.,* **7,** 3–42 (1993).
12. C. Chang, J. Combs, and R. Stachowitz, "A report on the expert systems validation associate," *Expert Syst. Appl.,* **1,** 217–230 (1990).