PROTÉGÉ-II: A Suite of Tools for Development of Intelligent Systems from Reusable Components

Mark A. Musen, M.D., Ph.D., Henrik Eriksson, Ph.D., John H. Gennari, Ph.D., Samson W. Tu, M.S., and Angel R. Puerta, Ph.D.

Section on Medical Informatics, Stanford University School of Medicine, Stanford, CA 94305-5479 USA

PROTÉGÉ-II comprises a set of tools that developers use to build intelligent software systems. One tool, called MAÎTRE, allows developers to browse through and edit domain models (ontologies). Another tool, called DASH, takes as its input a domain ontology, and generatres as its output a graphical knowledge-acquisition tool that application experts can use to enter the detailed content knowledge necessary to define new applications. Other tools in the PROTÉGÉ-II collection allow developers to define the problem-solving methods that automate application tasks by making selections from a library of reusable problem-solving-method building blocks. PROTÉGÉ-II has been used to create knowledge-acquisition tools and the associated knowledge bases for a number of applications, including the decision-support components of the T-HELPER computer-based patient record system.

There are multiple dimensions of knowledge sharing and reuse [1]. Although much work to date has concentrated on development of standardized syntaxes for knowledge representation, the development of large-scale intelligent systems requires attention not only to representation of propositions about the world being modeled, but also to the control knowledge that allows complex problem-solving to take place.

For the past several years, our research group has been building a development environment, known as PROTÉGÉ-II, that permits developers to reuse knowledge in multiple ways [1]. Unlike PROTÉGÉ-I [2], which assumed that the performance system that interprets a knowledge base is based on a fixed problem solver (namely, a domain-independent version of the ONCOCIN expert system), our new architecture produces custom-tailored problem solvers based on the method configurations defined by PROTÉGÉ-II users. In particular, PROTÉGÉ-II supports libraries of reusable *problem-solving methods* that define, in domain-independent terms, the manner in which domain knowledge may be used to solve application tasks. Developers may select from a library of well-known problem-solving methods such as *episodic skeletal-plan refinement* [3], and configure those methods to create running application programs.

To build a knowledge-based system, a system builder uses the MAÎTRE tool to develop the domain ontology and either selects from the library the method that most closely fits the requirements of the application task, or constructs a composite method from the available methods and mechanisms. The system builder then uses the DASH tool to generate a graphical knowledge-acquisition tool based on the ontology. Domain specialists can then use this custom-tailored knowledge-acquisition tool to create a knowledge base.

PROTÉGÉ-II offers a new approach to the construction of intelligent systems—namely, the use of reusable building blocks that are at various levels of abstraction. PROTÉGÉ-II has been used to create a number of clinical applications, including the decision-support modules of the T-HELPER computer-based patient record system [4]. PROTÉGÉ-II demonstrates that the process of engineering and maintaining complex software systems can be enhanced through the use of libraries of appropriate components. The automated generation of domain-specific knowledge-acquisition tools should allow health-care workers to have a direct role in the creation and maintenance of a variety of clinical software applications.

Acknowledgments

This work has been supported, in part, by NLM grants LM05157 and LM05305, and by gifts from Digital Equipment Corporation and from the Computer-Based Assessment Project of the American Board of Family Practice. Dr. Musen is recipient of NSF Young Investigator Award IRI-9257578.

References

[1]. M.A. Musen. Dimensions of knowledge sharing and reuse. *Computers and Biomedical Research* **25**: 435–467, 1992.

[2]. M.A. Musen. Automated support for building and extending expert models. *Machine Learning* **4**:349–377, 1989.

[3]. S.W. Tu, Y. Shahar, J. Dawes, J. Winkels, A.R. Puerta, and M.A. Musen. A problem-solving model for episodic skeletal-plan refinement. *Knowledge Acquisition* **4**:197–216, 1992.

[4]. M.A. Musen, R.W. Carlson, L.M. Fagan, S.C. Deresinski, and E.H. Shortliffe. T-HELPER: Automated support for community-based clinical research. In *Proceedings of the Sixteenth Annual Symposium on Computer Applications in Medical Care*. Baltimore, MD, pp. 719–723, 1992.

0195-4210/94/$5.00