Expert System Development through the World Wide Web

Brian R. Gaines and Mildred L. G. Shaw

Knowledge Science Institute
University of Calgary
Alberta, Canada T2N 1N4
{gaines, mildred}@cpsc.ucalgary.ca

Abstract

WebGrid-II is a knowledge acquisition and inference server on the World Wide Web that uses an extended repertory grid system for knowledge acquisition, inductive inference for knowledge modeling, and an integrated expert system shell for inference. This article describes WebGrid-II, and illustrates its use in eliciting and modeling knowledge

1 Introduction

The development of knowledge-based systems involves knowledge acquisition from a diversity of sources often geographically distributed. The sources include books, papers, manuals, videos of expert performance, transcripts of protocols and interviews, and human and computer interaction with experts. Expert time is usually a scarce resource and experts are often only accessible at different sites, particularly in international projects. Knowledge acquisition methodologies and tools have developed to take account of these issues by using hypermedia systems to manage a large volume of heterogeneous data, interactive graphic interfaces to present knowledge models in a form understandable to experts, rapid prototyping systems to test the models in operation, and model comparison systems to draw attention to anomalous variations between experts.

However, existing knowledge acquisition tools are largely based on personal computers and graphic workstations, and their use in a distributed community involves moving software and often computers from site to site. The process of building and testing knowledge models across a distributed community would be greatly expedited if wide-area networks could be used to coordinate the activities at different sites. The initial objective of the work reported in this paper hasbeen to use the World Wide Web to support distributed knowledge acquisition by porting existing knowledge acquisition tools to operate through the web. A further objective has been to use the web on a continuing basis to support distributed knowledge management in which the acquisition, representation and application of knowledge become an integral part of the activities of an organization.

Repertory grids based on personal construct psychology (Kelly, 1955) have been used for knowledge acquisition since the early years of knowledge-based system development (Shaw and Gaines, 1983; Boose 1984) and have been refined over the years to support increasingly complex knowledge structures (Boose Bradshaw, Koszarek and Shema, 1993; Ford, Bradshaw Adams-Webber and Agnew, 1993; Gaines and Shaw, 1993a). Since grid elicitation tools are used directly by experts it would be very useful to have them accessible through any personal computer or workstation with access to the web. This is feasible because the primary data input format is through rating scales, and this can be done effectively by using popup menus which are available through the HTML graphic user interface, and is the basis of WebGrid (Shaw and Gaines, 1995). WebGrid-II extends WebGrid with new data types and analysis techniques to support expert system development.

The next section gives an overview of the web architecture and protocols, and following sections illustrate how knowledge acquisition systems have been ported to the web to operate in a distributed client-server environment.

2 WebGrid-II: Repertory Grid Client-Server Implementation

WebGrid was developed in 1994 to offer repertory grid elicitation and analysis through the World Wide WebThe server is based on our RepGrid/KSS0 program(CPCS, 1993) which operates on an Apple Macintosh and offers a wide range of elicitation and analysis capabilities for single and multiple repertory grids and has been widely used in expert system development (Gaines and Shaw, 1993b; Shaw and Gaines, 1993; Shaw, Gaines and Linster, 1994). Offering access to RepGrid through the web made much of its functionality available to users throughout the world through standard web browsers operating on any computer.

WebGrid was intended primarily to introduce repertory grid techniques to new users and focused on basic capabilities. However, it rapidly came into widespread use for major research projects and interest grew in having a full range of knowledge acquisition tools available on the web. WebGrid-II extends WebGrid to provide integration with multimedia, additional data types in the grid, entailment analysis to induce rules from griddata, and an expert system inference engine to allow the rules to be tested on new cases as part of the elicitation process.

2.1 Elicitation

Figure 1 shows the initial screen of WebGrid-II. The HTML form requests the usual data required to initiate grid elicitation: user name; domain and context; terms for elements and constructs; default rating scale; data types allowed; and a list of initial elements. It also allows the subsequent screens to be customized with an HTML specification of a header and trailer—this capability to include links to multimedia web data is also used to allow annotation, text and pictures, to be attached to elements.

The problem domain is that of "selecting among graduate program applicants" andthe intention of the user is to develop a set of rules defining the criteria for selection. The system will incorporate the text typed in by the user into the later dialog, and the user has stated that the elements in the grid are to be called "students" and the constructs "qualities." The user has set the types of constructs to be "Ratings +Names +Categories +Numbers"—indicating that the user wants to use more advanced facilities for knowledge modeling. She has left the default rating scale to be 1-5 and has entered the initials of 4 applicants typifying the range of selection decisions normally made.

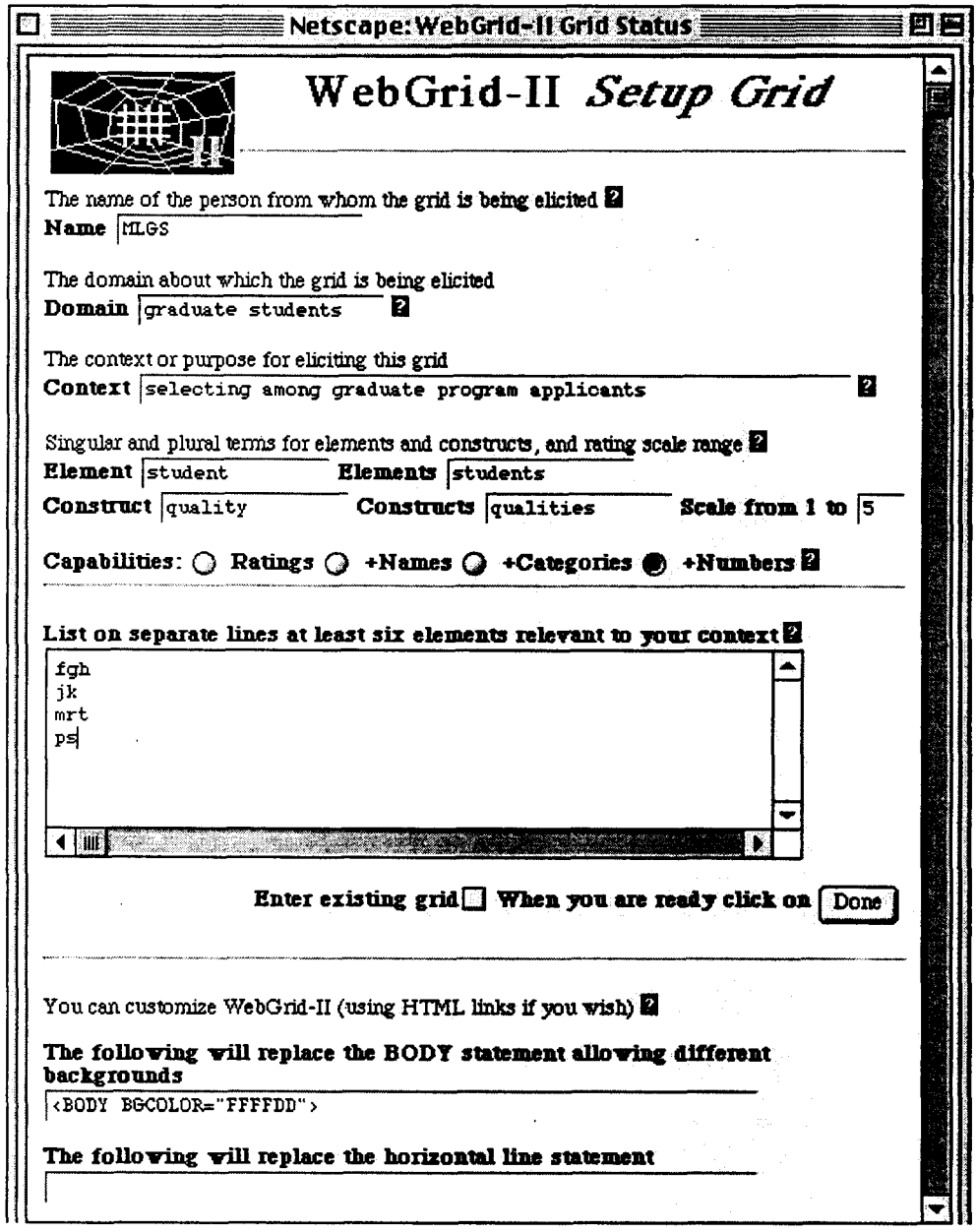

Figure 1 Repertory grid elicitation initial screen

When the user clicks on the "Done" button at the bottom, the browser transmits the data entered to the remote server which passes it through its common gateway interface to a specialist knowledge acquisition auxiliary server. The server processes the data and generates an HTML document that it returns to the browser resulting in the screen shown in Figure 2 eliciting a construct from a triad of elements.

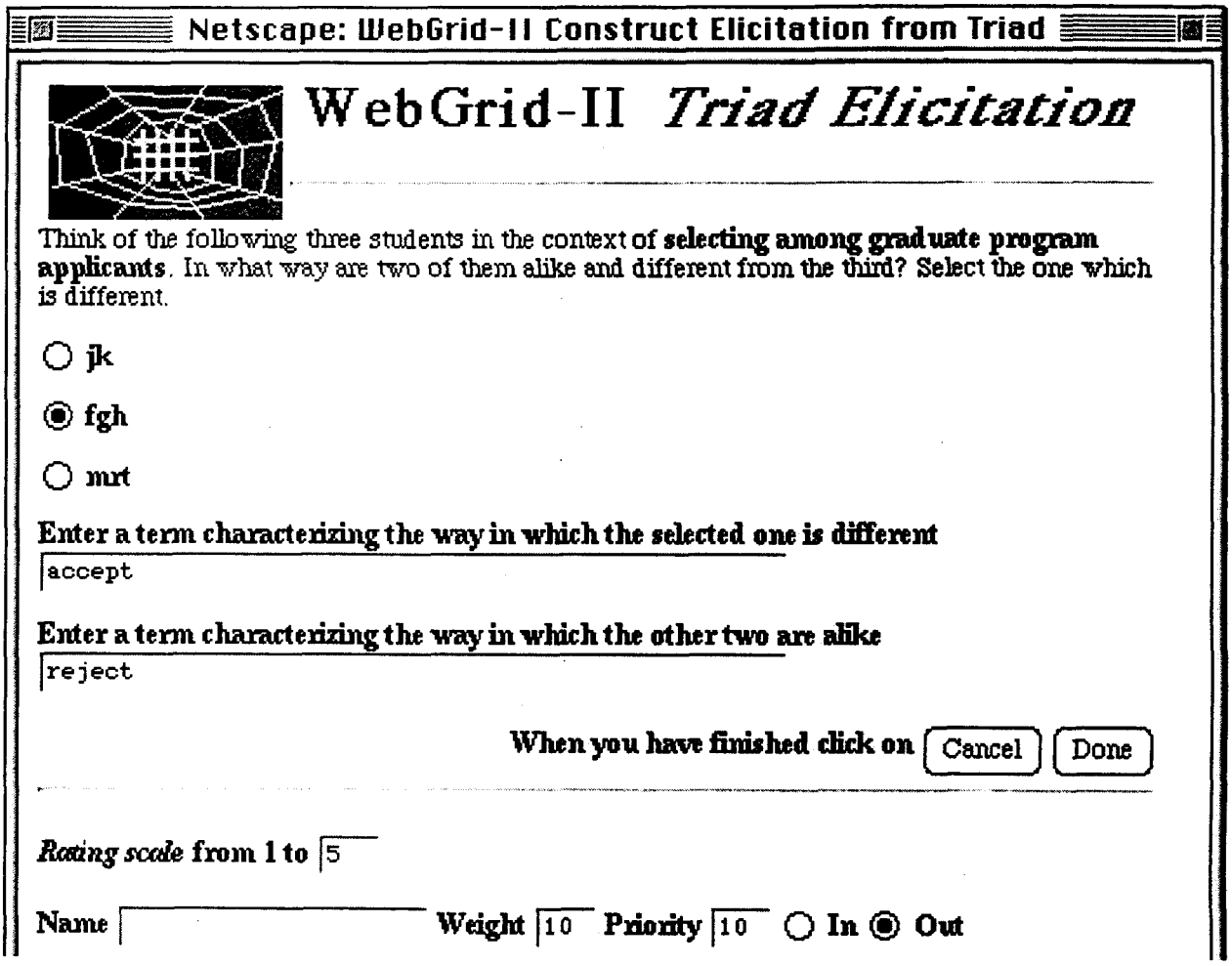

Figure 2 Construct elicitation from a triad

The user clicks on a radio button to select an element which she construes as different from the other two and enters terms characterizing the construct. The lower part of the screen is generated because advanced features wereIt allows the user to vary the rating scale for an individual construct, to name the construct, to give it a weight in clustering, a priority in asking questions in an expert system application, and to specify that the construct is an input for decision making purposes or an output to be anticipated. Since the construct entered is the main decision to be made, to accept or reject an applicant, and the user wants the system to anticipate the correct decision from the way the applicant is construed she has made it an output.

When the user clicks on "Done" the server generates the screen shown on the left of Figure 3 which places a popup menu rating scale alongside each element enabling the user to rate each one along the new construct as shown on the right. She is also able to change the terms used if they seem inappropriate in the context of all the elements, and to change the ratings of the already entered elements if appropriate.

Clicking on the "Done" button in Figure 3 sends the ratings back to the server which generates the status screen shown in Figure 4.

This shows the elements and constructs entered, allowing them to be selected for deletion, editing and so on. It also offers various suggestions as how to continue the elicitation based on the data entered so far, facilities for analysis, saving the grid, and so on.

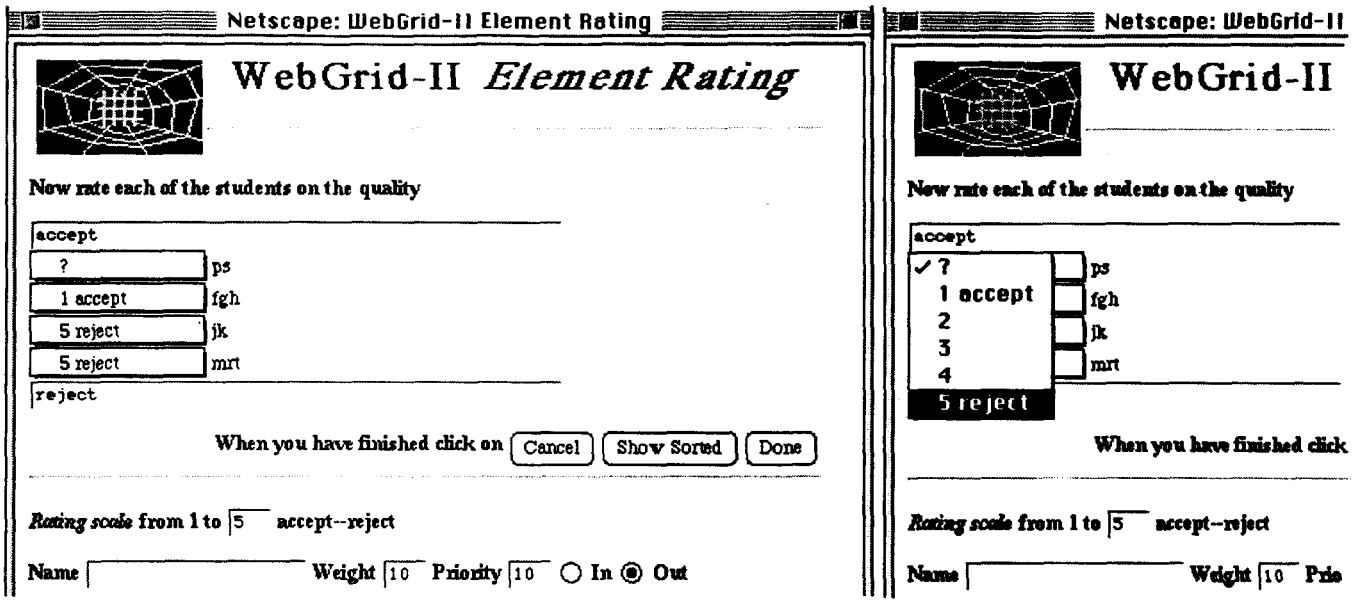

Figure 3 Rating elements on constructs

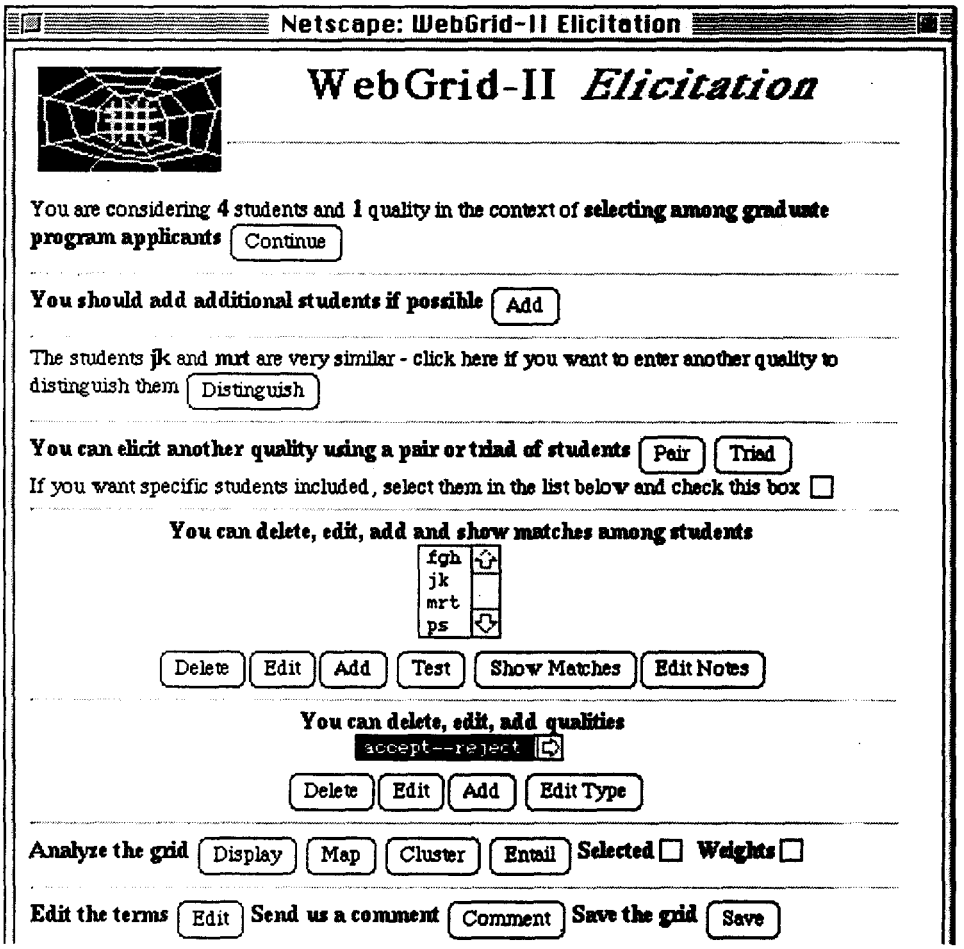

Figure 4 Status screen

2.2 New Data Types in Grids

The user decides to develop another two constructs from triads of elements and enters "highgpa—inadequate gpa" and "good references—inadequate references" since she knows that these are critical to acceptance. The applicant's gpa is a numeric rating between 0.0 and 4.0 and hence cannot be entered on a normal grid rating scale. It requires one of the new data types that we have used to extend repertory grids for knowledge acquisition.

Figure 5 shows the gpa categories being added as a new construct. WebGrid-II makes provision not only for rating scales but also for a set of categories, ordered or unordered, and for numerical rating scales of the type integer, float or date with prescribed ranges. The user has entered "inadequate", "adequate" and "high" as possible categories for gpa, has entered "gpa" as the name of the construct, and has noted that the value is a floating point number. The only difference the name entry makes is to shorten the process sinceWebGrid will generate the names "gpa = inadequate", etc, rather than have the user enter "gpa is inadequate", etc, as category names. The elements are rated on a purely categorical construct through popup menus and on a numeric construct through a numeric entry box.

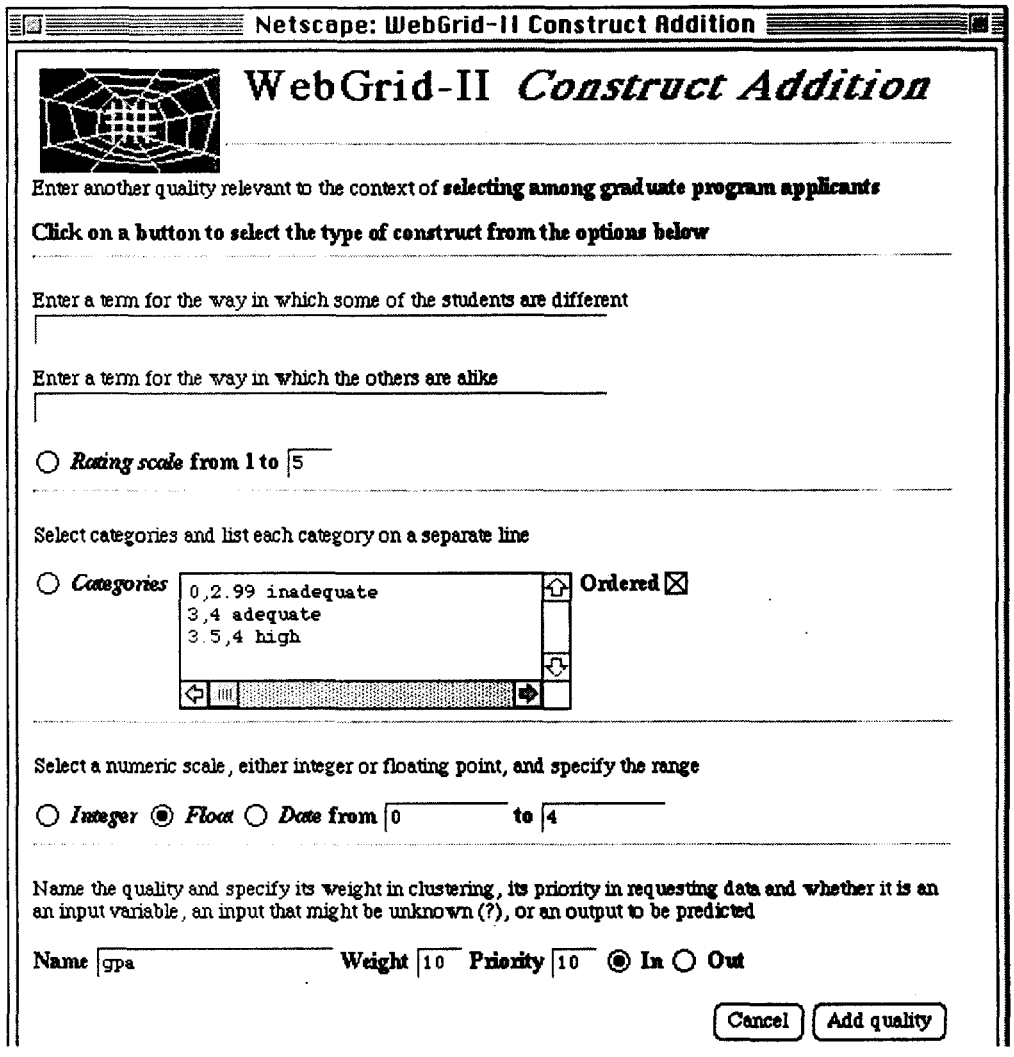

Figure 5 Adding a numeric construct with categories

2.3 Entailment Analysis through Rule Induction

The status screen now shows three constructs and the user clicks on the "Entail" button near the bottom of Figure 4 hoping the data is sufficient to generate a model of the way in which the constructs entered entail the "accept—reject" decision. The system develops a model using the Induct algorithm (Gaines, 1989; Gaines, 1991) for inductive derivation of entailments and generates the screen shown in Figure 6.

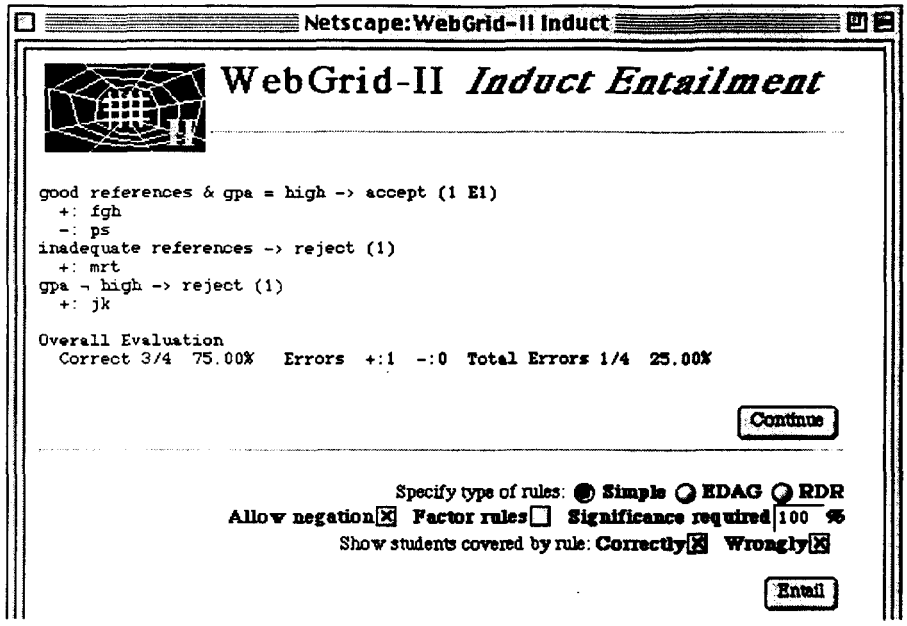

Figure 6 Rules entailed by the grid after 4 elements and 3 constructs

She is disappointed to receive the entailments shown which look sensible but work for only 3 cases in 4. The output consists of a set of rules defining the anticipatory entailments and each rule is followed by the positive elements where the correct decision is anticipated and the negative ones where that anticipated is incorrect. On thinking about it she realizes that the applicant "ps" has a high gpa and good references but was rejected because there was no-one willing to supervise him. She returns to the status screen, adds a construct "willing supervisor—no supervisor" and clicks on "Entail" again. This time the rules entailed by the grid returned as shown in Figure 7 correctly anticipate the decisions made for all 4 cases.

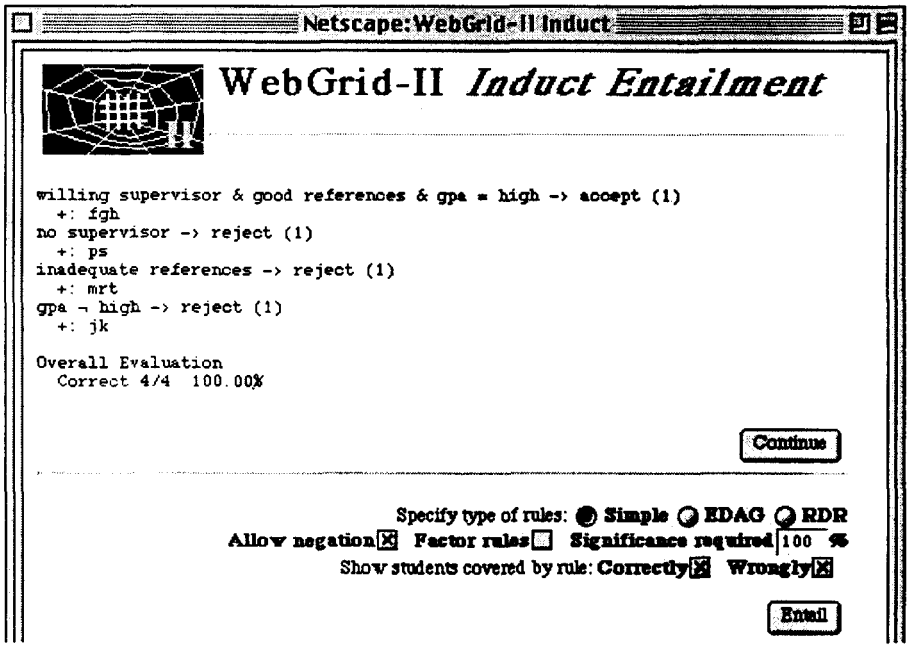

Figure 7 Rules entailed by the grid after 4 elements and 4 constructs

The user knows that Induct is also capable of producing default rules with exceptions(Gaines, 1994; Gaines, 1995) that are often more natural in their correspondence to psychological models, and runs Entail again with "EDAG" selected in Figure 7. The system generates the screen of Figure 8 which has the default rule to reject all applicants unless they have a high gpa, good references and a willing supervisor—a good initial characterization of the selection process.

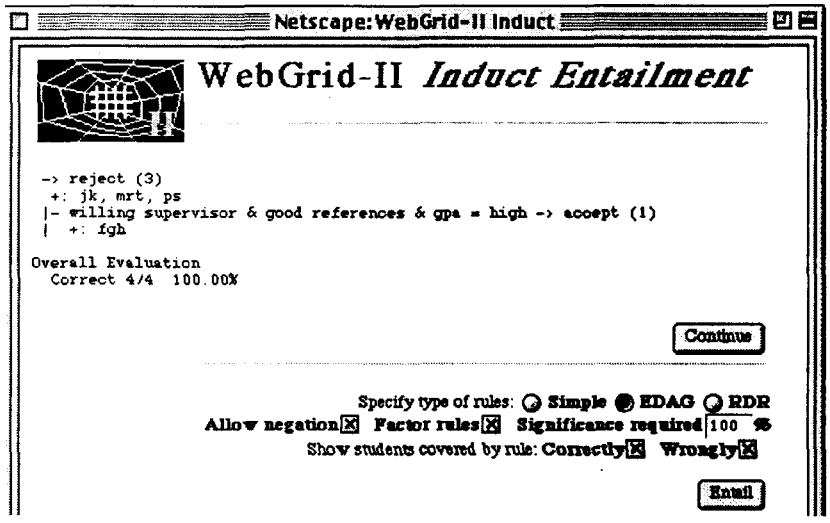

Figure 8 with exceptions entailed by the grid after 4 elements and 4 constructs

2.4 Inference through an Embedded Expert System Shell

The user decides to test the entailed rules on further cases and returns to the status screen of Figure4 which now shows four constructs. She clicks on the "Test" button under the list of elements and the system returns the screen shown in Figure 9 which allows her to specify a test case and click on "Infer" to see what decision is anticipated.

The user does this for several similar cases and the conclusion returned is correct but when she does so for the applicant "Iemp" as shown in Figure 10 the conclusion is "reject" when she knows the applicant has been accepted.

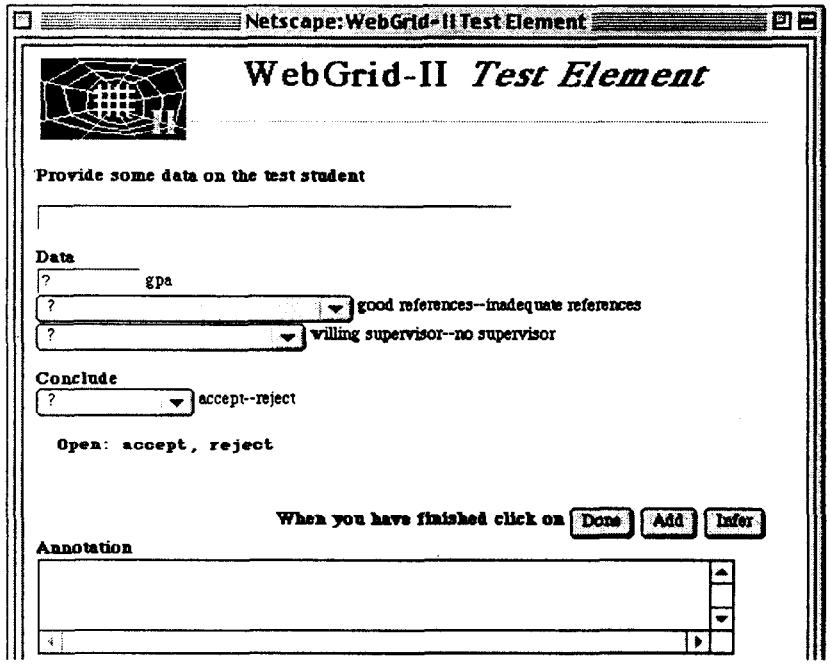

Figure 9 Entering a test case into the entailed expert system

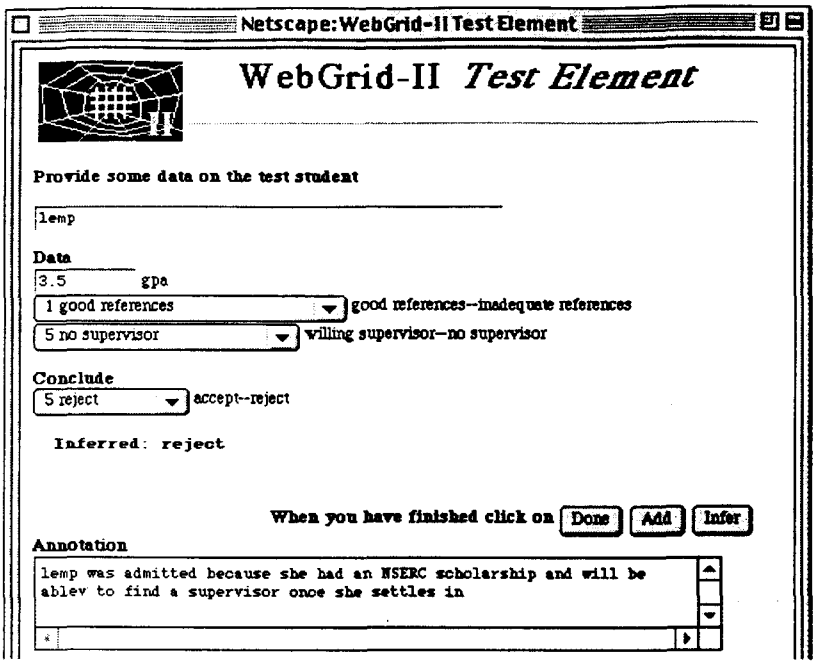

Figure 10 Conclusion anticipating the decision for a test case

The user realizes that applicant "lemp" was admitted because she had a major scholarship, corrects the conclusion to "admit", and clicks on "Add" to add this applicant as a new element in the gridThe system returns to the status screen of Figure 4 with "lemp added and the user clicks on "Add" under the construct list to enter the categorical "scholarship" is "major" or "none" She then clicks on "Entail" and the system generates the entailed rules with exceptions as shown in Figure 11 which now account for 'lemp" also. The default rule is to reject all applicants unless they have a major scholarship or a willing supervisor, high gpa and good references when they should be accepted unless they have no supervisor and no scholarship—a more refined characterization of the selection process. However, she also realizes there are applicants with major scholarships who have not been admitted because they have a poor gpa or inadequate references.

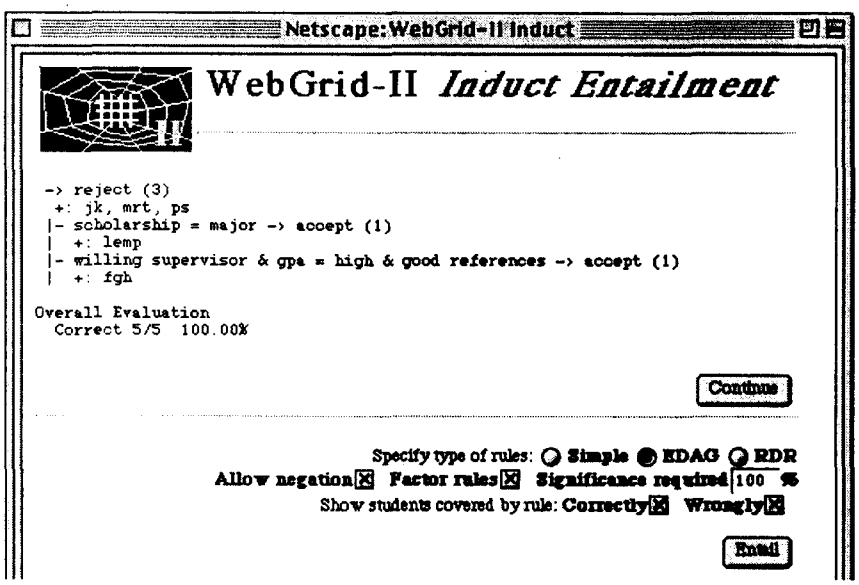

Figure 11 Rules with exceptions entailed by the grid after 5 elements and 5 constructs

The user enters some relevant examples, returns to test case entry again and this time come across a applicant who was admitted to a new program, an industry-based software engineering degree with the requirement that the applicant have substantial industry background. She enters some relevant cases and constructs leading to the entailed decision rule shown in Figure 12, that applicants should be rejected by default unless they have good references and in addition either have a highgpa but not no scholarship and no supervisor, or are a software engineering application with industry experience and an adequate gpa.

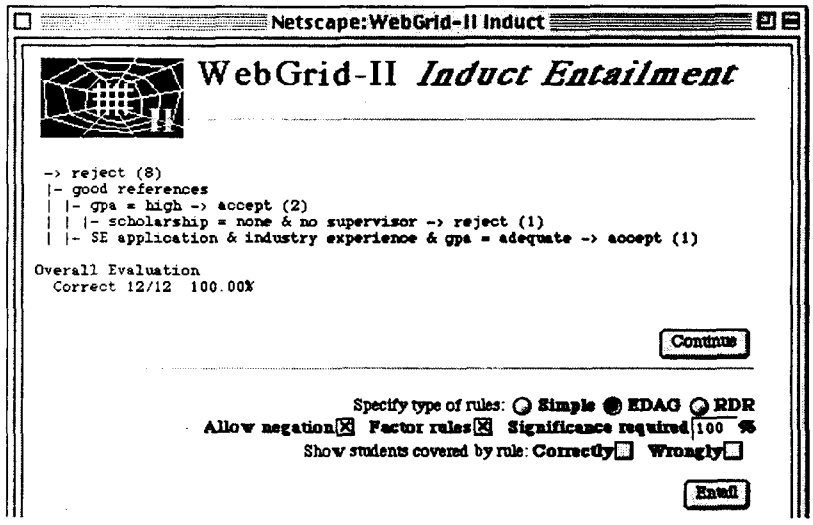

Figure 12 Rules with exceptions entailed by the grid after 12 elements and 7 constructs

Note how the decision rule entailed by the grid is becoming more refined as it encompasses more cases and also how the Induct algorithm generating it changes its structure to be as comprehensible as possible as it grows in complexity. This attempt to generate an anticipatory structure which is natural for people is an important feature of EDAG generation (Gaines, 1995). It is also noteworthy that a fairly small grid of only 12 elements and 7 constructs can be used to generate such a complex decision rule. This is because the repertory grid elicitation technique with PEGASUS-style (Shaw, 1980) feedback leads to the elicitation of typical elements and relevant constructs, allowing knowledge to be elicited highly efficiently compared with rule induction from a database where thousands of examples might be required to derive the same knowledge (Gaines, 1989).

3 Conclusions

The technologies, tools and protocols developed for the World Wide Web make it possible to provide knowledge acquisition, representation and inference systems internationally to anyone with access to the Internet. Such knowledge-based technologies integrate naturally with other applications and hypermedia systems to provide a distributed knowledge medium capable of supporting the knowledge processes of professional communities.

This paper has illustrated what is possible through examples of the development of knowledge modeling tools operating through the web to provide knowledge acquisition, representation and inference through repertory grids. These systems were developedpxrimarily for locally managed projects, and their availability has only been advertised through specialist list servers for the knowledge acquisition and personal construct psychology communities. However, in the period June 1995 to June 1997 WebGrid was accessed from over 5,000 different sites in some 60 countries, delivering over 200Mbytes of grid data in some 70,000 transactions. The web is essentially an anonymous medium and we do not track the activities of outside users. However, occasionally users contact us to discuss their applications and, if the local network goes down, we receive mail from serious users reporting the problem and requesting notification of when the knowledge acquisition programs will be available again. An interesting example of serendipitous use was a masters student in Holland studying operators' models of nuclear reactors who came across the grid elicitation system on the web in September 1996 used it with his experts to elicit their conceptual models, and included these in an additional chapter in his thesis for examination in October 1996—the web certainly accelerates research processes.

In conclusion, we note that the results presented in this paper demonstrate that the World Wide Web client server model operating through the Internet can effectively support the porting of existing knowledge acquisition, modeling and representation tools to operate in a distributed, open architecture environment.This opens up new possibilities for collaborative research in which the methodologies and tools developed by one group are made

available on the network for use by other research groups world-wide. It enables past studies of the integration of hypermedia, knowledge acquisition and inference tools within a local environment (Gaines and Linster, 1990; Gaines, Rappaport and Shaw, 1992) to be generalized to integration on a distributed basis across wide area networks. A major advantage of this approach is that tools can be operated at the site of the originators responsible for them, and their maintenance, upgrading and monitoring can be done very much more effectively than when the software is distributed to other sites.

We envision a collaborative research program internationally in which the first objective is to make tools available across the Internet in interoperable forms, and the second objective is to develop some integrated knowledge management systems that use these tools in a variety of configurations to provide a new generation of knowledge-based systems

Acknowledgments and URLs

Financial assistance for this work has been made available by the Natural Sciences and Engineering Research Council of Canada.

WebGrid can be accessed at
http://tiger.cpsc.ucalgary.ca/WebGrid/

Related papers on WebGrid and World Wide Web can be accessed through **http://ksi.cpsc.ucalgary.ca/articles/**

References

Boose, J.H. (1984). Personal construct theory and the transfer of human expertise. **Proceedings AAAI-84.** pp.27-33. California, American Association for Artificial Intelligence.

Boose, J. H., Bradshaw, J. M., Koszarek J.L. and Shema, D. B. (1993). Knowledge acquisition techniques for group decisions support. **Knowledge Acquisition** 5(4) 405-447.

CPCS (1993). **RepGrid 2 Manual** Centre for Person Computer Studies, 3019Underhill Drive NW, Calgary, AB, Canada T2N 4E4.

Ford, K.M., Bradshaw, J.M., Adams-Webber, J.R. and Agnew, N.M. (1993). Knowledge acquisition as a constructive modeling activity. **International Journal of Intelligent Systems** 8(1) 9-32.

Gaines, B.R. (1989). An ounce of knowledge is worth a ton of data: quantitative studies of the trade-off between expertise and data based on statistically well-founded empirical induction. **Proceedings of the Sixth International Workshop on Machine Learning.** pp.156-159. San Mateo, California, Morgan Kaufmann.

Gaines, B.R. (1991). Refining induction into knowledge. **Proceedings of the AAAI Workshop on Knowledge Discovery in Databases.** pp.1-10. Menlo Park, California, AAAI.

Gaines, B.R. (1994). Exception DAGs as knowledge structures. **AAAI-94 Workshop: KDD'94 Knowledge Discovery in Databases.** Menlo Park, California, AAAI.

Gaines, B.R. (1995). Transforming rules and trees into comprehensible knowledge structures. Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P. and Uthurusamy, R., Ed. **Knowledge Discovery in Databases II** pp.205-226. Cambridge, Massachusetts, AAAI/MIT Press.

Gaines, B.R. and Linster, M. (1990). Integrating a knowledge acquisition tool, an expert system shell and a hypermedia system. **International Journal of Expert Systems Research and Applications3**(2) 105-129.

Gaines, B.R., Rappaport, A. and Shaw, M.L.G. (1992). Combining paradigms in knowledge engineering.**Data and Knowledge Engineering 9** 1-18.

Gaines, B.R. and Shaw, M.L.G. (1993a). Basing knowledge acquisition tools in personal construct psychology. **Knowledge Engineering Review 8**(1) 49-85.

Gaines, B.R. and Shaw, M.L.G. (1993b). Eliciting knowledge and transferring it effectively to a knowledge-based systems. **IEEE Transactions on Knowledge and Data Engineering5**(1) 4-14.

Kelly, G.A. (1955). **The Psychology of Personal Constructs.** New York, Norton.

Shaw, M.L.G. (1980). **On Becoming A Personal Scientist: Interactive Computer Elicitation of Personal Models Of The World** London, Academic Press.

Shaw, M.L.G. and Gaines, B.R. (1983). A computer aid to knowledge engineering. **Proceedings of British Computer Society Conference on Expert Systems** pp.263-271. Cambridge, British Computer Society.

Shaw, M.L.G. and Gaines,B.R. (1993). Group knowledge elicitation over networks. Bramer, M.A. and Macintosh, A.L., Ed. **Research and Development in Expert Systems X. Proceedings of British Computer Society Expert Systems Conference** pp.43-62. Cambridge, UK, Cambridge University Press.

Shaw, M.L.G. and Gaines, B.R. (1995). Comparing constructions through the web. Schnase, J.L. and Cunnius, E.L., Ed. **Proceedings of CSCL95: Computer Support for Collaborative Learning** pp.300-307. Mahwah, New Jersey, Lawrence Erlbaum.

Shaw, M.L.G., Gaines, B.R. and Linster, M. (1994). Supporting the knowledge engineering life cycle. Bramer, M.A. and Macintosh, A.L., Ed. **Research and Development in Expert Systems XI** pp.73-86. Oxford, SGES Publications.