AN APPLICATION OF THE PROSPECTOR SYSTEM TO DOE'S NATIONAL URANIUM RESOURCE EVALUATION*

John Gaschnig
Artificial Intelligence Center
SRI International
Menlo Park, CA 94025

Abstract

A practical criterion for the success of a knowledge-based problem-solving system is its usefulness as a tool to those working in its specialized domain of expertise. Here we describe an application of the Prospector consultation system to the task of estimating the favorability of several test regions for *occurrence of uranium deposits*. This pilot study was conducted for the National Uranium Resource Estimate program of the U.S. Department of Energy. For credibility, the study was preceded by a performance evaluation of the relevant portion of Prospector's knowledge base, which showed that Prospector's conclusions agreed very closely with those of the model designer over a broad range of conditions and levels of detail. We comment on characteristics of the Prospector system that are relevant to the issue of inducing geologists to use the system.

1. Introduction

This paper describes an evaluation and an application of a knowledge-based system, the Prospector consultant for mineral exploration. Prospector is a rule-based judgmental reasoning system that evaluates the mineral potential of a site or region with respect to inference network models of specific classes of ore deposits. Here we describe one such model, for a class of "Western states" sandstone uranium deposits, and report the results of extensive quantitative tests measuring how faithfully it captures the reasoning of its designer across a set of specific sites (used as case studies in fine-tuning the model), and with respect to the detailed subconclusions of the model as well as its overall conclusions. Having so validated the performance of this model (called RWSSU), we then describe a pilot study performed in conjunction with the National Uranium Resource Evaluation (NURE) program of the U.S. Department of Energy. The pilot study applied the RwSSU model to evaluate and compare five target regions, using input data provided by DOE and USGS geologists (using the medium of a model-specific questionnaire generated by Prospector). The results of the experiment not only rank the test regions, but also measure the sensitivity of the conclusions to more certain or less certain variations in the input data. One interesting facet of this study is that several geologists provided input data independently about each test region. Since input data about each region varies among the responding geologists, so do the conclusions; we demonstrate how Prospector is used to identify and resolve the disagreements about input data that are most significantly responsible for differences in the resulting overall conclusions. This paper is a condensation of portions of a larger report [4].

2. Validation of the Model

The practical usefulness of an expert system is limited if those working in its domain of expertise do not or will not use it. Before they will accept and use the system as a working tool, such people (we shall call them the "domain users") usually expect some evidence that the performance of the system is adequate for their needs (e.g., see [8]). Accordingly, considerable effort has been devoted to evaluating the performance of the Prospector system and of its various models [2, 3]. In the present case, we first needed to validate the performance of the uranium model to be used in the pilot study for the U.S. Department of Energy.

The methodology used to evaluate Prospector's performance is discussed in detail elsewhere [2, 3]. For brevity, here we outline a few relevant factors. The Prospector knowledge base contains a distinct inference network model for each of a number of different classes of ore deposits, and a separate performance evaluation is performed for each model. Here we are concerned with one such model, called the regional-scale "Western states" sandstone uranium model (RWSSU), designed by Mr. Ruffin Rackley. Since there exist no objective *quantitative measures of the performance* of human geologists against which to compare that of Prospector, we instead use a relative comparison of the conclusions of a Prospector model against those of the expert geologist who designed it. To do so, first a number of test regions are chosen, some being exemplars of the model and others having a poor or less good match against the model. For each such case, a questionnaire is completed detailing the observable characteristics that the model requests as inputs for its deliberation. Prospector evaluates each such data set and derives its conclusion for that test case, which is expressed on a scale from -5 to 5. As a basis of comparison, we also independently elicit the model designer's conclusion about each test case, based on the same input data, and expressed on the same -5 to 5 scale. Then we compare Prospector's predictions against the target values provided by the model designer.

Table 1 compares the top-level conclusions of Prospector (using the RWSSU model) against those of the model designer for eight test regions.

---

*This research was supported by the U. S. Geological Survey under USGS Contract No. 14-08-0001-17227. Any opinions, findings, and conclusion or recommendations expressed in this report are those of the author and do not necessarily reflect the views of the U.S. Geological Survey.

Table 1. Comparison of RWSSU model with Designer for Eight Test Cases

Test Region	Designer's Target	Prospector Score	Difference
Black Hills	3.50	4.33	-0.83
Crooks Gap	4.70	4.26	0.44
Gas Hills	4.90	4.37	0.53
Shirley Basin	4.95	4.13	0.82
Ambrosia Lake	5.00	4.39	0.61
Powder River	4.40	4.40	0.00
Fox Hills	1.50	2.17	-0.67
Oil Mountain	1.70	3.32	-1.62
		Average:	0.69

Table 1 indicates that the average difference between the Prospector score and the corresponding target value for these eight cases is 0.69, which is 6.9% of the -5 to 5 scale.

Besides the overall conclusions reported above, quite detailed information about Prospector's conclusions was collected for each test case. One feature of the Prospector system is the ability to explain its conclusions at any desired level of detail. In its normal interactive mode, the user can interrogate Prospector's conclusions by indicating which conclusions or subconclusions he wishes to see more information about. The same sort of information is presented in Table 2 (using the Gas Hills region as an example), in the form of Prospector's overall evaluation, the major conclusions on which the overall evaluation is based, and the subconclusions that support each major conclusion. For brevity, each section of the RWSSU model represented in Table 2 is identified by its symbolic name, which is indented to show its place in the hierarchy of the model. For comparison, we first elicited from the model designer his target values for each section of the model listed in Table 2; these values are included in Table 2.

Table 2. Detailed Comparison of RWSSU Model with Designer for Gas Hills

	Designer's Target	Prospector Score	Difference
RWSSU	4.90	4.37	.53
FTRC	4.80	4.64	.16
TECTON	4.50	4.50	.00
AHR	5.00	4.95	.05
FAVHOST	4.80	5.00	-.20
SEDTECT	4.80	4.88	-.08
FAVSED	4.90	4.68	.22
FLUVSED	4.90	4.68	.22
MARINESED	-3.50	-2.07	-1.43
AEOLSED	-2.50	-2.10	-.40
FMA	4.95	4.41	.54
RBZONE	5.00	4.60	.40
AIZONE	4.00	4.77	-.77
MINZONE	5.00	5.00	.00

Average difference = 0.36
(Average of absolute values)

The data in Table 2 indicate that Prospector not only reaches essentially the same numerical conclusions as its designer, but does so for similar reasons. This detailed comparison was repeated for each of the eight cases, resulting in 112 distinct comparisons between Prospector's prediction and designer's target value (i.e., 8 test regions times 14 sections of the model). The average difference between Prospector's score and designer's target value over these 112 cases was 0.70, or 7.0% of our standard 10-point scale.[**]

Gaschnig [4] also reports sensitivity analysis experiments showing the models to be rather stable in their conclusions: for the RWSSU model, a 10% perturbation in the input certainties caused only a 1.2% change in the output certainties.

3. Results of the NURE Pilot Study

Having established the credibility of the RWSSU model by the test results just discussed, we then undertook an evaluation of five test regions selected by the Department of Energy. For this purpose USGS and DOE geologists completed questionnaires for this model. As a sensitivity test, several geologists independently completed questionnaires for each test region. For comparison, the model designer, R. Rackley, also completed questionnaires for the five test regions. The overall results are reported in Table 3, in which the abbreviations M.H., P.B., Mo., N.G., and W.R. denote the names of the test regions, namely Monument Hill, Pumpkin Buttes, Moorcroft, Northwest Gillette, and White River, respectively.

Table 3. Overall Conclusions for Five Test Regions

Geologist	A	B	C	D	USGS team	Rackley data	Range
M.H.	4.17	3.32	3.97			4.40	1.08
P.B.	4.20	3.30	4.19			4.40	1.10
Mo.			3.92	3.88		4.00	0.12
N.G.			3.64	0.10		3.42	3.54
W.R.					0.13	0.01	0.12

The results in Table 3 indicate that the Monument Hill, Pumpkin Buttes, and Moorcroft regions are very favorable, and about equally favorable, for occurrence of "Western States" sandstone uranium deposits. Northwest Gillette is scored as moderately favorable, whereas White River is neutral (balanced positive and negative indicators).

Note that each respondent has had different exposure to the target regions, in terms of both first-hand, on-site experience and familiarity with field data reported in the literature. These differences in experience are reflected in their answers on the questionnaires. Since different inputs yield different conclusions, one would expect a spread in the certainties about each region, reflecting the differences in input data provided by the various geologists. Inspection of Table 3 reveals, however, that the scores

---

[**] [4] also reports analogous results for two other uranium models called WSSU and ECSU, involving 112 and 171 distinct comparisons, respectively, resulting in average differences of 7.2% and 7.8%, respectively. Aggregating the tests of the three models, 395 distinct comparisons resulted in an average difference of 7.3%.

derived from different geologists' input data about the same region agree rather closely for each region except Northwest Gillette (see the column labeled "Range"). These generally close agreements reflect the capability of Prospector models to synthesize many diverse factors, mechanically ascertaining general commonalities without being unduly distracted by occasional disparities.

In cases such as Northwest Gillette in which a large difference in conclusions occurs, it is easy to trace the source of the disagreement by comparing the individual conclusions for different sections of the model (representing different geological subconclusions), as in Table 4.

Table 4. Comparison of Detailed Conclusions About Northwest Gillette

	Geologist C	D	Rackley data	Avg.
RWSSU	.10	3.66	3.42	3.56
FTRC	4.67	3.80	4.63	4.37
TECTON	4.90	4.50	4.50	4.63
AHR	4.95	1.03	4.94	3.64
FAVHOST	5.00	5.00	5.00	5.00
SEDTECT	4.98	4.33	4.78	4.69
FAVSED	.04	3.92	4.79	2.92
FLUVSED	.04	3.92	4.79	2.92
MARINESED	-4.60	3.34	.02	-.41
AEOLSED	-4.99	-2.10	-3.23	-3.44
FMA	.27	2.45	1.33	2.18
RBZONE	4.10	4.83	4.73	4.55
AIZONE	-3.29	2.40	0.00	-0.30
MINZONE	.41	2.82	2.59	1.94

Inspection of Table 4 reveals that the conclusions agree fairly closely for the FTRC section of the model, and less closely for the FAVSED and FMA sections. Tracing the differences deeper, one sees that of the three factors on which FMA depends, there is fairly good agreement about RBZONE, but larger differences in the cases of the AIZONE and MINZONE sections. In some cases, such a detailed analysis can isolate the source of overall disagreement to a few key questions about which the respondents disagreed. These can then be resolved by the respondents without the need to be concerned with other disagreements in their questionnaire inputs that did not significantly affect the overall conclusions.

Prospector has also been applied to several other practical tasks. One evaluated several regions on the Alaskan Peninsula for uranium potential [1], as one of the bases for deciding their ultimate disposition (e.g., wilderness status versus commercial exploitation). Another application was concerned with measuring quantitatively the economic value of a geological map, resulting in statistically significant results [7].

4. Discussion

We have measured Prospector's expertise explicitly and presented a practical application to a national project, demonstrating in particular how the Prospector approach deals effectively with the variabilities and uncertainties inherent in the task of resource assessment. This work illustrates that expert systems intended for actual practical use must accommodate the special characteristics of the domain of expertise. In the case of economic geology, it is not rare for field geologists to disagree to some extent about their observations at a given site. Accordingly, the use of various sorts of sensitivity analysis is stressed in Prospector to bound the impact of such disagreements and to isolate their sources. In so doing, we provide geologists with new quantitative techniques by which to address an important issue, thus adding to the attractiveness of Prospector as a working tool. Other domains of expertise will have their own peculiarities, which must be accommodated by designers of expert systems for those domains. A more mundane, but nevertheless important, example concerns the use of a questionnaire as a medium for obtaining input data to Prospector from geologists. Most geologists have little or no experience with computers; furthermore, access to a central computer from a remote site may be problematic in practice. On the other hand, geologists seem to be quite comfortable with questionnaires. Our point is simply that issues ancillary to AI usually have to be addressed to ensure the practical success of knowledge-based AI systems.

References

1. Cox, D. P., D. E. Detra, and R. L. Detterman, "Mineral Resources of the Chignik and Sutwick Island Quadrangles, Alaska," U.S. Geological Survey Map MF-1053K, 1980 in press.
2. Duda, R.O., P.E. Hart, P. Barrett, J. Gaschnig, K. Konolige, R. Reboh, and J. Slocum, "Development of the Prospector Consultation System for Mineral Exploration," Final Report, SRI Projects 5821 and 6415, Artificial Intelligence Center, SRI International, Menlo Park, California, October 1978.
3. Gaschnig, J. G., "Preliminary Performance Analysis of the Prospector Consultant System for Mineral Exploration," Proc. Sixth International Joint Conference on Artificial Intelligence, Tokyo, August 1979.
4. Gaschnig, J. G., "Development of Uranium Exploration Models for the Prospector Consultant System," SRI Project 7856, Artificial Intelligence Center, SRI International, Menlo Park, California, March 1980.
5. National Uranium Resource Evaluation, Interim Report, U.S. Department of Energy, Report GJO-111(79), Grand Junction, Colorado, June 1979.
6. Roach, C. H., "Overview of NURE Progress Fiscal Year 1979," Preprint of Proceedings of the Uranium Industry Seminar, U. S. Department of Energy, Grand Junction, Colorado, October 16-17, 1979.
7. Shapiro, C., and W. Watson, "An Interim Report on the Value of Geologic Maps," Preliminary Draft Report, Director's Office, U.S. Geological Survey, Reston, Virginia, 1979.
8. Yu, V.L., et al., "Evaluating the Performance of a Computer-Based Consultant," Heuristic Programming Project Memo HPP-78-17, Dept. of Computer Science, Stanford University, September 1978.