Prospector II

Richard B. McCammon

U.S. Geological Survey
Reston, VA 22092

ABSTRACT

Prospector II has been designed as a frame-based expert system and the successor to PROSPECTOR to assist the geologist in performing regional mineral resource assessments. The changes that have occurred relate to the role of descriptive mineral deposit models in resource assessment and to the arrival of graphics-oriented workstations. These changes have brought about an expansion of the volunteer mode, the glossary of geological terms, and the explanation facilities in PROSPECTOR. The information volunteered by the geologist during a session with Prospector II is used to identify the deposit model that best fits the observations and to provide an explanation of what is explained or not explained by the model, and to indicate what is missing in the observations with respect to the model. The advice that is offered gives the geologist greater confidence in choosing a model and in deciding what additional data are needed in order to perform an assessment in an area.

INTRODUCTION

Few other computer programs have stirred the imagination of those who would aspire to build an expert system than PROSPECTOR, developed at SRI, International during the mid-seventies [1]. PROSPECTOR was designed to aid the exploration geologist in the search for hidden mineral deposits. During the years of its development which lasted until 1983, the PROSPECTOR program underwent extensive field testing and evaluation and to this day is regarded as one of the more successful attempts to model a specific, real-world task.

Since 1983, Prospector II, the successor to PROSPECTOR, has been under development at the U.S. Geological Survey (USGS). A number of factors have been involved to bring this about. Among the most important have been: 1) change in the uses of mineral deposit models, 2) the emergence of a new user-community, 3) the arrival of graphics-oriented workstations, 4) a need to acquire a more comprehensive, economical and adaptable format. The purpose of this paper is to indicate these changes and of the need to re-design PROSPECTOR, and the implementation and run-time capabilities of Prospector II. Some

final comments will be made about future directions.

A CHANGE OF DIRECTION

What has changed are the ideas about how mineral deposit models can be applied in regional mineral resource assessment [2]. Mineral resource assessment represents the reconnaissance stage of mineral exploration. Two main ideas relating to regional mineral resource assessment have emerged. First, mineral deposit models offer a mechanism for communicating information to others. Communication is essential when discussing the undiscovered resources of any area. Second, no model is ever final. New information is an accepted fact in mineral resource assessment. Both ideas derive from the realization that the processes that form mineral deposits are not fully understood nor are the data about mineral deposits ever complete. As a result, a variety of different models about mineral deposits have evolved. What is proving most useful now are descriptive mineral deposit models.

A descriptive mineral deposit model consists of a systematic arrangement of information describing the essential attributes (properties) of a class of mineral deposit [3]. Whether attributes are understood in the genetic context, or merely known to have a strong degree of association with the deposit type, is immaterial. Descriptive models accommodate information of either kind. As an example, a portion of the Climax-Molybdenum deposit model is shown in Table 1. What is sig-

Table 1. Portion of the descriptive model of the Climax-Molybdenum deposit [4]

Description: Stockwork of quartz and molybdenite associated with fluorite in granite porphyry
Age Range: Mainly mid-Tertiary
Rock Types: Granite-rhyolite with > 75 percent silica
Form-Structure: Predominantly in veinlets and fractures; minor disseminations
Mineralogy: Molybdenite + quartz + fluorite + K-feldspar + pyrite + wolframite + cassiterite + topaz
Alteration: Intense quartz and quartz + K-feldspar veining in ore zone. Upper phyllic and propylitic zones
Geochemical: Mo, Sn, W, and Rb anomalies close above ore zones. Pb, Zn, F, and U anomalies in wall rocks up to a few kilometers distant. Cu anomaly external to Mount Emmons deposit. In panned concentrates, Sn, W, Mo, and F may be important
Associated Deposit Types: Ag-base-metal veins, Fluorspar

nificant about this model is that a single format applies. The result is that with this type of format, geologists have found it much easier to create models or modify existing ones. Such models differ markedly from the models in PROSPECTOR.

The models in PROSPECTOR were encoded as inference nets in which the relations between field observations and important geological hypotheses were directly connected [5]. The connections were made in such a way as to determine how a change in probability in one of the nodes of the net would affect the nodes in other parts of the net. Part of the original Climax-Molybdenum deposit model is shown in Figure 1. Because the models in PROSPECTOR did

Figure 1. Portion of the inference net of the Climax-Molybdenum deposit in PROSPECTOR [6]

not always distinguish between descriptive data and the geological hypotheses whose truthfulness could only be inferred from the data, it has proved difficult to keep models current as new data have been gathered and new insights about ore-forming processes gained. By 1983, for instance, 31 deposit models containing over 2,000 rules and a taxonomy of geological terms containing over 1,000 entries had accumulated [7]. At that time, the prospect loomed that PROSPECTOR might become unmanageable. As a result, late in 1984, two important decisions were made. The first was to redesign PROSPECTOR as a frame-based expert system. The second was to build the system around descriptive deposit models that were then being assembled in [3].

PROSPECTOR II

Prospector II has been implemented as a frame-based expert system. The system offers advice on how close the observations (findings) in an area match the attributes of one or more descriptive mineral deposit models. Currently, 88 models reside within the system. Representatives of the 88 models are listed in Table 2.

Table 2. Representative Deposit Models in Prospector II

Primary Metal	Deposit Model
Gold, Silver	Carbonate-hosted-gold-silver Comstock-epithermal-veins Creede-epithermal-veins Homestake-gold Low-sulfide-gold-quartz-veins Polymetallic-vein
Lead, Zinc	Appalachian-zinc Kipushi-copper-lead-zinc Sandstone-hosted-lead-zinc Sedimentary-exhalative-zinc-lead Southeast-Missouri-lead-zinc
Copper	Besshi-massive-sulfide Copper-skarn Kuroko-massive-sulfide Porphyry-copper
Molybdenum	Climax-molybdenum Porphyry-molybdenum-low-fluorine
Tin	Tin-greisen Tin-skarn Tin-vein
Nickel, Chromium	Bushveld-chromium Dunitic-nickel-copper
Uranium	Quartz-pebble-conglomerate Sandstone-uranium Volcanogenic-uranium

For any given model, the geologist is advised for which of a given set of findings can be explained, or not explained, or missing. This strategy for advising is similar to the strategy used in INTERNIST-I for making multiple and complex diagnoses in internal medicine [8]. This strategy relies heavily on the information volunteered by the geologist. Adopting such a strategy meant greatly expanding the volunteer mode in PROSPECTOR.

The geologist volunteers information using a glossary of geology. The glossary is in the form of a taxonomy that is stored as part of the knowledge base. It represents an expanded version of the taxonomy in PROSPECTOR [9]. Currently, the glossary contains terms for describing the geologic age, petrology, structure, mineralogy, geochemistry, geophysics, and associated mineral deposits of an area. Structuring the glossary in this way makes it possible to take advantage of the semantic

relationships that exist between elements. To illustrate, a portion of the hierarchical structure for describing felsic-plutonic igneous rocks is shown in Figure 2. Within an area, for

Figure 2. Representative deposit models in Prospector II

instance, the geologist may have specific knowledge about the type of granite that occurs. Suppose the type is leucogranite. In selecting leucogranite, the geologist is informing the system not only about leucogranite, but also about the presence of a felsic-plutonic rock. This serves to distinguish a broad class of deposit models all of which are associated with felsic plutonic rocks. Should granite be an element of one of these models, the presence of leucogranite would constrain even more the possible deposit models. Such a strategy for pattern-matching is a simplification of the strategy employed by the semantic network matcher in PROSPECTOR [10].

The top-level hierarchy for describing the findings in an area is shown in Figure 3. Each

Figure 3. Top-level of glossary of geology

box represents the top-level for a lower-level hierarchy. In the current hierarchical structure, there are 57 subgraphs containing 1011 nodes and 865 links.

Uncertainty about the presence of an element is handled more simply than in PROSPECTOR. The reason for a more simple representation lies in the distinction between mineral exploration and regional mineral resource assessment. In mineral exploration, the objective is to locate hidden deposits. In regional mineral resource assessment, the objective is to identify areas likely to contain hidden deposits. This difference is reflected in the kinds of data that are developed in addressing the two objectives. Data collected in regional mineral resource assessment are generally more areally extensive and less purged of information extraneous to a specific model. Consequently, an area is characterized in more generalized terms. Usually, it is enough to establish the approximate age of the rocks, the type of rocks and minerals present, the presence of any geochemical or geophysical anomalies, and any evidence of mineralization. The geologist collects an amount of data that are sufficient to characterize an area in these terms. The uncertainty lies most often in the degree of specificity in the description, as illustrated by the earlier example about rock type identification. This type of uncertainty is accommodated most easily by the pattern-matching algorithm inherent in the hierarchical structure of the glossary. There may be instances, however, when the geologist suspects that some attribute is present, but where direct evidence of the attribute is lacking. In these situations, the geologist can indicate that the attribute is "present(?)," which is interpreted to mean that should the attribute be one of the elements of a model, it will be considered a match, but considered of less importance than had the attribute been declared as being present unequivocally. The geologist can state also that an attribute is absent, in which case, the attribute was looked for but not found. For some models, the absence of an attribute is essential if the model is to be considered applicable.

Deciding which models best match the given findings in an area is based on the pattern-matching algorithm used in INTERNIST-I [8]. For this algorithm, the elements of a model are assigned scores depending upon their relative importance in deciding whether or not a particular model applies. In particular, a positive score is assigned if the element is present, and a negative score (often different), if it is missing. When the algorithm is applied to a model for a given set of findings, the scores (positive or negative) are applied. Models are then ranked according to their scores. Maximum scores are a measure of the size of the model just as the number of rules was a measure of the size of a model in PROSPECTOR [1].

Statements of subjective probability have been avoided in Prospector II. The decision as to which deposit model best fits the given set of findings in an area is based primarily on the essential attributes of the model that are

present. In building the 88 models in Prospector II, it has proved much easier to ask of the geologists contributing models to identify and rate the essential attributes rather than to elicit the subjective probabilities required of a probabilistic model. To date, the system has been successful in correctly classifying deposit types for a wide variety of geologic settings. In a recent experiment involving the geologic settings for 127 lode deposits in Alaska [11], Prospector II correctly identified 105 of the 127 different deposit types classified by a group of expert geologists. Of the 22 deposits classified differently, 16 of these were classified as the 2nd best choice. Of the 7 remaining deposits classified differently, the experts were uncertain about 4 of these deposits. Considering the subjective nature of classifying deposit types in general, it is significant that Prospector II correctly classified over 80 percent of the 127 deposits, or over 95 percent if 16 of the top two choices are considered. This represents a significant step forward towards the automated classification of mineral deposit occurrences.

The Climax-Molybdenum deposit model in Prospector II is shown in Table 3. The maximum

Table 3. Portion of the Prospector II model of Climax-molybdenum deposit

Description: Stockwork of quartz and molybdenite associated with fluorite in granite porphyry
Age Range (100 100): Tertiary
Rock Types (75 75): Felsic-plutonic (5 5) Granite (4 5) Rhyolite (4 3)
Form-Structure (0 0): Stockwork Veinlets
Mineralogy (75 75): Molybdenite (4 5) Fluorite (4 5) K-feldspar (2 5) Pyrite (2 4) Wolframite (2 3) Cassiterite (4 3) Topaz (4 3)
Alteration (400 400): Silicification (5 2) Potassic (5 2)
Geochemical Signature (75 75): Mo (4 5) Sn (4 3) W (4 5) Rb (3 5) Pb (2 4) Zn (2 3) F (3 5) Th (2 3) K (2 2) Cs (2 4) Li (2 4) Nb (2 3) Ta (2 3) Mn (2 4) Re (2 5)
Associated Deposit Types (400 400): Climax-molybdenum Fluorspar Ag-base-metal-veins

scores associated with the main categories are shown as a pair of values that applies to each element. The first number represents the maximum score that can be assigned if the element is present. The second number represents the maximum score (negative) that can be assigned if the element is absent. For many of the elements, there is a pair of values that reflects the model-specific importance of that element within each category. The first number is a measure of its specific importance when the element is present. The second number is a measure of its specific importance when it is absent. Both numbers are defined in the range of 0 to 5. The score that is assigned to the specific importance is based on a table defined for each category. The number 5 is associated with the maximum score for each category. Default scores are assigned

if the specific importance of an element is unspecified.

RUN-TIME CAPABILITIES

As a program, Prospector II accepts input from the geologist, offers advice, accepts more input, offers more advice and so on. Above all, the initiative remains with the geologist. The session is terminated when the geologist has no further input. The session, typically, lasts less than 30 minutes. This is the time it takes for the geologist to enter the findings for an area, to select an initial set of models, provide an explanation as to what is explained, not explained, or missing for a set of models, and the geologist acting on this advice to modify the findings, to obtain a revised set of models, to select a subset of these based on an examination of partial-scores or scores of individual elements, and finally, to decide which model or models best fit the findings. At the conclusion of a session, the geologist can decide whether or not to generate a hardcopy, or else to save the findings for a future session. Figures 4 and 5 show the input and sample output for a typical session.

Figure 4. Screen showing input of findings to Prospector II

Figure 5. Screen showing some of the different types of results during a consultation with Prospector II

FUTURE DIRECTIONS

Prospector II represents a continuing effort to build upon the ideas in PROSPECTOR and to take a step closer to the day when the geologist sitting at a workstation can retrieve geologic, geochemical, and geophysical maps of an area, bring up a knowledge base of mineral deposit models, apply the models, and identify areas likely to contain undiscovered mineral deposits, what types they might be, and how many there might be. For the present we need to add more models to the knowledge base, quantify the relationships that govern the occurrence or non-occurrence of deposits, and most important, develop computer-based systems capable of spatial reasoning.

REFERENCES

- [1] R. O. Duda, "The PROSPECTOR System for Mineral Exploration," Final SRI Project 8172, (April 1980).
- [2] S. M. Cargill and S. B., Green, ed., "Prospects for Mineral Resource Assessments on Public Lands: Proceeding of the Leesburg Workshop," U.S. Geological Survey Circular 980, 1986.
- [3] D. P. Cox and D. A. Singer, "Mineral Deposit Models," U.S. Geological Survey Bulletin 1693, 1986.
- [4] S. D. Ludington, "Descriptive Model of Climax Mo Deposits," in D. P. Cox and D. A. Singer, ed., Mineral Deposit Models, p. 73. U.S. Geological Survey Bulletin 1693, 1986.
- [5] R. Duda, J. Gaschnig, and P. Hart, "Model Design in the PROSPECTOR Consultant System," in D. Michie, ed., Expert Systems in the Microelectronic Age, pp. 153-167, University of Edinburgh Press, 1979.
- [6] R. Reboh and J. Reiter, "A knowledge-Based System for Regional Mineral Resource Assessment," Final Report SRI Project 4119, (February 1983).

- [7] R. B. McCammon, "Recent Developments in PROSPECTOR and Future Expert Systems in Regional Resource Evaluation," Proceedings on Spatial Information Technologies for Remote Sensing Today and Tomorrow, (Sioux Falls, 2-4 October 1984).
- [8] R. A. Miller, H. E. Pople, Jr., and J. D. Myers, "INTERNIST-I, An Experimental Computer-Based Diagnostic Consultant for General Internal Medicine," New England Journal of Medicine, 307(8), (August 1982).
- [9] Duda, R. O., P. E. Hart, N. J. Nilsson, R. Reboh, J. Slocum, and G. Sutherland, "Development of a Computer-Based Consultant for Mineral Exploration," Final Report, SRI Projects 5821 and 6415, (October 1977).
- [10] R. Reboh, "Knowledge Engineering Techniques and Tools in the PROSPECTOR Environment," Tech. Note 243 SRI Projects 5821, 6415, and 8172, (June 1981).
- [11] W. J. Nokleberg, T. K. Bundtzen, H. C. Berg, D. A. Brew, D. Grybeck, M. S. Robinson, T. E. Smith, and W. Yeand, "Significant Metalliferous Lode Deposits and Placer Districts of Alaska," U.S. Geological Survey Bulletin 1786, 1987.