Profile of PROSPECTOR

Tao Zhang (zt@wpi.edu)

Category 1 – General

Domain

This is a system in the geology domain.

Main General Function

PROSPECTOR is an expert system designed for decision-making problems in mineral exploration. It aids geologists in evaluating the favorability of an exploration site or region for occurrences of ore deposits of particular types.

Once a site has been identified, PROSPECTOR can also be used for drilling-site selection.

In addition, Geologists had remarked about its potential value as an educational tool. In this regard, the models in the system contain explicit, detailed information from the literature and the experience of expert explorationists, together with explanatory text that can be obtained upon request.

System Name

The system is called PROSPECTOR.

Dates

The system was developed between 1976 and 1981.

Researchers

The main developers of the system were Richard Duda, John Gaschnig, Peter Hart, Rene Reboh, Nils Nilsson.

Location

The original research was performed in SRI International located at Menlo Park, California 94025.

Language

PROSPECTOR is written in INTERLISP, an advanced dialect of the LISP language.

Machine

The system runs on a DEC PDP-10.

Brief Summary

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 deposit. Knowledge about a particular type of ore deposit is encoded in a

computational model representing observable geological features and the relative significance thereof.

Related Systems

The domain of PROSPECTOR is very specific. There's no other system in the field of geology that is directly

There is another system, CONPHYDE (stands for CONsultant for PHYsical property DEcisions) that also used the concept of Inference Networks. CONPHYDE was based on PROSPECTOR and used the framework of the PROSPECTOR. CONPHYDE is designed to aid an engineer in the selection of an appropriate vapor-liquid equilibrium method when performing various process calculations.

Category 2

Characterizations of givens

In the interactive consultation mode of operation, the system is first in the Antecedent mode. A typical consultation session begins with the user volunteering information in the form of simple assertions. A typical consequence of initial volunteering is that a few exact and a number of partial matches are made between volunteered evidence and the nodes in the inference network, and some changes occur in the probabilities of the top-level hypotheses. Second, once a top-level hypothesis H has been chosen, the program enters the consequent mode. Here it searches the inference network below H to determine what question to ask the user to help resolve the issue.

In the batch processing mode, PROSPECTOR can automatically generate questionnaires for any inference network. When the data entered in the questionnaire are later transferred to a file, the program can read the answer from the file as if it were in normal interactive mode.

When PROSPECTOR is used for the drilling site selection, it would also request map data. The user enters it by using a digitizing tablet with a contour editing and display program.

Characterization of Output

PROSPECTOR can reach a conclusion about a particular ore deposit. It gives a certainty value of the ore deposit. It as well provides the explanation text for the conclusion.

Like MYCIN, it allows the user to execute commands as well as to answer questions. For example, in response to a WHY command, the program accessed some explanatory text that the expert has previously prepared to explain why a particular piece of evidence is important. Other commands allow the user to do such things as trace internal inferences, change previous answers, change top-level goals, and obtain summaries of conclusions reached up to that point.

When PROSPECTOR is used for the drilling site selection, it could also produce the favorability map of the site.

Characterization of Data

The data in PROSPECTOR is matched against existing models. Most of PROSPECTOR's questions expect Yes/No or certainty answers, although some questions ask for quantities; in the latter case, the likelihood ratio for the rule is a function of that quantity. In some cases, the PROSPECTOR may request the user to input the digitizing map data for a drilling site.

Generic Tasks

Based on the paper of B.Chandrasekaran about generic tasks in Expert System design, there are three generic tasks that PROSPECTOR uses. These are Hierarchical classification, Hierarchical matching or assessment, Knowledge-directed information passing,

Theoretical Commitment

The extensive use of Inference Network is the central theme of PROSPECTOR. In addition, the utilizing of Semantic Network in PROSPECTOR is worth noting.

Reality

According to the information released by the contributors of PROSPECTOR system, it is a successful design and implementation. The average difference with the experts is about 6.9%, which is not so significant. However, this result may be biased because they have no mention of "double-blind" experiments. They admit that the models are incomplete.

Category 3

Completeness

The Inference Network of PROSPECTOR is completely functioning. However, the models developed are not complete in term of the coverage of the ore deposits.

Use

PROSPECTOR was used by the real-life task and business. Unlike other project like MYCIN, it is a project developed in a business company. Thus, the PROSPECTOR from the very beginning is targeting the practical use instead of pure research purpose.

Performance

According to the evaluations done by the PROSPECTOR system developers and researchers, the results concluded by PROSPECTOR in term of certainty value only differ from the results of the experts by 6.9%, which is not so significant. However, this result may be biased because they have no mention of "double-blind" experiments. They admit that the models are incomplete.

Category 4

Phases

PROSPECTOR has 3 modes of operations: Interactive consultation; Batch processing; Compiled execution.

For the Interactive consultation mode, there are 2 phases: 1.The antecedent mode. 2.The consequent mode. A typical consultation session begins with the user volunteering information in the form of simple assertions. A typical consequence of initial volunteering is that a few exact and a number of partial matches are made between volunteered evidence and the nodes in the inference network, and some changes occur in the probabilities of the top-level hypotheses. Second, once a top-level hypothesis H has been chosen, the program enters the consequent mode. It searches the inference network below H to determine what question to ask the user to help resolve the issue.

Subfunctions

KAS (Knowledge Acquisition System), Inference Network, Semantic Network, English Language Parsing, Explanation System, Console System.

Use of Simulation or Analysis

Simulation is not used in PROSPECTOR system.

System/Control Implementation Architecture

The control architecture involves the use of the Inference Network to control the information influence and probability propagation.

Category 5

Characterization of Structure Knowledge

PROSPECTOR structures the taxology knowledge of Geology by the Semantic Network. The expert knowledge about a particular ore deposit was encoded into each model. The assertions in a model are organized according to the requirement of the Inference Network.

Characterization of Process Knowledge

The process knowledge in PROSPECTOR is about the logical relations, plausible relations, and the context relations between assertions. These three kinds of relations are the basis of the Inference Network.

Deep or Surface

The models in PROSPECTOR represent some deep knowledge from the expert in the domain. However, in term of how the knowledge is utilized, the PROSPECTOR system is not considered as “deep” because it uses the knowledge in the Inference Network on the surface basis.

Category 6

Search Space

In the antecedent mode of an interactive consultation, the program is receiving information; it matches statements from the user against the assertions in the inference network and propagates probability changes up through the network. In the consequent mode, the program is attempting to establish (or rule out) a top-level hypothesis, and it searches the inference network for evidence nodes that are most effective for this task. The strategy that the program uses to choose goals and select evidence is called control strategy.

Space Traversal

A typical consequence of initial volunteering is that a few exact and a number of partial matches are made between volunteered evidence and the nodes in the inference network, and some changes occur in the probabilities of the top-level hypotheses. At this point PROSPECTOR turns to the problem of working interactively with the user to choose a top-level hypothesis for further refinement. For each top-level hypothesis, a score is computed that combines the certainty of that hypotheses and an average "effective" certainty of related nodes that were partially matched by volunteered evidence. The top-scoring hypotheses are revealed to the user, who is then given the choice of either following the program's recommendation or selecting a different hypothesis to pursuer. In either case, the system has a goal hypothesis to work on.

Search Control Category

Once a top-level hypothesis H has been chosen, the program enters the consequent mode. Here it searches the inference network below H to determine what question to ask the user to help resolve the issue. In doing so, PROSPECTOR attempts to satisfy the following criteria:

1. Effectiveness. The evidence sought should have the potential to make a difference in the conclusion.
2. Naturalness. The program should not jump from topic from topic to topic in a disorganized way that might confuse the user, but should follow a coherent line of reasoning.
3. Responsiveness. Interactive consultations should not be delayed by time-consuming optimization calculations.

Standard Search

PROSPECTOR's strategy for satisfying these requirements is based on backward chaining. Fundamentally, it recursively looks for the antecedent node that has the greatest effect on the current consequent node, stopping when it reaches a new node that it can ask about.

Search Control Characterization

PROSPECTOR's strategy for satisfying these requirements is based on backward chaining. Fundamentally, it recursively looks for the antecedent node that has the greatest effect on the current consequent node, stopping when it reaches a new node that it can ask about.

Subproblems

Subproblems in PROSPECTOR are identified by the "context" relations. PROSPECTOR allows any proposition C to be designated as a context. In the inference network, special arcs called context arcs link any dependent propositions A to C. A context arc blocks the upward propagation of any information about A if the proper context has not been established. Thus, if an important conclusion depends upon A, the inference network interpreter will set up the subgoal of first establishing the context C.

Search Control Representation

The search control is based on the three relations between assertions: the logical relation, the plausible relation (combined weight) and the context.

Search Control Strength

Like MYCIN, PROSPECTOR's search control is also domain independent. It has a number of advantages with this approach. The most important one is that the problem solving mechanisms and system facilities of the interpreter can be applied to similar problem domains by replacing the old knowledge base by a knowledge base for the new domain.

Category 7

Failure Method

The core of KAS (Knowledge Acquisition System) of PROSPECTOR is an "intelligent" network editor that can assist the user in building, testing, searching and maintaining these networks. It knows about the representation constructs and inference mechanisms employed by PROSPECTOR and can therefore protect the user against certain kinds of syntactic errors. It also includes a book-keeping system that keeps track of incomplete data structures.

Uncertainty

PROSPECTOR deals with uncertainty by using certainty value C(H|E). PROSPECTOR accepts user input that is uncertain with a probability.

During a consultation, the user can express his or her uncertainty about an input quantity E by specifying a certainty factor for various intervals. Probability distributions are computes at each node H in the inference network by combining the probability distributions associated with its antecedents E'. The exact combination methodology is determined by the type of numerical computation construct associated with H.

Management of Uncertainty

PROSPECTOR manages the uncertainty exactly by the utilizing of the Inference Network. It can therefore accurately prorogate the certainty value/uncertainty to the top-level hypothesis.

Management of Time

PROSPECTOR does not have any issues related to the management of time

Category 8

Knowledge Representation Method

PROSPECTOR system uses models and semantic networks to represents knowledge. Semantic Network is a network of nodes linked together by directed arcs to represent relevant knowledge like taxonomic relations among objects in the domain. Model is a body of knowledge about a particular domain of expertise encoded into the system, which the system can act.

Knowledge Representation Generality

All knowledge in the system is expressed in coded INTERLISP structures.

Knowledge Structuring

There are 3 kinds of Relations between the nodes: 1. Logical Relations. 2. Plausible Relations. 3. Contextual Relations.

Category 9

Alternative Representations

PROSPECTOR does not use alternative representations for the same piece of knowledge.

Alternative Solution Methods

PROSPECTOR does not use alternative solution methods to reach the same solution.

Optimization

During the last phase of model development, the developers involve the optimization of the model. Given the questionnaire data for a number of actual deposits, it is possible to make a serious quantitative evaluation of how well particular deposits match the model. This evaluation inevitably exposes various shortcomings of the model as encoded, requiring revision of the work done in Phases B and C.

Multiple Results

PROSPECTOR deals with multiple results by using certainty value. For each top-level hypothesis, a score is computed that combines the certainty of that hypothesis and an average "effective" certainty of related nodes that were partially matched by volunteered evidence. The top-scoring hypotheses are revealed to the user, who is then given the choice of either following the program's recommendation or selecting a different hypothesis to pursuer. In either case, the system has a goal hypothesis to work on.

Category 10

Interaction

The program begins by allowing the user to enter information about the significant types of rocks and minerals that have been observed or suspected to be present. This antecedent mode helps in the selection of models to pursue, which are scored and initially

ranked. Once the user selects a particular model, the program enters into the consequent mode and begins asking a series of questions.

Most of PROSPECTOR's questions expect Yes/No or certainty answers, although some questions ask for quantities; in the latter case, the likelihood ratio for the rule is a function of that quantity.

Data Collection

The inference network can be used in any of the three operating modes of PROSPECTOR (Interactive, batch, compiled). For the application to prospect evaluation, the interactive mode is natural. We can imagine the typical user to be an exploration geologist who has just spend several days making a surface examination of a prospect, who has found evidence that sparks his or her interest, and who would like to receive advice from a specialist in evaluating the degree to which the prospect matches a classical model.

The program begins by allowing the user to enter information about the significant types of rocks and minerals that have been observed or suspected to be present. This antecedent mode helps in the selection of models to pursue, which are scored and initially ranked. Once the user selects a particular model, the program enters into the consequent mode and begins asking a series of questions.

In some cases, the PROSPECTOR may request the user to input the digitizing map data for a drilling site.

Data Format

Most of PROSPECTOR's questions expect Yes/No or certainty answers, although some questions ask for quantities; in the latter case, the likelihood ratio for the rule is a function of that quantity. Other information is given to PROSPECTOR in a restricted subset of natural language form.

In some cases, the PROSPECTOR may request the user to input the digitizing map data for a drilling site.

Acquisition

PROSPECTOR provides a Knowledge Acquisition System (KAS) to enable easy acquisition of domain knowledge.

Learning

No learning ability appears in PROSPECTOR.

Explanation

PROSPECTOR provides an explanation facility that allows it to explain to the expert the behavior of PROSPECTOR.

Category 11

Strengths

- It uses a generic Inference Network to do the knowledge reasoning. This is more accurate and domain independent.
- The performance of the PROSPECTOR system is not bad.
- It uses the Semantic Network to help the Inference Network better understand the assertions. This is effective.

Weaknesses

- It is not easy to build a model. As a result, the models are incomplete.