CONTACT INFORMATION
Mining Records Curator
Arizona Geological Survey
416 W. Congress St., Suite 100
Tucson, Arizona 85701
520-770-3500
http://www.azgs.az.gov
inquiries@azgs.az.gov

The following file is part of the

James Doyle Sell Mining Collection

ACCESS STATEMENT

These digitized collections are accessible for purposes of education and research. We have indicated what we know about copyright and rights of privacy, publicity, or trademark. Due to the nature of archival collections, we are not always able to identify this information. We are eager to hear from any rights owners, so that we may obtain accurate information. Upon request, we will remove material from public view while we address a rights issue.

CONSTRAINTS STATEMENT

The Arizona Geological Survey does not claim to control all rights for all materials in its collection. These rights include, but are not limited to: copyright, privacy rights, and cultural protection rights. The User hereby assumes all responsibility for obtaining any rights to use the material in excess of "fair use."

The Survey makes no intellectual property claims to the products created by individual authors in the manuscript collections, except when the author deeded those rights to the Survey or when those authors were employed by the State of Arizona and created intellectual products as a function of their official duties. The Survey does maintain property rights to the physical and digital representations of the works.

QUALITY STATEMENT

The Arizona Geological Survey is not responsible for the accuracy of the records, information, or opinions that may be contained in the files. The Survey collects, catalogs, and archives data on mineral properties regardless of its views of the veracity or accuracy of those data.

JAN WILT

ASARCO

Southwestern Exploration Division

January 11, 1984

WILT

To: W. L. Kurtz

From: J. D. Sell

Preliminary Evaluation
Gold in Quartzites
Nevada

On November 11, 1983 I contacted Jan C. Wilt about doing a computer search for outlining the occurrence of gold in quartzite-type units in the Western United States.

Submitted as Attachments A and B are the General Literature Research Proposal (July 20, 1982) and the Literature Research Process (Sept. 19, 1983) which Ms. Wilt uses in such a search. After further discussion the search was restricted to gold in Nevada with emphasis on occurrences in quartzite. The search was to be "quick and easy" with no great in-depth study. Attachment C is her proposal estimate (dated November 15, 1983) and Steps 1 through 3 were carried forward.

Step 1 and Step 2 (partial) were completed by December 13, 1983 (Attachment D) after ½ hour of Dialog computer time revealed 714 references on Nevada metals with 399 references to Nevada gold occurrences. These were xeroxed on 8½ x 11 sheets as received from the computer and placed in a 3-ring binder (Tucson office), rather than being cut apart and placed on individual 3" x 5" cards as is normally done. When placed and indexed on individual cards, a number of duplicate references which the printer disgorges, is eliminated and the file is tidy and cross-indexed as you wish--i.e., title, metal, county, author, etc. However, for our purpose, the 3-ring binder gives the total mixed-up references to metals and gold in Nevada.

Attachment E is the format and item description which Jan used for the gold in quartzite search question. From this printout she added to the file the following references which are included in the 3-ring binder:

A) Bibliography of Graduate Theses of Nevada Geology to 1976. Nevada Bureau Report 31.

B) Bibliography of Geologic Literature of Nevada and Bibliography of Geologic Maps of Nevada Areas. Nevada Bureau Bull. No. 43.

C) Gold Excerpts and Reference List from Mineral and Water Resources of Nevada. Nevada Bureau Bull. No. 65.

D) Reference lists from various Nevada County Bulletins and other pertinent reference sources (15 sources).

As noted on Attachment D, Steps 1 and 2 (partial) cost $680.95.

From this data base and the use of Nevada Map No. 32--Gold-Producing Districts of Nevada--Jan proceeded to Step 3. Entered on the computer floppy disk was the following information to create a "Data Base Management System" (DBMS):

a. Gold map number from map 32
b. County
c. Mining district name, and alternate names 1,2,3
d. Commodities
e. Geologic age
f. Host rock
g. Deposit type
h. Production dollar
i. Tons ore produced
j. Gold ounces produced
k. Silver ounces produced
l. Gold ounces per ton
m. Silver ounces per ton
n. Gold/silver ratio
o. Notes
p. Pre-1940 dollar production value
q. Pre-1940 tons produced

From the DBMS input, a printout was made of the data a through q sorted by County, and patched together as a large roll (in Tucson office). Naturally, with the above base, any of the information can be retrieved in whatever form or sequence is desired.

Except for commodities (d), geologic age (e), and notes (o), data a through q were sorted and printed out by c, f, g, h, i, j, l, and n. Those districts having quartzite were also sorted and printed out by c, f, g, h, i, j, k, l, m, n, p, and q.

Step 3 was invoiced at $1280 (Attachment F) when the data were submitted on January 3, 1984.

I have not progressed further on utilizing the computer printouts for evaluating gold in quartzite units in eastern Nevada during 1984.

Jim

James D. Sell

JDS/cg

General Literature Research Proposal

July 20, 1982

PURPOSE-- To quickly focus on specific locations that are most likely to contain an economic ore deposit.

BENEFIT-- This selection process eliminates waste in costly field expenses. Rather than look at many prospects in a haphazard manner, only the more favorable are examined on a priority basis.

In addition, the field examination in specific localities will be more efficient because the geologist will have all the previously known information at hand and know what new observations to look for.

GENERAL PROCEDURE-- Literature research enables one to recognize significant geologic controls on mineralization and this knowledge leads to the discovery of new ore deposits.

The first step in literature research is a literature search, which obtains bibliographies, reference lists and copies of articles about the general topic and about specific occurrences.

The second step is data gathering, which obtains significant geologic parameters, such as mineralogy, host rock, age, production, etc. about each occurrence.

The third step is pattern analysis, in which specific geologic parameters are compared with production, etc., by means of maps and correlation charts to determine which geologic factors can best be used to predict production.

The fourth step is prediction, in which the geology of various areas is examined for the predictive factors discovered in the third step and in which specific areas are recommended for field examination or drilling.

GENERAL COST-- The amount of time spent and the cost depend upon the degree of detail desired and upon the abundance of the commodity. A detailed examination of one state with about 400 occurrences generally takes 2 to 3 months for 1 person for a cost of $12,000 to $20,000. This cost is minor when compared to field expenses or drill hole costs.

Jan C. Wilt
Consulting Geologist
3035 S. Shiela Ave.
Tucson, AZ 85706
(602) 883-6669

LITERATURE RESEARCH PROCESS

J. C. Wilt
September 19, 1983

PLANNING

What final product do they want?
- Who makes the decisions?
- What are the deciding factors in the decisions?

What are the necessary intermediate steps or products?

What is the most effective way to display or present the product?

What is the most efficient way to obtain the product?
- 80-20 rule = 80% of data is collected in 20% of time
- Prioritizes-completeness (detail) or time (money) most critical?

Write out plan, including the purpose, proposed products, and estimated cost.

PROCEDURE:

I. Literature Search

1. Obtain bibliographies through search of computer files in GEOREF and GEOARCHIVE and through published bibliographies of state geologic surveys and U.S.G.S. index to geologic mapping.
2. Obtain lists of references from the bibliographies in the papers found through the computer lists, in guidebooks to local areas, in landmark economic geology papers, and general commodity or area references of the U.S.G.S. or U.S.B.M.
3. Organize reference cards by photocopying these lists of references in the backs of the articles, reducing them if necessary, and cutting and taping each reference on a separate 3"x5" card.
4. Obtain copies of articles and books that were listed in the bibliographies from the University of Arizona library and other sources such as the U.S. Geological Survey, the state geological surveys, interlibrary loan, or through the computer ordering service that copies the articles on the GEOREF database. Divide the cards into piles acccording to the most important articles, less informative papers, and difficult to find cards. Then subdivide these into piles according to the location in the library so you can look up all the Econ. Geol. articles at once. Photocopy the most important articles first.

II. Data Gathering

A. Gathering information.

1. CRIB - computer printout from U. S. Geological Survey on geologic information about mineral occurrences.
2. MILS - computer printout from the U. S. Bureau

of Mines on locations and status of mines and prospects.

3. Data from literature obtained from county mine reports by the state geological surveys, from literature found in the bibliographic search, from metal production information in U.S. Bureau of Mines yearbooks, books on mineral of particular state, U.S.G.S. maps on various commodities, bulletins on mineral and water resources of each state done in the mid 1960's for each state, U.S.G.S. Prof. Paper 820 on U.S. mineral resources, and various directories and other reference works in the reference section of the Science Library.

B. Recording data

1. Record information on 3"x5" file cards with notches or file divisions, or in notebooks with dividers, or in file folders, or in large charts. It is important to cross-index these cards or charts with the name and date of reference, and other information that may be important, such as elements.
2. Record information in data base management system on computers, such as the DB Master for Apple II or Perfect Filer for CP/M systems. This allows you to rearrange the information in sequence for any item in the data base.
3. Record occurrences on geologic map or index map with colored dots or pins that can be coded to types of occurrences or amounts of production, etc.

III. Pattern Analysis

1. Analyze the information found in earlier steps by rearranging the data in various ways to discover correlations and patterns, either on the maps, or in coincidences of various geologic factors with production or high grade, etc. Then arrange the occurrences in order of priority for greatest potential for economic ore deposits.

IV. Prediction

1. Analyze the maps produced and information obtained in the data gathering step to find the previously undiscovered areas that have geologic conditions similar to the geology of the known economic occurrences delineated in the pattern analysis and data gathering steps. These new areas can then be listed in priority order so that the most promising areas cay be examined in the field first.

Jan C. Wilt
Consulting Geologist
3035 S. Shiela Ave.
Tucson, AZ 85746
602-883-6669

ATTACHMENT C

Jan C. Wilt
3035 S. Shiela Ave.
Tucson, Az. 85706
(602) 883-6669

November 15, 1983

James D. Sell
ASARCO Inc.
1150 N. 7th Ave.
Tucson, AZ 85703

Estimate for research on gold in quartzite in Nevada

Step 1 -	gather bibliographic material (from GEOREF, State bibliographies, and general articles such as County mine reports, Mineral & Water Resource Bull., other general geology articles & guidebooks)	$ 600	complete see 12/13/83
Step 2 -	organize bibliography onto cards (can be done by secretaries, etc. with some supervision-in house)	$ 200	partial 12/13 if xerox in 3-ring binder
Step 3 -	Make list of Nevada gold districts with emphasis on those containing quartzite or coarse clastic host rocks by reading through county mine reports - about 1 week at $25/hr.	$1000	partial 12/13 completed 1/3/84, see print-outs. Review
Step 4 -	Make bibliography for each mining district that has gold and quartzite by taking cards from step 1 & 2 plus additional references from references at back of those articles listed earlier -about 1 week at $25/hr.	$1000	
	Total	$ 2800	

For California, the cost would be approximately the same, although for projects lasting longer than 2 weeks, the cost decreases to $20/hr. The bibliographies for the mining districts in Arizona are already very complete, although it would take some time (1-1½ weeks) to determine which districts have gold in quartzite.

Sincerely yours,

Jan C. Wilt

Jan C. Wilt
Registered Geologist

ATTACHMENT D

Jan C. Wilt
3035 S. Shiela Ave.
Tucson, Az. 85706
(602) 883-6669

December 13, 1983

James D. Sell
ASARCO Inc.
1150 N. 7th Ave.
Tucson, AZ 85703

Cost for obtaining references on Nevada metals and Nevada gold

Cost of computer time Dialog ½ hr @ $82/hr	$ 30.59
Cost of 714 prints GEOREF on Nevada metals	142.80
Cost of 399 prints GEOREF on Nevada gold	98.07
Cost of xeroxing references - Alphagraphics	3.82
Cost of xeroxing references - Alphagraphics	35.67
	$ 310.95

Time to obtain these references and prepare notebook
18.5 hours at $20/hr $ 370.00

Total cost $ 680.95

Step 1 of 2

Thank you,

Jan C. Wilt

Jan C. Wilt
Registered Geologist #15664

OK for Payment
James D Sell
General Exploration

Pay as soon as possible.

ATTACHMENT E

User23640 Date:24nov83 Time:19:36:18 File: 89

Set	Items	Description
1	7376	NEVADA
2	7992	NEV?
3	35	NV
4	8025	1 OR 2 OR 3
5	8449	GOLD
6	3928	AU
7	295	PRECIOUS(W)METAL?
8	12396	5 OR 6 OR 7
9	399	4 AND 8
10	3175	QUARTZITE?
11	17423	SANDSTONE?
12	34	SS
13	4532	CONGL?
14	22897	10 OR 11 OR 12 OR 13
15	9	9 AND 14

Print 9/5/1-399

Search Time: 0.193 Prints: 399 Descs.: 11

A DIALOG* SEARCH FROM THE GEOREF DATABASE

For: _____

Address: _____

_____

If you have any questions, please call:

Telephone: _____

Topic of search: _____

Searcher: _____

Date: _____

The attached report is the result of a search of the GEOREF database using the Dialog Information Retrieval Service.

GEOREF provides comprehensive access to more than 4,500 international journals, plus books, conference papers, government publications, dissertations, theses, and maps concerned with all aspects of geology, geochemistry, geophysics, mineralogy, paleontology, petrology, and seismology. Approximately 40% of the indexed publications originate in the U.S. and the remainder from outside the U.S. Publications of international organizations make up about 7% of GEOREF.

SAMPLE RECORD

The positions of the key fields are shown in the following sample record.

AN 1012344 80-48454

TI **Manganese and copper geochemistry of interstitial fluids from manganese nodule-rich pelagic sediments of the northeastern equatorial Pacific Ocean**

AU Callender, E.; Bowser, C. J.

CS U. S. Geol. Surv., Reston, Va., USA; Univ. Wis., USA

JN PY Am. J. Sci. 280: 10, 1063-1096p., 1980

CO SN CODEN: AJSCAP ISSN: 0002-9599

SF Subfile: B

CP Country of Publ.: United States

DT BL Doc Type: SERIAL Bibliographic Level: ANALYTIC

LA Languages: English

LT LN Latitude: N000000; N200000 Longitude: W1800000; E1400000

DE Descriptors: *Pacific Ocean; *nodules; *manganese; *diagenesis; *metals; *sediments; *copper ; oceanography; geochemistry; pore water ; genesis; secondary structures; sedimentary structures; Equatorial Pacific; Northeast Pacific; remobilization ; solubility; desorption; precipitation

SH Section Headings: 07 .(MARINE GEOLOGY AND OCEANOGRAPHY)

(Copyright by the American Geological Institute, 1982.)

Key to Data Fields

AB	Abstract	JN	Journal Name
AN	GEOREF Accession Number	LA	Language
AU	Author	LN	Longitude
BL	Bibliographic Level	LT	Latitude
BN	ISBN	PU	Publisher
CL	Conference Location	PY	Publication Year
CO	CODEN	RN	Report Number
CP	Country of Publication	SF	Subfile
CS	Corporate Source	SH	Section Heading Code
CT	Conference Title	SL	Summary Language
CY	Conference Year	SN	ISSN
DE	Descriptor	TI	Title
DT	Document Type		

Data present in record depends on output format requested and type of record.

*DIALOG Information Services, Inc. Trademark
Reg. U. S. Pat. and Trademark Office

(Copyright DIALOG Information Services, Inc., 1982)

ATTACHMENT F

JAN C. WILT

Registered Geologist
Arizona Certificate No. 15664

J. C. Wilt & Co.
3035 S. Shiela Ave.
Tucson, Az. 85746
(602) 883-6669

EXPLORATION FOR ENERGY AND MINERAL RESOURCES

January 3, 1984

James D. Sell
ASARCO Inc.
1150 N. 7th Ave.
Tucson, AZ 85703

Compilation of information on Nevada gold districts in data base management system on computer	
54 hours @ $20/hr	$ 1080.00
Preparing and printing tables and reports from DBMS system on computer	
10 hours @ $20/hr	200.00
	$ 1280.00

Thank you,

Jan C. Wilt

Jan C. Wilt
Consulting Geologist

OK for Early Payment
James D Sell
General Exploration
1/5/84

JAN C. WILT

Registered Geologist
Arizona Certificate No. 15664

J. C. Wilt & Co.
3035 S. Shiela Ave.
Tucson, Az. 85746
(602) 883-6669

EXPLORATION FOR ENERGY AND MINERAL RESOURCES

January 25, 1984

William L. Kurtz
ASARCO Inc.
1150 N. 7th Ave.
Tucson, AZ 85703

Dear Bill,

I hope that 1984 will be very happy and prosperous for you! Although the last year and a half have not been terribly prosperous for most of us, I have kept busy compiling useful geologic data and references.

In addition to compiling bibliographies on interesting subjects, I have been copying the articles and abstracting them. One of these reference projects was a compilation of all articles concerning petroleum or natural gas in Arizona. Another project that I worked on with Stan Keith was a compilation of references on Laramide Tectonics in Arizona and surrounding areas. We now have a similar compendium on Tertiary Tectonics in Arizona, California, and Nevada. What makes our compilations so useful is the way we organize them into geological concepts and trace their historical development. That way you can decide for yourself who's 'out to lunch' and who's 'right on'. If you are interested in looking at any of these, give me a call.

See enclosed details.

Occasionally I work in association with Stan Keith and Monte Swan as MAGAMACHEM Associates. One of our current projects is a geochemical survey of the three major types of gold deposits in the Mojave region. We are preparing a detailed geological model (including geochemical zoning of 35 elements, as well as mineralogical and structural characteristics) for each of the three types - Carlin type, Oatman type, and Picacho type. If your company would be interested in joining this limited group project, call Stan (602-893-1434), Monte (303-674-1272) or myself.

I still have my Apple II+ computer and used it, along with a new letter quality printer, to put together the guidebook to the Arizona Geological Society field trip to western Arizona this November. The first day of that trip was to the Plomosa Mountains, where Bob Scarborough, Norm Meader, Stan Keith and I did a mapping project in 1981. That map and a report are available for $50 if you are interested in the geology of western Arizona.

Bonny surrounded by Fischer west I Norwood NE in District

also available (?) From Az Bu Mines (Bob Scarborough works for them)

xc: FJPK 2/8/84

RECEIVED
FEB 7 1984
EXPLORATION DEPARTMENT

You know how it is with computers - you always have to have a bigger and better one. So I recently acquired a Zenith Z-100 computer with 10 times the memory and speed of my Apple. One of the most useful programs is a data base management system called Lotus 1-2-3. I've used it to enter published information on production statistics, ages, and geologic information from Arizona mining districts. The program calculates ore grades and metal ratios and then, with a flick of a finger, I can sort the mining districts by any column - such as by gold grade, by gold:silver ratio, by geologic age, or alphabetically by district name. It will also graph any of the information. I'm giving a paper at a microcomputer conference in Houston in late February on using personal computers in exploration, so I've included my expanded abstract, as well as some examples from my Arizona mining district file that will illustrate how a data base management program (in this case Lotus 1-2-3) works.

There's another really useful program called Citation, which records references, abstracts, and cross-indexing and can be searched by keywords. I have over 1500 references on Arizona mining districts which are cross-indexed by mining district name and other keywords. If you're interested in a particular district or county, I can search the keywords and get a printout. I am currently working on similar lists of references for California and Nevada.

See Science Citation Index From Chem Abstracts

There is one project that I need your help with. I've started to collect stories that geologists tell, although they don't need to be limited to geologists. I'm looking for stories about adventures (and misadventures) that you or someone you know has had during exploration or mining activities. If you have heard any mining scam stories, I'm also collecting them for the mining club. If you think of any stories or someone who tells stories, call me and we can meet for lunch at the Mining Club of the Southwest and I can tape record them. We may have to change some names if the stories ever get published, but the collection should really be fun!

You must have lots of stories!

JHC?

Yours truly,

Jan Wilt

Jan C. Wilt
Consulting Geologist

MAGMACHEM ASSOCIATES

Regional Mineral Exploration and Property Evaluation

Suite 202, Ahwatukee Professional Building
10827 South 51st Street
Phoenix, Arizona 85044
(602) 893-1434

Stanley B. Keith • Monte M. Swan

REGIONAL GOLD EXPLORATION MODELS FOR THE MOJAVE REGION

Products:

1. In-depth report
2. Geochemical data base for 3 deposit types
3. Mineral district map
4. Mineral district data base
5. Igneous outcrop map
6. Time slice maps (tectonic, magmatic, and metallogenic)
7. Correlation chart for Cretaceous and Tertiary
8. Map of low-angle faults
9. Map of dike swarms and fractures

Got from Patty Dunning

The geochemical data base consists of multi-element analyses of at least one deposit in each of the three gold deposit types - alkalic (Oatman - 200 samples), calc-alkalic (Socorro Reef - 800 samples), and peraluminous (Tumco - 15 samples, Picacho - 25 samples, Mineral Ridge - 150 samples, and Mesquite - 8 samples). Multi-element data includes analyses for some or all of the following elements: Cu, Pb, Zn, Mo, Ag, Ni, Co, Mn, Fe, Cr, Hg, Bi, V, Au, U, W, F, As, B, Be, Sn, Sb, Ba, Rb, Sr, Nb, Y, Th, Se, Ta, Te, Li, Tl, S, Zr, and Cs. The value of this part of the report alone, if a company were to duplicate the analytical cost, would exceed $30,000.

The mineral district map locates all the mining districts at a scale of 1:1,000,000. The igneous outcrop map is also at a scale of 1:1,000,000.

The mineral district data base consists of summaries, production information, and reference lists for each of the more than 100 gold districts in the Mojave region. The information is also presented in tabular form sorted by various parameters, such as gold production, gold grade, gold:silver ratio, and total production.

The time slice maps consist of 1:2,000,000 scale maps at about 15 m.y. intervals from 105 m.y. to 43 m.y. They show the positions of the gold systems relative to the magmatism, structural features, and stratigraphic features of each time slice.

A correlation chart of Cretaceous amd Tertiary features, including stratigraphy, tectonics, and ore deposits, tracks the movement of the magmatic, metallogenic, and structural belts through time and space. A clear 'transgression' and 'regression' is obvious and can be related tectonically to the shallowing (speeding up) and later steepening (slowing down) of the subducting slab beneath western North America.

A structural habit common to many low-grade disseminated gold deposits is their occurrence at structural intersections of low-angle features, such as thrust faults, detachment faults, or stratigraphic phenomena, with high-angle

See full ABS Field Trip Guide + file notes

features, such as fractures and dikes. For this reason, the project includes maps of both low-angle and high-angle features. The 1:1,000,000-scale map of low-angle faults shows the late Tertiary detachment faults, as well as thrust faults of Sevier age, of early Laramide age, and of late Laramide age. The map of dike swarms and high-angle fractures is also at a scale of 1:1,000,000.

COSTS AND REPORT DELIVERY

The total cost of the package is $15,000.00. To offset production costs MAGMACHEM Associates would like an advance payment of $7,500.00 with the remaining $7,500.00 payable upon delivery of the final product. With the initial payment, the buyer will receive any accumulated data in the project to date. As of January 29, 1984, available data include: paleotectonic-metallogenic time slice maps from 105-43 m.y., correlation diagram for 105-43 m.y. stratigraphic, tectonic, and metallogenic data, and Bondar-Clegg multi-element geochemistry. Delivery of the remaining maps, data, and report is anticipated in March 1984.

REGIONAL GOLD EXPLORATION MODELS FOR THE MOJAVE REGION

by Stanley B. Keith, Monte M. Swan, and Jan C. Wilt, 1984

Gold occurrences are widespread throughout the Mojave region, which includes southeastern California northeast of the San Andreas fault and south of the Garlock fault, western Arizona, and southernmost Nevada. The region has yielded consistent gold production since the 1860's, mostly from high grade veins. However, the announcement by Goldfields of a major gold deposit of low grade and large tonnage in east-central Imperial County, California, underscores the previously underestimated potential of the region for disseminated gold deposits that are amenable to open-cut mining techniques. Consequently, this report reviews salient aspects of the geology and gold deposits in the Mojave region with the goal of targeting areas of bulk-gold potential.

Data obtained included information about igneous rock outcrops, dike swarms, and low-angle faults, all of which have known affinities with gold mineralization in the Mojave region. Known gold districts were also screened to establish what types of gold mineralization were present and which of these types were more favorable for the occurrence of disseminated gold deposits. Mining district data and igneous outcrop data were compiled on 1:1,000,000 maps and synthesized according to the chemical classification of mineral deposits and magma chemistry that is being developed by S. B. Keith. Three types of gold systems were found to exist in the Mojave region: calc-alkalic, alkalic, and peraluminous gold systems. The stratigraphic position and the chemical, mineralogical, and metallogenic characteristics of the three types of gold deposits are discussed comprehensively within the report.

Of the three types of gold deposits, the peraluminous and calc-alkalic gold systems are currently considered to have the most potential for disseminated, bulk-gold deposits. It is recommended that the search for these systems be concentrated in areas of low-angle thrust faults of late Mesozoic to early Tertiary age that are widespread throughout the Mojave region. Exploration should be particularly directed to areas where peraluminous or calc-alkalic magmatism (especially dikes) coincides with areas of low-angle thrust faults. Specific areas that contain such targets are elaborated in the report.

We conclude from our analysis that the Mojave region has excellent potential for disseminated gold deposits and is a worthy candidate for a serious regional gold exploration program.

START EXPLORATION WITH LITERATURE RESEARCH AND A MICROCOMPUTER

by Jan C. Wilt, Tucson, AZ, 1984

Efficient exploration begins with literature research, which consists of: Literature Search, Data Gathering, Pattern Analysis, and Prediction.

The purpose of a LITERATURE SEARCH is to obtain references. The fastest and cheapest way to get into the geologic literature is with computerized bibliographic searches through a large database supplier, such as Dialog Information Services or SDC Orbit. These systems are accessible to anyone with a microcomputer, modem, and account number and contain hundreds of databases covering topics that range from A (Agricola) to Z (Zoological Record). The most useful databases for geologic literature are GEOREF and GEOARCHIVE, with COMPENDEX, Dissertation Abstracts, NTIS, and DOE ENERGY also quite useful. Other databases that are sometimes useful for exploration are EI Engineering Meetings, METADEX, Water Resources Abstracts, PTS Prompt, Chem Abstracts, Find/SVP Reports, Nonferrous Metals Abstracts, Paperchem, Surface Coatings Abstracts, Government Printing Office Publications, National Newspaper Index, Federal Register, and Commerce Business Daily.

Literature Searching also involves locating these references, either in the library or by ordering them through the computer. The references at the end of these articles, combined with those additional references found through published bibliographies of state geologic surveys, U.S.G.S. indices to geologic mapping, guidebooks to local areas, landmark economic geology papers, and general commodity or area references of the U.S.G.S. or U.S.B.M., can more than double the total reference list. It is most convenient if these references are placed on 3"x5" cards so they can be sorted by availability, significance, or location in the library. The most efficient method is to download (with modem or communications software programs) the computerized reference list from Dialog onto a disk in a file which can later be manipulated with a word processing program by adding the additional references and printing the list with margins for index cards or manuscripts. These references can also be cross-indexed and searched by keywords in a specialized bibliographic database such as Citation.

The purpose of DATA GATHERING is to obtain significant geologic parameters, such as mineralogy, host rock, age, production data, type of deposit, and other information about each mining district or occurrence. There are several computerized data bases of geologic or mining information, but the cheapest and most complete are those by governmental agencies. CRIB (now called MRDS for Mineral Resource Data System) is supplied by the U. S. Geological Survey and contains geological information concerning mineralogy, structure, ore control, and host rocks, as well as location and commodity information. MILS (Mineral Industry Location System) is supplied by the U. S. Bureau of Mines and contains primarily location coordinates and commodities. Additional information is obtained from the literature found in the first step, county mine reports of the state geological surveys, U.S. Bureau of Mines yearbooks, U.S.G.S. maps on commodities, bulletins on the mineral and water resources of each state done in the mid 1960's, and various directories and other reference works.

Part of the process of gathering data is recording it. The old methods of recording the data on 3"x5" cards or in notebooks, file folders or large charts still work, although it is timeconsuming and cumbersome to rearrange the data according to various parameters. It is much more efficient to enter, manipulate, and rearrange the information with a data base management software program on a microcomputer. Examples of data base management systems include DBMaster for the Apple II+ computer, dBase II and Perfect Filer for CP/M computers such as the Kaypro, and Lotus 1-2-3 for 16-bit computers such as the Zenith 100, Compaq, and IBM PC. If the data is largely numeric, such as production information, it is more efficient to enter the data into a file with a spreadsheet program such as Visicalc or Multiplan, have that software program do the calculations, and then transfer the data to the data base file. The Lotus 1-2-3 software program is particularly useful because it functions as spreadsheet, data base management, and graphing programs without cumbersome transfering of files.

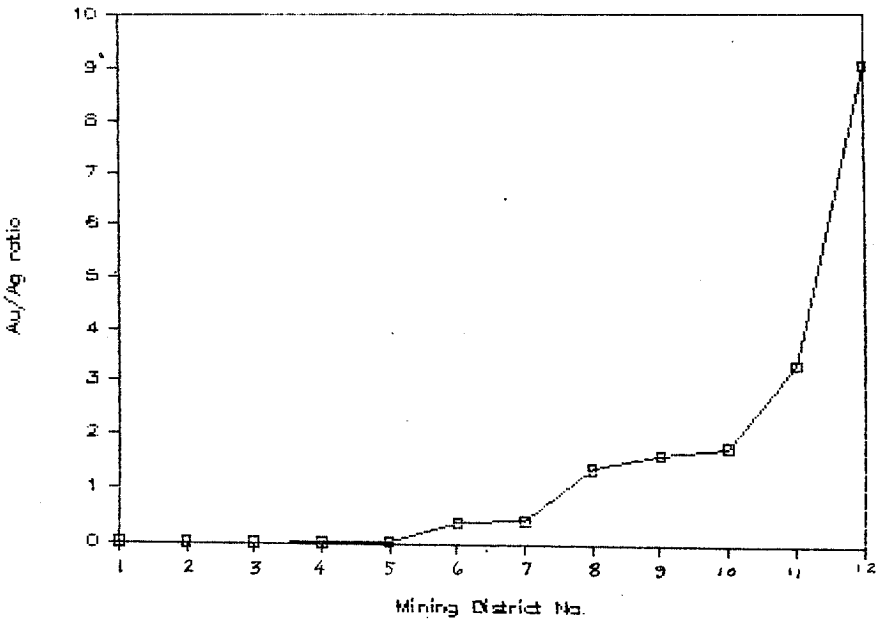
The purpose of PATTERN ANALYSIS is to discover correlations between various items of geologic data found and recorded in the previous step. Graphs, histograms, cross-sections, and maps are still effective means of analyzing patterns, and, with the appropriate software programs, these can be done on a microcomputer. The sorting function of the data base management system can rearrange the information by any of the fields, for example alphabetically by mining district or numerically by total production, by gold grade, or by gold:silver ratio. Comparing these sorted

printouts can point out where certain geologic factors, such as age of ore deposit, coincide with type of deposit, such as porphyry copper deposits, or with grade and tonnage. Overlay maps or graphs of each geologic factor, such as age or ore geochemistry, are particularly effective in isolating which geologic factors correlate with each other.

The purpose of PREDICTION, the final step in literature research, is to predict which geologic factors are the best fingerprints for the particular type of ore deposit needed. Literature research in exploration goes beyond this to predict previously undiscovered areas that contain geologic conditions similar to the those found in the pattern analysis and data gathering steps. All of these potentially economic areas, both presently known deposits found in steps two and three and those predicted in this fourth step, can then be listed in priority order so that the most promising areas can be examined in the field first. Although spreadsheet programs such as Visicalc or Multiplan have been very useful in financial predictions and planning, geological information is much more complex, less quantitative, and more equivocal. Computers can't deal with this ambiguity, so the final step of exploration with literature research rests with an experienced geologists' reasoning and intuition.

Ariz. mining dist. by Au/Ag ratio

Mining District	Au/Ag
1 Silver	0.000076
2 Silver Bell	0.000376
3 Pima	0.000482
4 Castle Dome	0.011811
5 Swansea	0.015151
6 Harquahala	0.383561
7 Union Pass (Katherine)	0.408945
8 Planet	1.333333
9 Little Harquahala	1.588888
10 Oatman	1.714036
11 Gold Basin	3.241379
12 Fortuna	9.034482



JAN C. WILT

3035 S. Shiela Ave.
Tucson, Az. 85746
(602) 883-6669

Ariz. Mining Dist. by Geol. Age

Mining District	Geol. Age	Deposit Type	Au oz	Au oz/Ton	Au/Ag
Fortuna	Paleocene	peg Au	131000	0.856	9.03448
Gold Basin	Paleocene	peg Au	9400	0.47	3.24137
Pima	lt Cretac	por cu-moly	27200	0.00003	0.00048
Silver Bell	lt Cretac	por cu-moly	2200	0.00002	0.00037
Swansea	mid-Tert	detach Cu-Au	500	0.0009	0.01515
Planet	mid-Tert	detach Cu-Au	400	0.0004	1.33333
Harquahala	mid-Tert	poly met Au	2800	0.13	0.38356
Little Harquahala	mid-Tert	poly met Au	143000	0.9	1.58888
Castle Dome	mid-Tert	vn Ag	3000	0.025	0.01181
Silver	mid-Tert	vn Ag	100	0.001	0.00007
Oatman	mid-Tert	vn Au	1966000	0.48	1.71403
Union Pass (Katherine)	mid-Tert	vn Au	128000	0.18	0.40894

Ariz. mining dist. by Au/Ag ratio

Mining District	Geol. Age	Deposit Type	Au oz	Au oz/Ton	Au/Ag	Age
Silver	mid-Tert	vn Ag	100	0.001	0.00007	
Silver Bell	lt Cretac	por cu-moly	2200	0.00002	0.00037	67 m.y.qtz monz
Pima	lt Cretac	por cu-moly	27200	0.00003	0.00048	60-69 m.y.stocks
Castle Dome	mid-Tert	vn Ag	3000	0.025	0.01181	19-20 m.y.dk swarm
Swansea	mid-Tert	detach Cu-Au	500	0.0009	0.01515	72 m.y.gran
Harquahala	mid-Tert	poly met Au	2800	0.13	0.38356	23-28 m.y.dks
Union Pass (Kathermid-Tert		vn Au	128000	0.18	0.40894	20-14 m.y. volc
Planet	mid-Tert	detach Cu-Au	400	0.0004	1.33333	
Little Harquahala	mid-Tert	poly met Au	143000	0.9	1.58888	23-28 m.y.dks
Oatman	mid-Tert	vn Au	1966000	0.48	1.71403	20-14 m.y.volc
Gold Basin	Paleocene	peg Au	9400	0.47	3.24137	
Fortuna	Paleocene	peg Au	131000	0.86	9.03448	

Ariz. mining districts by Au oz/Ton

Mining District	Geol. Age	Deposit Type	Au oz	Au oz/Ton	Au/Ag	Commodities
Silver Bell	lt Cretac	por cu-moly	2200	0.00002	0.00037	Cu,Mo,Ag,Zn,Pb,Au,(F,Ba)
Pima	lt Cretac	por cu-moly	27200	0.00003	0.00048	Cu,Mo,Ag,Au,Pb,Zn
Planet	mid-Tert	detach Cu-Au	400	0.0004	1.33333	Cu,Au,Ag
Swansea	mid-Tert	detach Cu-Au	500	0.0009	0.01515	Cu,Ag,Au,(Fe,F)
Silver	mid-Tert	vn Ag	100	0.001	0.00007	Ag,Pb,Zn,Au,Cu,(F,Ba)
Castle Dome	mid-Tert	vn Ag	3000	0.025	0.01181	Pb,Ag,Au,Zn,Cu,(F,Ba)
Harquahala	mid-Tert	poly met Au	2800	0.13	0.38356	Au,Cu,Ag,Pb,(F)
Union Pass (Katherine)	mid-Tert	vn Au	128000	0.18	0.40894	Au,Ag
Gold Basin	Paleocene	peg Au	9400	0.47	3.24137	Au,Ag,Pb,Cu
Oatman	mid-Tert	vn Au	1966000	0.48	1.71403	Au,Ag,Cu,Pb
Fortuna	Paleocene	peg Au	131000	0.86	9.03448	Au,Ag,Cu
Little Harquahala	mid-Tert	poly met Au	143000	0.9	1.58888	Au,Ag,Pb,Cu

Examples of a file sorted various ways in a Data Base Management System
from the Arizona Mining District Files of Jan C. Wilt, Tucson, AZ 883-6669

ASARCO

Southwestern Exploration Division

June 13, 1983

FILE NOTE

Computer Reference

Jan Wilt, geologist-computer user, along with Stan Keith, has put together the following tome (collected for the AGS Field Trip, Spring 1983):

> Recent References for Laramide & Sevier Tectonism in Southwestern North America in their Historical Perspectives, by J.C. Wilt and S.B. Keith, Tucson, AZ, March 27, 1983.

Attached, p. 2, is the CONTENTS of this compilation. Except for the initial volume which gives a rundown on the system used and why, it also lists all the references according to the three breakdowns in the CONTENTS and then an alphabetical author listing. All references in the CONTENTS are Xerox articles from published sources. No new or correlating thoughts by JCW or SBK -- just the articles as found.

The total work, 4 volumes, contains some 1000 pages containing some 100 references -- articles. For this the sale price is $580.00 ($500.00 time, $80.00 Xerox).

This is just one of the types of services that Jan offers with her computer search in the mineral-geology fields.

James D. Sell

James D. Sell

JDS:mek
Att.

cc: W. L. Kurtz
FRK/HGK/GJS/JRS

CONTENTS

I. CONTRASTING INTERPRETATIONS OF LARAMIDE-SEVIER TECTONIC STYLES

A. Lineament Tectonics (1902-1977)
B. Differential vertical uplift and/or extensional models (1963-1982)
C. Stratigraphic or Paleogeographic Approaches (1968-1983)
D. Wrench fault models (1902-1977)
E. Large-scale Thrust Fault Models (1947-present)
F. Small-scale Compressive Fold-thrust Models (1976-1982)

II. PLATE TECTONIC MODELS (1969-present)

A. Early Plate Tectonic Models (1969-1972)
B. Low-angle Subduction Models (1974-1980)
C. Variable-dip Subduction Models (1977-present)
D. Application of Variable Dip Subduction Models to Ore Deposit Genesis (1978-present)

III. MODERN TECTONIC MODELS (1981-present)

A. The Model (1981-present) and Evidence for Decretionary Tectonics (1973-1983)
Evidence for Decretionary Subduction
a. Insights Gleaned from the Franciscan Problem (1973-present)
b. Peninsula Range Batholith Data (1975-1982)
c. Reduced K-Ar Age Data (1975-1982)
d. Metamorphic core complexes and their tectonic significance (1977-present)
e. The 2-mica Granite Problem (1977-present)
f. Results from 1978-1982 "overthrust" petroleum play in Arizona
g. Evidence from the Colorado Plateau and Colorado-Wyoming Rockies (1981-present)
h. Recognition of Extensive Mesozoic Low-angle Faulting in SE California and western Arizona (1982-present)
i. Southwest-directed thrusting (1968-present)
B. Accretionary tectonics (1978-present)
C. Plate Rotation (1981)

ASARCO

Southwestern Exploration Division

June 13, 1983

FILE NOTE

Computer Reference

Jan Wilt, geologist-computer user, along with Stan Keith, has put together the following tome (collected for the AGS Field Trip, Spring 1983):

> Recent References for Laramide & Sevier Tectonism in Southwestern North America in their Historical Perspectives, by J.C. Wilt and S.B. Keith, Tucson, AZ, March 27, 1983.

Attached, p. 2, is the CONTENTS of this compilation. Except for the initial volume which gives a rundown on the system used and why, it also lists all the references according to the three breakdowns in the CONTENTS and then an alphabetical author listing. All references in the CONTENTS are Xerox articles from published sources. No new or correlating thoughts by JCW or SBK -- just the articles as found.

The total work, 4 volumes, contains some 1000 pages containing some 100 references -- articles. For this the sale price is $580.00 ($500.00 time, $80.00 Xerox).

This is just one of the types of services that Jan offers with her computer search in the mineral-geology fields.

James D. Sell

James D. Sell

JDS:mek
Att.

cc: W. L. Kurtz
FRK/HGK/GJS/JRS

CONTENTS

I. CONTRASTING INTERPRETATIONS OF LARAMIDE-SEVIER TECTONIC STYLES

A. Lineament Tectonics (1902-1977)
B. Differential vertical uplift and/or extensional models (1963-1982)
C. Stratigraphic or Paleogeographic Approaches (1968-1983)
D. Wrench fault models (1902-1977)
E. Large-scale Thrust Fault Models (1947-present)
F. Small-scale Compressive Fold-thrust Models (1976-1982)

II. PLATE TECTONIC MODELS (1969-present)

A. Early Plate Tectonic Models (1969-1972)
B. Low-angle Subduction Models (1974-1980)
C. Variable-dip Subduction Models (1977-present)
D. Application of Variable Dip Subduction Models to Ore Deposit Genesis (1978-present)

III. MODERN TECTONIC MODELS (1981-present)

A. The Model (1981-present) and Evidence for Decretionary Tectonics (1973-1983)
Evidence for Decretionary Subduction
a. Insights Gleaned from the Franciscan Problem (1973-present)
b. Peninsula Range Batholith Data (1975-1982)
c. Reduced K-Ar Age Data (1975-1982)
d. Metamorphic core complexes and their tectonic significance (1977-present)
e. The 2-mica Granite Problem (1977-present)
f. Results from 1978-1982 "overthrust" petroleum play in Arizona
g. Evidence from the Colorado Plateau and Colorado-Wyoming Rockies (1981-present)
h. Recognition of Extensive Mesozoic Low-angle Faulting in SE California and western Arizona (1982-present)
i. Southwest-directed thrusting (1968-present)
B. Accretionary tectonics (1978-present)
C. Plate Rotation (1981)

3/26

CIRCULATE:

~~S. A. CATLIN~~
~~J. H. COURTRIGHT~~
~~R. B. CRIST~~
~~F. B. KOUK~~
~~H. G. KREIS~~
W. L. KURTZ
~~F. MICHEL~~
~~B. MASON~~

~~W. D. PAYNE~~
~~G. W. PICKARD~~
~~J. D. SELL~~
~~G. J. STATHIS~~
~~P. G. VIKRE~~
~~N. P. WHALEY~~

WDP — circulate to geologists

Jan C. Wilt
3035 S. Shiela Ave.
Tucson, Az. 85706
(602) 883-6669

February 15, 1982

Computer Literature Search

William L. Kurtz
ASARCO Inc.
1150 N. 7th Ave.
Tucson, AZ 85703

Dear Bill,

I hope you have a very happy and prosperous New Year! The year and a half since I left the Arizona Bureau of Geology and Mineral Technology have been very busy for me. I've expanded my professional library to include the other western states and have recently acquired a small computer that will help me access scientific literature faster, more thoroughly and more efficiently.

The best computer geologic bibliographies are GEOREF and GEOARCHIVE. GEOREF now contains over 700,000 geologic references, mostly since 1967, and will contain over a million entries by June, 1982, when the pre-1967 references are added. In my experience computer printouts from the library give about 30 to 50% of the references important to a particular subject with traditional search methods obtaining the remainder.

With my own computer I am able to increase that percentage by continuing to access GEOREF data throughout the research. For example, there are 265 references in GEOREF that contain both the key words gold and California. Most librarians would stop at this point and give you a list of those references. By reading some of the important articles and searching the data bases for the mining districts that produced gold and for terms such as epithermal veins or precious metals, many more references are found. Having a geologist do the searching gives a more effective result.

In addition to GEOREF and GEOARCHIVE I can access 170 other data bases ranging from agriculture to zoology. The data bases that I expect to be most useful include Chemical Abstracts, Engineering Information, Disclosure (SEC reports such as 10-K), DOE Energy, Environmental Bibliography, Federal Register, Foundation Directory, Government Printing Office Publications, Magazine Index, NTIS (National Technical Information Service of the U. S. Department of Commerce), and PTS Predicasts Overview of Markets and Technology. It is also possible to order photocopies of any article or book whose reference is listed in the data bases. When ordered through the computer, many of them arrive within 3 to 7 days.

RECEIVED
FEB 23 1982
EXPLORATION DEPARTMENT

This last year I've been dealing mainly with precious and strategic metals, some petroleum, and a little uranium and nonmetallics. Although I've done some mapping, structural studies, and stratigraphy, most of my projects have involved the same type of research that I did at the Arizona Bureau of Geology, but in all the western states and some foreign countries.

In the study of molybdenum occurrences in Arizona for the Bureau, I first surveyed the general mineral and metal occurrence literature to obtain a quick list of pertinent localities and mining districts. Then I obtained bibliographies for each of these mining districts and read the most important articles, recording important geologic factors in a chart. This I can now do on my computer.

The finished list of molybdenum in Arizona contained ten times the number of occurrences of any previous list. The initial data gathering was followed by the most important step - in which I analyzed the data and discovered a correlation between certain geologically significant factors and economic occurrences. (See Fieldnotes article of July, 1980.) With this type of correlation I can predict areas where other economic occurrences could be expected. The descriptive information on Arizona molybdenum occurrences will be published as a U.S. Geological Survey open-file report, as soon as I finish editing the final draft.

Although the last year has been very busy, there are always methods to expand. This year I've been able to do that even when I'm busy by hiring technically qualified people, such as graduate students or other scientific professionals to assist in gathering data under my supervision. This results in a lower cost for my clients and allows me more opportunity to analyze the occurrences and search for the patterns that predict potential discoveries, which is the part of the research that I find most exciting.

Sincerely yours,

Jan C. Wilt

Jan C. Wilt

P.S. I've included my updated resume for your files.

Jan C. Wilt
3035 S. Shiela Ave.
Tucson, Az. 85706
(602) 883-6669

RESUME

GEOLOGIC EXPERTISE

Mineral Exploration - strategic minerals, precious metals, molybdenum, porphyry copper, massive sulfides, fluorite - in western U.S.
Energy Exploration - uranium, oil, gas, and coal - in Arizona.
Bibliographies and sources of geologic information, - U. S. and foreign.
Mineralogy, paleontology, tectonics and geochronology of Arizona.
Permian and Cenozoic stratigraphy and paleoenvironments in Arizona.

GEOLOGIC EXPERIENCE

Consultant: AMAX, Ariz. Dept. Mineral Resources, Ariz., Public Service, Essex, J. David Lowell, New Jersey Zinc, Noranda, Superior, U.S. Borax; 1974-77, 1980-81.
Assistant Geologist: Arizona Bureau of Geology and Mineral Technology (formerly Ariz. Bur. of Mines); 1978-80, 1969-71.
Associate Curator: Mineral Museum, Geosciences Dept., University of Arizona, 1975-79.

GEOLOGIC SKILLS

Analyze and interpret geologic data
Recommend exploration targets
Acquire geologic data and analyze its significance
Prepare list of references for exploration program
Edit and publish manuscripts and project reports
Visualize geologic events in space and time
Communicate essential information efficiently and effectively

TEACHING EXPERIENCE

Geology: Univ. of Ariz., 1977-78; Pima Comm. Coll., 1975-77.
Science: Pueblo High School, Tucson, 1967-69.

PROFESSIONAL ORGANIZATIONS

Treasurer, Southwestern Mineral Exploration Association, 1981.
Organizer of Tucson chapter of Association of Women Geoscientists, Vice-President, 1980; Finance Chairman, 1981.
Officer, Arizona Geological Society, Councillor, 1979-80; Secretary, 1976-77; Assistant Secretary, 1975-76.
Vice-president, Tucson Miniature Society, 1981.
Tucson Public Service Network, Metropolitan Amateur Radio Club, license N7BLJ (formerly KØYTX), 1979-80.
Board of Directors, Secretary, Tucson Museum of Science and Industry, 1977-79.
President, Millstone Manor Property Owners Association, 1969-77.

EDUCATION

University of Arizona, Tucson, Arizona
B.S. Geology - May 1965; M.S. Geology - January, 1969.

Jan C. Wilt
3035 S. Shiela Ave.
Tucson, Az. 85706
(602) 883-6669

PUBLICATIONS

MINERAL RESOURCES

Molybdenum in Arizona, 1980, by Jan C. Wilt and Stanley B. Keith: Fieldnotes from the State of Arizona, Bureau of Geology and Mineral Technology, v. 10, no. 3, p. 1-3, 7-9, 12.

Molybdenum Occurrences in Arizona, Preliminary Report, 1982 in press, by Jan C. Wilt and Stanley B. Keith: U.S. Geological Survey open-file report.

Arizona molybdenum minerals as keys to metallogenic types (abs.), 1980, by J. C. Wilt: Geol. Soc. America, Abstracts with Programs, v. 12, no. 6, p. 309.

ENERGY RESOURCES

Cenozoic sediments, volcanics and related uranium in the Basin and Range Province of Arizona, 1981, by J. C. Wilt and R. B. Scarborough: Am. Assoc. Petroleum Geologists Studies in Geology No. 13, Philip C. Goodell and Aaron C. Waters, editors, Uranium in volcanic and volcaniclastic rocks.

A study of uranium favorability of Cenozoic sedimentary rocks, Basin and Range Province, Arizona, Part I, General geology and chronology of pre-late Miocene sedimentary rocks, 1979, by R. B. Scarborough and J. C. Wilt: U.S. Geol. Survey open-file report, 79-1429, 101 p..

Coal, oil, natural gas, helium, and uranium in Arizona, 1970, by H. W. Peirce, J. C. Wilt, and S. B. Keith: Arizona Bureau of Mines, Bull. 182, 289 p., maps.

FIELD TRIPS

Road logs for New Mexico Geological Society Guidebook, 1978, Day 1 by S. B. Keith, J. C. Wilt, E. Deal, R. Clemons, D. Lynch, and J. Forrester; Day 2 by S. B. Keith and J. C. Wilt; Exit log by S. B. Keith and J. C. Wilt; Supplemental log to Nogales by S. B. Kieth, N. E. Lehman, J. Sell, and J. C. Wilt: Land of Cochise, 29th guidebook.

Colorado Plateau field trip, 1979, by J. C. Wilt: Ariz. Bur. Geology and Mineral Technology, Fieldnotes, v. 9, no. 4, p. 3, 8.

EDITOR

Land of Cochise, the geology of southeastern Arizona, 1978, edited by J. Callender, J. C. Wilt, and R. Clemons: New Mexico Geol Soc., 29th guidebook, 372 p..

Tectonic Digest, 1976, edited by J. C. Wilt and J. P. Jenney: Ariz. Geol. Soc. Digest, v. 20, 430 p.

BIBLIOGRAPHY

Bibliography of the geology and mineral resources of Arizona, 1965-1970, 1974, by J. C. Wilt and J. S. Vuich: Ariz., Bur. Mines Bull. 190, 155 p..

STRATIGRAPHY AND SEDIMENTATION

Petrology and stratigraphy of the Colina Limestone (Permian) in Cochise County, Arizona, 1969, By J. C. Wilt: Univ. Ariz., unpub. M.S. thesis, 117 p..

Fossils of Arizona, 1971, by J. C. Wilt: Ariz. Bur. Geology and Mineral Technology, open-file report 71-1, 400 p..

Sedimentologic studies in the Willcox Playa area, Cochise County, Arizona, 1972, by J. F. Schreiber, Jr., G. L. Pine, B. W. Pipkin, R. C. Robinson, and J. C. Wilt, in C. C. Reeves, Jr., ed., Playa Lake Symposium: International Center for Arid and Semi-arid Land Studies, Lubbock, Texas, Pub. No. 4, p. 133-184. Some features of wind deposits near Wilcox Playa, Cochise County, Arizona, 1965, by J. C. (Rasmussen) Wilt: Univ. Ariz., unpub. Honors thesis, 41 p..

and we have other data on this but filed where?

ASARCO

Northwestern Exploration Division
John C. Balla
Manager

May 28, 1982

Mr. W.L. Kurtz, Manager
Western U.S.A.
ASARCO Incorporated
P.O. Box 5747
Tucson, AZ 85703

PROSPECTOR - A Computer
Based Consultation System
for Mineral Exploration

Dear Mr. Kurtz:

Fortune magazine has been running a series of articles on Thinking Machines. The May 31, 1982 issue of Fortune contains an article on artificial intelligence. As an example of a successful program in artificial intelligence, Fortune published a brief article entitled "Putting a Prospector into the Computer," copy attached. I am familiar with the Prospector program, having attended various meetings held by SRI International on Prospector, the developer of the program.

Reference is made to the attached article, and the mention that it allegedly predicted the location of a molybdenum deposit at Mt. Tolman, Washington. (Prior to the property being called Mt. Tolman, it was generally known as the San Poil project). This allegation is either false, or represents the ability to predict the location of a mineral deposit, knowing beforehand its location.

As you know, I am particularly familiar with the Mt. Tolman deposit, having been in charge of the San Poil project (as Mt. Tolman was then called) for Bear Creek Mining Company from 1967 to 1969. Up to 1976, when the Prospector program was initiated, Bear Creek had drilled 105 drill holes into the Mt. Tolman deposit, with expenditures being in excess of one million dollars. A major low grade molybdenum deposit had been outlined.

RECEIVED
JUN - 1 1982
EXPLORATION DEPARTMENT

In 1978 the Bear Creek data was made available by the Colville tribe to interested parties who might be interested in leasing the property. Mr. Vic Hollister and Dr. Allen Campbell (son of Neil Campbell) took the Bear Creek data, with the company's approval, and developed for SRI International "The Campbell/Hollister Porphyry Molybdenum Drilling-Site Selection Models." (For

detailed discussion, see the Final Report by SRI entitled "The Prospector System for Mineral Exploration," dated April, 1980).

Based upon the above historical review, it is clear that Prospector did not predict a new deposit, but merely predicted what was already known to exist, using the data from that deposit.

The importance of the article, and the Prospector program, lies in the application of the program. This is a matter of real concern, as outlined in my letter of November 16, 1978. There is no need to reiterate my previous comments, but we should be careful if the U.S.G.S. is using Prospector to evaluate the mineral resources of wilderness areas.

Very truly yours,

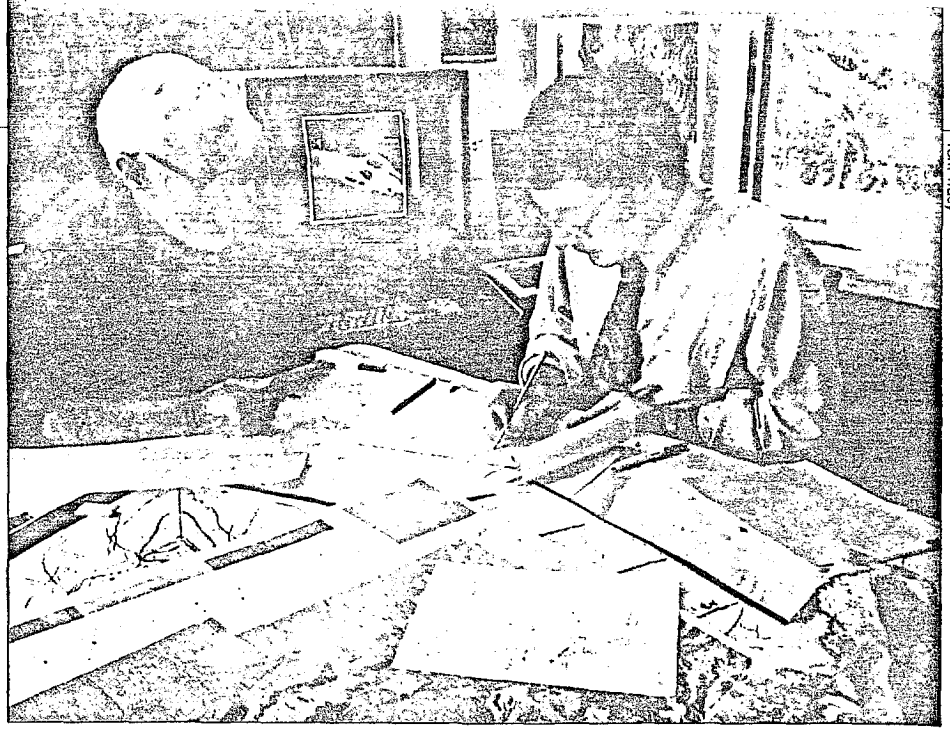
John C. Balla

John C. Balla

JCB/dt
Enclosure

cc: R.L. Brown with enclosure
J.D. Sell " "
D.M. Smith " "
Peter Vikre " "

Putting a Prospector into the Computer



Tom Tracy

Geologist Dennis Cox (left) explains the secrets of his trade to knowledge engineer Rene Reboh, who is constructing an expert system to prospect for minerals. Reboh is incorporating into the program's "inference network" the subjective rules Cox follows in looking for volcanically formed copper deposits.

A valuable contribution of artificial intelligence may be to help stem the loss of one of society's most priceless assets, expert knowledge. Much of that knowledge can't be passed along to posterity in classrooms because it comes only from years of on-the-job experience. The most proficient experts tend to be those with the fewest productive years left.

Builders of expert systems believe they can preserve that knowledge by encoding it into computer programs. Then anyone with access to a computer terminal could consult a program that combined expertise from dozens of leading authorities in any given field.

One of the most ambitious—and apparently successful—of these projects is Prospector, which has been under development at SRI Institute since 1976. The U.S. Geological Survey is now financing the project, with the aim of using it to help in the huge job of assessing the mineral resources in the nation's wilderness areas. Prospector has so far tapped the expertise of some 20 leading specialists in economic geology, the science—or art—of divining underground mineral deposits from surface indications. The system has already gained a measure of fame by predicting the location of a molybdenum deposit on Washington's Mt. Tolman—a prediction that has since been confirmed by drilling.

The unreadable book

A whole new profession, "knowledge engineering," has emerged to extract the deep understanding experts call on in making judgments or following hunches. Knowledge engineers convert this understanding into thousands of rules a computer can follow. In theory, they could write the rules in a book, but who could stand to read and memorize it?

Participating in the knowledge-cloning process can be simultaneously agonizing and stimulating for the expert. One who has recently undergone the ordeal is USGS geologist Dennis Cox, 52, an authority on porphyry copper deposits in Arizona, Alaska, and Puerto Rico. Cox recently spent several days at SRI being interviewed by knowledge engineer Rene Reboh, 38, now head of the Prospector project.

For the expert, the painful part lies in trying to translate the thought patterns and subjective judgments acquired in a lifetime of hard-rock prospecting into simpleminded rules digestible by a computer. Rarely is anything certain in prospecting: virtually all the rules must be assigned levels of probability, and special cases often crop up where rules don't apply. Therefore geologists and knowledge engineers perfect the systems through a long process of rewriting and adding rules, running the program on trial data, and judging whether the computer's prediction matches the expert's.

Most experts also regard the experience as intellectually rewarding. Says Cox, "As the model is worked and reworked, all kinds of inconsistencies and contradictions in your own thinking pop out. You have to rethink, refine your own reasoning process, examine yourself again. It really clarifies your thought."

Far from resenting the possibility that Prospector may make his hard-won knowledge available to anyone with a computer terminal, Cox says he'll be happy if it relieves him from having to sit in the rain in Alaska making elementary assessments. "This way I disseminate my knowledge to lots more people," he says, "and it will give me a lot more time for productive research—for answering the *hard* questions."

Dons – other brochures in Tucson describing programs if you want to see them

~~took to PTG~~

(1)

ASARCO

copy for Dons 1) Balla 16
2) Exploration Program
3) Balla 16 May 19, 1978

RECEIVED
NOV 20 1978
EXPLORATION DEPARTMENT

Northwestern Exploration Division
John C. Balla
Manager

November 16, 1978

Mr. W.L. Kurtz, Manager
Western USA
ASARCO Incorporated
P.O. Box 5747
Tucson, Arizona 85703

PROSPECTOR - A Computer Based Consultation System for Mineral Exploration

Dear Sir:

On November 2, I attended a seminar at S.R.I. International on Prospector. The purpose of the seminar was to review the status of Prospector, after two years of development, and consider the future utilization of the program. For a review of Prospector, see my letter of May 16, 1978 to you.

Prospector is an outgrowth of a computer-based consultation system developed for medical diagnosis (based upon blood samples). It was originally funded by the U.S.G.S., National Science Foundation, and Exxon. Future funding will be solely from the U.S.G.S., Office of Resource Analysis.

So far, eight or so scientists have worked on Prospector over the last two years, all on a part-time basis. The total cost to date has been about $550,000. For the next two year period, the budget is $700,000. It should be mentioned that there are no geologists at S.R.I. who are working on Prospector, they are all computer scientists. All geological input is from outside sources (professors, consultants).

Dr. Richard F. Meyer is in charge of the Prospector program for the Office of Resource Analysis of the U.S. Geological Survey at Reston, Virginia. He is not a minerals geologist. Dr. Meyers stated that prospector was, and is, being developed to fulfill the needs of Washington (Administration, Congress). The program is in fact being developed to allow geological appraisal and political decisions on land use planning. This potential was described in my letter of May 16, 1978 to you. Dr. Meyer anticipates that Prospector will be operational in about 10 years. At that time, as the scenario is envisioned, Prospector will be able to "rate" the mineral potential of any given area (such as the B.L.M.'s roadless area), based upon the geologic characteristics of the mineral deposits stored in its "memory", and the geologic characteristics of the area.

There are two fundamental weak spots in the scenario. The first is in regards to the geologic models presently within prospector:

1. Kuroko-type massive sulfide deposits
2. Pb-Zn in carbonate rocks
3. Porphyry copper deposits
4. Cu in sandstone
5. Komatiite Ni deposits
6. U in sandstone

Quite obviously, there are a large number of geologic models which are not yet in Prospector, plus the variations on these models. In addition, new types of deposits keep being found, and eventually new models are developed for these deposits. Thus, Prospector does not have, and in all probability, will not have, most of the geologic models that are presently known, or being designed, much less all of the possible models. Hence, if Prospector is used, it will downgrade the mineral potential of an area because it does not have all of the geologic models in it.

Could work both ways on resource (not reserve) appraisals.

The second fundamental weakness is in regards to the geologic data base of the area being considered. What will be done is that the best available geologic maps of that area, no matter how old or on what scale, will be digitized into the computer, and interfaced with the geologic models. Prospector will then rate the mineral potential of that area, based upon the geologic characteristics of the models in its memory.

We are all familiar with the uneven quality of the geologic mapping which has been done by the U.S.G.S. over the years, and it need not be elaborated upon here. In order, in part, to rectify the problem, the U.S.G.S. has embarked upon the CUSMAP Program. This program is designed to "... gather, interpret, and disseminate information on non-energy mineral resources on an areal basis for land-use planning and resource management." Attached is a U.S.G.S. report entitled "Essential Goals and Minimum Product of the CUSMAP Program", which will be of interest. At the present time, there are 12, 2° quadrangles being mapped and compiled. It will take three years to complete each 2° sheet.

The U.S.G.S. does not know how many of the CUSMAP sheets will be completed. If it is not completed, then we will still have the same problem of uneven geologic data. If it is completed, and if all 2° sheets in the mountainous West are done, at the present rate, then it will take 60 years to complete the program (there are about 180 2° sheets covering the area from Denver to the Pacific Ocean, and from the Canadian border to the Mexican border). I personally doubt that the program will, in fact, be completed.

In conclusion, the Prospector program is a very sophisticated program which, unfortunately, has two fundamental defects. These two defects

notwithstanding, Prospector may well be utilized by the Government, or Congress, in future land use planning. Should this occur, we should point out to those involved in making decisions the deficiencies inherent in the program.

Yours very truly,

John C. Balla

John C. Balla

JCB/mc
Enclosures

cc: TCOsborne w/enc.

ATTENDEES EXPECTED AT PROSPECTOR SEMINAR
SRI INTERNATIONAL
NOVEMBER 2, 1978

Dr. Samuel S. Adams
3030 Third Street
Boulder, Colorado 80302

Mr. George Argall
World Mining
500 Howard Street
San Francisco, California 94105

Dr. Howard Austin
Schlumberger-Doll Research Center
Schlumberger
Old Quarry Road
Ridgefield, Connecticut 06887

Mr. Paul A. Bailly
Occidental Minerals Corporation
777 South Wadsworth Blvd.
Lakewood, Colorado 80226

Mr. John C. Balla
ASARCO Incorporated
E. 920 Wolverton Court
Spokane, Washington 99207

Ms. Phyllis Barrett
Artificial Intelligence Center
SRI International
333 Ravenswood Ave.
Menlo Park, California 94025

Mr. W. B. Beatty
SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

Tom Bell
Department of Geology
University of California
Berkeley, California 94720

Henry C. Berg
U.S. Geological Survey
345 Middlefield Road
Menlo Park, California 94025

Jesus A. Bilbao
Petroleos De Venezuela, S.A.
Apt. 169
Caracas 105
VENEZUELA

Thomas J. Blair
Regional Economic Evaluation Unit
U.S. Geological Survey
Denver, Colorado 80225

David M. Blanchfield
U.S. Department of Energy
P.O. Box 2567
Grand Junction, Colorado 81501

Byron Brandt
CONOCO
1000 South Pine
Ponca City, Oklahoma 74601

David A. Brew
U.S. Geological Survey
345 Middlefield Road
Menlo Park, CA 94025

Dr. Peter L. Briggs
Atlantic Richfield Company
P.O. Box 2819
Dallas, Texas 75221

Prof. William Brigham
Petroleum Engineering
Stanford University
Stanford, CA 94305

Dr. Alan N. Campbell
Box 2542
Smithers, British Columbia V0J 2N0
CANADA

Prof. Haber Cinco
Petroleum Engineering
Stanford University
Stanford, California 94305

Mr. Ritchie B. Coryell
National Science Foundation
1800 G Street, NW Room 1149
Washington, DC 20550

Mr. Merle E. Crew
Resource Assessment
Department of Energy
P.O. Box 2567
Grand Junction, Colorado 81501

Dr. John T. Cumberlidge
Hanna Mining Co.
100 Erieview Plaza
Cleveland, Ohio 44114

Dr. Richard O. Duda
Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

Frank M. Eckerson
Department of Energy
P.O. Box 2567
Grand Junction, Colorado 81501

Mr. Paul I. Eimon
Chevron Resources Company
P.O. Box 599
Denver, Colorado 80201

Dr. Marco T. Einaudi
Department of Earth Sciences
Stanford University
Stanford, California 94305

Mr. Richard Engelder
Bendix Field Engineering Company
P.O. Box 1569
Grand Junction, Colorado 81501

Ms. Joan Engles
Engineering Sciences Laboratory
SRI International
333 Ravenswood Ave.
Menlo Park, California

Mr. Warren I. Finch
U.S. Geological Survey
P.O. Box 25046, MS 916
Denver, Colorado 80225

Dr. Martin J. Fischler
Artificial Intelligence Center
SRI International
333 Ravenswood Ave.
Menlo Park, California

Dr. Michael Foose
Eastern Mineral Resources
U.S. Geological Survey
Reston, Virginia 22092

Mr. H. S. Peter Fowler
6633 Colton Boulevard
Oakland, California 94611

Mr. William Frawley
Schlumberger-Doll Research Center
Schlumberger
Old Quarry Road
Ridgefield, Connecticut 06887

Mr. Peter Friedland
Department of Computer Science
Stanford University
Stanford, California 94305

Dr. John Gaschnig
Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

Mr. R. Grabowski
Woodward-Clyde Consultants
3 Embarcadero Center, Suite 700
San Francisco, California 94111

K. A. Grace
David S. Robertson & Associates, Inc.
1658 Cole Blvd., #200
Golden, Colorado 80401

Mr. Roger Harris
Texasgulf Inc.
High Ridge Park
Stamford, Connecticut 06904

Dr. Peter E. Hart
Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

Mr. L. W. Heiny
Continental Oil Company
555 17th St.
Denver, Colorado 80202

Mr. J. J. Henley
U.S. Geological Survey
12201 Sunrise Valley Drive
Reston, Virginia 22092

Mr. John Houghton
Massachusetts Institute of Technology
545 Technology Square
Cambridge, Massachusetts 02139

Mr. Victor F. Hollister
Duval International Corporation
844 West Hastings Street
Vancover, British Columbia V6C 1C8
CANADA

Mr. Stewart A. Jackson
Houston Oil & Minerals
222 Milwaukee
Denver, Colorado 80206

Mr. William R. James
U.S. Geological Survey
Reston, Virginia 22092

Daniel A. Jobin
Conservation Division
U.S. Geological Survey
P.O. Box 25046, MS 602
Denver, Colorado 80225

Ms. Maureen Johnson
U.S. Geological Survey
345 Middlefield Road
Menlo Park, California 94025

Mr. William Jordan
Houston Oil and Minerals Corporation
222 Milwaukee Street
Denver, Colorado 80206

Mr. G. F. Koehler
Cominco American Inc.
E. 15120 Euclid Road
P.O. Box 3087
Spokane, Washington 99216

Mr. Kurt Konolige
Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

Mr. John O. Kork
Office of National Resources
U.S. Geological Survey
Lakewood, Colorado 80225

Prof. Robert L. Kovach
School of Earth Sciences
Stanford University
Stanford, California 94305

Dr. Frederick C. Kruger
145 Wildwood Way
Woodside, California 94062

Mr. Ram Kulkarni
Woodward-Clyde Consultants
3 Embarcadero Center, Suite 700
San Francisco, California 94111

Dr. Philippe Lacour-Gayet
Schlumberger-Doll Research Center
Old Quarry Road
Ridgefield, Connecticut 06877

Dr. Robert D. Leighty
U.S. Army Engineering Topographic Laboratories
Fort Belvoir, Virginia 22060

Mr. Robert Lupe
U.S. Geological Survey
P.O. Box 25046, MS 916
Denver, Colorado 80255

Dr. Edward MacKevett
U.S. Geological Survey
345 Midlefield Drive
Menlo Park, California 94025

Mr. Richard P. McCammon
U.S. Geological Survey
Reston, Virginia 22092

Mr. T. P. McCann
Shell Oil Company
P.O. Box 2099, Suite 1160
Houston, Texas 77001

Mr. Gregory E. McKelvey
Cominco American Incorporated
E. 15120 Euclid
SpokAne, Washington 99216

Mr. David Menzie
U.S. Geological Survey
345 Middlefield Road
Menlo Park, California 94025

Dr. Richard F. Meyer
Office of Resource Analysis
U.S. Geological Survey
Reston, Virginia 22092

Dr. Donald Michie
Department of Computer Science
Stanford University
Stanford, California 94305

Mr. David Mickle
Bendix Field Engineering Corporation
P.O. Box 1569
Grand Junction, Colorado 81501

Dr. Anthony J. Naldrett
Department of Geology
University of Totonto
Toronto, Ontario M5S 1A1
CANADA

Dr. Nils J. Nilsson
Artificial Intelligence Center
SRI International
Menlo Park, California 94025

Ing. Jesus Ojeda-Rivera
Cia. Minera Autlan, S.A. de C.V.
Mariano Escobedo 510
Mexico 5, D.F.
MEXICO

Mr. Thomas Ovenshine
U.S. Geological Survey
345 Middlefield Road
Menlo Park, California 94025

Mr. James Padian
Petroleos De Venezuela, S.A.
Apt. 169
Caracas 105
VENEZUELA

Mr. Harry M. Parker
Fluor Utah, Inc.
177 Bovet Rd.
San Mateo, California 94402

Mr. William M. Pennell
Getty Oil Company
3810 Wilshire Blvd.
Los Angeles, California 90010

Mr. William Peppin
Mackay School of Mines
University of Nevada
Reno, Nevada 89557

Mr. John S. Philips
Chevron Resources Company
P.O. Box 599
Denver, Colorado 80201

Dr. Ruffin I. Rackley
2651 South Chase Lane
Denver, Colorado 80227

Prof. Henry Ramey
Department of Petroleum Engineering
Stanford University
Stanford, California 94305

Mr. Rene Reboh
Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

Dr. Charles A. Rosen
Artificial Intelligence Center
SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

Prof. Subir Samyal
Department of Petroleum Engineering
Stanford University
Stanford, California 94305

Mr. Sky Schaff
Woodward-Clyde Consultants
Three Embarcadero Center
San Francisco, California 94111

Mr. George B. Secor
Mineral Resources Group
The Anaconda Company
Denver, Colorado 80217

Dr. Richard P. Sheldon
Resource Systems Institute
U.S. Geological Survey
Honolulu, Hawaii 96848

Dr. Allistair Sinclair
Department of Geology
University of British Columbia
Vancouver, British Columbia
CANADA

Dr. Donald A. Singer
U.S. Geological Survey
345 Middlefield Road
Menlo Park, California 94025

Mr. Jonathan Slocum
SRI International
333 Ravenswood Avenue
Menlo Park, California 94025

Mr. Shea Clark Smith
Houston Oil and Minerals Corporation
222 Milwaukee Street
Denver, Colorado 80206

Mr. David A. Sterling
Liberty Center Complex, Suite 540
350 South Center Street
U.S. Department of Energy
Reno, Nevada 89501

Ing. Eugenio Tavera-Amezcua
Cia. Minera Autlan, S.A. de C.V.
Mariano Escobedo 510
Mexico 5, D.F.
MEXICO

Mr. Paul Taylor
Houston Oil and Minerals Corporation
222 Milwaukee Street
Denver, Colorado 80206

Dr. J. Martin Tenenbaum
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

Ted Theodore
U.S. Geological Survey
345 Middlefield Road
Menlo Park, CA 94025

Mr. Edwin W. Tooker
U.S. Geological Survey
345 Middlefield Road
Menlo Park, California 94025

Yakov Vinkovetsky
Exxon Production Research Co.
P.O. Box 2189
Houston, Texas 77001

Dr. Donald E. Walker
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

Dr. K. D. Watson
Department of Earth and Space Sciences
University of California
Los Angeles, California 90024

Mr. David White
Exxon Production Research Company
3120 Buffalo Speedway
P.O. Box 2189
Houston, Texas 77001

B. Michael Wilber
SRI International
333 Ravenswood Avenue
Menlo Park, CA 94025

Mr. John M. Worden
EXXON Production Research Company
3120 Buffalo Speedway
P.O. Box 2189
Houston, Texas 77001

Publications Concerning the SRI PROSPECTOR System

The technical papers listed below concern the Prospector project. To receive copies of any of these reports, please complete the following and return to the box indicated for this purpose in the rear of the seminar room.

Name: _____

Address: _____

_____

_____

Check

1. R. O. Duda, P. E. Hart, and N. J. Nilsson, "Subjective Bayesian Methods for Rule-Based Inference Systems," National Computer Conference 1976 (AFIPS Conference Proceedings Vol. 45), pp. 1075-1082, 1976. ____

   ---Description of the subjective Bayesian inference procedures used in Prospector. (31 pages)

2. R. O. Duda, P. E. Hart, N. J. Nilsson, and G. L. Sutherland, "Semantic Network Representations in Rule-Based Inference Systems," Technical Note 136, Artificial Intelligence Center, SRI International, Menlo Park, California, March 1977. Appears also in Pattern Directed Inference Systems, D. A. Waterman and F. Hayes-Roth (eds.), Academic Press, New York, 1978. ____

   ---General description of partitioned semantic networks as a computer formalism for representing judgmental knowledge, as applied to the Prospector system. (31 pages)

3. R. O. Duda, P. E. Hart, and R. Reboh, "A Rule-Based Consultation Program for Mineral Exploration," Proceedings of the Lawrence Symposium on Systems and Decision Sciences, pp. 306-309, Berkeley, California, October 2-4, 1977. ____

   ---Brief description of the Prospector system for system scientists.

4. R. O. Duda, P. E. Hart, N. J. Nilsson, R. Reboh, J. Slocum, and G. L. Sutherland, "Development of a Computer-Based Consultant for Mineral Exploration," Annual Report, Artificial Intelligence Center, SRI International, Menlo Park, California, October 1977. ____

   ---Extensive description of the Prospector system and the probabilistic theory on which it is based. Complete documentation of three ore deposit models: a Kuroko-type Massive Sulfide model, a Mississippi-Valley-Type Carbonate Lead/Zinc model, and a Type-A Porphyry Copper model. (202 pages)

5. P. E. Hart, R. O. Duda, and K. Konolige, "A Computer-Based Consultant for Mineral Exploration," Second Semiannual Report, Artificial Intelligence Center, SRI International, Menlo Park, California, April 1978.

---Documentation of a Komatiite-Nickel-Sulfide Model and a "Drilling Site Selection" model for porphyry copper. Description of system improvements since the October 1977 Annual Report. (17 pages plus appendices)

6. R. O. Duda, P. E. Hart, P. Barrett, J. G. Gaschnig, K Konolige, R. Reboh, and J. Slocum, "Development of the Prospector Consultation System for Mineral Exploration," Final report, Artificial Intelligence Center, SRI International, Menlo Park, California, October 1978.

*** Available in rear of Seminar room ***

---Documentation of substantial revisions to the Porphyry Copper model, the Nickel Sulfide model, and the Drilling-Site Selection model. Application of the latter to two test sites. Preliminary performance analysis of Porphyry Copper and Nickel Sulfide models. Description of several system improvements and extensions, including map input and network compiler facilities used with the Drilling-Site Selection model. Description of the model building process. (193 pages)

7. P. E. Hart and R. O. Duda, "PROSPECTOR -- A Computer-Based System for Mineral Exploration," to appear in the Journal of the International Association of Mathematical Geology, Vol. 10, Nos. 5-6.

---Early general description of the Prospector system.

8. P. E. Hart, R. O. Duda and M. T. Einaudi, "A Computer-Based Consultation System for Mineral Exploration," to appear in Computer Methods for the 80's, Society of Mining Engineers of the AIME, 1978.

---Current general description of the Prospector system. (37 pages)

ESSENTIAL GOALS AND MINIMUM PRODUCT OF THE CUSMAP PROGRAM

The immediate objective of the Mineral Resources Assessment Program is to gather, interpret, and disseminate information on non-energy mineral resources on an areal basis for land-use planning and resource management. The Conterminous States element of this Program provides systematic assessments of the mineral resource potential of selected 2° 1:250,000 scale (1 inch equals approximately 4 miles) quadrangles in the conterminous United States. Quadrangles selected for study are chosen according to the probable importance of their contained mineral resources, the amount of Federally owned land they contain, and, particularly, the urgency of the need for resource information on them; for example, areas likely to undergo federal withdrawal or urbanization. After incorporation of all pertinent data already available on a quadrangle, supplementary mapping and research will then produce a broad spectrum of information that will be made available to users in the form of geologic, geochemical, and geophysical maps, and interpretive mineral potential maps with appropriate discussion. An interpretive report may or may not be necessary.

The broader objective of the Conterminous Program is regional geologic analysis and synthesis. This is of equal importance to the assessment work noted above because it expresses the longer range scientific rationale of USGS work on a regional scale; namely, problems of crustal evolution and the controls on regional metallogenesis. This type of program should and would be underway regardless of whether or not CUSMAP existed. CUSMAP is, therefore, a valuable vehicle for getting needed regional geologic work done and for doing it cooperatively with a maximum of communication on the job between different specialists. This is the kind of approach that will result in the best type of product.

Geologic studies and mapping.

A principal objective of the CUSMAP Program is the preparation of a geologic bedrock map published as an MI series colored geologic map and accurately showing the geometry of rock bodies intersecting the earth's surface. In areas where the project coordinator and branch chief deem necessary, maps of surficial geology will also be made. The geologic maps are expected to be high quality professional maps and not reconnaissance maps; ideally, contacts are to be located at 0.5 mm planimetrically and within a half contour interval. The number and kinds of units to be shown on the maps depend on the purpose and scale of mapping and should reflect current stratigraphic, structural, and genetic concepts. The geologic maps should be considered as an essential tool required to make an accurate mineral appraisal of a two degree sheet and are not an end in themselves. Although the geologic map is likely to be used for many purposes, the highest priority for emphasis in mapping will be to delineate rock units and structures that control known and potential mineral and energy resources.

The time necessary for completion of geologic mapping of a two degree sheet must depend on the amount of available work already completed, as well as other physical parameters such as length and time of field season and terrane accessibility. Under ideal conditions, but with no previous work done within a quadrangle, geological mapping of bedrock should involve approximately 8 to 10 man years of study. Because difficult areas will involve more man years of effort, the need for an adequate team of regional, economic, and specialist geologists (surficial geologists, volcanologists, paleontologists, isotope geologists, etc.) is strongly emphasized. If this is not done, the product map will certainly be of a reconnaissance type.

Topical studies on aspects of the geology and mineralization of a given district will be carried out if such studies clearly enhance resource assessments. Such studies could include a wide spectrum of disciplines ranging from redefinition of critical stratigraphic sequences, the chronology of emplacement of plutonic rocks, the time and nature of deformation and metamorphism, a detailed study of a critical hydrothermal system, and analyses of structures related to mineral deposits. Because of the wide range of program commitments, topical studies must be subject to review by the appropriate branch chiefs and quadrangle coordinators.

Geophysical studies.

For every CUSMAP quadrangle, geophysical surveys including aeromagnetic, gravity, and LANDSAT data will provide a common basis for regional analysis. The precision of the regional products (aeromagnetic and Bouguer gravity anomaly maps and LANDSAT imagery) must be sufficiently high to help significantly in obtaining an understanding of the geologic setting of known deposits and for estimating the potential for discovery of mineral deposits that span a wide range in size. The basic products will be the result of extensive computer processing in conjunction with the color compositing of imagery, and will require relatively small flight line spacings for airborne data, and adequate station spacing and terrain corrections for gravity and other ground data. The regional gravity and magnetic surveys will be directed toward the production of 1:250,000 scale maps with contour intervals adequate to define the anomalies of interest. Uniform digital data sets will be prepared for both gravity and magnetic data. Acquisition, processing, and analysis of the regional data must be accompanied by measurements of the physical properties of selected rock samples sufficient to ensure optimal interpretation. In addition to the regional gravity and magnetic data, CUSMAP will utilize any other available regional geophysical techniques that will contribute toward a better understanding of the regional geology and thus the mineral resources of the quadrangle under investigation. In some quadrangles, other regional surveys such as

gamma ray, electromagnetic, magnetotelluric, and audiomagnetotelluric may be necessary to meet CUSMAP objectives.

In order to facilitate the regional assessments, more detailed surveys of limited areas may be required to determine the characteristics of specific settings of certain mineral deposits. Ideally, these would include studies in districts which are well drilled and characterized so that we know the three dimensional geology and mineralogical patterns with an unusual degree of completeness. These more detailed surveys may include one or a combination of airborne, ground, and borehole investigations using electromagnetic, induced polarization, electric resistivity, very low frequency radio wave, gamma ray, multispectral reflection and emission, self-potential, telluric, magnetotelluric, audiomagnetotelluric, and seismic methods. Well-characterized mineralized districts in the CUSMAP quadrangles offer opportunities as test areas for other geophysical techniques being developed within other mineral resource programs. Technique development projects will be encouraged to take advantage of these opportunities.

CUSMAP provides an opportunity to improve the effectiveness of geophysical studies in interdisciplinary programs. This can be accomplished by encouraging personnel to adopt creative approaches to the diverse problems that are encountered. All geophysicists assigned to CUSMAP will be urged to employ the most advanced techniques in presenting and analyzing the geophysical data. The use of techniques that enhance the usefulness of geophysical data in the solution of geologic problems such as projection, filtering, automatic interpretation, and modeling, and those directed more particularly toward resource appraisal, such as pattern recognition, should be encouraged.

Geochemical studies.

The essential character of the geochemical investigations for CUSMAP will be a regional sampling program aimed at a definition of broad geochemical patterns as expressed by the background geologic and petrologic context, and within that framework a delineation of anomalous patterns and trends resulting from mineralization features and activity in the region. Fundamental geochemical information on crustal abundances as related to rock type, age, and structure will thus be forthcoming from CUSMAP yielding geochemical information that should be a significant contribution to regional geological analysis. Anomalous or high concentration patterns and trends related to mineralization are superimposed on such data and can be properly understood only in terms of this more fundamental type of information. A number of interesting questions that have not been fully pursued and resolved can be cited, such as the detailed geochemical character of a so-called

metallogenic province (defined roughly by mining activity) and the extent to which other types of geochemical areas or provinces may lie within, extend, or be antithetic to metallogenic regions.

The sampling procedures and density required to meet the need on the 250,000 scale are not easily defined and will vary from quadrangle to quadrangle. Perhaps between 5,000 and 10,000 will be a reasonable number. Some bootstrap research on this very question will be an inherent part of every CUSMAP quadrangle effort. A minimum product for CUSMAP might be stated as: 1:250,000 scale maps illustrating variation in critical elements or groups of elements in some of the significant types of sample media (stream sediments, rocks, soil, vegetation, water, etc.) over the quadrangle for the purpose of characterizing the background petrologic framework and mineral potential within that framework. Other basic goals that will be pursued as appropriate within the time, manpower, and funding frame include: (1) Development of geochemical methods for discovering concealed mineral deposits buried in premineral rocks and/or under postmineral cover; (2) Development of a better understanding of processes controlling the migration of elements in primary and secondary environments; (3) Possible development of a national grid as an outgrowth of regional work such as CUSMAP.

Mineral Resource Appraisal.

As noted earlier, the immediate purpose of the CUSMAP Program is to provide mineral assessment information that can readily be used by federal, state, and local government agencies, as well as by parties in the private sector, for land-use decisions and planning. The broader purpose of the Program is to develop commodity related and purely scientific information concerning the Nation's mineral resource data base. This effort is of great importance to us in maintaining a long-term capability in responding to mineral resource questions.

The minimum mineral resource assessment product will be one that fulfills the immediate purpose just noted. This product should be a map showing areas of varying degrees of favorability based upon the observed geology, the observed occurrence of ore deposits and mineralization, and our geological understanding of the environments of occurrence of the types of commodities actually or potentially involved. The map will be accompanied by an explanatory text describing how mineral resource data were used to produce the map. Areas of differing favorability for different commodities will be designated in as realistic and non-arbitrary manner as possible. The geochemical and geophysical data maps, as well as other research in the Program, will contribute critically to the Mineral Potential map.

A data file will be developed and a spot map compiled as an index. The basic component of the data file will be a tabulation of descriptions of known deposits and production figures, from the literature. To this

should be added tables of geochemical data, where available, and any previously unpublished data that can be released from exploration companies. All components of the data file will reside in appropriate Survey computer data banks (GRIB, RASS, etc.) Proprietary data that cannot be released will be kept confidential. In addition to identifying well-known geologic environments for various commodities, an attempt should be made to identify new environments that might be important in the future, but such environments need not be assessed in any quantitative way.

Beyond the above minimum requirement, various levels of probabilistic assessment are possible or desirable depending upon time and manpower available and whether a more detailed assessment appears necessary. The initial approach here is to estimate the number of undiscovered deposits of each type present for which there is reasonable probability (i. e. about 15% or greater) that at least one new deposit might be found. Whether a range of numbers of deposits or a single number is given, they should be accompanied by probability estimates. Obviously, if known deposits of a given type have only very small tonnages relative to typical economic deposits of the major commodities involved, and there is no reason to believe that new discoveries may be any larger, attempting this level of assessment will not be worthwhile. The amount of data required is much larger at this level than is the case for a minimum assessment.

The next level of probabilistic assessment involves predicting amounts of metal present in certain types of deposits. A prediction will only be possible if reliable data on economic and subeconomic reserves in known deposits are available. In addition, estimates of amounts of metal must also be assigned to each of the probabilistic estimates of numbers of undiscovered deposits. The latter procedure is only possible if grade-tonnage data are available for many known deposits throughout the same region and geologic environment. In an area the size of a 2° quadrangle, this level of assessment will probably be unattainable for more than one or two deposit types. Industry cooperation will be essential at this level.

The highest level of probabilistic assessment involves estimation of amounts of a commodity available at various probability levels regardless of deposit type. The amount of data, effort, and explanation required for a meaningful assessment of commodities in a 2° quadrangle at these more detailed levels of treatment would likely require a separate report.

Several problems will be unavoidable in making these assessments. There will always be difficulties in classifying some deposits, in defining a single deposit (i. e., in the case of scattered small occurrences, how should they be grouped into deposits?), and in evaluating

various unique, unusual deposits. Carefully prepared data tables backed by a computer file can ameliorate these problems to the extent that anyone interested can obtain the basic data used and decide for themselves whether they agree with the decisions made. Another difficult problem is depth of cover; that is, at what depth is there essentially no potential for significant deposits? The answer to this question will vary from one quadrangle to another depending upon many factors including quantity and quality of subsurface data and the deposit types in question. Here, narrow guidelines are not appropriate or desirable; these problems must be solved within individual 2° quadrangle projects.

Finally, a very serious uncertainty involved is the justification with which a given area or province can be used as a model or yardstick for assessment of another area of apparently similar environment. All such mineralized areas, despite basic similarities, are in fact discrete tectonic-geologic entities, and these will defy, to a greater or lesser degree, any statistical treatment. Thus, the probabilistic work can be considered an addition to the minimum type of product described above, but it will not necessarily prove more definitive. Decades of exploration and drilling will be required before the accuracy is adequately tested, and many undoubtedly will be found to be considerably in error. This is a research approach, however, that merits continuing active development in the interest of greater predictive capacity and a more accurate future national resource inventory.

CUSMAP project personnel should be urged to consider the requirements for mineral assessment early in the development of each project, because areas where new information is needed should be identified as soon as possible. If time permits, all important commodities should be identified and the minimum evaluation and first level probabilistic assessments attempted on a trial basis to help determine where effort should be directed. If no attempt is made until the end of the project, critical data may never be collected.

Support Investigations.

In addition to the above work in individual quadrangles, research must also be undertaken on studies pertinent to mineralization problems common to several quadrangles or clusters thereof. One of these programs would consist of regional metallogenic synthesis involving larger blocks of terrane, up to a few states and extending on up to continental scale. Clues as to the existence of entirely new provinces, as well as extensions of known districts and provinces, can arise from such work.

Another area of effort is case studies on mineral deposits or types of deposits found in the quadrangles investigated. Economic geologists, process geochemists, isotope geochemists, and other specialists will focus on studies aimed at a better understanding of environments of formation of mineral deposits and therefore at improved concepts and

methodology of exploration. These programs will be genetically organized and oriented. Mapping programs involving the resource specialists (commodity geologists) should be a part of this effort. The hydrothermal-late-magmatic research proposal is an example of one major effort that we hope to get underway. It embraces classical hydrothermal, porphyry-greisen-skarn, and volcanogenic massive sulfide processes.

Another similar lab-field program should be initiated in the early magmatic Cr-Pt-Ni-Cu-Ti environment as well as a broad effort on base metal stratobound deposits in sediments, embracing the platform carbonates, the black shale deposits of the German and Australian types, and the sandstone Cu and U deposits of the western United States. Suggestions and initiatives in these directions will be welcome. These research thrusts do not rule out appropriate investigative work on other types of resources, for example, the non-metallics. Also, low temperature studies embracing weathering and supergene processes are needed.

These two broad lines of attack (metallogenic synthesis and case studies) are larger than CUSMAP itself and would involve a large portion of the several Branch efforts. This is to be expected as CUSMAP will probably be providing much of the funding and much of the subject matter for our future investigations. Coherent national programs must be developed that will make maximum scientific use of our limited personnel.

Because availability of adequate scientific expertise from within the USGS is and will remain a very serious problem in our development and conduct of CUSMAP, plans for involvement of scientists from State Geological Surveys and the academic world, as appropriate, must be pursued. Participation by scientists from outside the USGS can help strengthen the Program and the products. Anyone participating in the Program must be brought into the project on which he/she is working as a full participating member, whether that person is a member of the USGS or not and regardless of scientific discipline.

Authority in CUSMAP.

The individual in authority on a CUSMAP quadrangle is the project coordinator. Through conferences, periodically held with the entire group (at least twice a year), work should be planned and coordinated. The coordinator should initially have all contributors' project descriptions in hand as a basis for planning and for resolution of possible conflicts. Though it may not be possible to assign all contributors to a single quadrangle project number for the appropriate time, at least that should be taken as the implicit organizational intent. Geochemical, geophysical, and geological work must be closely integrated. These are not independent items. Only in this way can the best product be obtained. The coordinator does not dictate what is done, but because he is given the responsibility for accomplishing the goals of the project, his views

are more important than those of anyone else in conducting the Program. If any conflicts or defaults cannot be resolved in the conferences, the problem should go to the two Branch Chiefs involved. Presumably, it would never have to go to the Office Chiefs or Chief Geologist. The Office of Mineral Resources is the lead office for this Program, and the Chief of that office has the responsibility for the overall planning and conduct of the Program.

A COMPUTER-BASED CONSULTATION SYSTEM FOR MINERAL EXPLORATION

By: Peter E. Hart and Richard O. Duda
SRI International
Menlo Park, California

Marco T. Einaudi
Stanford University
Stanford, California

May 24, 1978

Acknowledgemnt: This work was supported in part by the Office of Resource Analysis of the U. S. Geological Survey under Contract No. 14-08-0001-15985, and in part by the RANN Division of the National Science Foundation under Grant AER77-04499. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors, and do not necessarily reflect the views of either the U. S. Geological Survey or the National Science Foundation.

333 Ravenswood Ave. • Menlo Park, California 94025
(415) 326-6200 • Cable: STANRES, Menlo Park • TWX: 910-373-1246

A COMPUTER-BASED CONSULTATION SYSTEM FOR MINERAL EXPLORATION*

Peter E. Hart and Richard O. Duda**
Marco T. Einaudi***

I INTRODUCTION

The search for mineral deposits is a difficult, high-risk enterprise whose likelihood of success is influenced by a great many factors. Among the most important of these factors are the knowledge and judgment brought to bear on an exploration problem. If, for example, a panel of expert, highly experienced geologists could be convened to examine every prospect, it is at least plausible that fewer deposits would be overlooked in the exploration process. Unfortunately, whether a project is undertaken by a private company or a government agency, it is not usually feasible to convene such a panel.

Our purpose in this paper is to describe an experimental system of computer programs called PROSPECTOR that is intended ultimately to provide expert consultation on problems of mineral exploration and

---

* This work was supported in part by the Office of Resource Analysis of the U. S. Geological Survey under Contract 14-08-0001-15985, and in part by the RANN Division of the National Science Foundation under Grant AER77-04499. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors, and do not necessarily reflect the views of either the U. S. Geological Survey or the National Science Foundation.

** SRI International, Menlo Park, California

*** Stanford University, Stanford, California

resource evaluation. Ideally, PROSPECTOR would serve as a surrogate for a panel of expert economic geologists. However, because the cost of computation is very low-- and is projected to drop by a factor of about 100 over the next two decades-- systems like PROSPECTOR could be made available to the exploration enterprise at large.

When consulting about a prospect, the PROSPECTOR system first gives the field geologist an opportunity to tell the program about the most significant features of the site: the major rock types, minerals, and alteration products. The system then compares these observations with a number of stored models of various types of ore deposits and forms alternative hypotheses about the types of deposits that might be present. The system next engages the geologist in a dialogue in order to obtain additional data that would be useful in resolving the multiple hypotheses. The geologist can interrupt the system at any time to obtain clarifications of questions, volunteer additional data, request the reason a particular question has been asked, or obtain a general summary of the state of the consultation process.

The less skillful the poorer his data and the less the PROSPECTOR can help although the less skillful geologist needs more help. A Catch 22 b

The more skillful he is the less he needs the computer, still a case of GIGO

Our goal is to provide the geologist with a service comparable to providing him telephone access to authorities on many different kinds of ore deposits. The field geologist must still make all the observations; the more skillful he is, the greater the likelihood of his detecting an ore deposit present on the prospect. However, the expert consultants can play a crucial role by alerting him to unsuspected possibilities, by suggesting that he seek data that he might otherwise have not considered

important, and by using their own judgment to evaluate the significance of the data that have been gathered. In addition, the educational value of such a consultation session may comprise an invaluable secondary benefit.

While we are developing PROSPECTOR as a tool for consulting on individual prospects, it may also be ultimately useful as a tool for evaluating regional resource potential. For example, the importance of maintaining databases of properties with mineral potential is becoming generally recognized. PROSPECTOR could in principle be adapted to screening such databases, either to select for particular commodities or to form estimates of aggregate resources. We can think of the database in this mode as furnishing the geological characteristics of a prospect that would otherwise have been provided by the field geologist.

The ability of PROSPECTOR to provide expert consultation rests on a base of knowledge about economic geology. This "knowledge base" has several components, the most important of which are the models that contain geological (and eventually geochemical and geophysical) information relevant to exploration for various types of ore deposits. The models are stored in the computer in a special way (to be described later) that enables PROSPECTOR to utilize them to draw conclusions from geological evidence. An important feature of the overall design is its modularity; models can modified or augmented incrementally, so that the competence of the system can be continually improved. The models themselves are obtained by interviewing recognized authorities on various types of ore deposits.

Mineral exploration is perhaps as much an art as a science, and the state of this art does not admit the construction of models as rigorous and complete as, say, those of Newtonian mechanics. This situation has two important effects on the design of PROSPECTOR. First, the system must accommodate plausible or probabilistic styles of reasoning in addition to rigorously logical styles. Second, the models often reflect the subjective judgments of expert economic geologists more than objectively derivable facts. Of course, the use of subjective judgments and probabilities to make technical evaluations is not unique to PROSPECTOR. Subjective probabilities have been used in resource evaluation [Harris, et al., 1970; see Harris, 1977, for a comprehensive treatment], while panels of experts have been frequently used in Delphi studies [Linstone and Turoff, 1975; Ellis et al., 1975] to forecast technological events. What is unique here is PROSPECTOR's particular combination of plausible and logical reasoning, using a knowledge base supplied by experts to provide a computer-based consultation service.

To the best of our knowledge, PROSPECTOR represents the first serious attempt to build a computer system able to consult actively on problems of mineral exploration. The general notion of a computer-based consultant system, however, has been explored before. Procedures for performing the required plausible and logical reasoning have been developed through computer science research on artificial intelligence [Nilsson, 1971; Raphael, 1976; Winston, 1977]. These procedures have been applied in several fields, with the most advanced development having taken place in medicine. In particular, computer-based

consultation systems have been developed for infectious diseases [Shortliffe, 1976], for glaucoma [Weiss, et al., 1977], and for a substantial portion of internal medicine [Pople, 1977]. While these systems are rather recent and still in the experimental stage, some have already attained levels of competence in their specialties that rival the performance of expert physicians. These successes provide additional motiviation for us to achieve comparable levels of performance in the area of mineral exploration.

The remainder of this paper is organized as follows. The next section illustrates the operation of PROSPECTOR with an excerpt from an actual, relatively lengthy consultation session. The two subsequent sections outline the method we use for representing model information in the system and the mechanisms for plausible reasoning that are employed. The final section summarizes the current status of our work and our plans for future development. We should make the observation at this point that PROSPECTOR is a rather complex system, and that the present paper is intended as no more than a brief introduction to it. A more complete description is available in [Duda, et al., 1977].

II A CONSULTATION EXAMPLE

This section illustrates the current version of PROSPECTOR by presenting an annotated excerpt from a consultation session. The geological data supplied to the program are representative of an actual site, but are not necessarily to be considered complete.

Annotations are set off to the right and are enclosed in square brackets.

```
PROSPECTOR ...... Mark II ...... Version 4

3-May-78 21:36:38

Do you want instructions? yes
```

In using PROSPECTOR, you will be asked to answer questions about the existence of various kinds of evidence. Answers to most questions will be numbers from -5 to 5. These numbers indicate the certainty of your response, with 5 indicating that the requested evidence is definitely present, -5 that it is definitely absent, and 0 that you have no opinion one way or the other.

In addition to supplying answers, you can request information at any time with the following commands:

? -- Prints a rephrased version of the question, often with additional comments or suggestions.

WHY -- Prints the reason that PROSPECTOR asked the question.

SUMMARIZE -- Prints a summary of the conclusions at this point in the consultation.

VOL -- Lets you volunteer information to PROSPECTOR.

CHANGE -- Lets you change your answer to a previous question.

SHOW -- Lets you see a previously asked question and your answer.

COMMENT -- Lets you type comments which will appear in your transcript file.
Two consecutive carriage returns will end the comment.

TRACE -- Lets you see PROSPECTOR's partial conclusions.

NOTRACE -- Stops the trace.

WAIT -- Interrupts the session to allow modification of the program. (Should be used by qualified users only.)

QUIT -- Terminates execution of the program.

HELP -- Prints this explanation.

Do you want to volunteer any information? yes

When you are finished volunteering, please type DONE
Do you want to see the instructions? yes

PROSPECTOR can accept certain simple English statements that describe the prospect. You can name the types and ages of the rocks, minerals, and alteration products that are present, suspected, or even absent. You can indicate uncertainty either through wording such as 'might be,' 'certainly is,' or 'probably isn't,' or by an explicit score such as '3'. Currently, PROSPECTOR can accept information about age only in terms of geologic eras, periods and epochs. Pronouns are not understood, nor are any references to previous sentences.

Some example sentences are:

THERE IS DOLOMITE. [default certainty = 5]
ANDESITE IS VERY LIKELY. [certainty = 4]
THERE MAY BE QUARTZ-MONZONITE. [certainty = 2]
IT IS DOUBTFUL THAT THERE ARE STOCKS. [certainty = -3]
THE LIKELIHOOD OF LATE PALEOZOIC RHYOLITE PLUGS IS RATHER LOW. [cert. = -2]
MY CERTAINTY OF ROCKS FROM THE PRECAMBRIAN ERA IS HIGH. [cert. = 3]
Or, more simply:
5 DOLOMITE
QUARTZ MONZONITE 2
STOCKS ARE -3
etc.
Again, please type DONE when you are (for the moment) finished volunteering.

1: There is a granodiorite batholith

(GRANODIORITE BATHOLITH) (5)

[PROSPECTOR gives a terse restatement and a certainty symbol to indicate that it has understood the input.]

2: There are quartz-monzonite dikes

(QUARTZ-MONZONITE DIKE) (5)

3: There are Mesozoic quartz-monzonites

(MESOZOIC QUARTZ-MONZONITE) (5)

4: Quartz veins are present

(QUARTZ VEIN) (5)

5: Chalcopyrite and pyrite and bornite and magnetite

(CHALCOPYRITE) (5)

(PYRITE) (5)

(BORNITE) (5)

(MAGNETITE) (5)

9: There is biotite and k-feldspar

(BIOTITE) (5)

11: Sericite

(SERICITE) (5)

12: Done [This completes the initial evidence volunteered by the user.]

I am considering the possibility of a
(Type-A porphyry copper deposit -- PCDA)
in the target area.

[A "Type A" deposit is Prospector's version of a certain type of continental margin porphyry deposit.]

Presently, the possible deposits are:

1 - Komatiitic Nickel Sulfide Deposit . (0.0)
2 - Type-A porphyry copper deposit -- PCDA. (0.0)
Current hypothesis
3 - Mississippi Valley type Lead/Zinc deposit -- MVTD. (-4.8)
4 - Kuroko-type massive sulfide deposit -- MSD. (.001999)

Which of the above would you like to rule out?
(type ? for available options).....? none

[The program lists the four model types that (in this version) it knows about, and allows the user to rule out some of them if he wishes.]

The following questions are intended to establish the nature of the petrotectonic setting.

You may skip over this part of the model if you wish, in which case I will assume that the corresponding conditions are established.
Do you want me to skip? No

[The program allows the user to 'short-cut' major parts of the model if he wishes.]

12--To what degree do you believe that:
(* The target area is in a continental margin mobile belt) ? 5

...
...
...

[After several more questions, PROSPECTOR has finished considering the petrotectonic setting, and is ready to move on to questions dealing with the regional environment.]

Changing the certainty of (* Right petro-tectonic setting)
to 3.182

The following questions are intended to determine the nature of the regional environment.

You may skip over this part of the model if you wish, in which case I will assume that the corresponding conditions are established.
Do you want me to skip? no

18--To what degree do you believe that:
(* There is a granitic intrusive system in the region) ? 5

Changing the certainty of (* Right regional environment)
from .03238 to .1414

...
...
...

[After several more questions, the user answers question 23 with a -3, indicating the probable absence of the evidence requested.]

23--To what degree do you believe that:
(* The region contains intrusive breccias) ? -3

26--To what degree do you believe that:
(* There is a throughgoing fault system in the region) ? ?

[The user responds to the question with "?" to request a fuller explanation of its meaning.]

*
Zones of crustal weakness or major linear fault zones on a regional scale, which have been active over a considerable span of geologic time, may serve to guide the emplacement of stocks in the upper part of the crust. Such lineaments might be recognized by the presence of:

1) abrupt sedimentary facies or thickness discontinuities if the faults acted as basinal controls during sedimentation in the early stages of mobile belt development;
2) unusual alignment of stocks, breccia pipes, volcanic deposits;
3) aeromagnetic or gravity anomalies.

To what extent do you believe that such a lineament or lineaments exist in the region of interest? 4

[PROSPECTOR repeats the question, and the user responds with 4. At this point the system has completed its consideration of the regional environment and is ready to proceed further.]

Changing the certainty of (* Right regional environment)
to 2.162

The following questions are intended to determine the existence of the right kind of an intrusive system in the target area.

You may skip over this part of the model if you wish, in which case I will assume that the corresponding conditions are established. Do you want me to skip? no

---

...
...
...

[After some additional questions, the system reminds the user of a previous statement to "prime" him for a question. While not strictly necessary, this technique is useful because the user is reassured that the system is tracking him. If the user were to contradict a previous response, PROSPECTOR would note that and would give him an

opportunity to change either the previous or current response.]

You told me about:
- FELDSPAR. (5.0)

36--To what degree do you believe that:
(* The phenocrysts are composed of quartz, feldspar, and biotite or hornblende) ? 5

37--To what degree do you believe that:
(* The groundmass texture is aplitic) ? why

[The user responds to this question by asking PROSPECTOR to explain the reason for asking it.]

I am trying to establish (* Right groundmass)
*
A sugary-textured groundmass of quartz and feldspar, usually with a high alkali-feldspar/plagioclase-feldspar ratio, is a characteristic feature of porphyries associated with porphyry copper deposits. The texture and abundance of K-feldspar may be due in part to alteration and recrystallization, but more likely is due to pressure quenching.

37--To what degree do you believe that:
(* The groundmass texture is aplitic) ? 4

...
...
...

[The system ends its consideration of the main aspects of the intrusive system, and reports its conclusion.]

Changing the certainty of (* Favorable intrusive system)
to 4.258

[PROSPECTOR now begins a detailed consideration of zoning phenomenon.]

Porphyry copper deposits are often composed of characteristic zones of alteration and mineralization. (In some cases these can be looked upon as different stages superimposed upon one another.) I shall call the

most prominent such zone (or stage) Zone 1.

Even though you may have already identified this zone as, say, the propylitic zone or the sericitic zone, I need to consider other possibilities as well. Thus, the following questions are intended to determine the possibilities for Zone 1.

Hypothesis: Zone 1 is the (Barren core zone)

42--To what degree do you believe that:
(* There are fresh-looking feldspars) in Zone 1 ? 5

You told me about:
- BIOTITE. (5.0)

43--To what degree do you believe that:
(* There is hornblende with little or no secondary biotite) in Zone 1 ? 5

...
...
...

[After further questioning, PROSPECTOR assesses the likelihood that Zone 1 is the barren core zone, and goes on to consider alternative hypotheses about the identity of Zone 1.]

Conclusion: my certainty in (Barren core zone) is now: 3.375

Hypothesis: Zone 1 is the (Propylitic zone)

52--To what degree do you believe that:
(* There is moderate to major epidote) in Zone 1 ? -4

...
...
...

[PROSPECTOR effectively rules out the likelihood that Zone 1 is the propylitic zone and continues.]

```
    Conclusion: my certainty in (Propylitic zone) is now: -4.253
--------------------
                    ...
                    ...
                    ...

                          [PROSPECTOR goes on to consider the
                           hypothesis that Zone 1 is potassic or
                           sericitic, and effectively rules out these
                           possibilities after asking a few questions.

                           The system announces its conclusion about
                           Zone 1 and then asks whether any other zones
                           have been delineated.]

  My best guess for this Zone is:
                    - Barren core zone. (3.375)

Are there any other zones or stages present in the target area? yes

--------------------
    Hypothesis: Zone 2 is the (Barren core zone)
--------------------

                    ...
                    ...
                    ...

                          [The consultation continues in the same
                           manner as before.  In this example the
                           user has delineated a number of zones, each
                           of which the system attempts to identify.
                           We omit the details of this rather lengthy
                           sequence of interactions and proceed directly
                           to the conclusions.]

                    ...
                    ...
                    ...

--------------------

    On a scale from -5 to 5 ,
    my certainty in (PCDA) is now: 1.697
```

In summary, the following have been observed:

The major favorable conclusions were:
- Right petro-tectonic setting. (3.182)
- Right regional environment. (2.162)
- Right PCDA intrusive system. (1.694)

The encouraging observations were:
- The throughgoing fault system is older than the intrusive system. (5.0)
- The continental margin mobile belt contains granitic intrusives. (5.0)

...
...
...

[PROSPECTOR has concluded that the certainty of a porphyry copper deposit in the target area is 1.697, a moderate but decidedly positive assessment.

At this point, the system invites the user to request additional information about its reasons for this assessment. However, a truly perceptive elucidation of the entire consultation session is beyond PROSPECTOR's capability at present.]

...
...
...

I can not do any more with this hypothesis,
Do you want to pursue another one? no

III REPRESENTATION OF MODELS

For PROSPECTOR to make use of models such as the one mentioned above, each model must be stored in the computer in a way that allows inferences to be drawn from an examination of the parts of the model and their interrelationship. Descriptions of models in textbooks or articles take the form of conventional prose interspersed with diagrams; this is certainly inappropriate for a system like PROSPECTOR. An alternative would be to attempt to represent a model through some sort of checklist that would enumerate the geological evidence required of a prospect if the model is to be satisfied. We have avoided this approach because it does not allow the logical richness inherent in modern models of ore deposits.

We have elected to structure model information as a collection of relations between observable pieces of evidence and those hypotheses that the evidence tends to confirm or refute. These relations link evidence and hypotheses together to form an information structure termed an <u>inference net</u>. Inference nets are most easily described by enumerating the several types of relations upon which they are based.

Logical Relations

The standard logical connectives of AND, OR, and NOT are often useful for representing the relation between evidence and hypothesis. The use of AND implies that a set of evidence must be established in its

entirety to confirm the hypothesis. For n pieces of evidence E-1, E-2, ..., E-n and a hypothesis H we would write

IF	E-1 AND E-2 AND . . . AND E-n,
THEN	H.

To illustrate this, consider one such relation that occurs in the porphyry copper model included in the example section. The relation defines the characteristics of the phenocrysts in the porphyry texture as

IF	The grain size of the phenocrysts is between 0.5 and 3.0 mm
AND	The phenocrysts are subhedral or euhedral
AND	The phenocrysts are composed of quartz, feldspar, and biotite or hornblende
THEN	The phenocrysts satisfy the model requirements.

In the next section we shall discuss the measures of certainty that are associated with every piece of evidence and describe our methods for reasoning from uncertain evidence. Here we shall state only that the rule for combining certainties is derived from the methods developed for dealing with so-called "fuzzy sets" [Zadeh, 1965].

We represent AND relations graphically as shown in Figure 1.

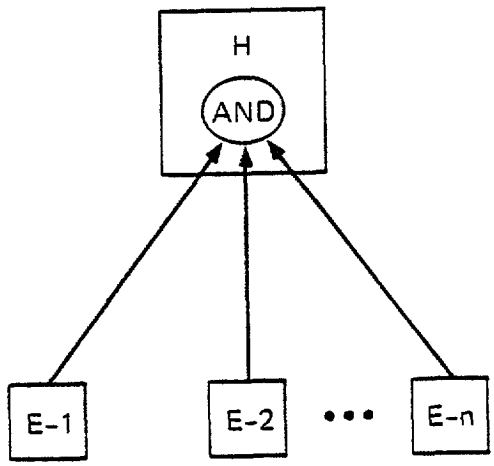


FIGURE 1 Representing the AND relation

The OR relation is used whenever any of several pieces of evidence serves to establish a hypothesis. In a manner analogous to the AND relation, we represent the OR relation as

IF E-1 OR E-2 OR . . . OR E-n,
THEN H.

To illustrate this, the following relation in the porphyry copper model concerns the sulfide assemblages that help to establish the appropriate alteration-mineralization zones or stages:

IF There is pyrite less than chalcopyrite

OR There is chalcopyrite

OR There is bomite-chalcopyrite

OR There is bomite

THEN The mineralization is favorable for the potassic zone.

The rule for combining certainties of "OR's" is derived from the methods of fuzzy sets. The OR relation is represented graphically as shown in Figure 2.

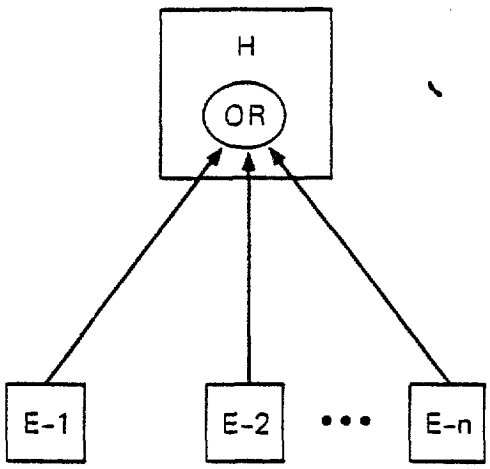


FIGURE 2 Representing the OR relation

The third logical connective, NOT, is used whenever the absence of a piece of evidence establishes a hypothesis. For a single piece of evidence E, we would write

IF NOT E

THEN H.

The certainty of the conclusion is defined as the complement of the certainty of the evidence; i.e., H is certain to the extent that the absence of E is certain.

Plausibility Relations

The logical relations provide natural ways of expressing certain kinds of relations between evidence and hypothesis. An important and characteristic case occurs when several pieces of evidence independently influence the certainty of the hypothesis, with each piece of evidence casting a "vote" of a different weight. To handle this commonly occuring situation, we must have the means both to express a "weighted vote" mathematically and to combine the votes of several pieces of evidence. We accomplish this by using a type of relation called an <u>inference rule</u>. A single inference rule has the form

IF E

THEN (to degree LS, LN) H.

The two numbers LS and LN specify the degrees to which the evidence is necessary and sufficient for the hypothesis, and thus measure the strength or "weight" of the rule. A precise definition of LS and LN, together with a description of how they are used in the reasoning process, will be presented shortly.

As an illustration, the following rule concerns evidence for the propylitic zone of a porphyry copper deposit.

IF Magnetite or pyrite in disseminated or veinlet form is present

THEN (2, -4) There is favorable mineralization and texture for the propylitic stage.

(As we shall see, the number 2 means that the presence of the evidence is mildly encouraging for the hypothesis, while the -4 means that the absence of the evidence is strongly discouraging for the hypothesis.)

When several pieces of evidence bear on the same hypothesis, we represent the situation graphically as shown in Figure 3.

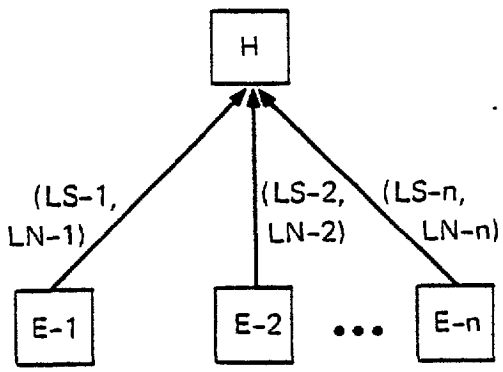


FIGURE 3 Representing inference rules

Contextual Relations

In addition to logical relations and rules of inference, it proves useful in specifying models to state that a piece of evidence can be considered only after some other condition has been satisfied. For example, one part of our model of a particular type of nickel deposit uses a number of inference rules to define a favorable magmatic association and local environment. However, none of the rules in this section of the model are applicable (or in fact make any sense at all) until the occurrence of a sequence of mafic to ultramafic igneous rocks has been established. We therefore require the existence of such an igneous sequence as a condition or "context" for the rules.

In general, we use contexts as a way of expressing a condition that must obtain before a piece of evidence can be used in the reasoning process. Having established the necessary context for a piece of evidence, the evidence itself can influence some other hypothesis. We indicate the context for a piece of evidence graphically by a dotted arrow as shown in Figure 4.

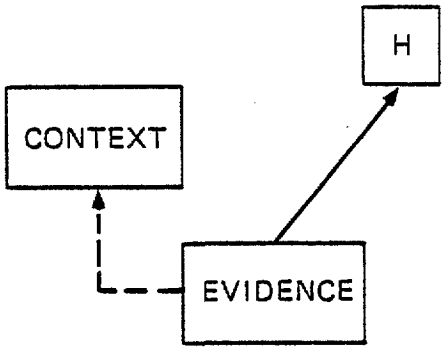


FIGURE 4 Representing contextual dependencies

Inference Nets

The basic types of relations mentioned above-- logical relations, plausibility relations, and contextual relations-- form the building blocks from which exploration models of ore deposits are constructed. In a typical model these relations link together many different hypotheses and pieces of evidence to form an inference net. An abstract example of a simple inference net is shown in Figure 5. For the sake of simplicity, no distinctions have been indicated among the various types of logical relations and plausibility relations. Each box in the net is called a space, following the terminology established in [Hendrix, 1975].

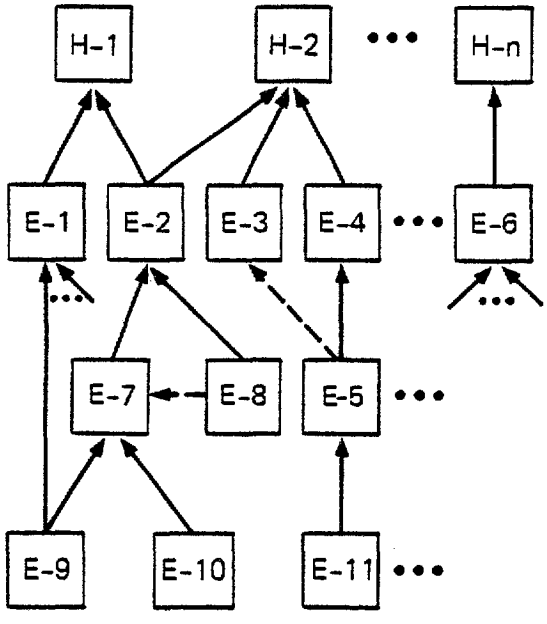


FIGURE 5 An inference net

The top-level hypotheses H-1,...,H-n correspond to the possible occurrence of the various types of ore deposits modeled by the system. Spaces with no incoming arrows must always correspond to observable data that, in principle, could be furnished by the user (although in any given consulting session the user may not actually have the data available.) Intermediate spaces can be viewed in either of two complementary ways: they can be thought of as evidence related to the spaces above them, or they can be thought of as subhypotheses to be resolved by the spaces below them.

Inference net representations of exploration models comprise the major portion of the knowledge base used by PROSPECTOR. Reasoning about such models is, at least conceptually, quite simple. Whenever a piece of evidence is given the system by the user, the consequences of that evidence are "propagated" along outgoing arrows to all related hypotheses. (The mathematical form of this propagation process will be discussed in the next section.) For example, with reference to Figure 5, suppose the user provides the information that evidence E-10 is probably present on the prospect. This information will affect the likelihood of E-7, which will in turn affect first E-2 and finally H-1 and H-2.

A more complete discussion of inference nets, with a clarification of some important details that are beyond the scope of the present paper, may be found in [Duda, et al., 1978.]

Taxonomies

While models of ore deposits form the main part of the knowledge base of PROSPECTOR, they by themselves do not facilitate certain elementary but pervasive forms of reasoning that are important in a consultation system. For example, the occurrence of pyrite certainly establishes the presence of sulfides; analogously, if intrusives are known to be absent, it is pointless to ask the user whether dikes are present.

Elementary reasoning of this type is most readily accomplished by incorporating a set of taxonomies within PROSPECTOR. A mineral taxonomy, for example, indicates that pyrite is a particular iron sulfide, and that iron sulfides as a class are a subset of the sulfide minerals. Similarly, a taxonomy of morphological descriptors shows that dikes are a type of intrusive. In general, PROSPECTOR uses taxonomies to reason about subset-superset relations.

IV PLAUSIBLE REASONING

Any system for reasoning about geological situations and events must accommodate the uncertain nature of geological observation and interpretation. In PROSPECTOR two different types of uncertainty are represented and exploited. First, the field geologist may be uncertain as to whether a sought piece of evidence is present or absent. Second, the interpretation of a piece of evidence within a model is often a matter of influencing probabilties, rather than establishing the absolute truth of a hypothesis.

Communicating Beliefs

As we saw in the illustrative consulation session, the field geologist using PROSPECTOR communicates his belief about the certainty of a piece of evidence by stating a number between -5 (meaning the evidence is definitely absent) and +5 (meaning the evidence is definitely present.) A value of 0 signifies that the user has no opinion about either the presence or the absence of the evidence.

The certainty scale used by the field geologist is immediately converted by PROSPECTOR into a probability number for internal computation, with the two extreme points corresponding to probabilties of 0 and 1 respectively. The zero point on the certainty scale deserves special attention. A certainty of 0 for a piece of evidence, E, is associated with the <u>prior probability</u> P(E) for that evidence. The prior

probability is simply the probability that the evidence is present, given that no actual observations relating to the evidence have been made. Certainty values between 0 and 5 are linearly interpolated into probabilities between P(E) and 1, while certainties between -5 and 0 are interpolated into probabilities between 0 and P(E).

We have employed certainty measures for communicating with the user precisely because they allow statements relating a piece of evidence E to P(E). If, for example, the user suspects (but is not sure) that E is present, we want the system to respond by increasing P(E) by a small amount. As we shall see shortly, the value of P(E) is specified by the expert geologist defining the model, and the user normally is not burdened with that information. Thus, if PROSPECTOR communicated with the user in terms of ordinary probabilities, the user might unintentionally prejudice the outcome by associating with E a probability far from P(E) when, in fact, he meant to indicate only that the presence of E was not very likely.

All information communicated between PROSPECTOR and a user regarding degree of belief is couched in terms of certainties. All internal computation regarding degree of belief is done in terms of probabilities, using the general approach described below.

Probabilistic Inference

The probability of a logical combination of evidence E-1,...,E-n is defined by applying the fuzzy-set relations alluded to previously.

Specifically, if a hypothesis H is defined by H = E-1 AND ... AND E-n, then PROSPECTOR takes P(H) to be the minimum of P(E-1),...,P(E-n). Similarly, if H = E-1 OR ... OR E-n, then PROSPECTOR takes P(H) to be the maximum of P(E-1),...,P(E-n). If H = NOT E, then we set P(H) = 1 - P(E).

Inference rules involve probability calculations that are somewhat more complicated than the simple definitions associated with logical combinations. Propagation of probabilities in inference rules is accomplished using a form of reasoning known as Bayesian decision theory [Raiffa, 1968].

Bayesian inference is based on an elementary theorem of probability known as Bayes rule. For our purposes, the so-called "odds-likelihood" form of the rule is most convenient. This form relates three quantities involving E and H: the prior odds O(H) on the hypothesis, the posterior odds O(H|E) on the hypothesis given that E is observed to be present, and a measure of sufficiency LS. By Bayes rule, we can write

$$O(H|E) = LS * O(H). \qquad (1)$$

The odds for any evidence (or hypothesis) are simply the ratio of the probability in favor of the evidence to the probability against the evidence. Odds and probabilities are therefore freely interchangeable through this simple relation. The sufficiency measure LS is a standard quantity in statistics called the likelihood ratio, and is defined by

$$LS = \frac{P(E|H)}{P(E|\sim H)} \qquad (2)$$

where ~H means "not H."

Equation (1) prescribes a means for updating the probability (or odds) on H, given that the evidence E is observed to be present. We call the quantity LS a sufficiency measure because of the effect it has when it is large. If LS is much greater than unity, the occurrence of E has the effect of transforming neutral prior odds on H into large posterior odds in favor of H. In other words, an inference rule for which LS is large means that the observation of E is encouraging for H-- in the extreme case, E is sufficient to establish H in a strict logical sense. On the other hand, if LS is much less than unity, then the observation of E is discouraging for H, inasmuch as the observation of E diminishes the odds on H.

A complementary set of equations describes the case in which E is known to be absent-- that is, when ~E is true. In this case, we can use Bayes rule to write

$$O(H|\sim E) = LN * O(H) \qquad (3)$$

where

$$LN = \frac{P(\sim E|H)}{P(\sim E|\sim H)} . \qquad (4)$$

The quantity LN is called the necessity measure. If LN is much less than unity, the known absence of E transforms neutral prior odds on H into very small posterior odds in favor of H. In the extreme case, E is logically necessary for H, so that the absence of E is discouraging for H. On the other hand, if LN is large, then the absence of E is encouraging for H.

It is intuitively clear that we cannot have a situation in which either the presence or absence of E is encouraging for H. Similarly, the presence or absence of E cannot both be discouraging. Interestingly enough, it can be shown mathematically that LS and LN cannot both be either large or small.

Equations (1) and (3) provide the basis for updating the probability of the hypothesis of an inference rule when the evidence is observed to be either definitely present or definitely absent. They also suggest the information the expert must furnish at the time the model is defined. For the inference rule

IF E

THEN (to degree LS, LN) H,

the expert must specify the measures LS and LN as well as E and H. He must also specify prior odds O(H), since the updating equations contain this term.

In the general case, the user may not be able to state that E is either definitely present or definitely absent. The updating formulas

(1) and (3) cannot then be applied directly, but can be extended to accommodate the uncertainty in the evidence. Under quite reasonable assumptions, the extension involves a linear interpolation between the extremes of E's being definitely present and definitely absent. A more complete description of this extension, together with a discussion of some problems in the use of subjective probabilities, is given in (Duda, et al., 1976).

V CONCLUDING DISCUSSION

We have summarized our work on creating a computer-based system that can perform the role of expert consultant on certain types of problems in economic geology. The PROSPECTOR system is still in a strictly developmental stage, and it would be premature to attempt a systematic experimental evaluation of its performance. We can, however, describe the status of the system in general terms and outline our plans for future development.

Current Status

Currently, PROSPECTOR has 5 models of ore deposits. These are: a Kuroko-type massive sulfide model, provided by Prof. Charles F. Park, Jr. of Stanford University; a Mississippi-Valley-type lead/zinc model, provided by Dr. Neil Campbell, an independent consultant; a near-continental-margin porphyry copper model, provided by Prof. Marco T. Einaudi of Stanford University; a Komatiitic nickel sulfide

model, provided by Prof. Anthony Naldrett of the University of Toronto, and; a sandstone uranium model, provided by Mr. Ruffin Rackley, an independent consultant.

The models vary considerably in size and sophistication. As a group, the models require an inference net of several hundred spaces for their representation. The porphyry model, however, is approximately ten times the size and complexity of the lead/zinc model. (We should emphasize, of course, that size per se does not necessarily make a model either more or less useful.) The taxonomy supporting the models currently has approximately 1000 entries, many of which are synonyms for rocks and minerals.

To some extent, the differences in size and complexity among the models reflect the degree to which various types of deposits have been studied. Much more significant in our work has been the evolution of our methodological approach to the problem of eliciting models from experts and representing them in PROSPECTOR. The earliest models we acquired had a form suggestive of a somewhat elaborate checklist (although still represented as an inference net.) The most recently acquired models show far greater logical complexity; they also evince a common general structure, in which the models proceed from broad regional considerations through progressively greater detail to the inclusion of such fine-grain features as mineralization and alteration products.

The most elaborate model we have acquired to date has necessitated about 50 hours of direct contact with the expert consultant, a total we do not deem unreasonable. Some additional time is also required to furnish ancillary information about the model, and to run initial tests designed to detect significant errors.

We have not attempted to run PROSPECTOR consultations on a large number of actual cases. However, we have demonstrated the system in an experimental mode to many exploration geologists of varying levels of experience. An interesting, if informal, observation has been that virtually all geologists who have experimented with PROSPECTOR have learned something new about ore deposit models during the consultation session. This suggests that PROSPECTOR may have significant value as an educational tool.

An important consideration in using PROSPECTOR either for educational purposes or for consulting on an exploration project is the cost of necessary computer time. Our experimental version of the system has been implemented on a third generation DEC PDP-10 computer that supports up to a hundred time-shared users simultaneously. The main programming language used is INTERLISP, a high-level language especially designed for ease of experimentation with systems like PROSPECTOR. In these circumstances a normal interactive session would cost approximately $10.

Plans for Further Development

We plan to extend PROSPECTOR by working, as we have done to date, on two main topics: extension of the knowledge base and of the basic capabilities of the program.

We plan to extend the knowledge base by adding new models of ore deposits, by extending and refining the models that already exist in the system, and by considerably extending and refining the taxonomy of rocks, minerals and morphological descriptors. One problem we expect to confront in this activity is the interactions among models (and submodels) that can be expected to multiply as the number of models grow. Among the types of deposits we plan to model are chromite deposits, Archean massive sulfides, porphyry molybdenum deposits, and porphyry skarn deposits.

The models developed to date lend themselves to consultations that are typical of an early stage in an exploration venture. We plan to extend the usefulness of PROSPECTOR to later exploration stages. Specifically, we plan to allow the user to furnish the system with sketch maps summarizing initial geological, geophysical, and geochemical data. The map would be interactively provided using a digitizing tablet. PROSPECTOR's knowledge base will be augmented so that the favorability of candidate drill sites can be inferred from the data in the sketch map.

We expect to improve the basic capabilities of the PROSPCTOR program in several ways to make the knowledge-acquisition phase of the work easier, and to provide the user of the system with more directly

relevant information. One valuable prospective addition would be to improve the ability of the system to explain its reasoning processes upon request. In principle, the approach we have adopted in PROSPECTOR is particularly well-suited to satisfying this need, since reasoning proceeds through a well-defined sequence of specific steps. As we saw in the illustrative example, the system currently possesses a limited ability to summarize the status of a consultation session, but is not yet able to respond to a wide range of detailed questions. An improved ability to explicate the reasoning process is essential because, in the final analysis, a system like PROSPECTOR cannot make ultimate decisions about the course of an exploration program. Its function must always be to aid the professional-- it obviously cannot replace him.

VI ACKNOWLEDGEMENTS

The work described in this paper is the result of contributions by many economic geologists and computer scientists. We wish particularly to acknowledge the efforts of Alan Campbell, Neil Campbell, V. F. Hollister, Anthony Naldrett, Charles Park, Jr., Ruffin Rackley, Phyllis Barrett, John Gaschnig, Kurt Konolige, Nils Nilsson, Rene Reboh, Jonathan Slocum, and B. M. Wilber.

REFERENCES

1. Duda, R. O., P. E. Hart, and N. J. Nilsson, 1976, "Subjective Bayesian Methods for Rule-Based Inference Systems," National Computer Conference 1976, AFIPS Conference Proceedings Vol. 45, pp. 1075-1082.

2. Duda, R. O., P. E. Hart, N. J. Nilsson, and G. L. Sutherland, 1978, "Semantic Network Representations in Rule-Based Inference Systems," Technical Note 136, Artificial Intelligence Center, Stanford Research Institute, Menlo Park, California (March 1977) , to appear in Pattern-Directed Inference Systems, D. A. Waterman and F. Hayes-Roth, eds. Academic Press, New York, New York, in press.

3. Duda, R. O., P. E. Hart, N. J. Nilsson, R. Reboh, J. Slocum, and G. L. Sutherland, 1977, "Development of a Computer-Based Consultant for Mineral Exploration," Annual Report, SRI International Projects 5821 and 6415, SRI International, Menlo Park, California.

4. Ellis, J., D. P. Harris, and N. Van Wie, 1975, "A Subjective Probability Appraisal of Uranium Resources in the State of New Mexico," U.S. ERDA, Grand Junction, Colorado, p. 97.

5. Harris, D. P., G. Berry, and A. Freyman, 1970, "The Methodology Employed to Estimate Potential Mineral Supply of the Canadian Northwest-- an Analysis Based on Geologic Opinion and Systems Simulation," Proc. Ninth Int. Symposium on Decision Making, APCOM, p. 51.

6. Harris, D. P., 1977 Mineral Endowment, Resources, and Potential Supply: Theory, Methods for Appraisal, and Case Studies>, in press. (Preliminary 1977 edition available from the author.)

7. Hendrix, G. G., 1975, "Expanding the Utility of Semantic Networks Through Partitioning," Proc. 4th International Joint Conference on Artificial Intelligence, pp. 115-121, Tbilisi, Georgia, USSR.

8. Linstone, H. and M. Turoff (eds.), 1975, The Delphi Method-- Techniques and Applications (Addison-Wesley Publishing Co., Reading, Massachusetts).

9. Nilsson, N. J., 1971, Problem Solving Methods in Artificial Intelligence (McGraw-Hill Book Co., New York, New York),

10. Pople, H., 1977, "The Formation of Composite Hypotheses in Diagnostic Problem Solving: An Exercise in Synthetic Reasoning," Proc. 5th International Joint Conference on Artificial Intelligence, pp. 1030-1037, MIT, Cambridge, Massachusetts.

11. Raiffa, H., 1968, Decision Analysis, (Addison-Wesley Publishing Co., New York, New York).

12. Raphael, B., 1976, The Thinking Computer (W. H. Freemand and Co., San Francisco, California.

13. Shortliffe, E. H., 1976, Computer-Based Medical Consultations: MYCIN (American Elsevier Publishing Co., New York, New York).

14. Weiss, S. M., C. A. Kulikowski, and A. Safir, 1977, "A Model-Based Consultation System for the Long-Term Management of Glaucoma," Proc. 5th InternAtional Joint Conference on Artificial Intelligence, pp. 826-833, MIT, Cambridge, Massachusetts.

15. Winston, P. H., 1977, Artificial Intelligence, (Addison-Wesley Publishing Company, Reading, Massachusetts).

16. Zadeh, L. A., 1965, "Fussy Sets," Information and Control, Vol. 8, pp. 338-353.

DEVELOPMENT OF THE PROSPECTOR CONSULTATION SYSTEM FOR MINERAL EXPLORATION

Final Report
Covering the period October 1, 1976 to September 30, 1978

SRI Projects 5821 and 6415

USGS Contract No. 14-08-0001-15985
NSF Grant No. AER77-04499

October 1978

By: Richard O. Duda, Peter E. Hart, Phyllis Barrett,
John G. Gaschnig, Kurt Konolige, Rene Reboh,
and Jonathan Slocum

Artificial Intelligence Center
Computer Science and Technology Division

Prepared for:

The Office of Resource Analysis
U.S. Geological Survey
Reston, Virginia 22092

and

The Mineral Resource Alternatives Program
The National Science Foundation
Washington, D.C. 20550

Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors, and do not necessarily reflect the views of either the U.S. Geological Survey or the National Science Foundation.

SRI International
333 Ravenswood Avenue
Menlo Park, California 94025
(415) 326-6200
Cable: SRI INTL MNP
TWX: 910-373-1246

SRI International

DEVELOPMENT OF THE PROSPECTOR CONSULTATION SYSTEM FOR MINERAL EXPLORATION

Final Report
Covering the period October 1, 1976 to September 30, 1978

SRI Projects 5821 and 6415

USGS Contract No. 14-08-0001-15985
NSF Grant No. AER77-04499

October 1978

By: Richard O. Duda, Peter E. Hart, Phyllis Barrett,
John G. Gaschnig, Kurt Konolige, Rene Reboh,
and Jonathan Slocum

Artificial Intelligence Center
Computer Science and Technology Division

Prepared for:

The Office of Resource Analysis
U.S. Geological Survey
Reston, Virginia 22092

and

The Mineral Resource Alternatives Program
The National Science Foundation
Washington, D.C. 20550

Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors, and do not necessarily reflect the views of either the U.S. Geological Survey or the National Science Foundation.

Approved:

Peter E. Hart, Director
Artificial Intelligence Center

David H. Brandin, Executive Director
Computer Science and Technology Division

SRI International

ABSTRACT

This report describes the results of a two-year research effort to develop a computer program, called Prospector, that can serve as a consultant for certain problems in mineral exploration. The program uses special network structures to encode various types of geological knowledge, including models of different types of ore deposits. When provided with information about a particular prospect, the program matches that information against its models, requests additional information, evaluates the likelihood of the presence of an ore deposit, and provides an explanation for its reasoning.

In the report the theoretical principles underlying Prospector's operation are reviewed. These include the use of semantic networks and inference networks to draw logical and plausible conclusions from the data. The status of the model development is described, with particular emphasis on the porphyry copper and the nickel sulfide models. The results of preliminary tests of these models are also given. A new application of Prospector to the selection of drilling sites is presented as an example of using the system with spatial data. Finally, system improvements during the last year are described, the model construction methodology is presented, and recommendations are made for future work.

CONTENTS

ABSTRACT iii

LIST OF ILLUSTRATIONS vii

LIST OF TABLES ix

ACKNOWLEDGMENTS 1

1 INTRODUCTION AND OVERVIEW 3

1.1 Prospector and its Applications 3

1.2 Overview of the Report 5

2 REPRESENTATION AND USE OF GEOLOGICAL KNOWLEDGE 7

2.1 System Overview 7

2.2 Representation of a Model as an Inference Network . . 10

2.3 Probabilistic Reasoning 13

2.4 Probability, Certainty, and Contexts 17

2.5 Semantic-Network Representations of Spaces 19

3 MODEL DEVELOPMENT 23

3.1 Model Status 23

3.2 Revision of the Porphyry Copper Model 26

3.3 Revision of the Nickel Sulfide Model 45

4 PRELIMINARY PERFORMANCE ANALYSIS OF THE KNSD AND PCDA MODELS 59

4.1 Testing Prospector Models 59

4.2 Methodology of Experiments 63

4.3 Performance Analysis of the KNSD Model 65

4.4 Performance Analysis of the PCDA Model (Version 1) . . 77

4.5 Performance Analysis of the PCDA Model (Version 2) . . 87

4.6 Conclusions and Future Work in Performance Analysis . 94

5 RULES FOR DRILLING-SITE SELECTION 101

5.1 Introduction 101

5.2 Revision of Hollister's Rules 102

5.3 Test Examples for Hollister's Rules 114

6 SYSTEM IMPROVEMENTS 133

6.1 Quantitative Rules 133

6.2 Spaces with Variables 138

6.3 Graphic Input 139

6.4 The Network Compiler 140

6.5 Savemode and Batchmode 148

6.6 Model Implementation Aids 149

7 THE MODEL-BUILDING PROCESS: TECHNIQUES AND TOOLS 151

7.1 Introduction 151

7.2 Procedure for Model Construction 153

7.3 General Observations 167

8 CONCLUSIONS AND RECOMMENDATIONS 169

REFERENCES 173

APPENDICES

A INPUT DATA FOR THREE KNOWN DEPOSITS FOR THE KNSD MODEL . . 177

B INPUT DATA FOR THREE KNOWN DEPOSITS FOR THE PCDA MODEL . . 185

ILLUSTRATIONS

1 System Diagram for Prospector 8

2 A Hypothetical Inference Network 14

3 A Simple Semantic Net 21

4 Inference Network for PCDA 31

5 Inference Network for KNSD 48

6 Inference Network for Drilling-Site Selection 103

7 Geological Map of Island Copper Deposit 116

8 Data for Island Copper 117

9 Intermediate Results for Island Copper 120

10 Final Results for Island Copper 122

11 Geological Map of Bell Copper 123

12 Data for Bell Copper 124

13 Final Results for Bell Copper 125

14 Geological Map of Gibraltar 128

15 Data for Gibraltar 129

16 Final results for Gibraltar 131

17 Program Organization for Graphical Data Processing 141

TABLES

1 Statistics for the Prospector Models 24

2 Status of the Implementation 25

3 Statistics for the First PCDA Implementation 27

4 Statistics for the Second PCDA Implementation 30

5 Statistics for the First KNSD Implementation 46

6 Statistics for the Second KNSD Implementation 47

7 Comparison of Prospector Scores with Naldrett's Estimates . . 69

8 Relative Error (E) of Prospector Scores as Predictors of Naldrett's Estimates (Derived from data in Table 7) 72

9 Prospector Scores for KNSD Model Vary With the Degree of Certainty of the Inputs ("Standard" data from Table 7) . . . 73

10 Sensitivity Measure (S) for the KNSD Model (derived from Data in Table 9) 76

11 Prospector Scores for Several Levels of the PCDA Model (Version 1) 80

12 Zone Interpretations in the PCDA Model (Version 1) 82

13 Prospector Scores for PCDA Model (Versionl) Vary with degree of Certainty of Inputs ("Standard" Data from Table 11) . . 84

14 Sensitivity Measure (S) for the PCDA Model (Version 1) (derived from data in Table 13) 86

15 Prospector Scores for Several Levels of the PCDA Model (Version 2) 90

16 Relative Error (E) of Prospector Scores as Predictors of Einaudi's Estimates (derived from data in Table 15) 92

17 Zone Interpretations in the PCDA Model (Version 2) 93

18 Parameters for Favorability Functions 113

ACKNOWLEDGMENTS

We are indebted to many people who have contributed to the development of the Prospector system. In particular, Alan N. Campbell, the late Neil Campbell, Marco T. Einaudi, Victor F. Hollister, Anthony J. Naldrett, and Charles F. Park, Jr. were instrumental in providing the geological knowledge that we have tried to incorporate into the system. The assistance and advice of Gerald Askevold, William Beatty, Joseph Botbol, Sheldon Breiner, Allen Clark, Norman Page, Donald Singer, Edwin Tooker, and John Worden are also greatly appreciated.

We would like to take this opportunity to give special credit to individual members of the SRI staff for their respective contributions to the development of Prospector. Project leadership was shared by Peter E. Hart and Richard O. Duda, who also implemented the porphyry copper model. Rene Reboh produced the overall system implementation. Kurt Konolige implemented the nickel sulfide model, and developed both the network compiler and the PARSEFILE program. John Gaschnig devised and executed the model evaluation experiments. Jonathan Slocum developed the LIFER grammar for volunteering information. Phyllis Barrett created the graphics input and output programs. Finally, during the first year, Nils J. Nilsson and Georgia L. Sutherland made major contributions to the design and implementation of the system.

1 INTRODUCTION AND OVERVIEW

1.1 Prospector and its Applications

This report describes the results of a two-year research effort to develop a computer-based consultation system for certain problems in mineral exploration. Jointly funded by the Office of Resource Analysis of the U.S. Geological Survey, the Nonrenewable Resources Section of the National Science Foundation, and the Exxon Corporation, this work has led to the development of an interactive consultation program called Prospector.

Prospector was designed primarily to assist exploration geologists in interpreting and evaluating data on specific mineralized sites or prospects. In this mode of operation the program accepts simple English sentences describing a prospect, matches these observations against its specially coded models of known classes of ore deposits, requests additional information that could help establish a better match, and provides an evaluation of the available data. Since many prospects must be investigated to find one ore body, the main purpose of the program is not to classify a discovered ore body, but rather to reduce the risk of overlooking unfamiliar possibilities and to identify specific additional information which, if obtained, would be most useful for reaching more definite conclusions.

The practical applicability of such a program clearly depends upon the number of ore deposit models it contains, as well as upon the quality of each model. Since Prospector is part of a research effort, its coverage of the known classes of ore deposits is far from complete. Our intent has been to encode models that, to the extent we were able to determine, represent the best available information about each class of deposits modeled.

Prospector currently contains models of four different classes of ore deposits: a Kuroko-type massive sulfide model, a Mississippi-Valley-type carbonate lead/zinc model, a near-continental-margin porphyry copper model, and a Komatiitic-type nickel sulfide model.* Each model has been developed by interviewing an experienced exploration geologist who has special, expert knowledge about that deposit class. Thus, Prospector can be viewed as a system for providing a field geologist access to the specialized knowledge of a panel of authorities on a variety of types of ore deposits.

While the use of Prospector as an interactive consultation system continues to be our primary interest, there are several other potential applications that deserve mention. In Section 5 of this report we show how Prospector, by using different kinds of models to evaluate the favorability of candidate drilling sites, may help a geologist to locate an ore body. A natural extension of this work would be to help estimate its size and grade. Applied to regional-scale maps, this approach might be of value for locating regions that are most favorable with respect to a particular ore-deposit model and could assist in the task of regional resource evaluation. Another possible application of Prospector is to the screening of computerized data bases. In principle, whenever a new model is developed large files on mineralized properties could be systematically scanned to select promising candidates for subsequent detailed examination.

In addition, there are a number of indirect benefits from this research effort. Most of these are a consequence of the discipline required to express ore-deposit models in a formalism that can be used by a computer program. Since the observational data relating to a model and the way in which the data are combined must be stated precisely and unambiguously, even rudimentary quantitative comparison and evaluation of ore-deposit models become possible for the first time. The precise statement of data requirements could also be helpful to Survey personnel who have the responsibility for gathering field data of general use for mineral exploration.

* A Western-states sandstone uranium model is also being developed under a separate contract.

Finally, most geologists who have witnessed a session with the program have commented that it is a potentially valuable teaching tool. Although a different mode of operation would be needed to turn Prospector into an effective educational system, its detailed and specific models should prove very helpful to anyone who wants to learn more about such deposits.

1.2 Overview of the Report

This report is addressed primarily to readers who are interested in the applications of advanced computer science techniques to problems in economic geology. The work draws upon the accumulated geological knowledge of ore deposits that can be expressed in terms of exploration models. It also draws upon a body of computer science techniques that have been developed as a consequence of research in artificial intelligence. We believe that this effort will in turn contribute to both of these disciplines.

One problem that faces an interdisciplinary project of this kind is that few people possess a deep understanding of both disciplines. There is a consequent danger that, to understand the work, economic geologists may fear that they must learn too much about computer science techniques, and computer scientists may fear that they must learn too much about the geology of ore deposits.

We have not attempted to provide background material on either field--primarily because brief overviews, while helpful for orientation, are inadequate for an understanding in depth. We have tried to make this report sufficiently self-contained to enable it to be understood by economic geologists who have some appreciation of the role of computers in economic geology. For those readers who would like more background information on artificial intelligence, Raphael (1976) provides a readable general introduction, while Waterman and Hayes-Roth (1978) present current research that is closely related to the Prospector project.

The remainder of this report is organized as follows. Section 2 describes the structure of the program and the basic principles behind its operation. While much of this information is contained in our earlier papers and reports, it is included here in the interest of comprehensiveness.

Section 3 describes the status of our work on model development. Because most of our model work during the second year of this project has been devoted to refining the porphyry copper model and developing the nickel sulfide model, these models receive special attention.

Section 4 describes our initial results in subjecting models to quantitative tests. Rigorous testing of models of ore deposits is a new enterprise, engendering some difficult philosophical as well as practical problems. Consequently, a general discussion of model testing is included.

Section 5 presents the results of using Prospector for drilling site selection. While this is an interesting application in its own right, it also can be viewed more generally as an example of the way in which Prospector can be used with graphical input data.

Section 6 describes various technical improvements that have been made to the system during the second year. These include new procedures for utilizing contexts, handling quantitative data, treating spatially dependent data, compiling inference networks, and acquiring and editing models.

During the past two years we have evolved an informal methodology for interviewing exploration geologists and working with them to encode models in the Prospector formalism. We have also developed, and plan to continue to develop, various systems facilities to support this model construction process. For the benefit of those who would like to construct their own models, Section 7 discusses the steps in this process. Finally, Section 8 summarizes the current state of the research and makes recommendations for future work.

2 REPRESENTATION AND USE OF GEOLOGICAL KNOWLEDGE

2.1 System Overview

Figure 1 shows a diagram of the basic components of the Prospector system. The heart of the system is the subsystem called the Inference Network. Described in more detail in Section 2.2, the Inference Network expresses the various ore deposit models as a network of interconnections between field observations--such as "Hornblende has been altered to biotite"--and significant hypotheses--such as "There is favorable alteration for the potassic zone of a porphyry copper deposit."

A user of Prospector communicates with the system through a simple subsystem called the Executive. The Executive accepts commands from the user, interprets them, passes control to other subsystems, and returns responses to the user.

In a typical session, the user begins by volunteering relevant facts about the significant rocks, minerals, and alteration products present at the prospect. This is done through a series of simple English sentences, such as "There is a quartz monzonite intrusive," "The host rock is Tertiary granite," "Chalcopyrite is present," and "There is probably some biotite."

The Executive routes this information to the English Analyzer, which makes use of an SRI-developed system called LIFER to analyze each sentence (Hendrix, 1977). Those sentences that are successfully analyzed are represented in a network form (described in Section 2.5), and a subsystem called the Matcher performs the function of matching the observations against the assertions in the Inference Network.

Associated with each assertion in the Inference Network is a probability that the assertion is true. When a match is made between an

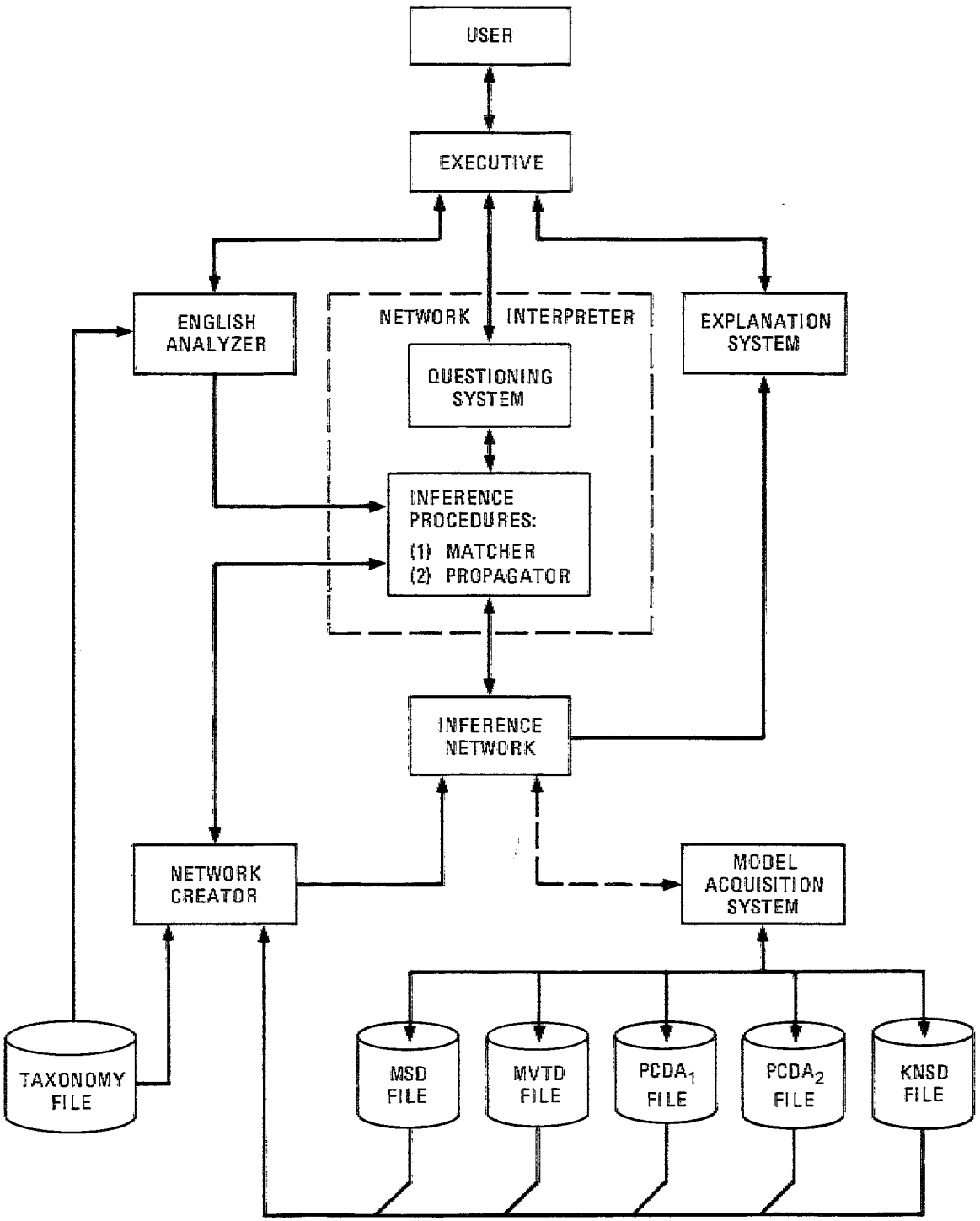


FIGURE 1 SYSTEM DIAGRAM FOR PROSPECTOR

analyzed observation and an assertion, the probability is changed and the consequences of that change are propagated throughout the network according to the procedures in the Propagator. These procedures are summarized in Section 2.3.

When the user has finished volunteering observations, the Executive passes control to the Questioning System. This subsystem examines the state of the Inference Network and selects the "best-matching" model. Since the user typically has limited observational data and has mentioned only a portion of that, the degree of match with even the best-matching model is usually rather low. Thus, the Questioning System tries to select the particular piece of unmentioned observational evidence that will be most effective in confirming or refuting a match.*

What typically follows is a cycle of questions and answers in which the Questioning System selects a piece of relevant evidence, the Executive asks the user about it, the user provides an answer (which may indicate uncertainty, including a total lack of information), the Propagator assesses the consequences of the answer, and the Questioning System then selects another piece of relevant evidence to inquire about.

Instead of answering a question, the user can take the initiative at any time to do such things as, for example, changing the answer to an earlier question, volunteering additional evidence, asking for a clarification of the question, or requesting an explanation. This latter request is handled by the Explanation System and takes one of two forms--an explanation of the reason for that particular question or a summary of the principal conclusions of the system at this point in the consultation. This ability of the system to examine the Inference Network and to produce explanations for its conclusions is quite important, and constitutes a major justification for calling Prospector a consultation program.

To provide its consultation services Prospector draws upon two kinds of geological knowledge--knowledge about specific ore-deposit

---

* The exact procedure used by the Questioning System is described in our last annual report (Duda et al., 1977).

models that is represented in the Inference Network, and general geological knowledge (such as the classifications of rocks, minerals, ages, and forms) that is represented in the Taxonomy. Both the coded models and the Taxonomy are separated from the Prospector program per se and reside on external disk files. The Network Creator reads these files to build the Inference Network. Thus, the geological knowledge in the system can be updated by editing these files and reexecuting the Network Creator.

We currently use a conventional text editor to edit these files. While workable, this approach has two disadvantages: (1) it is sometimes difficult to comprehend the interconnections in a network model when doing editing and to appreciate all the consequences of changes, and (2) it is time-consuming to run the Model Builder, which must process all the models even if only a few changes are made. For this reason, we are currently developing a Model Acquisition System that will make the process of creating and modifying model files much more convenient.

2.2 Representation of a Model as an Inference Network*

To be used in Prospector, an ore-deposit model must be represented in the form of an inference network that relates field observations or evidence to the hypotheses that are the important constituents of the model. This network is equivalent to a collection of rules of plausible inference, termed more simply "inference rules." In general, an inference rule has the form

IF E THEN (to degree LS, LN) H.

The rule is interpreted as meaning "The observed evidence E suggests (to some degree) the hypothesis H." A probability of truth is associated with every observation and hypothesis, and the inference rules specify how the probability that an hypothesis is true is changed by the observation of evidence. The way the two parameters LS and LN change

---

* Sections 2.2 through 2.5 provide background information that updates the corresponding material in Section I of our last annual report (Duda et al., 1977).

the probability and thus establish the "strength" of the rule is described in Section 2.3.

As a simple example, consider one rule in our porphyry copper model relating to the potassic zone of a porphyry deposit. This rule states

IF there are abundant quartz sulfide veinlets with no apparent alteration halos

THEN (LS, LN) there is alteration favorable for the potassic zone.

According to the model, observation of this evidence is quite encouraging--though not conclusive--that there is alteration characteristic of the potassic zone of this class of porphyry copper deposits. On the other hand, even the known absence of this evidence is only somewhat discouraging for this conclusion. In general, we need to be able to say how encouraging it is to find the evidence present, as well as how discouraging it is to find it absent; this is why two numbers--LS and LN--must be supplied by the expert for each rule.

Usually information about several pieces of evidence must be combined to reach a conclusion about any hypothesis. Prospector provides two basically different ways to combine evidence: (1) logical combinations, and (2) multiple inference rules.

When several inference rules share the same hypothesis, each one influences the probability of the hypothesis independently through its rule strength. In essence, the logarithms of the LS and LN values yield votes for or against the hypothesis, and the pieces of evidence are combined by adding these votes.

Different pieces of evidence can also be combined logically to form a single, compound piece of evidence. Arbitrary logical combinations are composed of simpler elements by means of the primitive operations of conjunction (AND), disjunction (OR) and complementation (NOT). For example, suppose we let E_a, E_b and E_c stand for the following pieces of basic evidence:

E_a : plagioclase has been altered to albite;

E_b : plagioclase has been altered to minor sericite;

E_c : plagioclase has been altered to major epidote.

Then the porphyry copper model combines this basic evidence in employing the logical expression $E = ((E_a \text{ OR } E_b) \text{ AND } (\text{NOT } E_c))$ to define the kind of plagioclase alteration that is favorable for the hypothesis of a barren-core zone.

If the probabilities associated with E_a, E_b, and E_c are P_a, P_b, and P_c, respectively, then the probabilities for logical combinations are computed as follows:

$$Pr(E_a \text{ OR } E_b) = \max \{P_a, P_b\}$$

$$Pr(E_a \text{ AND } E_b) = \min \{P_a, P_b\}$$

$$Pr(\text{ NOT } E_c) = 1 - P_c .$$

Thus, in the above example we obtain $P(E) = \min \{\max \{P_a, P_b\}, 1-P_c\}$.

It frequently happens that the hypothesis of one rule mentions the evidence of another. For example, our porphyry copper model also includes the following two rules:

Rule 1: IF volcanic rocks in the region are contemporaneous with the intrusive system (coeval volcanic rocks)

THEN (LS_1, LN_1) the level of erosion is favorable for a porphyry copper deposit.

Rule 2: IF the level of erosion is favorable for a porphyry copper deposit

THEN (LS_2, LN_2) there is a favorable regional environment for a porphyry copper deposit.

Because the rules mention each other, they form a "chain" from the evidence for Rule 1 to the hypothesis for Rule 2. In general, the rules in Prospector interconnect in various ways--through chains, through several pieces of evidence bearing on the same hypothesis, and through the same piece of evidence bearing on several different hypotheses. This is how the collection of rules forms an inference network, such as

the one shown in Figure 2. In this diagram assertions that correspond to hypotheses or pieces of evidence are shown in boxes, and the rules or logical relations that link these assertions are indicated as arrows. Following the semantic-network terminology introduced by Hendrix (1975), we refer to the boxes as "spaces" and the arrows as "arcs." The spaces at the tail of a rule are called antecedent spaces, while those at the head of a rule are called consequent spaces. The spaces at the top of the network represent hypotheses about the existence of a particular type of ore deposit and are called the top-level hypotheses.

It should be pointed out that there we have two different ways to treat rules that involve spaces that are similar but not equal. If one space describes a situation that is a special instance of another, then, using the procedure described in Section 2.4.3 of our last annual report (Duda et al., 1977), we create new rules that ensure that presence of the special case implies presence of the general case and that, conversely, absence of the general case implies absence of the special case.* If a space describes a situation that may exist simultaneously in several places on the prospect, we must be able to assign different probabilities to each instance. For example, if chalcopyrite is thought to exist in one zone of a porphyry copper deposit, but not in another, then a high probability should be assigned to the one instance and a low probability to the other. The methods currently used to treat spaces containing location variables are described in Section 3.2.

2.3 Probabilistic Reasoning

An inference network of the form shown in Figure 2 is used to represent the general part of the knowledge that Prospector has about various types of ore deposits. When engaged in consultation about a particular prospect, the system also needs to record the specific geological evidence as it is furnished by the user. This evidence is stored as probabilities that are associated with each space in the network. Suppose that the Questioning System asks the user about the

---

* We use the semantic-network representation described in Section 2.5 to detect such relations between spaces automatically.

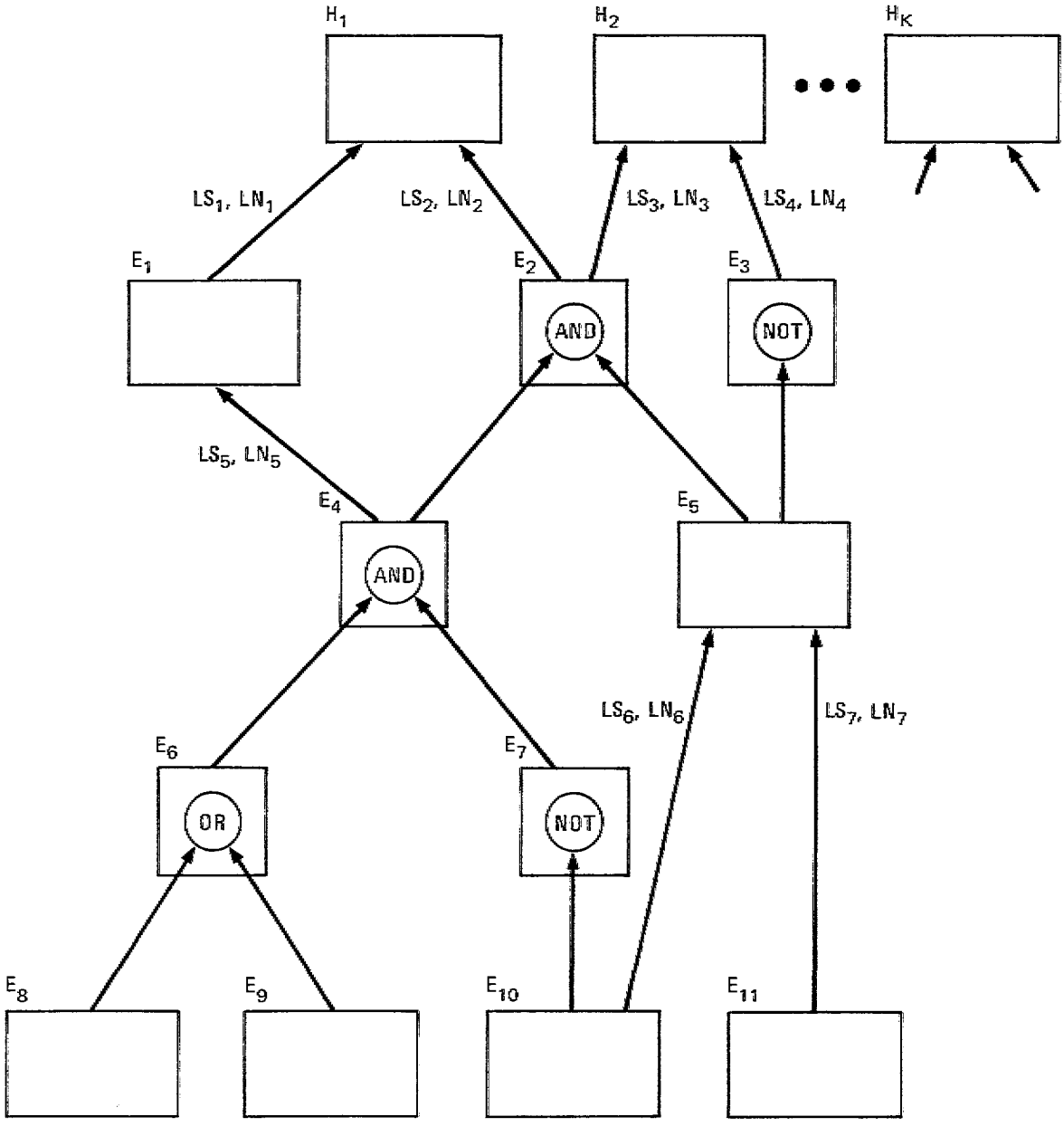


FIGURE 2 A HYPOTHETICAL INFERENCE NETWORK

evidence represented by a particular space, and that the user replies that the evidence is present. Prospector records this by setting the probability of the corresponding space to one. At any point in the consultation some spaces will not have been assigned a probability based on the user's observations. These spaces have "default probabilities" that are initially provided by the expert geologist at the time the model is constructed. Such probabilities, known technically as prior probabilities, are available for every space in the network.

The principal form of reasoning in Prospector is the propagation of probabilities through the Inference Network. As an example, suppose in Figure 2 that the user provides some geologic evidence regarding space E_5 by changing the probability of that space to a new value. This change should have an effect on the probabilities of spaces E_2 and E_3 which, in turn, might change the probabilities of the top-level hypotheses H_1 and H_2.

Propagation of probabilities is accomplished through the application of a form of reasoning known as Bayesian decision theory (Raiffa, 1968). This theory prescribes a method for propagating a probability from evidence E of a rule to hypothesis H. Propagation throughout the Inference Network is then a matter of iterating this procedure.

The basis for the procedure is an elementary theorem of probability theory called Bayes' rule. For our purposes, the so-called odds-likelihood form of the rule is most convenient. This form relates three quantities involving E and H: the prior odds* O(H) on the hypothesis, the posterior odds O(H|E) on the hypothesis (given that the evidence E is observed to present) and the sufficiency measure LS mentioned previously. By Bayes' rule, we can write

$$O(H|E) = LS * O(H) .$$

---

* The odds on any evidence (or hypothesis) are just the ratio of the probability in favor of the evidence to the probability against the evidence. Probabilities and odds are therefore freely interchangeable through this simple relation.

The quantity LS has a standard interpretation in statistics and is called the likelihood ratio. Thus, Bayes' rule tells how the odds on the hypothesis of a rule are updated by observing the presence of the evidence for the rule: the prior odds are simply multiplied by the likelihood ratio LS. An analogous formula tells how the odds on the hypothesis are updated if the evidence is observed to be absent; in this case, the prior odds are multiplied by the necessity measure LN:

$$O(H|\sim E) = LN * O(H) .$$

Direct application of Bayes' rule leads, therefore, to simple formulas for updating the probability of a hypothesis, provided that the user observes either the definite presence or absence of the evidence. In actual practice the user is often unable to make such definite statements. Typically, the user is prepared only to indicate a degree of confidence that the evidence sought is present. In this case, a formula for updating the probability of the hypothesis can be derived that effectively interpolates between the two extreme cases of perfect certainty.* Let E' denote the observations that cause the user to suspect the presence of the evidence E, and let O(H|E') denote the interpolated odds produced by this procedure; the effective likelihood ratio L' is then defined by

$$L' = \frac{O(H|E')}{O(H)} .$$

Clearly, the effective likelihood ratio ranges from LS when E is definitely present to LN when E is definitely absent, being unity when H is at its prior odds. When several rules share the same hypothesis, the posterior odds are computed by multiplying the prior odds by the product of the effective likelihood ratios. These Bayesian procedures, therefore, allow us to propagate the consequences of any observation throughout the inference network.

---

* This formula is given and discussed in Section 2.2 of our last annual report (Duda et al., 1977).

2.4 Probability, Certainty, and Contexts

When the user is certain about the presence or absence of requested evidence, the Bayesian procedures are directly applicable. When the user is uncertain, however, we face the problem of measuring that uncertainty. For example, suppose we ask the user a question of the form "Is E present on the prospect?" and suppose that the user seriously doubts its presence, yet cannot rule it out. Ideally, we would like the user to indicate this state of affairs by giving a probability to the system, say 0.1. Unfortunately, this might be higher than the prior probability assigned to this observation by the expert, and the system could interpret the user's response as indicating the possible presence of the evidence. In general, neither the expert nor the user can be relied on to assign accurate probability values to situations, particularly when the situations are rare events. For instance, is the prior probability that an ore body exists on a prospect one in a thousand, one in a million, or something else entirely?

To overcome this problem, we have adopted a method previously employed in a medical diagnosis program (Shortliffe, 1976), and we communicate with the user in terms of certainties instead of probabilities. By our definition the certainty C(H|E) is related to the probability P(H|E) through the formula

$$C(H|E) = \begin{cases} 5\,\dfrac{P(H|E) - P(H)}{1 - P(H)} & \text{if } P(H|E) > P(H) \\[2ex] 5\,\dfrac{P(H|E) - P(H)}{P(H)} & \text{otherwise.} \end{cases}$$

Thus, the certainty is a continuous, piecewise linear function of the probability. It ranges from -5 when H is definitely false to + 5 when H is definitely true, the value 0 corresponding to the no-information situation in which P(H|E) = P(H). Thus, if the user is asked a question and, if he or she has no opinion one way or the other, a response of

zero certainty will automatically be converted into the prior probability used internally by the program in its Bayesian calculations.

Many of the spaces in a model correspond to abstract hypotheses. These spaces are marked as being "unaskable," and a user is never asked about them directly. Sometimes it does not make sense to ask the user about a particular piece of evidence until a proper context for the question has been established. For example, it would be pointless to ask about the size of an intrusive before its existence has been established. Let E_1 and E_2 denote the following situations:

E_1 : An intrusive system is present

E_2 : The diameter of the youngest major stock in the intrusive system is less than 5000 feet.

If E_1 must be established before asking about E_2, then E_1 is a context for E_2. In our inference net diagrams, we denote this by a dotted arrow from E_2 to E_1. Our standard assumption is that if E_1 is known with any probability that is greater than its prior value, (i.e., with any positive certainty), then that is sufficient to establish E_1 as a context. However, we always require the system to try all available rules for establishing E_1 before declaring it established, so that, unless the user changes his or her answers, the status of E_1 will not change during the session.

We have generalized this idea by allowing the model builder to specify a certainty interval that establishes the context. The standard interval is (0,5]. If we want to ask about E_2 only if E_1 is not true, we could use [-5,0) for the interval. Similarly, if we want to ask about E_2 only if E_1 is unknown, we could use (-2, 2), say, for the interval. As a special case, we sometimes use the interval [-5, 5] merely to compel the Questioning System to ask about E_1 before considering E_2.

2.5 Semantic-Network Representations of Spaces

The foregoing sections describe how Prospector represents models of ore deposits as collections of inference rules, and outlines the Bayesian computation of probabilities that enable the effect of a piece of evidence to be propagated through an inference network. If these were the only mechanisms employed by the system, it would be seriously deficient in several ways. Many of these deficiencies would relate to the fact that the system would have no "understanding" of the content of the rules, much less of the whole collection. Of course, "understanding" is a subtle concept to come to grips with (for people or for computers), but a modest start can be made by noticing that each rule can be broken down into parts and the parts related to one another.

For a simple example, consider the evidence requested in the following hypothetical rules:

Rule 5: IF pyrite in veinlets is present
THEN H5,

and

Rule 6: IF sulfides are present
THEN H6.

Obviously, a user able to observe pyrite as requested in Rule 5 has surely also observed the presence of sulfides, while a user unable to observe any sulfides will surely be unable to observe pyrite. Prospector needs two mechanisms if it is to deal with this sort of elementary but pervasive reasoning. A taxonomy of minerals will clearly be needed in order to infer that pyrite is a sulfide. Less obviously, the rules must be stored in such a way as to reveal the meaning of their parts. In this example the internal representation of "pyrite in veinlets" must permit the system to perceive that "pyrite" is part of the statement. (Note, incidentally, that it would be unsatisfactory for the system merely to scan for the keyword "pyrite," for incorrect inferences would then be made from such statements as "absence of pyrite" or "has the same color as pyrite.")

We address the general set of problems alluded to here by using a network structure to represent the spaces for the rules. This kind of representation, called a "semantic network," has been developed through work in artificial intelligence for just such purposes. Viewed in their entirety, semantic networks comprise such a complicated subject that even a modest exposition of the topic would be beyond the scope of this report. Nevertheless, it is easy to see from Figure 3 how a semantic network can be used to solve our example problem, and thus to appreciate the way in which semantic networks are employed in Prospector.

Each heavily outlined space to the left in Figure 3 corresponds to a space of an inference network, as discussed up to this point. Now, however, each of these spaces has an internal structure. The evidence space for Rule 6 shows that the evidence sought is the presence of anything composed of sulfide minerals. The evidence space for Rule 5 shows that what is sought is the presence of anything composed of pyrite in the form of veinlets. To the right in the figure is a portion of a taxonomy of minerals, which allows the system to make the elementary deduction that pyrite is a member of the sulfide group. The structure displayed, therefore, allows Prospector to draw the inferences needed in this example. The procedures used to adjust probabilities in accordance with these deductions are described in Section 2.4.3 of our last annual report (Duda et al., 1977).

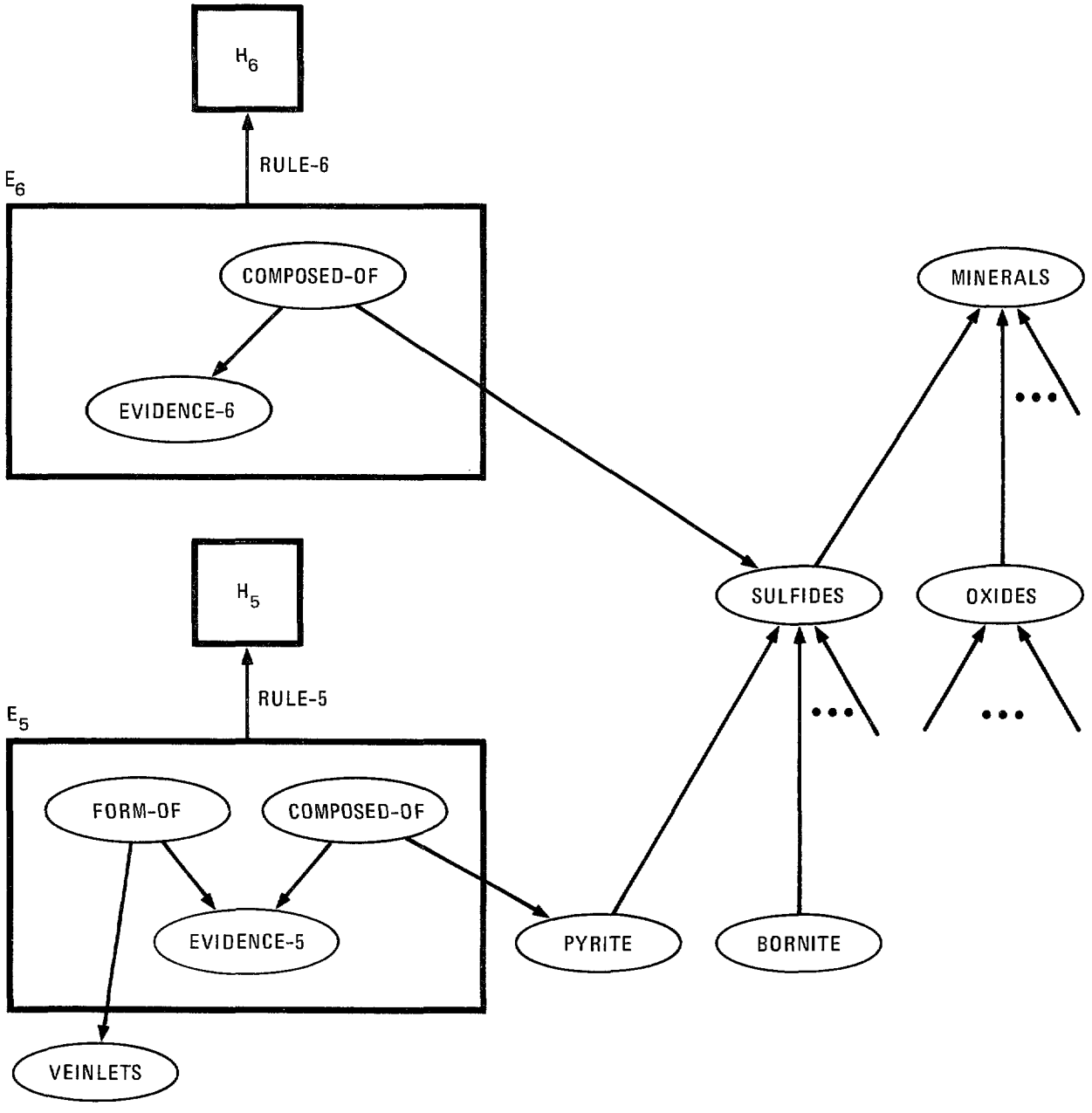


FIGURE 3 A SIMPLE SEMANTIC NET

3 MODEL DEVELOPMENT

3.1 Model Status

As mentioned previously, Prospector currently contains models of four different classes of ore deposits:

MSD -- a Kuroko-type massive sulfide model, contributed by Prof. Charles F. Park, Jr., of Stanford University,

MVTD -- a Mississippi-Valley-type deposit model, contributed by the late Dr. Neil Campbell,

PCDA -- a near-continental-margin porphyry copper model, contributed by Prof. Marco T. Einaudi of Stanford University, and

KNSD -- a Komatiitic nickel sulfide model, contributed by Prof. Anthony J. Naldrett of the University of Toronto.

The number of spaces and the number of rules in a model are simple measures of its size and complexity. These statistics for the four models are given in Table 1. From these we see that the porphyry copper model (PCDA) is by far the largest, and that the nickel sulfide model (KNSD) is about twice the size of either the massive sulfide model (MSD) or the Mississippi-Valley model (MVTD).

Table 1

Statistics for the Prospector Models

Model	Number of Spaces	Number of Rules
MSD	39	34
MVTD	28	20
PCDA	187	91
KNSD	75	49

The implementation of a model typically takes place in several phases. The first phase is the construction of the basic inference network. This involves identification of the important hypotheses, the relevant field evidence, and the structure of the net. The second phase is the determination of the parameter values--the likelihood ratios and the prior probabilities. At the end of this phase the program can be run, although without the semantic networks it cannot accept volunteered information, and will not make deductions involving the taxonomy. The third phase is creation of the semantic networks, the fourth phase composition of the explanatory text associated with spaces and rules, and the final phase testing and verification. While the earlier phases must be brought to some state of completion before work can start on the later phases, the implementation process is not strictly sequential; work on the later phases often leads to modifications of the earlier phases.

Table 2 shows the status of the implementation in terms of these developmental phases. Although much of the explanatory text is still missing, its incorporation is a routine matter, and all of the models

are otherwise fully functional. Testing, however, remains in a preliminary state, with our results to date discussed in Section 4.

Table 2

Status of the Implementation

Model	Inference Networks	Parameter Values	Semantic Networks	Descriptive Text	Testing
MSD	Complete	Complete	Complete	Missing	None
MVTD	Complete	Complete	Complete	Complete	None
PCDA	Complete	Complete	Complete	Partial	Preliminary
KNSD	Complete	Essentially Complete	Complete	Missing	Preliminary

The construction of these models took place over a two-year period, during which, through trial and error, we evolved a general approach to encoding models.* Because our methods changed during that time, the earlier models are not only smaller than the later ones, but they also have a different structural organization.

The MSD model was the first one to be implemented. Its encoding--given in Appendix C of our last annual report (Duda et al., 1977)--has not been changed. It is the least structured of the four models, with the majority of the rules going directly to the top-level hypothesis. Were we to revise that model, we would probably change its structure to a form more like that of the KNSD model.

The MVTD model was the second one implemented. Like the MSD model, its encoding--given in Appendix D of the aforementioned report--has not undergone revision. While revision would no doubt be beneficial, the

---

* This methodology is described in Section 7.

structural organization of that model is closer to the form that we are now using. In particular, the model has simple but clearly defined subsections devoted to the regional environment, the host rock, the mineralization, and the alteration, which is also the general form of the later models.

The PCDA model was the third model implemented. An initial encoding of that model was given in Appendix E of our last annual report (Duda et al., 1977). That model was also the first to receive any systematic testing and evaluation, and these tests lead to a substantial revision of the encoding. The reasons for and the results of that revision are described in Section 3.2. Suffice it to say that the basic structure of the encoding has remained the same, but that many detailed aspects, such as the rule strengths, have been significantly changed.

The initial encoding of the KNSD model was described in Appendix A of our Second Semiannual Report (Hart et al., 1978). The basic structure of the model had been established at that time, but the encoding was not actually complete. Since then the initial implementation was completed, and a first revision of that implementation, described in Section 3.3, is largely complete.

One of the evident conclusions is that model development is a continuing major activity. The rule-based organization of the system greatly simplifies the process of revising old models and constructing new ones. However, model building is still a time-consuming and intellectually demanding activity. Any tools that facilitate the model acquisition process are extremely valuable.

3.2 Revision of the Porphyry Copper Model

The Type-A porphyry copper model (PCDA) of Prof. Marco T. Einaudi describes a class of post-Paleozoic, near-continental-margin deposits in which the intrusive system and the wall rocks possess the same bulk chemical composition. The geology of such deposits was discussed in Appendix A of our last annual report (Duda et al., 1977), and the inference network for our first version of the model was given in

Appendix E of that report. In this section we summarize that first encoding, indicate the problems that were discovered when that version was tested, and describe the current version of the model.

Both versions of the model have the same overall structure: a subdivision into sections dealing with the petrotectonic setting, the regional environment, and the intrusive system. The section on the intrusive system is further subdivided into three subsections treating the compositional characteristics, the size, morphology, and rock texture characteristics, and the alteration and mineralization zones. More than half the model is devoted to describing the characteristics of the four most common alteration zones--the barren-core (or root) zone, the propylitic zone, the potassic zone, and the sericitic zone. A breakdown of the number of spaces and rules devoted to each part of the first version of the model is given in Table 3.

Table 3

Statistics for the First PCDA Implementation

Model Section	Number of Spaces	Number of Rules
PCDA: Total	204	64
Petrotectonic setting	9	0*
Regional environment	19	11
Intrusive system	175	53
Composition	16	3
Size, morphology & texture	16	6
Alteration & mineralization	142	44
Barren-core zone	27	5
Propylitic zone	19	7
Potassic zone	42	11
Sericitic zone	37	14
Zone combinations	16	7

---

* Although there are 9 spaces in this section of the model, they participate only in logical combinations, not as antecedents or consequents of inference rules.

While the implementation of the section of the model devoted to the alteration zones has undergone only minor revision, all other parts of the model were thoroughly revised. The main structural differences are that conjunctive combinations of evidence used in the first version have often been replaced by multiple inference rules in the second. For example, in the first implementation it was required that a favorable petrotectonic setting, a favorable regional environment, and a favorable intrusive system all be present for the evidence to match the model. While all three of these considerations must be confirmed for a perfect match, experimental tests quickly revealed that insisting that the overall match cannot be good unless all three aspects match well was too strong a requirement. In particular, note that lack of any knowledge about the petrotectonic setting would prevent favorable evidence about the intrusive system from influencing the certainty of the final evaluation--which was not the intended way to combine the evidence.

Another structural change was the systematic introduction of context spaces (see Section 2.4) to influence the questioning sequence. For example, since the questions about regional environment all concern the characteristics of the regional intrusive systems, we use the existence of granitic intrusives as a context for asking about the regional environment. In the first version this was accomplished by having the existence of granitic intrusives be the first conjunct in the conjunction defining the regional environment; the use of contexts--which was, in fact, necessitated by the replacement of conjunctions with inference rules--more naturally expresses the role of the intrusives.

The other major change was a revision of values for the prior probabilities and likelihood ratios. The values used in the initial version were obtained by using verbal classifications, such as "rare" for a prior probability of 0.01, and "very suggestive" for a likelihood ratio of 100. While these quantized values seemed satisfactory for individual pieces of evidence, they too frequently failed to capture the desired interactions between several pieces of evidence bearing on the same hypothesis. Values were needed that would be between those

provided by the verbal classifications. Rather than expand the set of verbal descriptions, we decided simply to employ numerical values as needed.

The most difficult problem encountered in determining the parameter values was that of choosing prior probabilities. In general, this problem was most difficult when there were few or no established contexts. Thus, for example, it was much more difficult to estimate the prior probability of having a Type-A porphyry copper deposit than it was to estimate the "prior" probability that the groundmass has aplitic texture, given the context of the existence of an intrusive system containing porphyry rock.

A special problem arose for two consequent spaces having very small estimated prior probabilities--the top-level space and the space asserting the existence of a favorable intrusive system. To overcome these low values, the incoming rules had to have very large likelihood ratios. This had the undesired result that two or more weakly believed pieces of evidence could combine to produce an excessively certain conclusion. In each case this problem was circumvented by introducing a fictitious consequent space that was assigned a prior probability sufficiently large to produce the desired interactions among different pieces of evidence. Probabilities for this fictitious space were transformed to probabilities for the true consequent space through a single inference rule with an LS of "infinity" and an LN of zero. This ad hoc mapping preserves certainties, maintains the desired interactions among different pieces of evidence, and guarantees the correct numerical results when the probability is zero, one, or its prior value.

The statistics for the revised porphyry copper model are presented in Table 4, and the actual inference network is shown in Figure 4. In that figure prior probabilities are shown only for spaces that are the consequents of rules; the numerical values of the likelihood ratios LS and LN are shown associated with each rule. Note that for certain rules (such as the rule used to infer the favorability of the petrotectonic setting from the age of the continental-margin mobile belt) the

likelihood ratio is a function of a parameter (in this example, the age). The treatment of uncertainty in this parameter value is explained in Section 5.1.

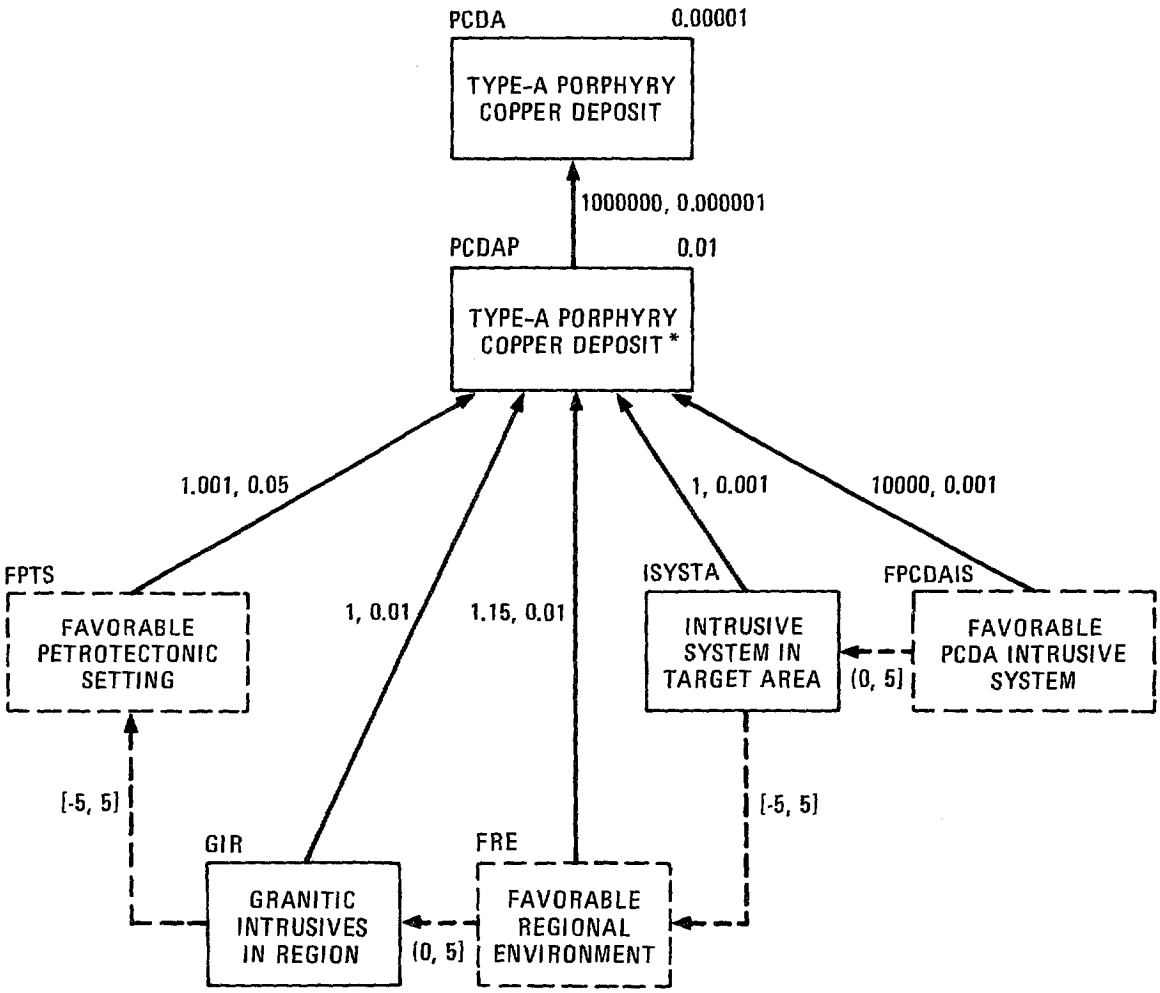
Table 4

Statistics for the Second PCDA Implementation

Model Section	Number of Spaces	Number of Rules
PCDA: Total	185	91
Petrotectonic setting	5	4
Regional environment	16	13
Intrusive system	164	69
Composition	15	6
Size, morphology & texture	15	13
Alteration & mineralization	130	45
Barren-core zone	31	6
Propylitic zone	14	8
Potassic zone	33	10
Sericitic zone	36	14
Zone combinations	13	7

It should be pointed out that all of the spaces for the four types of alteration zones contain "location variables," so that each situation being described can occur more than once on the prospect. Thus, associated with each such space is not a single posterior probability but a list of probabilities, one for each zone of alteration observed on the prospect.

When the Questioning System reaches the point of trying to establish the existence of favorable alteration zones, it invokes a special procedure to handle the location variables. The user is asked whether any alteration zones are present, and, if there are, the program calls the most "prominent" zone Zone 1. The four hypotheses corresponding to barren-core, potassic, propylitic and sericitic zones,



*Same as PCDA, except for scaling of probabilities.

FIGURE 4 INFERENCE NETWORK FOR PCDA

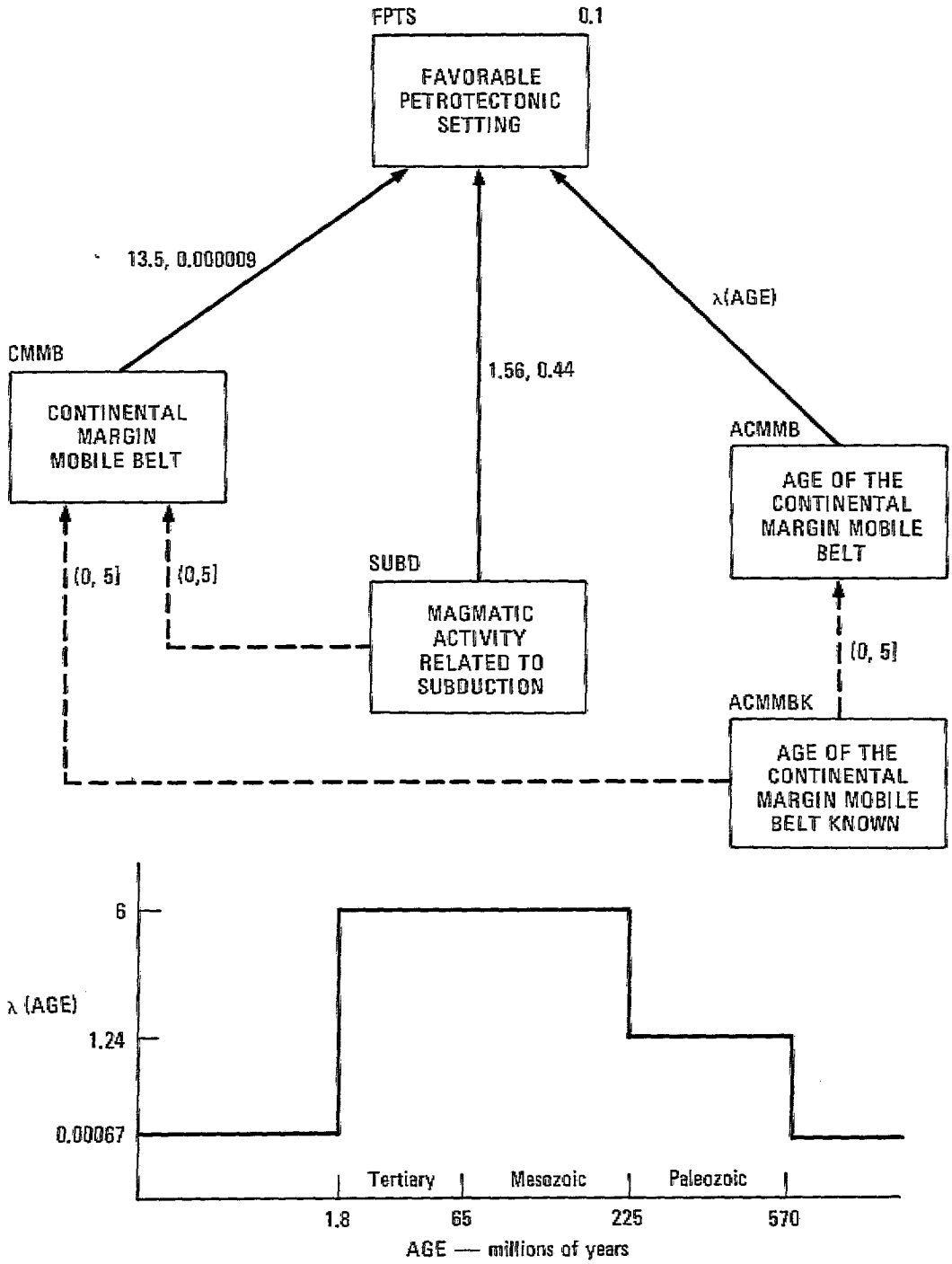


FIGURE 4 INFERENCE NETWORK FOR PCDA (Continued)

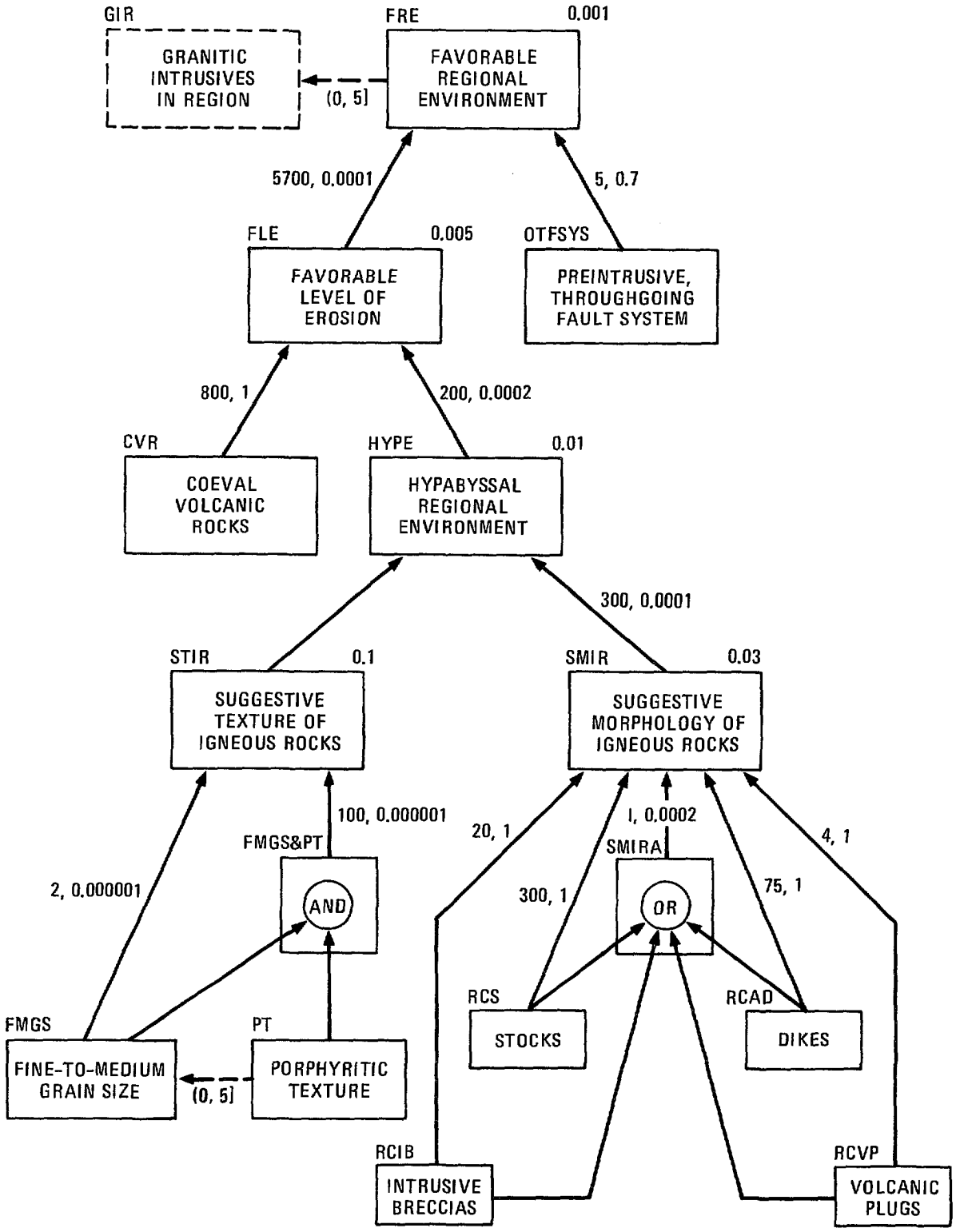
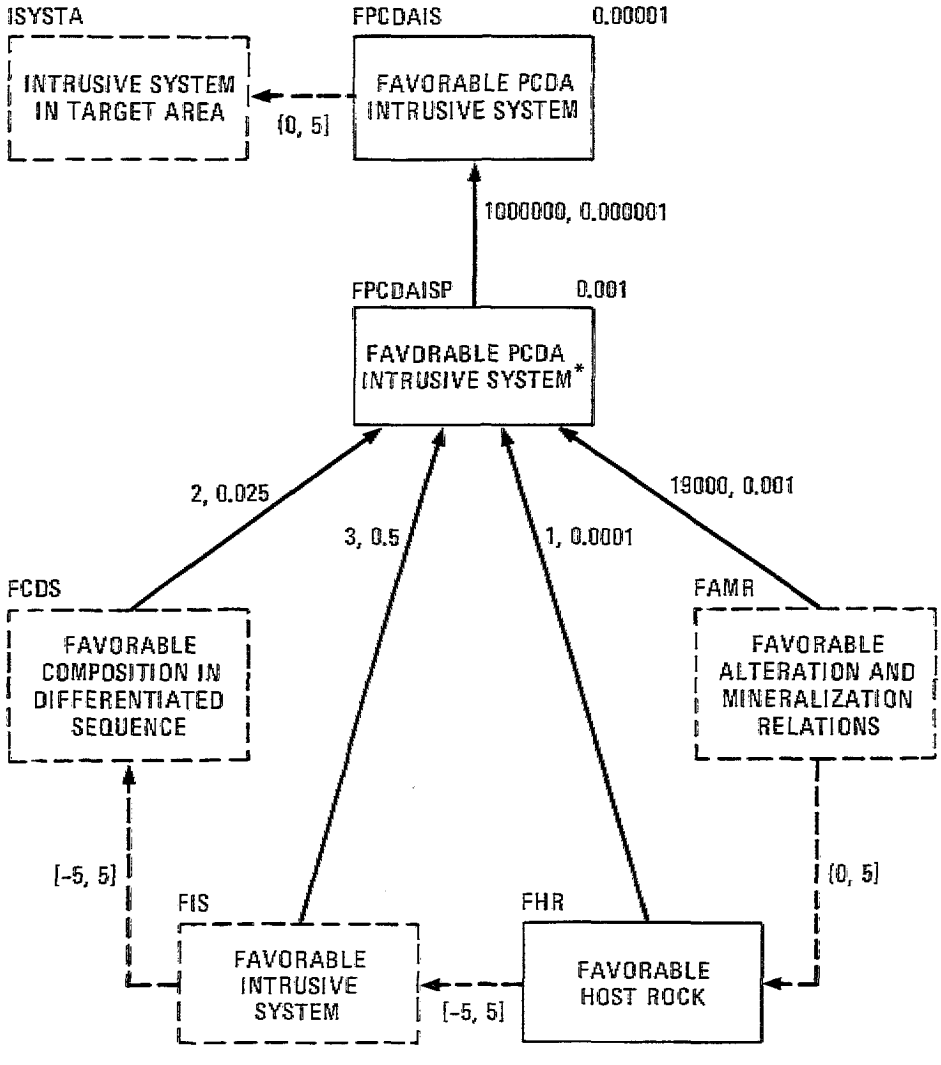


FIGURE 4 INFERENCE NETWORK FOR PCDA (Continued)



*Same as FPCDAIS, except for scaling of probabilities.

FIGURE 4 INFERENCE NETWORK FOR PCDA (Continued)

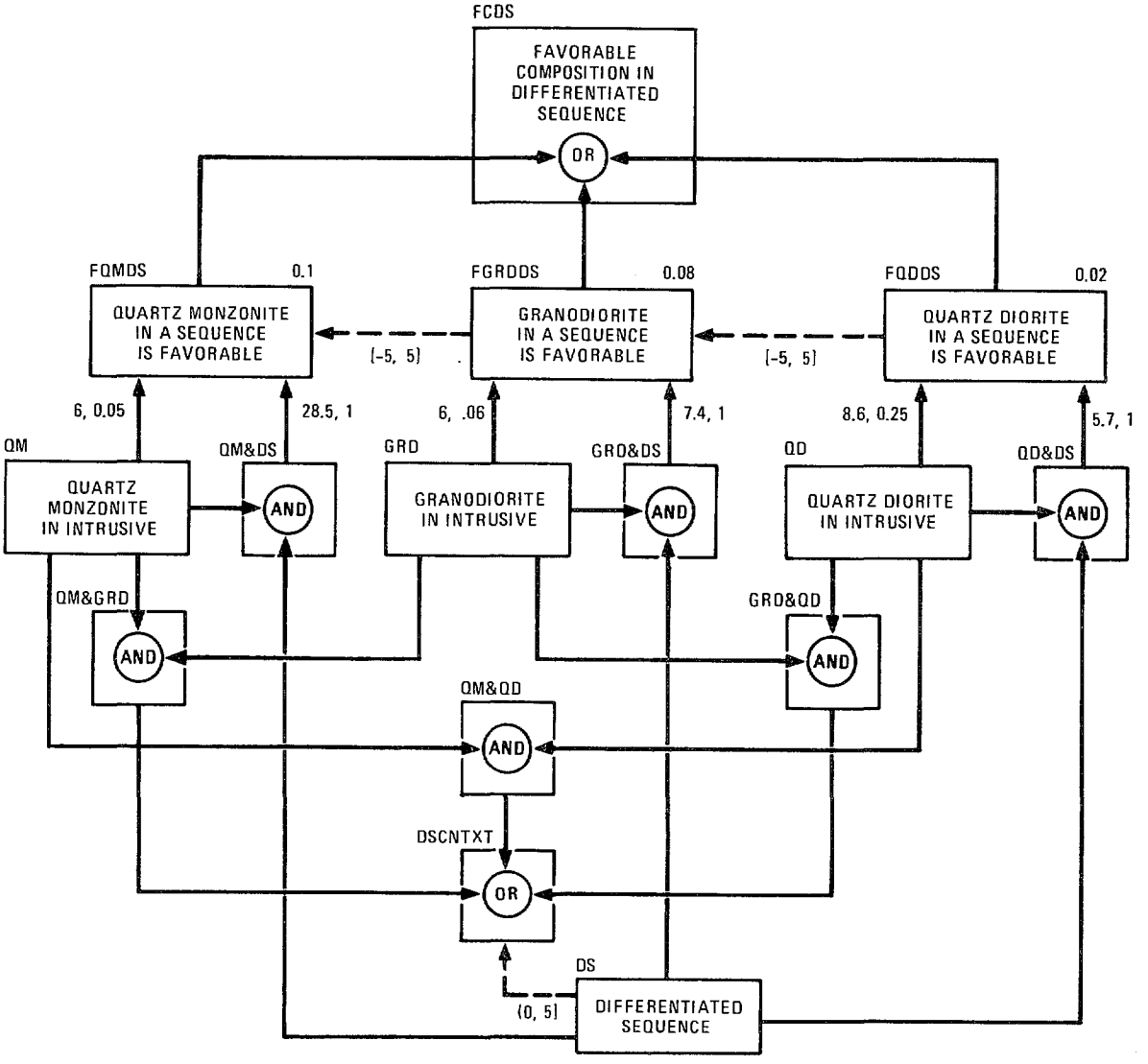


FIGURE 4 INFERENCE NETWORK FOR PCDA (Continued)

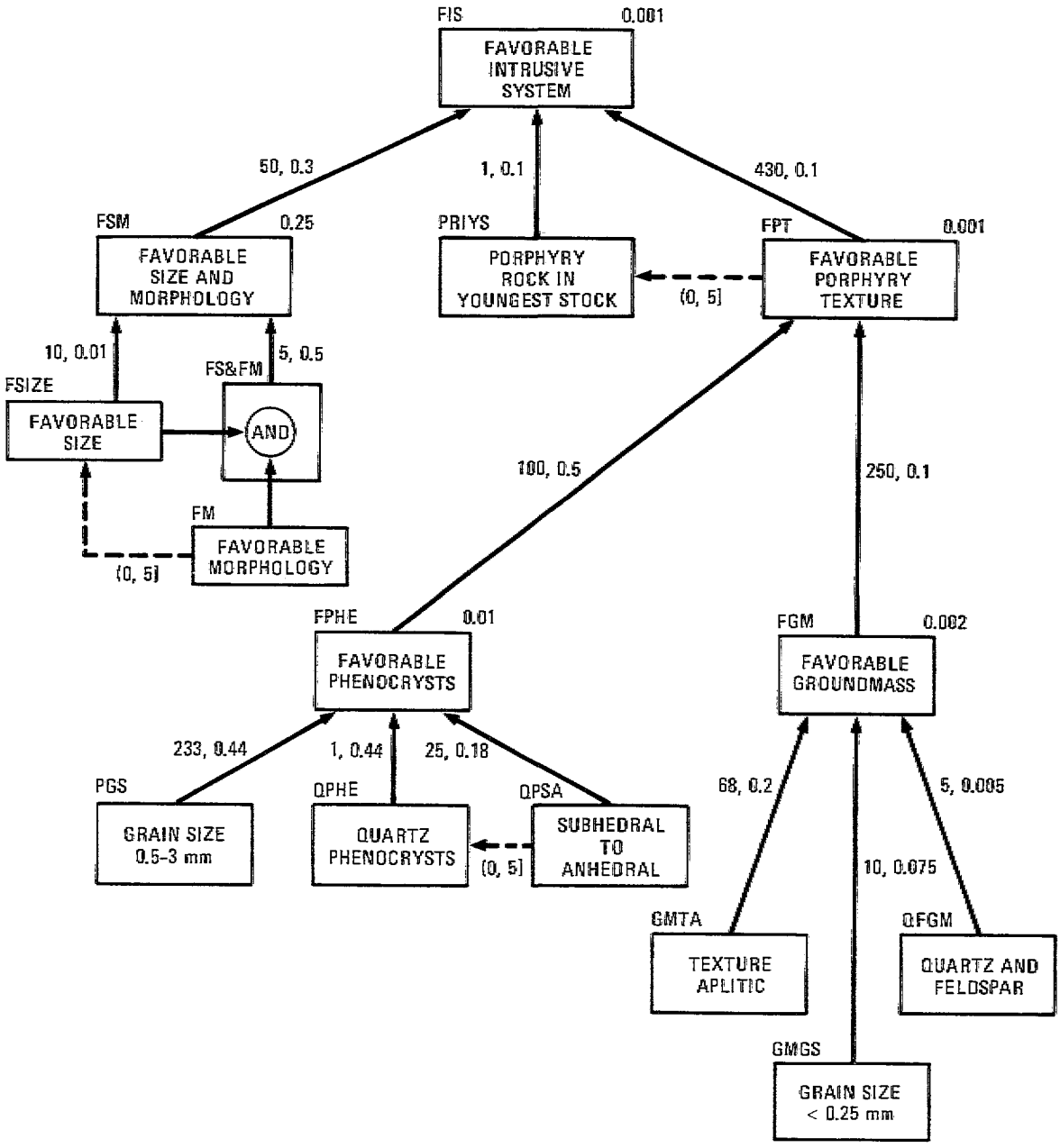


FIGURE 4 INFERENCE NETWORK FOR PCDA (Continued)

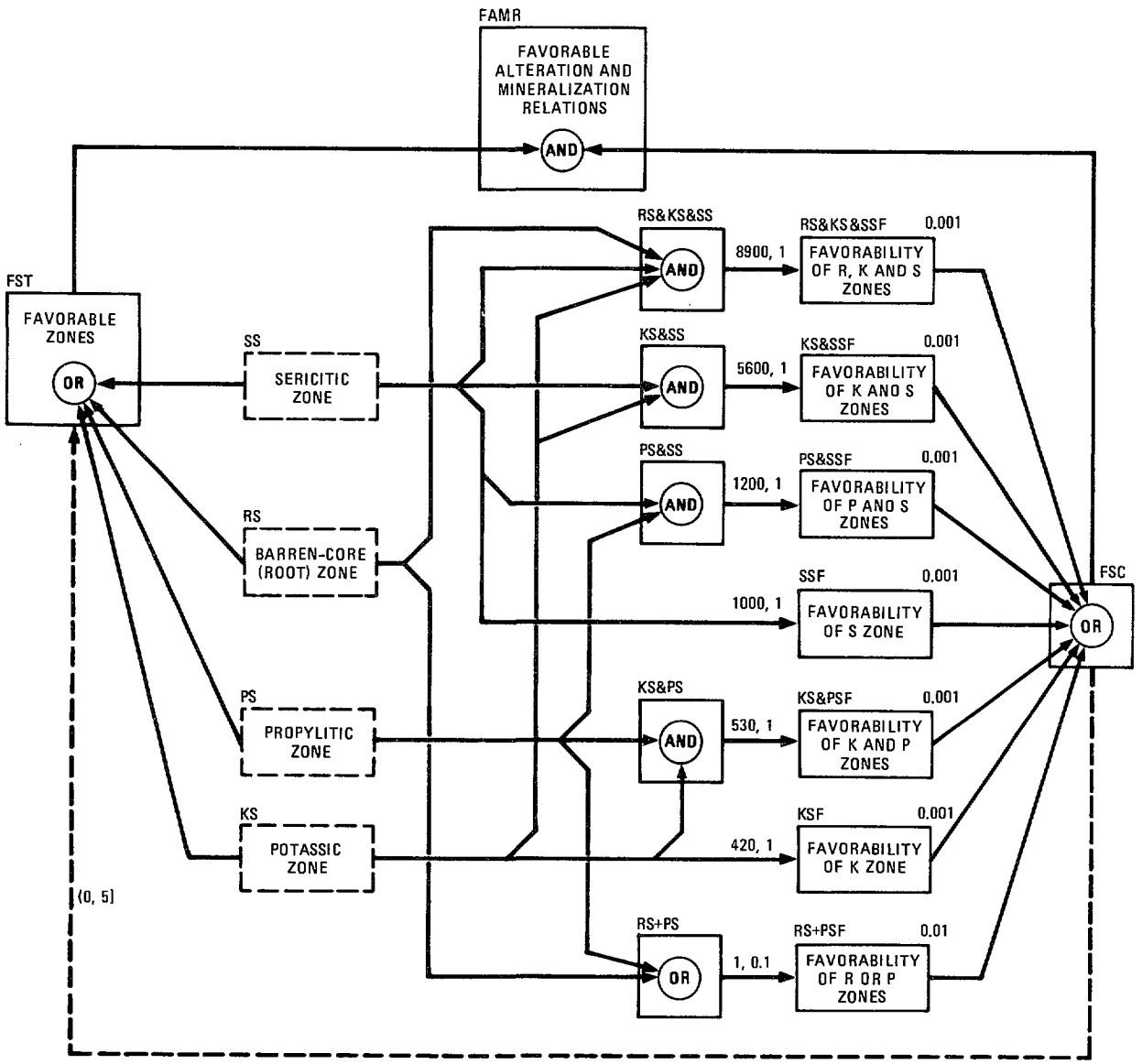


FIGURE 4 INFERENCE NETWORK FOR PCDA (Continued)

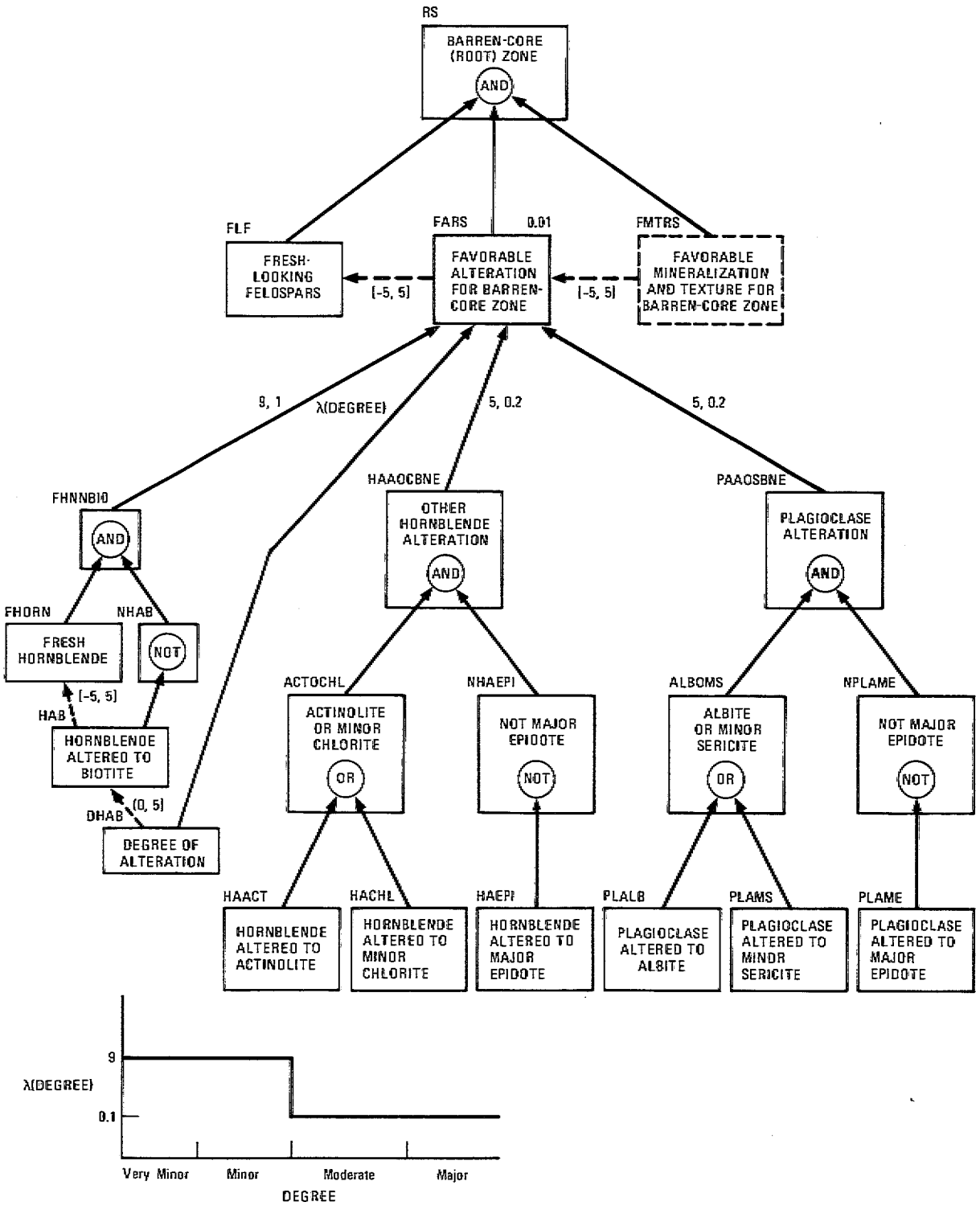


FIGURE 4 INFERENCE NETWORK FOR PCDA (Continued)

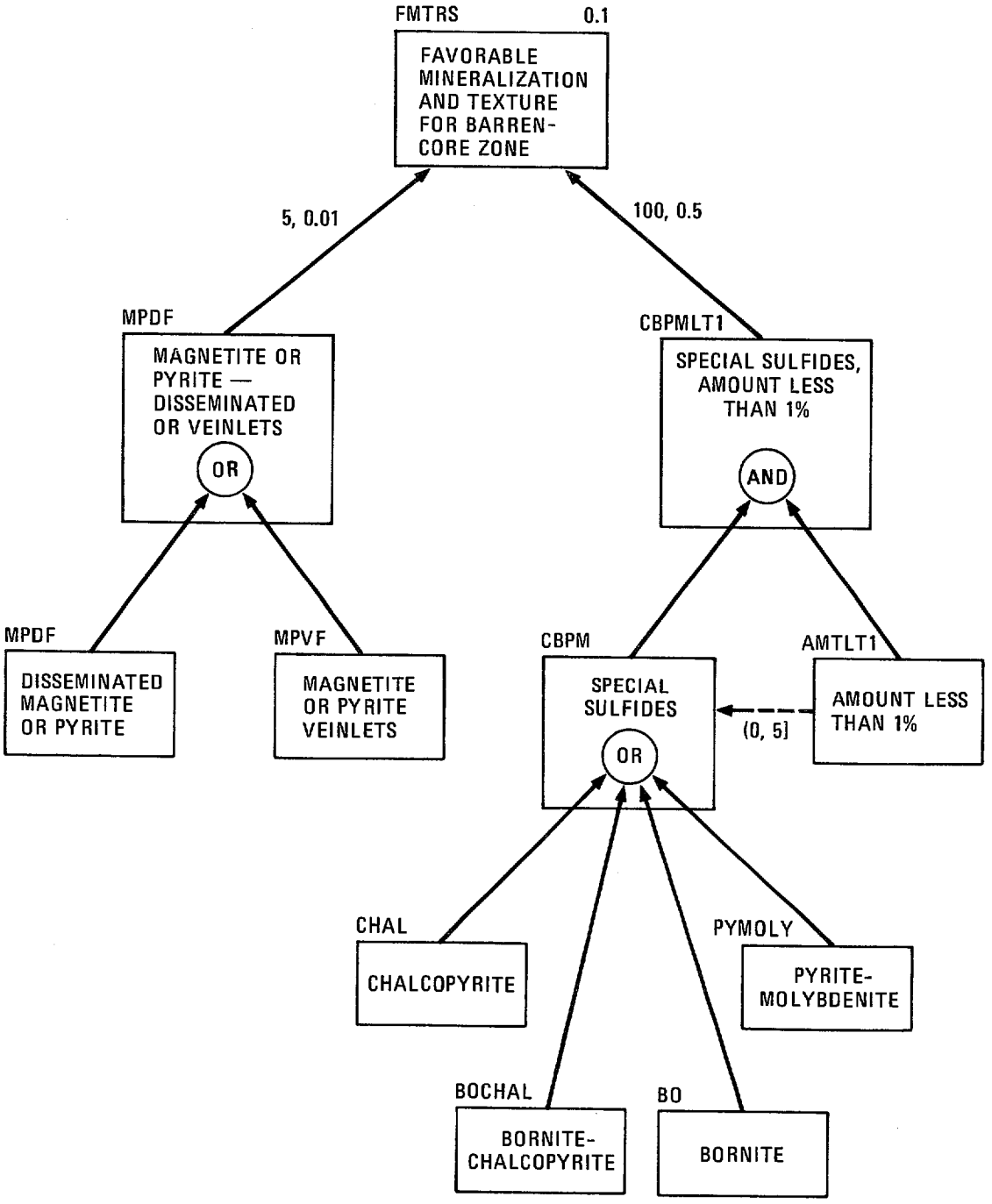


FIGURE 4 INFERENCE NETWORK FOR PCDA (Continued)

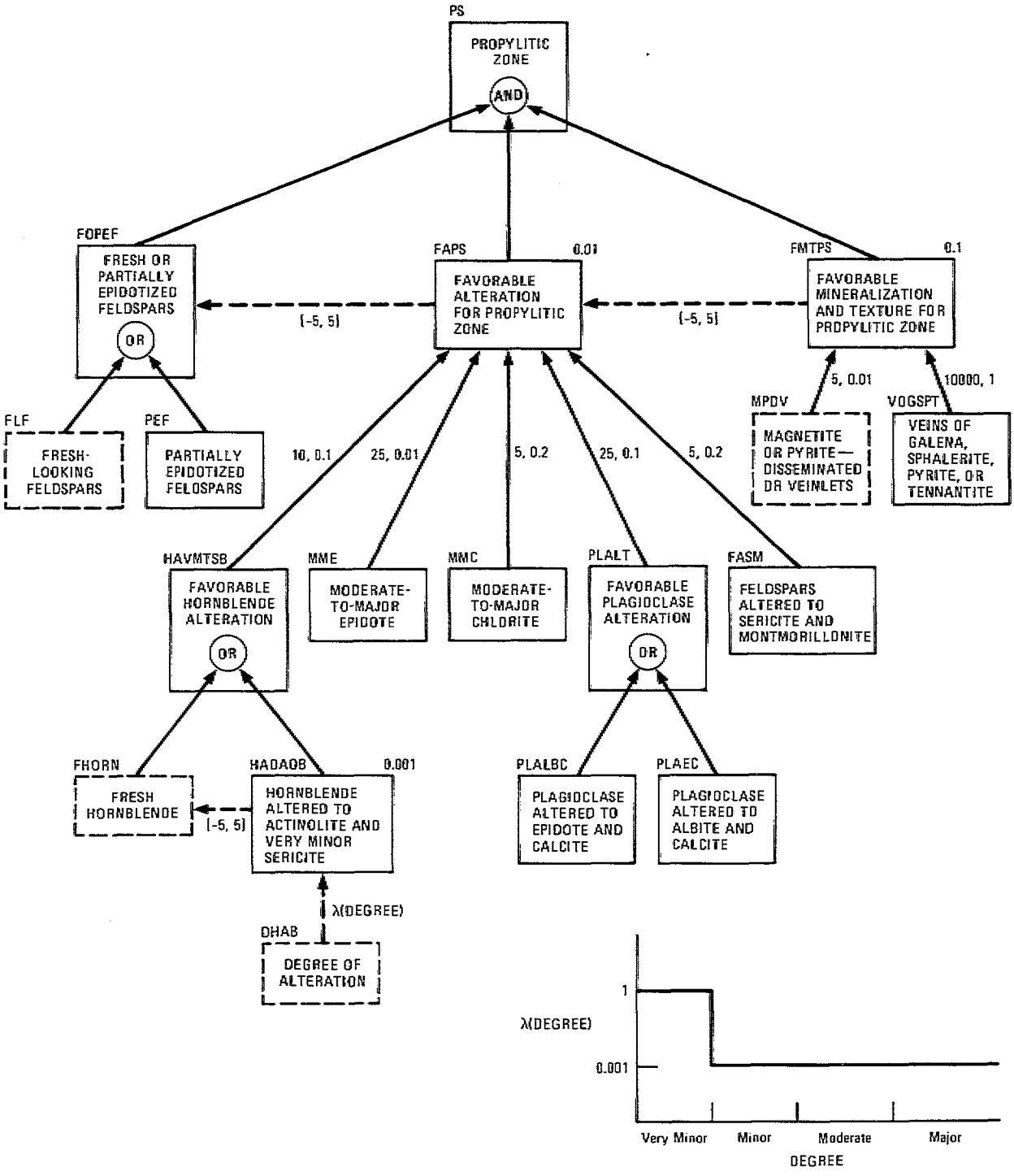


FIGURE 4 INFERENCE NETWORK FOR PCDA (Continued)

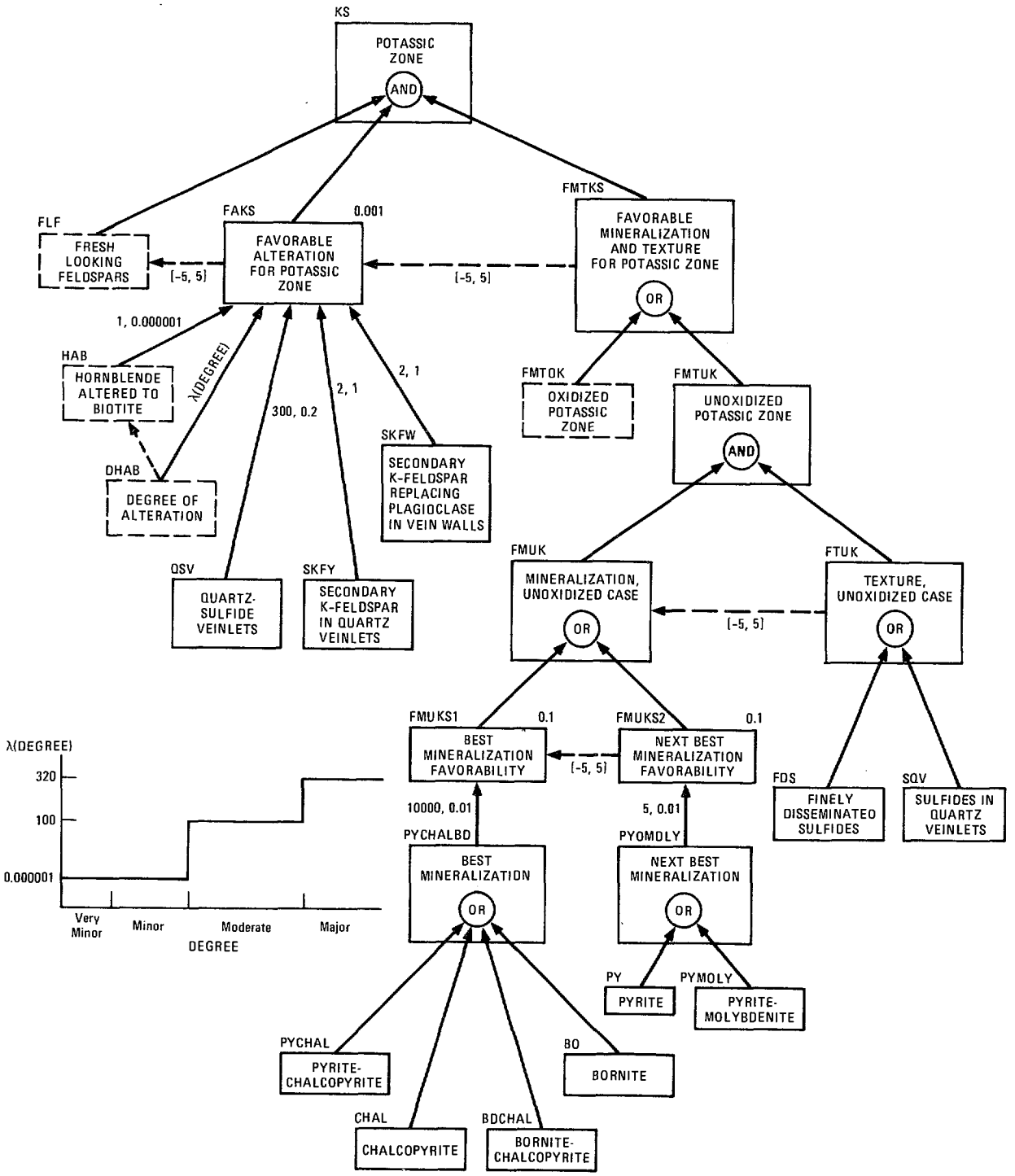


FIGURE 4 INFERENCE NETWORK FOR PCDA (Continued)

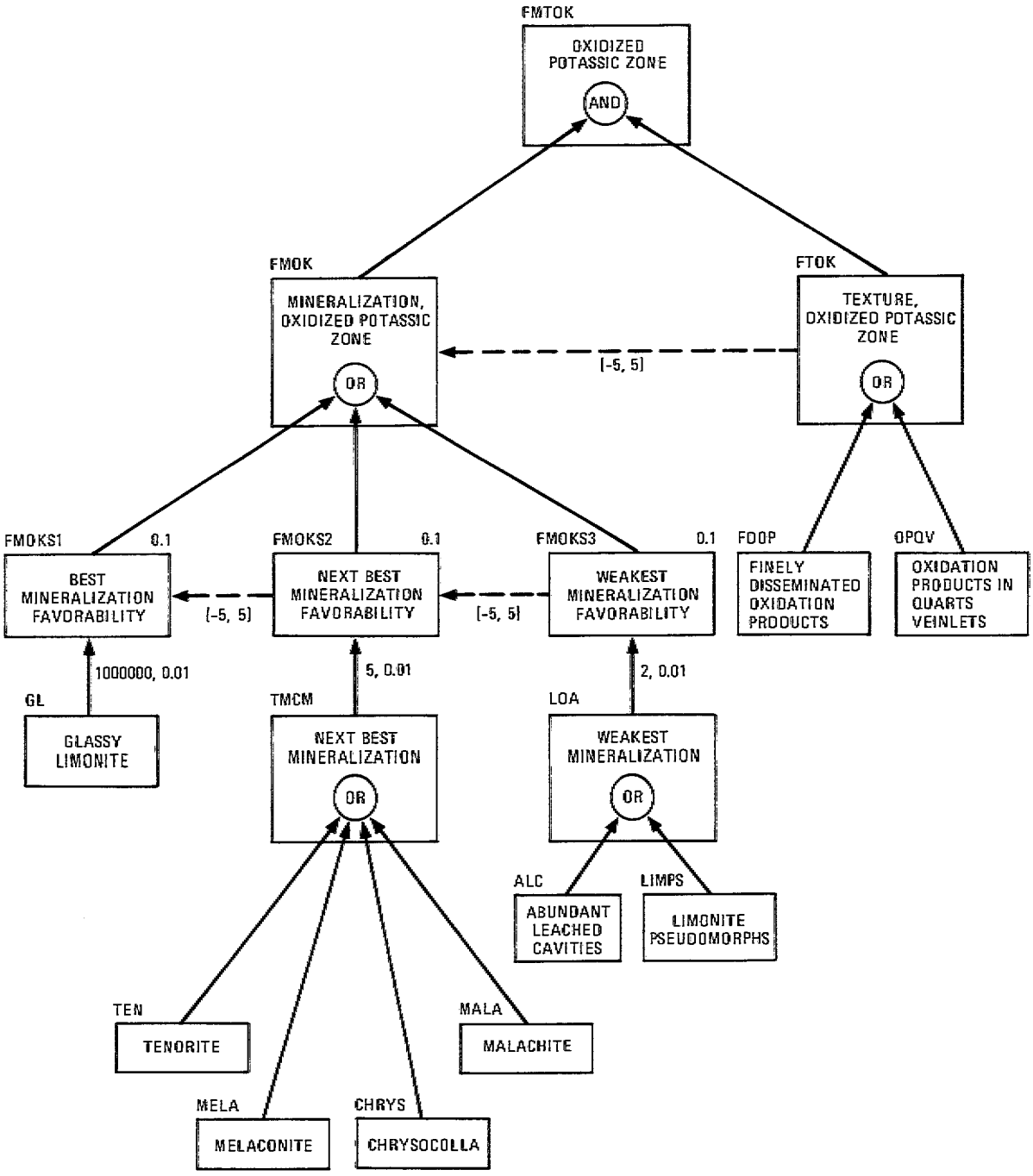


FIGURE 4 INFERENCE NETWORK FOR PCDA (Continued)

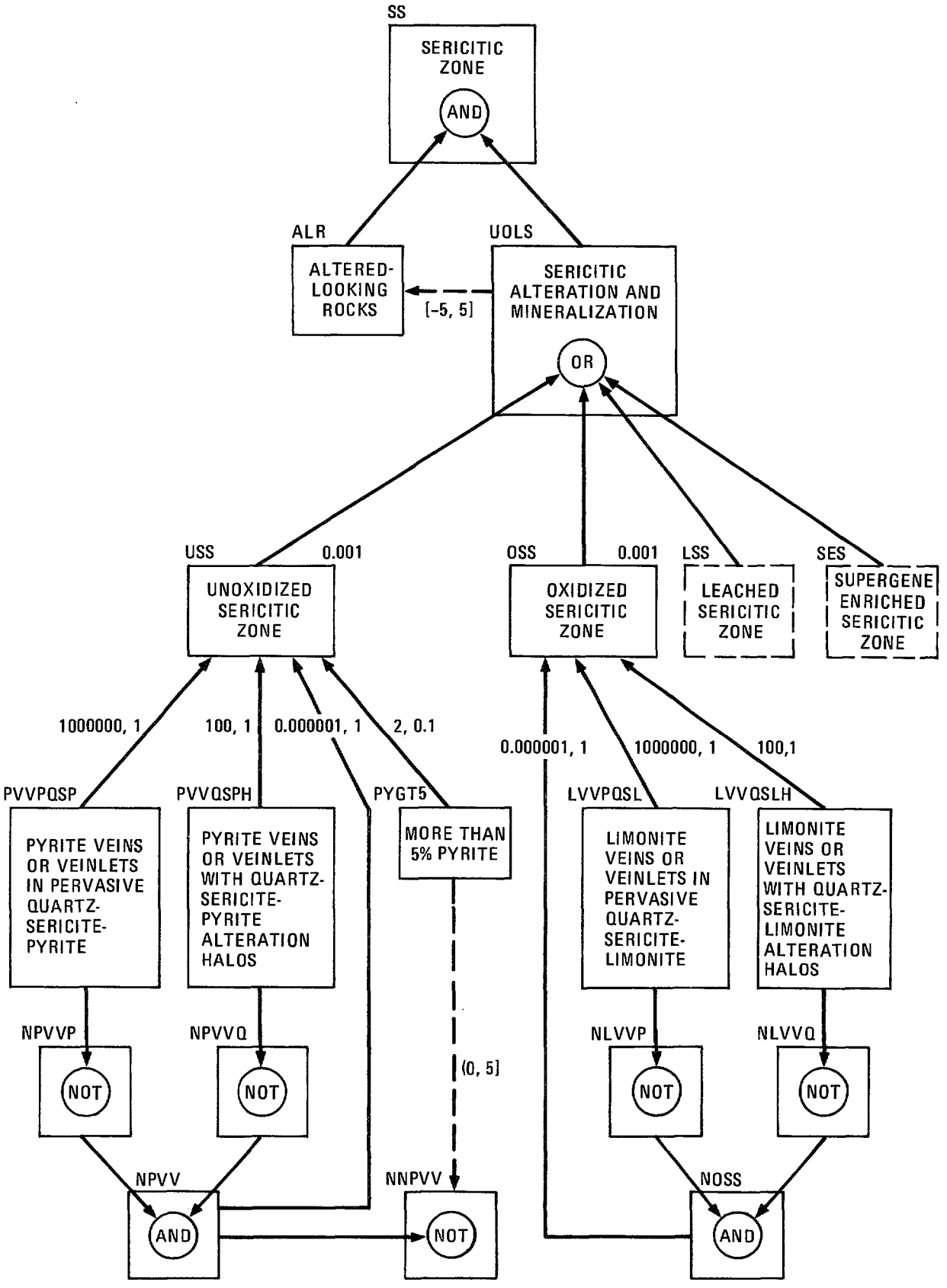


FIGURE 4 INFERENCE NETWORK FOR PCDA (Continued)

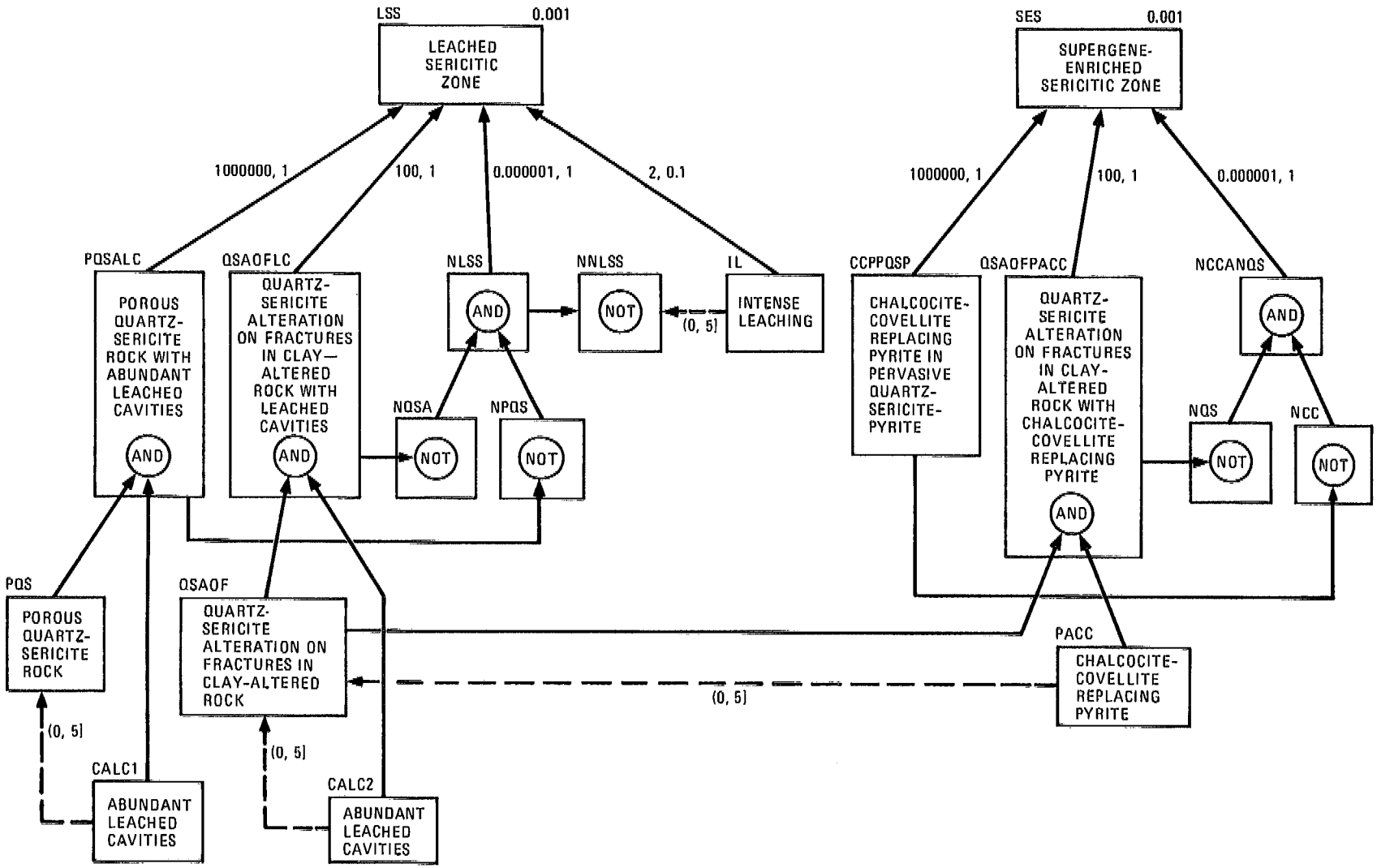


FIGURE 4 INFERENCE NETWORK FOR PCDA (Concluded)

effectively become new top-level hypotheses for the program to resolve. Subsequent questions--such as "To what degree do you believe that there are fresh looking feldspars?"--are modified by appending the phrase "in Zone 1." This same procedure is repeated for the other zones, and a list of probabilities is obtained for the four zone-identity hypotheses.

At any time the probability of a specific zone existing on the prospect (such as a potassic zone) is computed as the maximum of the probabilities in the list for that zone. In effect, each of the inference networks for the four alteration zone types is replicated for each existing alteration zone, and the hypothesis "There is a zone of Type X" is computed as "(Zone 1 is of Type X) OR (Zone 2 is of Type X) OR ... (Zone N is of Type X)."

3.3 Revision of the Nickel Sulfide Model

The Komatiitic-nickel-sulfide model (KNSD) of Prof. Anthony J. Naldrett has a general structure that resembles that of the porphyry copper model. It is subdivided into major sections treating the petrotectonic setting (a Precambrian greenstone belt), the magmatic association and local environment (a basaltic base of a sequence of mafic and ultramafic igneous rocks), the favorable ore-bearing unit, the favorable portion of the ore-bearing unit, and the characteristic mineralization.

An initial implementation of this model was described in Appendix A of our Second Semiannual Report (Hart et al., 1978). The statistics for that version are given in Table 5.

Table 5

Statistics for the First KNSD Implementation

Model Section	No. of Spaces	No. of Rules
KNSD: Total	94	44
Petrotectonic setting	6	5
Magmatic association and local env.	48	15
Ore-bearing unit	11	9
Proximity to ore	14	9
Mineralization	14	6

As with the porphyry copper model, this first implementation has been thoroughly revised. Because the required mineralization characteristics are usually not known in an exploration stage, this section has been isolated from the other four, which have been grouped under the heading of "Favorable Conditions, Excluding Mineralization." Some parts of the model have been moved from one section to another. For example, consideration of major paleofaulting was transferred from the section on the ore-bearing unit to the section on magmatic association and local environment.

The model section that has undergone the greatest change is the one that defines a Komatiite rock suite. The revised version focuses on the desired characteristics of the olivine rocks and the pyroxene rocks, with the former being the more important. Because it is difficult to find and identify volcanic feeders during exploration, the sharp distinction between flows and feeders that was made in Version 1 has been softened in Version 2; this accounts for most of the changes made to the model in the sections on the ore-bearing unit and proximity to ore. Finally, all the parameter values have been systematically

revised; the only ones still to be established are those in the section on mineralization.

The statistics for the revised nickel sulfide model are presented in Table 6, and the actual inference network is shown in Figure 5. It is interesting to see how this encoding determines the way Prospector pursues evidence in trying to ascertain whether or not a particular prospect might fit the Komatiitic nickel sulfide model. The following excerpt from a Prospector run is included as an illustration of this process. For this run no information was volunteered, so that the run consists of a sequence of questions asked by the system. The response of the user is typed in uppercase (when it is not numerical), and commentary on the run is enclosed in square brackets; symbolic names--such as KNSD and FCEM--that appear in the commentary identify spaces in the inference network shown in Figure 5.

Table 6

Statistics for the Second KNSD Implementation

Model Section	No. of Spaces	No. of Rules
KNSD: Total	75	49
Favorable conditions	62	43
Petrotectonic setting	11	8
Magmatic assn. & local env.	27	19
Ore-bearing unit	11	8
Proximity to ore	12	8
Mineralization	12	6

The program initially has the goal of establishing the certainty of KNSD, which is a conjunction of FCEM (Favorable Conditions, Excluding Mineralization) and RM (Right Mineralization). Since there is a chain of contexts from RM to FPTS (Favorable Petrotectonic Setting), the first

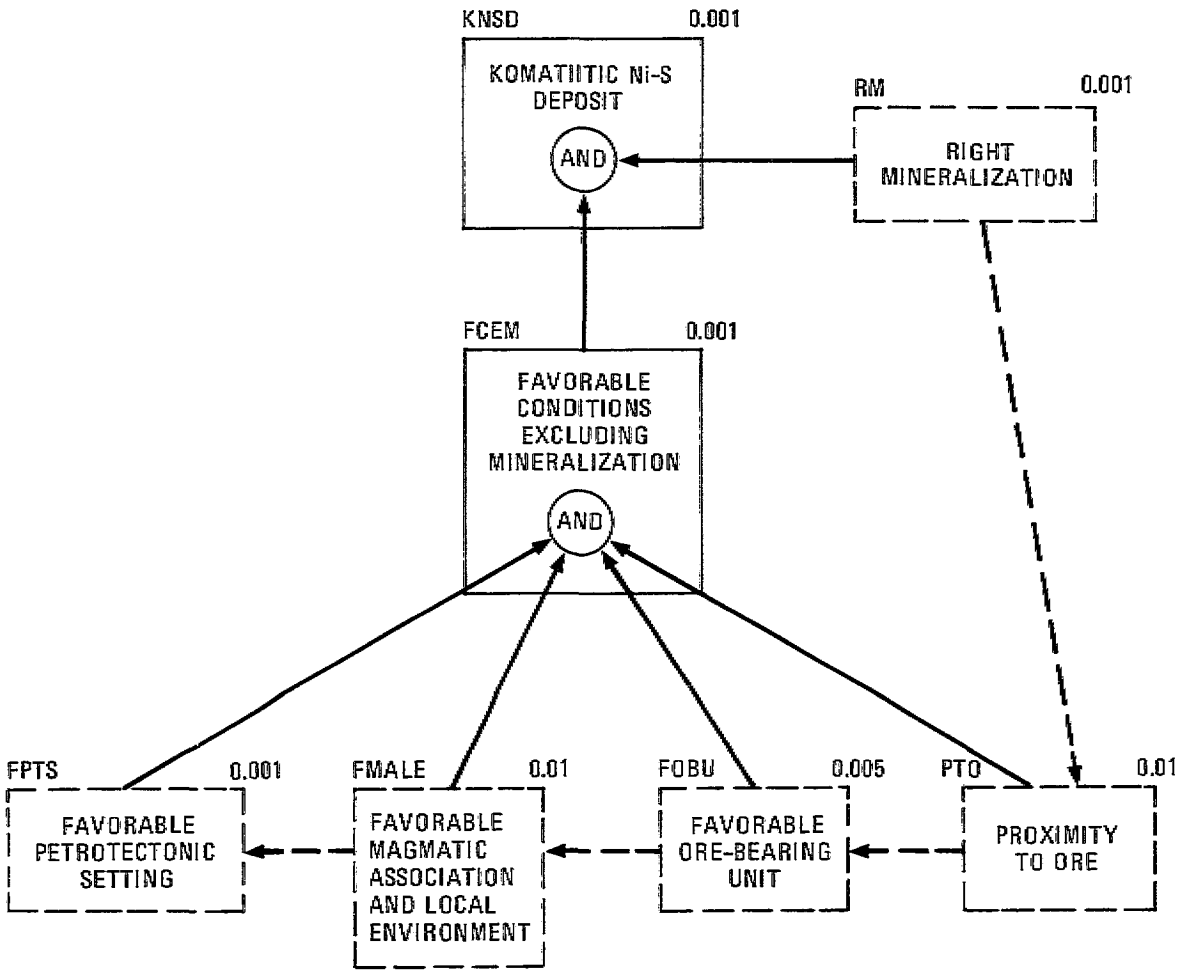


FIGURE 5 INFERENCE NETWORK FOR KNSD

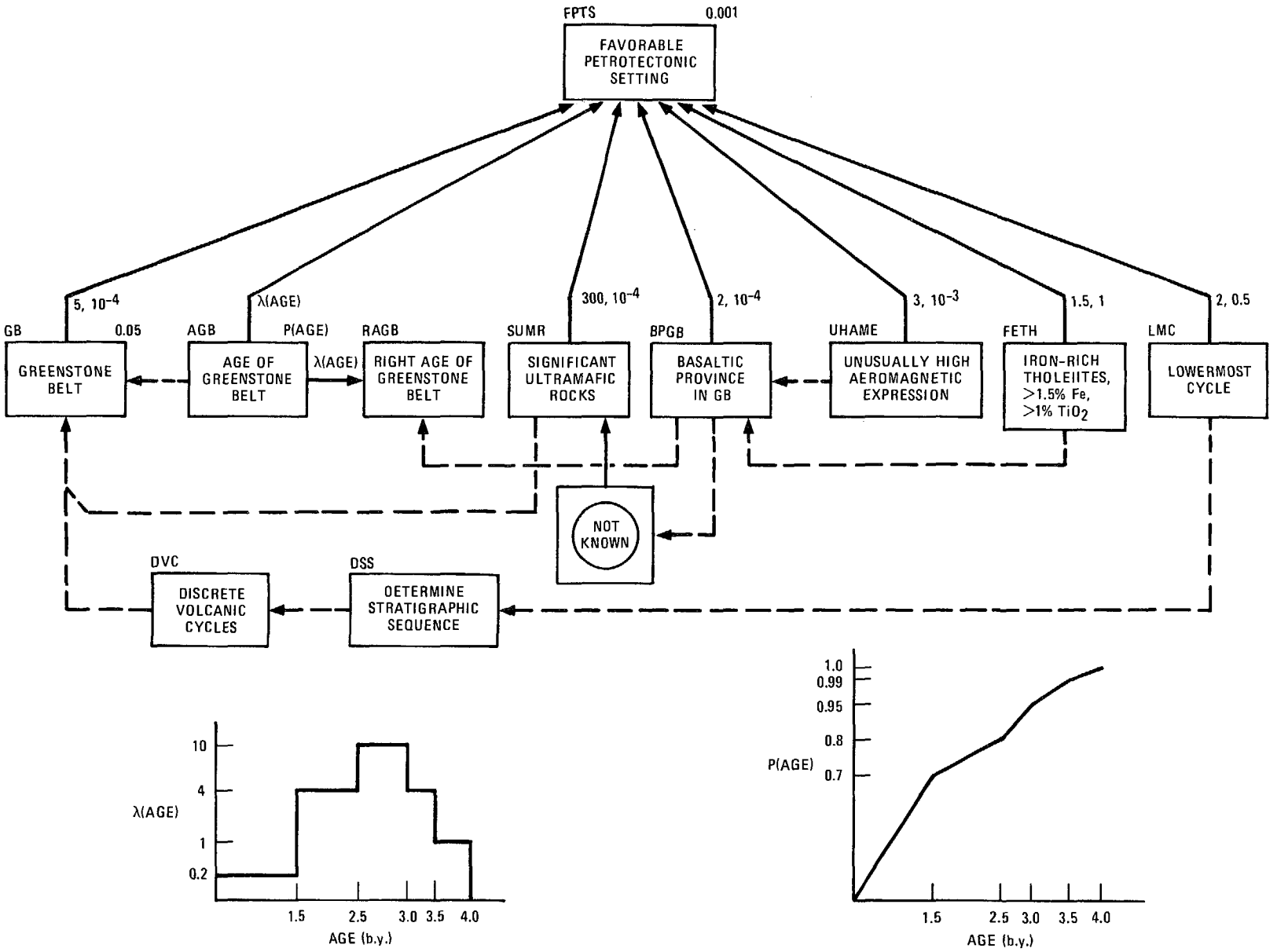


FIGURE 5 INFERENCE NETWORK FOR KNSD (Continued)

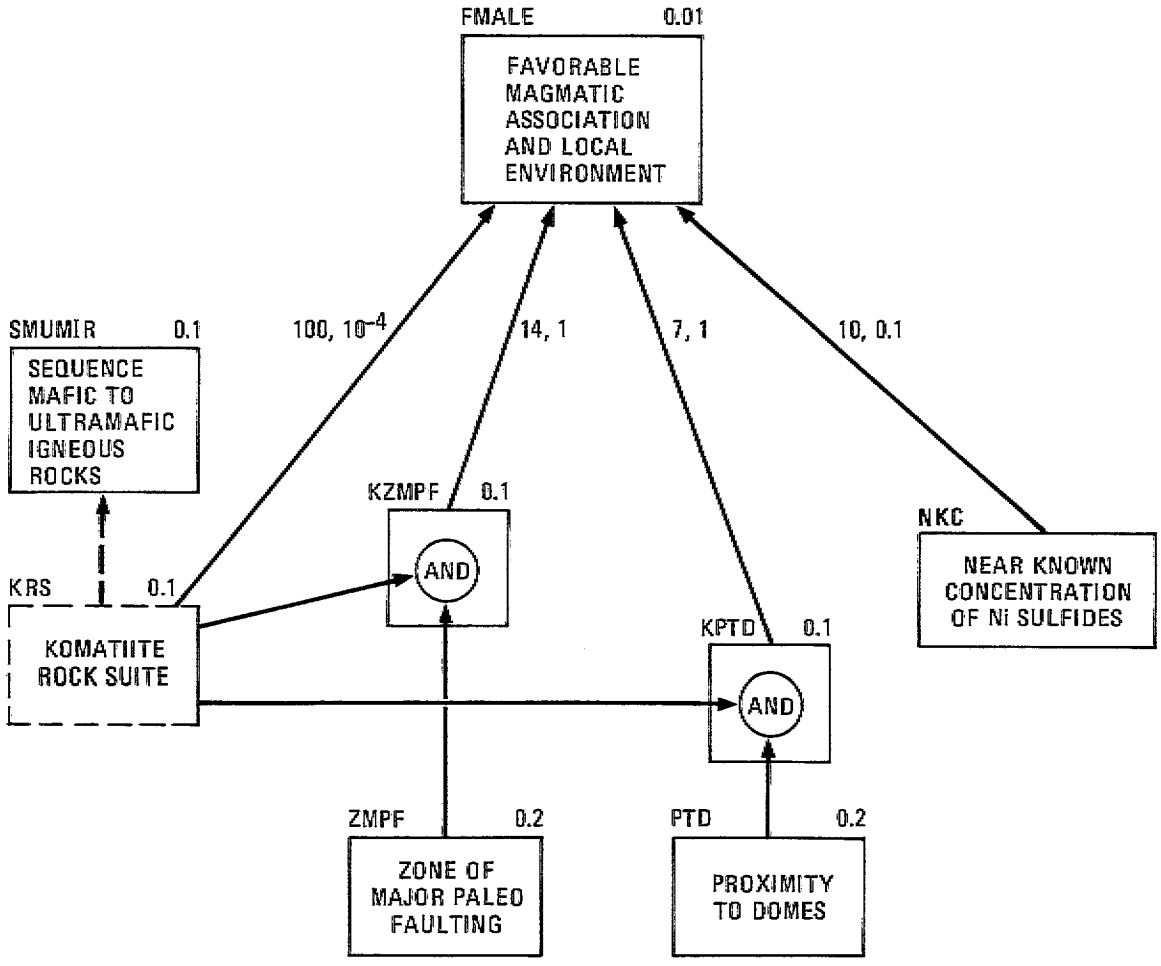


FIGURE 5 INFERENCE NETWORK FOR KNSD (Continued)

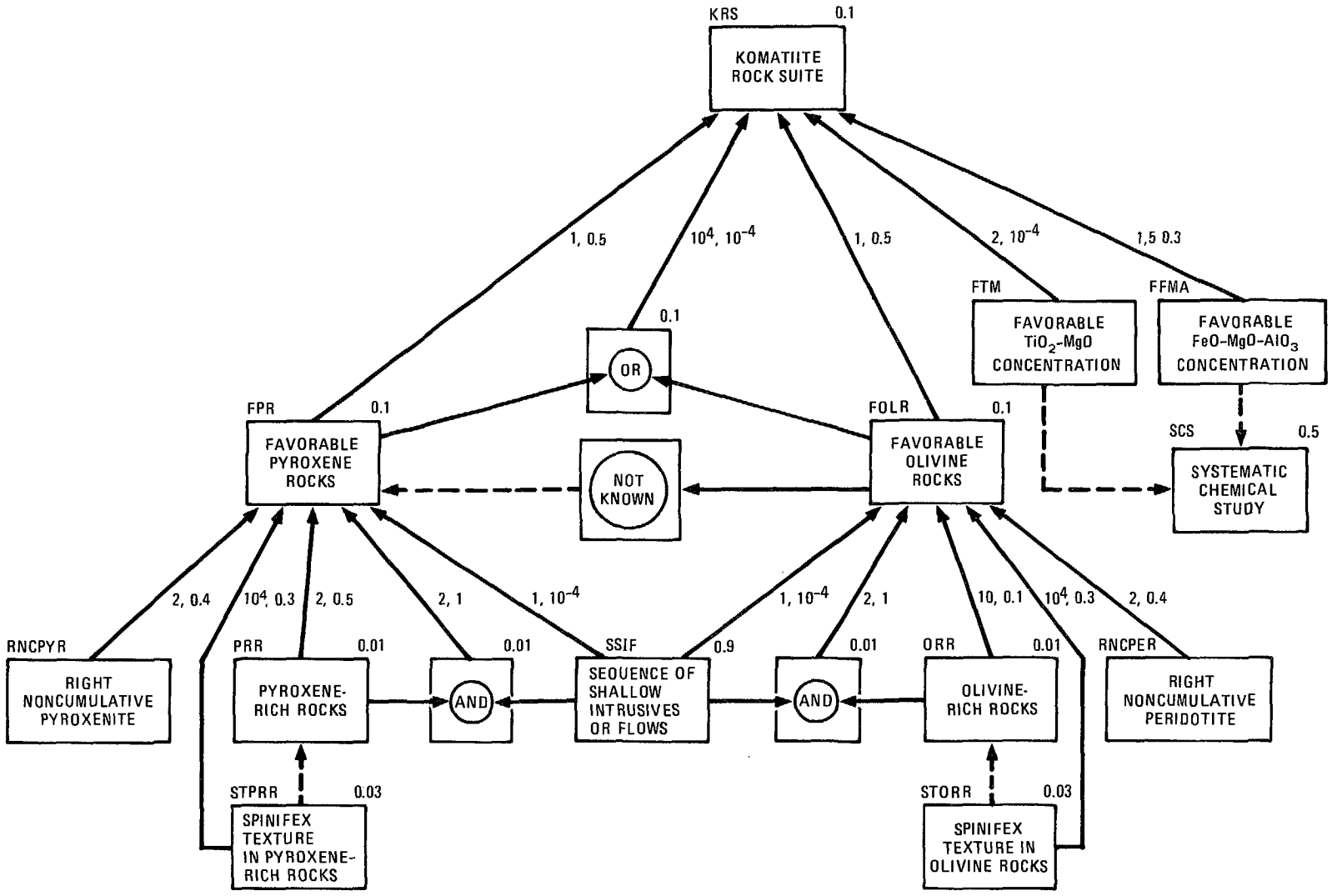


FIGURE 5 INFERENCE NETWORK FOR KNSD (Continued)

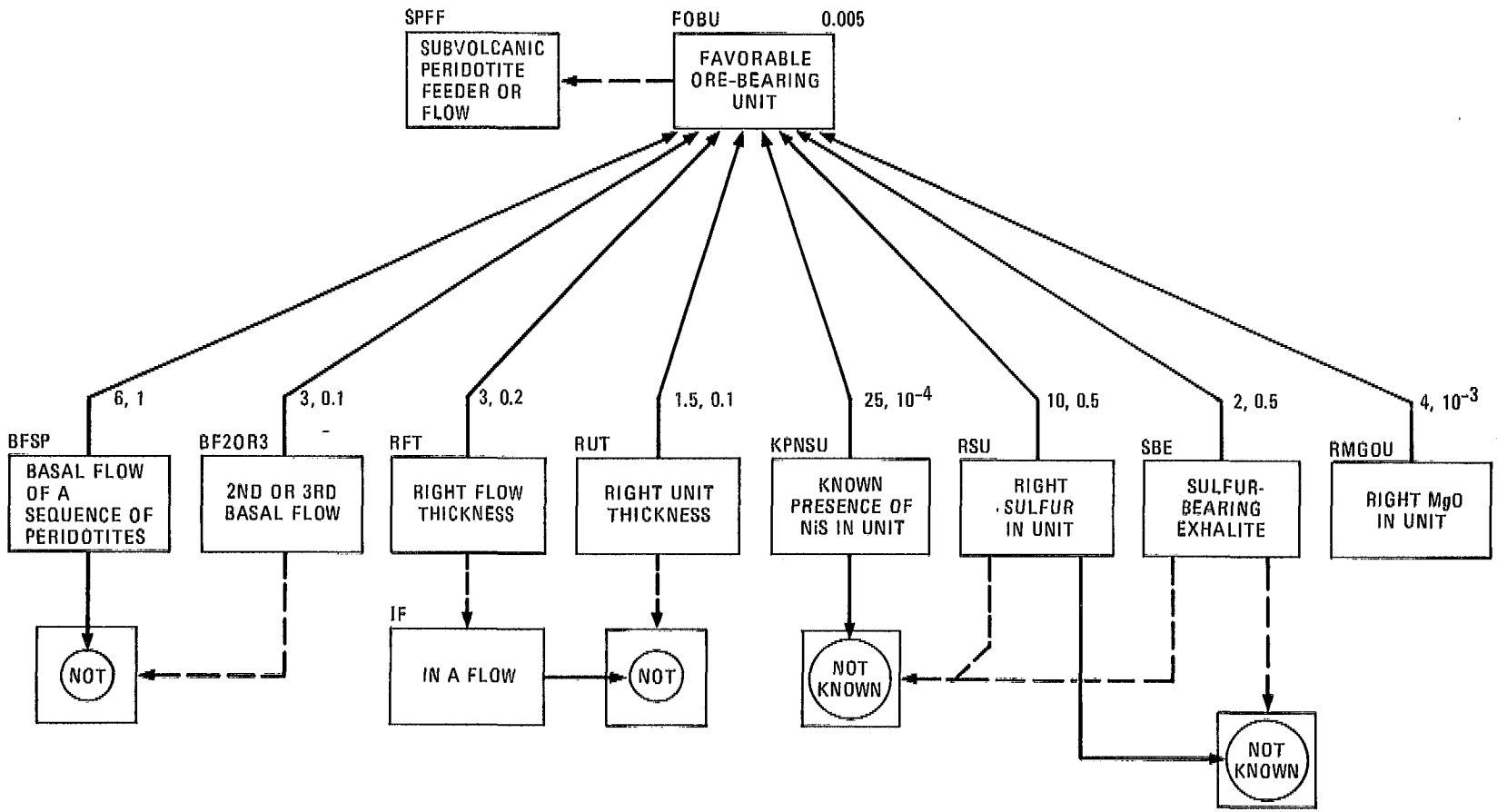


FIGURE 5 INFERENCE NETWORK FOR KNSD (Continued)

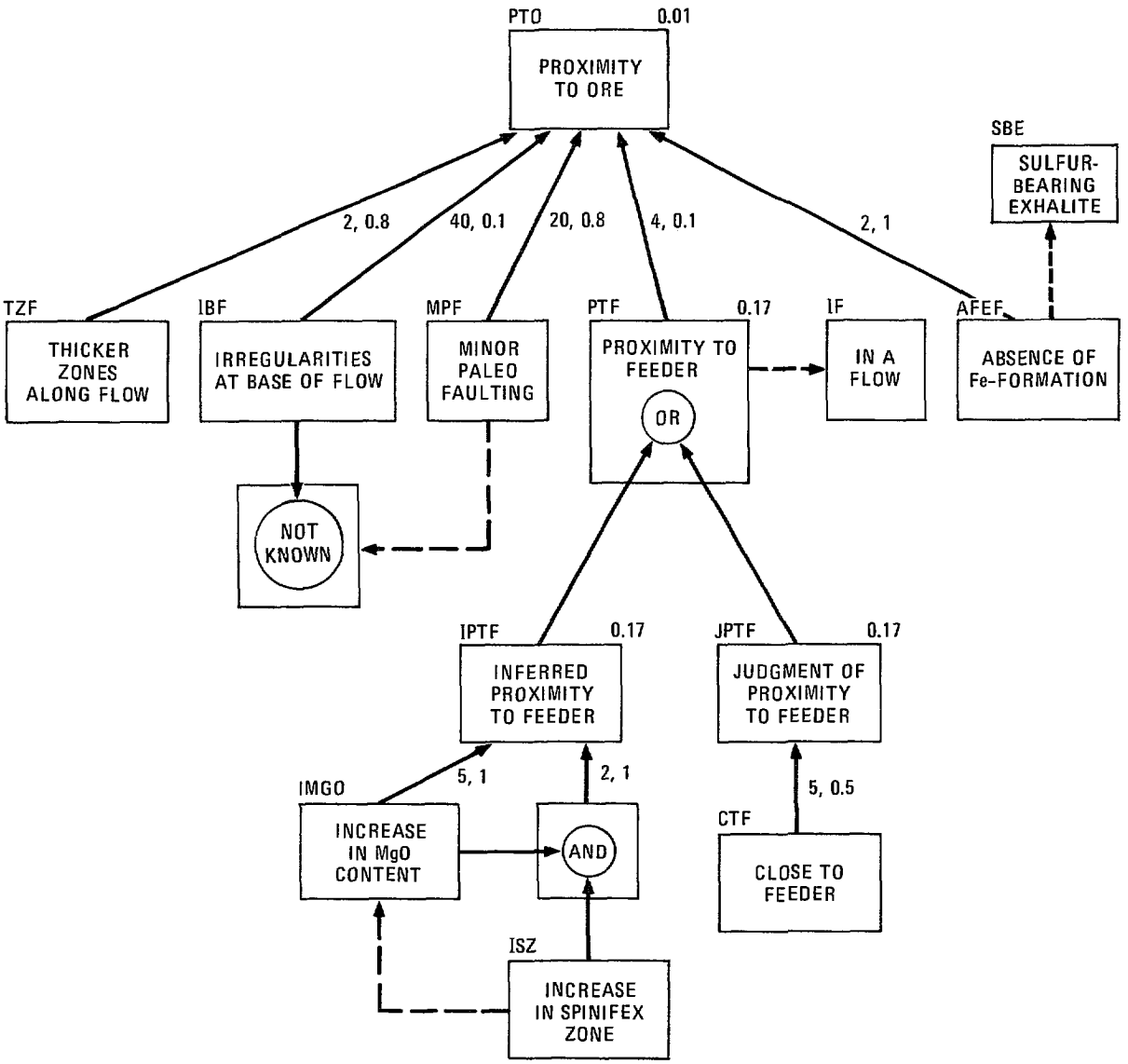


FIGURE 5 INFERENCE NETWORK FOR KNSD (Continued)

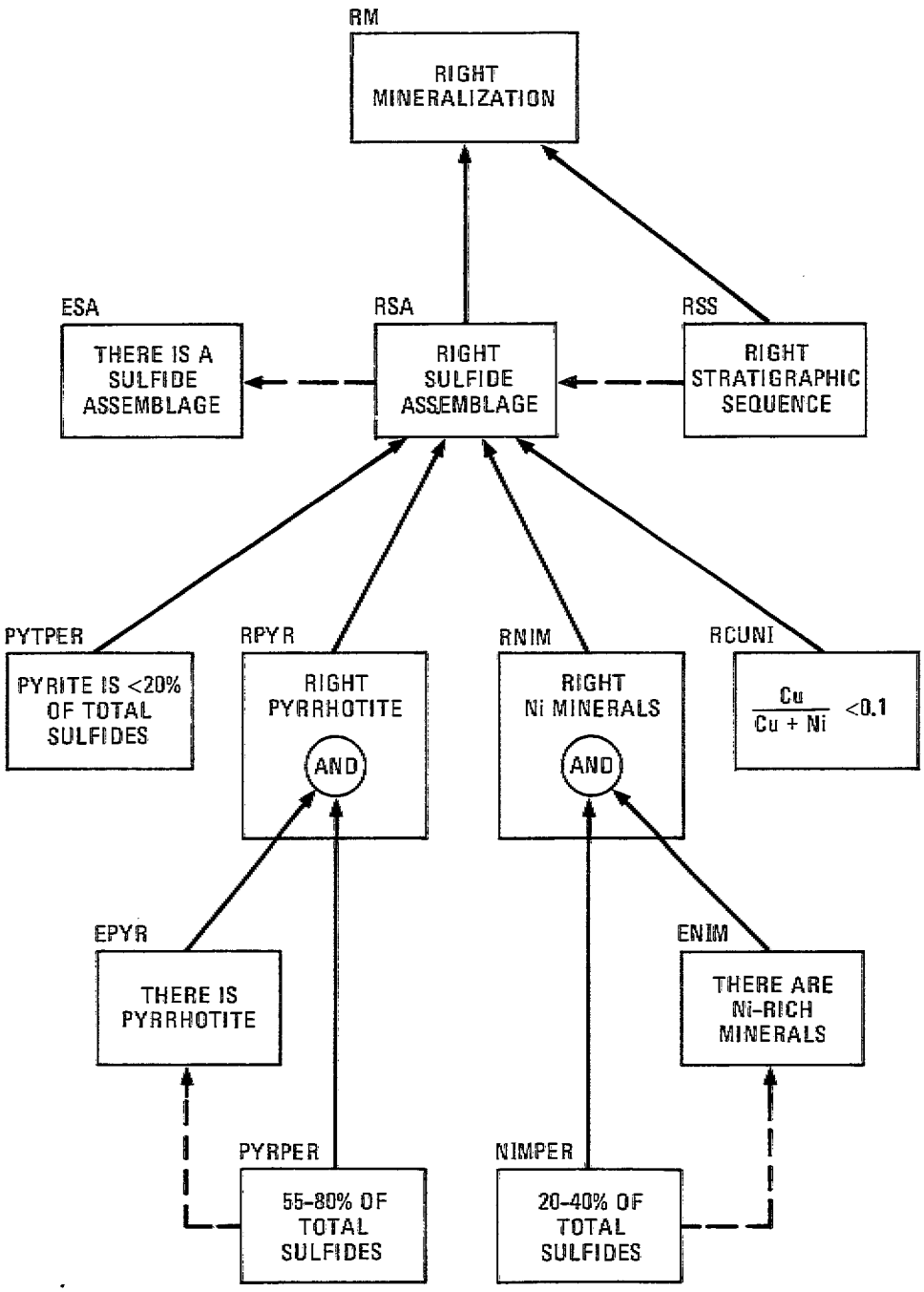


FIGURE 5 INFERENCE NETWORK FOR KNSD (Concluded)

subgoal is to establish the petrotectonic setting. The best of the seven rules for establishing FPTS comes from SUMR (Significant Ultramafic Rocks), but the latter has GB (Greenstone Belt) as a context. Thus, the first question asked is about the greenstone belt.

1--To what degree do you believe that:
(The target area lies in a greenstone belt) ? 5

2--To what degree do you believe that:
(The target area is in a region containing significant volume of ultramafic rocks) ? 5

[Question 1 establishes the context for Question 2, and the answer given raises the certainty of FPTS slightly. With the favorable answer to Question 2 about SUMR the certainty of FPTS rises from 0.02 to 3. Of course, since FPTS is only one of four low-probability conjuncts for determining FCEM, only a small increase in the certainty of FCEM occurs. Note that, had the answer to Question 2 indicated considerable uncertainty, the system would have gone on to inquire about BPGB (a basaltic province) and other weaker evidence, but this less certain line of questioning is now unnecessary.]

3-- What is the age of the greenstone belt ? ARCHEAN
What is your confidence in this estimate ? 4

[Here we have a rule in which the likelihood ratio is a function of an answer supplied by the user. The exact procedure used for such rules is described in Section 6.1. The effect in this case is to increase the certainty about FPTS from 3 to 4.6.]

4-- Can you identify more than one discrete volcanic cycle within the greenstone belt ? NO

[This turns off the context for the remaining questions about the petrotectonic setting, and the system now turns to seeing if there is a favorable magmatic association and local environment, FMALE. Of the four relevant rules, the most important concerns the presence of a Komatiitic rock suite, KRS. Question 5 establishes the context for KRS.]

5--To what degree do you believe that:
(The prospective ore body lies in a sequence of mafic to ultramafic igneous rocks) ? 4

6--To what degree do you believe that: (This sequence of mafic to ultramafic rocks forms a Komatiitic rock suite) ? ?

This KNSD model uses the extended definition of Komatiite rock suite as proposed by Arndt et al. (1977), rather than the other more restrictive definitions formulated by Viljoen and Viljoen (1969) and Brooks and Hart (1974). In case you are uncertain as to whether you are dealing with a Komatiite rock suite, many of the following questions are designed to clarify this point. To what degree do you believe that the sequence of mafic to ultramafic igneous rocks forms such a Komatiite rock suite? 0

[Here the user responded to Question 6 about KRS with a request for clarification. Since no opinion is expressed about KRS, the system turns to the five rules having KRS as a consequent. No significant conclusions are reached until Question 10 has been answered.]

7--To what degree do you believe that:
(Some olivine-rich rocks are present as part of this sequence of mafic to ultramafic igneous rocks) ? 0

8--To what degree do you believe that:
(This sequence of mafic to ultramafic igneous rocks forms a sequence of shallow-seated intrusives or flows) ? 3

9--To what degree do you believe that:
(This sequence of mafic to ultramafic igneous rocks contains noncumulate textured peridotite) ? 0

10--To what degree do you believe that:
(This sequence of mafic to ultramafic rocks contains pyroxene rich rocks) ? 5

[This answer increases the certainty of a Komatiitic rock suite to 1.1; the answer to the next question about spinifex texture increases it further to 4.3.]

11--To what degree do you believe that:
(These pyroxene-rich rocks have spinifex texture) ? 3

...

[Following two more questions about KRS that do not change any certainties, the system returns to investigating the other rules for establishing FMALE, which now has a certainty of 3.7.]

14--To what degree do you believe that:
(The target area is within 10 miles of known concentrations of nickel-bearing sulfides associated with Komatiites) ? 5

[This favorable answer pushes the certainty of FMALE up to 4.8. At this point, the system is more concerned with negative evidence that might reduce the score significantly. The only rule with potential for definitely reducing the score comes from KRS, which has not been definitely established. Thus, the system returns briefly to try to establish the context for FFMA and FTM. However, this attempt fails.]

15-- Has a systematic chemical study of this sequence of mafic to ultramafic igneous rocks been done ? NO

16--To what degree do you believe that:
(The target area is close to domes with granitic intrusive cores) ? 2

My certainty in (Favorable Magmatic Association and Local Environment) is now: 4.94

Do you want to pursue this further? NO

[Note that the system could still ask about major paleo faulting, but since that could only push the score still higher, the system leaves the decision up to the user. At this point, the certainty of FPTS is 4.6 and the certainty of FMALE is 3.7. The questioning now turns to establishing whether or not the target is in a favorable ore-bearing unit, FOBU.]

17--To what degree do you believe that:
(The prospective ore body lies in a peridotite subvolcanic feeder or flow) ? 3

[This establishes the context for investigating FOBU. The system now tries the eight rules for establishing FOBU, but only weak evidence is obtained.]

18--To what degree do you believe that:
(There are known concentration of nickel-sulfides within the unit) ? 0

19--To what degree do you believe that:
(The concentration of MgO in the central portion of the unit is greater than 40%) ? 0

20--To what degree do you believe that:
(The target area is the feeder to or in the lowermost flow of a sequence of peridotite flows) ? 0

21--To what degree do you believe that:
(The target area is in the second or third lowermost flow of a sequence of peridotite flows) ? 0

22--To what degree do you believe that:
(The target area lies in a flow) ? 4

23--To what degree do you believe that:
(The thickness of the flow is greater than 10 meters) ? 5

24--To what degree do you believe that:
(The concentration of sulfur in the central portion of the unit is greater than 1000 parts per million) ? WHY

Experience has shown that bodies of peridotitic Komatiite which contain ore bodies are characterized by a higher overall sulfur content, usually greater than a thousand parts per million, than Komatiites that do not contain such ore bodies.

24--To what degree do you believe that:
(The concentration of sulfur in the central portion of the unit is greater than 1000 parts per million) ? 3

[Here the user responded to Question 24 with a request for an explanation as to why the question was being asked. Although the final answer was encouraging, neither it by itself nor in combination with the next answer provided sufficient evidence.]

25--To what degree do you believe that:
(The target area is surrounded by sulfur-bearing exhalite) ? 3

At this point all applicable rules have been used and the certainty for FOBU is only 0.64. The system does go on to consider the other major parts of the model, namely, Proximity to Ore and Right Mineralization, but the very weak favorability of the ore-bearing unit limits the certainty of the overall conclusion. Since all major features of the encoding have been illustrated, the example is terminated at this point in the consultation.

4 PRELIMINARY PERFORMANCE ANALYSIS OF THE KNSD AND PCDA MODELS

4.1 Testing Prospector Models

The Prospector system is intended to emulate the reasoning process of an experienced exploration geologist in assessing a prospect site for its likelihood of containing an ore body of the type represented by the model he or she designed. The question naturally arises as to how well it performs that task. This section reports the results of preliminary experiments for the KNSD and PCDA models addressing this question. Because performance analysis has many aspects, we begin with a discussion of the major considerations.

As explained elsewhere, the basis for a Prospector computation is a set of ore deposit models formulated by expert geologists. Since each model represents the reasoning process of a different expert geologist, we must perform a separate evaluation for each model. When we refer to Prospector in this section, we will always mean the Prospector system executing some particular model.

To determine how well Prospector performs we must first define the term "performance," i.e., we must decide what performance measurements to make and against what scale to compare them. This immediately raises fundamental methodological issues: we wish to compare Prospector to human geologists in terms of the accuracy with which the presence or absence of a particular deposit class can be predicted, but there exist no standard quantitative measures of human performance. The lack of a common absolute scale for comparing Prospector with geologists inevitably imposes constraints on the design of the present experiments.

Lacking an absolute scale of comparison, we employ instead a relative scale; our principal objective is to measure how closely Prospector's conclusions agree with those of the model designer.

Expressing this numerically as Prospector's relative error (E) in predicting the model designer's estimates, we can state that Prospector's predictive accuracy is, say, 75% for one model and 89.7% for another. In those cases in which Prospector and the model designer disagree, we have no way of determining objectively whose conclusions most closely match reality. On the one hand, the model contains only a subset of its designer's knowledge; on the other hand, the uniformity of the Prospector mechanism may make its behavior more consistent for a number of prospect sites than that of the model designer. Since direct comparison of Prospector with its model designers on an absolute scale must await the development of a suitable common performance standard, it will not be discussed further here.

To compare Prospector's conclusions with those of its model designer, their conclusions must be expressed in common terms. Since many of the ways that the model designer can express his or her conclusions, such as writing reports, are fundamentally beyond Prospector's capabilities, we instead ask the model designer to express his or her conclusions on the same -5 to 5 scale used by Prospector.

Next we must specify what Prospector and the model designer should express their respective conclusions about. To impart some realism to the evaluation, we choose as the objects of analysis several known deposits (three sites for each of the two models tested here) with which the model designer is familiar. The data concerning each deposit on which the evaluation is based consist of a set of answers to all the questions asked in the model. These answers are supplied by the model designer.

Of course, this approximation of a realistic test of Prospector leaves other dimensions unexplored. To preclude the possibility of bias, it would be preferable to use data supplied by someone other than the model designer. Likewise, it would be useful to include cases in which the data reflect incomplete information about a prospect site, as would be the case, for example, if the prospect site were at an early stage of exploration. Hence the present experiments provide no evidence

regarding Prospector's performance when dealing with limited information supplied by someone other than the model designer.

Furthermore, since all the test cases used here are exemplars of their respective models, we have no data yet concerning the possibility that, even though Prospector may rate a prospect site favorably, the model designer may rate it unfavorably. Accordingly, it is desirable to include cases of known deposits that match a model, but less closely than the exemplars studied here, as well as "near miss" cases, i.e., prospect sites that are partially favorable, but lack certain characteristics that are critical to establishing overall favorability for the model.

Moreover, the number of cases tested limits the reliability of the results; little can be concluded with certainty from just three case studies per model. In particular, the sample sets are not large enough to permit the application of standard statistical techniques. (Of course, the application of statistical methods may be inherently limited by the relatively small number of known deposits of a given type of ore body throughout the world.) Conclusions drawn from the current results consequently, must be regarded as tentative, subject to possible qualification or refinement based on more extensive subsequent experiments.

These limitations indicate the preliminary nature of the present case studies. There is virtually no literature on objective quantitative tests of the predictive power of geological models of ore deposits. More common are attempts to systematize the available data concerning a particular type of ore deposit into a general descriptive model with some genetic interpretations based on intrinsic properties, e.g., (Lowell and Guilbert, 1970). Therefore the numerical approach used here may be of some interest. In addition, Section 4.6 enumerates methodological refinements that may be applied to subsequent experiments.

Besides measuring the degree of concurrence between Prospector and the model designer, the present experiments attempt to satisfy several

other objectives as well. We record Prospector's scores and the model designer's corresponding conclusions not only for overall favorability of the deposit, but also for each of the model's major sections. In this way we can determine not only the extent of agreement in detail, but also whether they agree on some sections of the model more closely than on others. Hence we demonstrate the usefulness of experimental performance analysis results as a diagnostic tool indicating specifically which sections of a model can most benefit from revision.

Another diagnostic purpose is served by performing analogous performance analyses of two versions of the PCDA model (Sections 4.4 and 4.5). In this way we measure quantitatively the extent to which the revisions achieve the specific objectives that motivated them. To appreciate the need for objective measurements of improvements to a model, one must realize that any given Prospector model lies along a continuum of possible versions. Two models that differ only in the value of the likelihood ratio of a single rule are distinct models and, in general, will perform differently. Thus, revising a model constitutes a search in a vast space of possible models. (This illustrates a common technical notion in computer science, sometimes referred to as "hill climbing," signifying "hills" of high performance and "valleys" of low performance in the space of related models.) The objective measurements we report here serve to supplement the model designer's geological experience and intuitions regarding the topography of this space.

Other experiments (Sections 4.3.2 and 4.4.2) measure the robustness of Prospector, i.e., the sensitivity of its conclusions to perturbations or uncertainties in the field observations on which such conclusions are based. This is important because, despite the model designer's efforts to make the questions asked by a model as objective and unambiguous as possible, it is nevertheless likely that two geologists could give different answers to some questions about the same known deposit or prospect site. Certain studies indicate that geologists do in fact disagree among themselves to some extent in their observations, at least

with respect to rock names and descriptions (D'Allesandro 1977). The sensitivity analysis experiments measure the degree to which such variations in the input data can effect changes in Prospector's conclusions.

4.2 Methodology of Experiments

Here we describe the sequence of steps followed in the experiments that are reported in subsequent sections. The same methodology was applied to both the KNSD model (Section 4.3) and to two versions of the PCDA model (Sections 4.4 and 4.5). Since our methodology is general, it may be applied to the performance analysis of any model. Also described below is a computer program we devised to facilitate the experiments.

To make it easier for the model designer to answer the questions asked by the model for each of the known deposits tested, we created a computer program called QUESTIONNAIRE that automatically generates a questionnaire listing these askable questions. A person who wishes to document a particular known deposit or prospect site can do so, of course, simply by running Prospector in the normal fashion and saving a transcript of the run. The use of a questionnaire for documentation purposes offers two advantages over this latter approach: first, the user may answer every question askable for that model, whereas Prospector may not ask certain questions if the replies would not appreciably affect the final conclusion in that run. (For the purpose of the sensitivity analysis experiments, it is important to receive answers to all questions.) Second, filling out a questionnaire does not require access to the Prospector system, hence the use of a questionnaire decreases the delay in obtaining these data and increases the number of potential suppliers of data.

The next step after the questionnaire data have been obtained is to execute Prospector, using the answers given in each questionnaire. To do so the experimenter would normally have to sit at the computer terminal with questionnaire in hand, consulting it to find the proper response to each question asked by Prospector. This procedure is

awkward for several reasons: the order in which Prospector asks questions depends, as a rule, on the answers to previous questions, necessitating some searching by the experimenter to locate each question on the questionnaire; in addition the sensitivity analysis experiments require several executions based on the replies made on each completed questionnaire, so that the time required of the experimenter is multiplied severalfold.

Accordingly, we added a facility to the QUESTIONNAIRE program that enables a user to input interactively at the terminal the answers written on a questionnaire. In this mode the program writes a computer disk file containing the answers in a format that may be used as input with the BATCHMODE feature (see Section 6.5). This greatly facilitates converting answers written on a hard-copy questionnaire into machine readable form. The questionnaire is printed on the terminal in a format identical to that appearing on the questionnaire, so that one can simply read the answers off a hard-copy questionnaire one by one, typing them in when prompted.

Having created an input data file containing the answers given on each questionnaire, we execute Prospector once for each such file. A permanent transcript is made of the run and the tracing facility turned on during the entire run, so that the final certainty values assigned to each space in the inference network during the run may be determined later by inspecting the transcript. After the final certainty values assigned to each space in the topmost levels of the inference network have been recorded, these values are analyzed in various ways described in detail in subsequent sections. In addition, the Prospector scores for each known deposit are compared with analogous estimates provided by the model designer (see Tables 7 and 15).

The next step is to make two additional executions of Prospector for each set of questionnaire data, to be used in sensitivity analysis. In general, sensitivity analysis measures the extent to which the output of a transformation varies as a result of changes in the inputs. In applying this general notion to Prospector, we see that the input values

are the user's replies to questions posed by Prospector (or, as in the present experiments, the questionnaire data supplied by the model designer). The output values are the certainty values assigned to the spaces of the inference net of a model. Our objective is to determine the amount of change in output (Prospector score) resulting from a unit change in input (certainty of the questionnaire answers).

Accordingly, we create two variants of each set of questionnaire data. The new input data sets systematically shift the questionnaire answers by one unit of certainty--in one case less certain, in the other case more. These new sets are referred to as the "less certain" data set and the "more certain" data set, respectively. The original questionnaire data are referred to as the "standard" data set. For example, if a given question is answered with a 3 in the "standard" set, the corresponding value in the "less certain" set is 2, while in the "more certain" set it is 4. Similarly, -3 in the "standard" set becomes -2 in the "less certain" set, -4 in the "more certain" set. The exact formulas used in these transformations are defined below.

"Standard"	"Less Certain"	"More Certain"
if X = 0	0	0
X > 0	X-1	minimum(X + 1, 5)
X < 0	X+1	maximum(X - 1, -5)

The certainty values obtained from the "standard," "less certain" and "more certain" runs are then compared in various ways (e.g., see Tables 9, 10, 13, and 14).

4.3 Performance Analysis of the KNSD Model

4.3.1 Performance analysis for three known deposits

The development of the Komatiitic-nickel-sulfide-deposit model (KNSD), having passed the stage of implementation and initial revisions, is now ready to be revised for "fine tuning" of its performance. The performance analysis reported in this section is intended to provide an objective measure of its current performance and to serve as a guide to

future revisions. The input data for this evaluation consist of questionnaire answers (of the kind described in Section 4.2) for each of the following three known deposits:

* the Langmuir deposit, located in the vicinity of Timmins, Ontario
* the Alexo deposit, located in the vicinity of Timmins, Ontario
* the Lunnon deposit, located in Kambalda, Western Australia

These data were supplied by Prof. Anthony Naldrett of the Department of Geology of the University of Toronto, who developed the KNSD model. In each instance we asked Prof. Naldrett to identify the source of his knowledge and the maturity of the field observations underlying his data. As regards the Langmuir and Alexo deposits, Prof. Naldrett's data are based on extensive first-hand knowledge and reflect a detailed three-dimensional model (i.e., the results of several field seasons of observations, plus significant drilling and mining exposures). In the case of the Lunnon deposit, his data are based on minor first-hand knowledge and the descriptions published by Woodall and Travis (1969), Ross and Hopkins (1975), and Ewers and Hudson (1972). The questions listed in the questionnaire and the answers to them supplied for each of the three sites by Prof. Naldrett are contained in Appendix A.

Our first experiment in evaluating the KNSD model was simply to execute Prospector, using the data for each of the three sites in turn, to see how well these sites are scored by the KNSD model (on the -5 to 5 scale of certainty described in Section 2.4). The certainty scores observed were as follows:

	Langmuir	Alexo	Lunnon
Certainty:	3.465	1.435	1.958

Since these three sites are exemplars of the KNSD model, one might have expected their scores to be closer to 5.0 than they actually are. On the other hand, since the exact behavior of Prospector depends

on the interaction of dozens of numerical values that have not yet been fine-tuned for optimum performance, one might have been equally surprised had these initial tests revealed a close match.

In view of the fact that the overall scores are not close to 5.0, it is now our objective to account for the observed values. The overall score is derived from the scores of two spaces representing high-level conclusions in the model. These spaces are called "Favorable Conditions, Excluding Mineralization" (FCEM) and "Favorable Mineralization" (FM). For the three sites tested the following certainty scores were observed:

	Langmuir	Alexo	Lunnon
KNSD (overall)	3.465	1.435	1.958
FCEM	3.465	1.35	1.958
FM	5.0	5.0	5.0

These data serve three purposes: to pinpoint those sections of the model that are not performing as intended, to assess the consistency of performance across different sites, and also to illustrate the manner in which certainty scores at one level of abstraction in the model are combined to determine the score of a higher level space. The data show a perfect score for each of the three sites for the FM section of the model, whereas the scores for the FCEM section are lower than anticipated in each case tested. This indicates that efforts to revise the model should be focused on the FCEM section, rather than on the FM section.

Since the topmost space of KNSD is a logical conjunction of the FCEM and FM spaces (see Section 3.3 and Figure 5), the probability value associated with the KNSD space is computed as the minimum of the probability values associated with the FCEM and FM spaces. Although probability values are treated differently from certainty values in Prospector, there is a close mathematical relation between the two (see Section 2.4), with the result that the overall certainty score for KNSD

in these cases is in fact the minimum of those for the FCEM and FM spaces, as indicated above.

The data tabulated above suggested that we focus attention on the FCEM portion of the model, so we determined the certainty scores for each of the spaces on which the score of FCEM is based. The hierarchical structure of the top two levels of the KNSD model is given below. Included at the right in this enumeration is the total number of questions askable by Prospector for each section of the model, thus showing the relative distribution of these questions.

	Total Number of Questions Defined in KNSD Model		
Komatiitic nickel sulfide deposit (KNSD)	51		
Favorable conditions, excluding mineralization (FCEM)		43	
Favorable petrotectonic setting (FPTS)			9
Favorable magmatic association and local environment (FMALE)			17
Favorable ore-bearing unit (FOBU)			10
Proximity to ore (PTO)			7
Favorable mineralization (FM)		8	
Favorable sulfide assemblage (FSA)			7
Favorable stratigraphic sequence (FSS)			1

In an attempt to calibrate the KNSD model, we asked Prof. Naldrett to estimate the overall certainty score that should be assigned to each of the three deposits, as well as the score that should be assigned to each of the major components named above. Each such assignment is made on the same -5 to 5 scale used by Prospector, and is intended as a target number for the purpose of future fine tuning of the KNSD model. The estimates are given either in the form of a single number, or as two numbers establishing an upper and lower bound on a certainty interval. The answers are shown on the left in Table 7 for each site in turn, with the scores as determined by execution of Prospector indicated on the right. Prof. Naldrett was informed of the right-hand values on the right only after he had given us the ones on the left.

Table 7

Comparison of Prospector Scores with Naldrett's Estimates

Langmuir Deposit

Space Name	Naldrett's Estimate	Prospector Score
KNSD	4.5 - 5.0	3.465
FCEM	4.5 - 5.0	3.465
FPTS	3 - 4	4.412
FMALE	5.0	4.978
FOBU	4 - 5	3.459
PTO	4 - 5	3.547
FM	4.0 - 5.0	5.0
FSA	5.0	5.0
FSS	3	5.0

Alexo Deposit

Space Name	Naldrett's Estimate	Prospector Score
KNSD	4.0 - 4.5	1.435
FCEM	4.0	1.435
FPTS	3 - 4	4.479
FMALE	5.0	4.93
FOBU	3 - 4	1.656
PTO	4	1.403
FM	5.0	5.0
FSA	5.0	5.0
FSS	5.0	5.0

Lunnon Deposit

Space Name	Naldrett's Estimate	Prospector Score
KNSD	4.5 - 5.0	1.958
FCEM	5.0	1.958
FPTS	5	4.661
FMALE	5	4.978
FOBU	5	4.5
PTO	5	1.931
FM	5.0	5.0
FSA	5	5.0
FSS	5	5.0

Inspection of Table 7 reveals that PTO is the limiting factor in depressing the score of FCEM (which is defined as a logical conjunction), and hence of KNSD (also defined as a logical conjunction). This indicates that revisions of the FCEM section for the purpose of fine tuning should focus on PTO before any of the other constituents of FCEM.

The model is quite consistent in this deficiency across the three sites tested. The certainty scores for FPTS, FMALE, FOBU, and PTO show that FMALE gets the highest score, followed by FPTS, FOBU, and PTO in that order, and that this ordering is observed for each of the three sites (except that for Langmuir FOBU is 3.46 and PTO is 3.55).

Comparison of Prospector's scores with Prof. Naldrett's expectations provides additional insight. First, among the spaces FPTS, FMALE, FOBU, and PTO Prospector agrees with Prof. Naldrett's estimates most closely on FMALE, followed in order by FPTS, FOBU, and PTO (with the exception that for Langmuir the value for FOBU is slightly higher than for PTO). Hence, we have observed Prospector to be quite consistent across different sites for this section of the model, both in the actual scores and in its agreement with Prof. Naldrett's estimates.

Second, the definition of the KNSD space as the logical conjunction (AND) of FCEM and FM has the effect of propagating disagreements between Prospector and Prof. Naldrett in lower sections of the model (mainly PTO) to the highest level of the model. This indicates the desirability of replacing conjunctions by rules in the top several levels of the KNSD model, as was done in the PCDA model with encouraging results (reported in Section 4.5).

Third, for the FCEM section of the model Prospector's scores are in most cases somewhat less than those of Prof. Naldrett. For the FM section of the model, Prospector adheres closely to Prof. Naldrett's expectations. (However, the known deposits tested do not include any in which the mineralization is unfavorable.)

These conclusions can be expressed quantitatively by first identifying the values in Table 7 with a concise notation, then defining a simple formula for the relative error of Prospector in predicting Prof. Naldrett's estimates. Thus:

Let C(X, Y, Z) = Certainty score given to space Z by agent X for site Y, where X denotes either Prospector or Naldrett

so C(Prospector, Langmuir, FCEM) = 3.465. Whenever Prof. Naldrett gave an interval of values instead of a single value, we use the midpoint of the interval as the value of C. Then the error measure is given by:

$$\text{Let } E(Y, Z) = \frac{C(\text{Naldrett}, Y, Z) - C(\text{Prospector}, Y, Z)}{C(\text{Naldrett}, Y, Z)}$$

For example, E(Langmuir, FCEM) = (4.5 - 3.465) / 4.5 = .271, i.e., Prospector predicts Prof. Naldrett's estimate in this case to within 27.1%. Since Table 7 gives values for nine spaces for each of three known deposits, we can compute the value of E(Y,Z) for each of the 27 different combinations of deposit site (Y) and space in the model (Z). The average value of E for these 27 cases is 25.0%.

For convenience, we list in Table 8 these 27 values of E(Y,Z), expressed as percentages. The fourth column in Table 8 presents the average of the absolute values of the entries in the first three columns. These data indicate explicitly where additional revisions would be beneficial. That such "fine tuning" can dramatically improve the prediction ability of a Prospector model is demonstrated in Sections 4.4 and 4.5, in which we compare an early version of the PCDA model to a revised version.

Table 8

Relative Error (E) of Prospector Scores as Predictors of Naldrett's Estimates (Derived from data in Table 7)

	Langmuir	Alexo	Lunnon	Average of Absolute Values
KNSD	27.1 %	66.2 %	56.5 %	49.9 %
FCEM	27.1	64.1	60.1	50.4
FPTS	-26.1	-28.0	6.8	20.3
FMALE	-0.4	1.4	0.4	.7
FOEU	23.1	52.7	10.0	28.6
PTO	21.2	64.9	61.4	49.2
FM	-11.1	0	0	3.7
FSA	0	0	0	0
FSS	-66.7	0	0	22.2

4.3.2 Sensitivity analysis

As discussed in Section 4.2, it is important to determine how Prospector's performance varies as a function of the degree of certainty reflected in a user's answers to Prospector's questions. Following the methodology outlined in Section 4.2, we created two variants of each of the three sets of questionnaire data supplied by Prof. Naldrett. The new input data sets change Prof. Naldrett's answers systematically by one unit of certainty--in one variant making it more certain, in the other less certain. The objective is to determine the amount of change in output (Prospector score) resulting from a unit change in input (certainty of the questionnaire answers). The results for the three sites tested are shown in Table 9.

Table 9

Prospector Scores for KNSD Model Vary With the Degree of Certainty of the Inputs ("Standard" data from Table 7)

Langmuir Deposit

Space Name	Less Certain	Standard	More Certain
KNSD	2.798	3.465	3.465
FCEM	2.798	3.465	3.465
FPTS	4.119	4.412	4.412
FMALE	4.95	4.978	4.978
FOBU	2.799	3.459	3.459
PTO	2.778	3.547	3.682
FM	5.0	5.0	5.0
FSA	4.991	5.0	5.0
FSS	4.0	5.0	5.0

Alexo Deposit

Space Name	Less Certain	Standard	More Certain
KNSD	1.157	1.435	identical to "standard" values
FCEM	1.157	1.435	
FPTS	4.146	4.479	
FMALE	4.844	4.93	
FOBU	1.241	1.656	
PTO	1.122	1.403	
FM	4.999	5.0	
FSA	4.956	5.0	
FSS	4.0	5.0	

Lunnon Deposit

Space Name	Less Certain	Standard	More Certain
KNSD	1.482	1.958	2.101
FCEM	1.482	1.958	2.101
FPTS	4.134	4.661	4.752
FMALE	4.95	4.978	4.978
FOBU	4.062	4.5	4.5
PTO	1.45	1.931	2.075
FM	5.0	5.0	5.0
FSA	4.989	5.0	5.0
FSS	4.0	5.0	5.0

The fact that Prof. Naldrett's questionnaire answers are mostly "5"s and "-5"s, rather than less certain values, accounts for the fact that Prospector's scores for the "more certain" runs are close to those for the "standard" runs. The values in the "more certain" questionnaire data set are identical in most cases to the corresponding values in the "standard" questionnaire data set.

Comparison of the "standard" scores with the "less certain" scores indicates that several pieces of favorable evidence can combine to establish a cogent hypothesis, even when none of the evidence is absolutely certain. For example, consider the certainty values for the FMALE space for the Langmuir deposit. The space FMALE represents an intermediate hypothesis based on the answers to questions 10 through 24 of the questionnaire reproduced in Appendix A. Since those answers are mostly "5"s, a score of 4.978 results for FMALE in the "standard" run. When these "5"s are changed to "4"s in the "less certain" run, this evidence is still sufficient to establish the FMALE hypothesis with a certainty of 4.95.

Inspection of Table 9 supports the conclusion that the KNSD model is rather insensitive to unit changes in the certainties of a user's answers. Thus, the score given by Prospector may vary with the geologist who supplies the inputs, but not by a large amount if the geologists agree in their observations to within one unit of certainty. In actuality, the intersite differences are generally small, suggesting consistency in the behavior of the KNSD model across different sites.

Moreover, among the spaces FPTS, FMALE, FOBU, and PTO the relative disparity in score between "standard" and "less certain" parallels the relative ordering that was observed earlier; FMALE has the smallest relative difference, followed by FPTS, FOBU, and PTO, in that order. (The Alexo and Lunnon deposits observe this ordering partly.) This provides further evidence of consistency of the KNSD model across different sites. It also suggests that three measures--absolute score (C), Prospector's relative error (E) in predicting Prof. Naldrett's estimates (E), and sensitivity (S)--appear somewhat related. This suggests the possibility that revising a section of the model to achieve closer concurrence with Prof. Naldrett might have the fortuitous effect of decreasing the sensitivity of that section as well, although this remains to be established.

These conclusions can be expressed quantitatively by defining a sensitivity measure, S, analogous to the relative error measure E defined in Section 4.3.1.

Let C(W, Y, Z) = Certainty score given to space Z for site Y during run with uncertainty setting W, where W denotes either "standard" or "less certain"

so C("less certain", Langmuir, FCEM) = 2.798. Then the sensitivity value is given by:

$$\text{Let } S(Y, Z) = \frac{C(\text{"standard"}, Y, Z) - C(\text{"less certain"}, Y, Z)}{C(\text{"standard"}, Y, Z)}$$

For example, S(Langmuir, FCEM) = (3.465 - 2.798) / 3.465 = .192, i.e., a one unit change in the certainty of the input data causes a 19.2% change in the resulting score. Since Table 9 gives values for nine spaces for each of three known deposits, we can compute the value of S(Y,Z) for each of the 27 different combinations of deposit site (Y) and space in the model (Z). The maximum of these 27 values is 25%, and 11 of the 27 values are less than 1%. The average value of S for these 27 cases is 12.4%. For convenience, Table 10 lists these 27 values of S, expressed as percentages.

Table 10

Sensitivity Measure (S) for the KNSD Model
(derived from Data in Table 9)

	Langmuir	Alexo	Lunnon	Average
KNSD	19.2 %	19.4 %	24.3 %	21.0 %
FCEM	19.2	19.4	24.3	21.0
FPTS	6.4	7.3	11.3	8.3
FMALE	.6	1.7	.6	.97
FOBU	19.1	25.1	9.7	18.0
PTO	21.7	20.0	24.9	22.2
FM	0	.02	0	.007
FSA	.2	.9	.2	.4
FSS	20.0	20.0	20.0	20.0

The "less certain," "standard," and "more certain" runs differ not only in the certainty values computed by Prospector, but also in the number of questions Prospector asks during execution. Prospector does not ask a question bearing on a certain hypothesis in the model if the answers already given to other questions also bearing on that hypothesis are sufficient to establish the hypothesis with high certainty. In addition, some questions or sections of the model are asked only if another space has achieved a certainty falling within an interval specified by the model designer. Hence, changing the certainty of the inputs can have the effect of switching sections of a model on or off.

As a general rule, the less certain the answers given by the user, the more questions Prospector will ask. We obtained a precise measurement of this correlation by counting the numbers of questions asked during the "less certain" and "standard" runs. For the Langmuir and Lunnon deposits Prospector asked 25 questions during the "standard" run and 28 during the "less certain" run. For the Alexo deposit 25 questions were asked during the "standard" run, 26 during the "less

certain" run. Thus, the decrease in certainty has the effect of increasing the number of questions asked by 4% to 12%.

4.4 Performance Analysis of the PCDA Model (Version 1)

4.4.1 Performance Analysis for Three Known Deposits

4.4.1.1 Analysis of Model Excluding Zones

We now evaluate Version 1 of the PCDA model, using essentially the same methodology as was used in the preceding section to evaluate the KNSD model. In this section we evaluate the version of the PCDA model described in our last annual report (Duda et al., 1977). As elucidated in Section 3.2 of the present report, the PCDA model has since been extensively revised. One purpose of this evaluation is to measure the divergence in performance between the two versions. The revised version of the PCDA model (referred to as Version 2) is evaluated in Section 4.5. In the present section we use the term "PCDA model" to refer to Version 1 of the model.

The input data for evaluating the Type-A porphyry copper deposit model (PCDA) consist of questionnaire answers (of the kind described in Section 4.2) for each of the following three known deposits:

* the Yerington deposit, located in Weed Heights, Lyon County, Nevada
* the Bingham Canyon deposit, located in Salt Lake County, Utah
* Kalamazoo deposit, located in the San Manuel district, Pinal County, Arizona

These data were supplied by Prof. Marco Einaudi, of the Department of Applied Earth Sciences of Stanford University, who developed the PCDA model. In each case we asked Prof. Einaudi to identify the source of his knowledge and the maturity of the field observations his data reflect. With regard to the Yerington deposit, Prof. Einaudi reports that his data are based on considerable first-hand knowledge and reflect extensive mining and drilling exposures. In the

case of the Bingham deposit, his data are based on moderate first-hand knowledge and considerable reading and discussion with geologists who have first-hand knowledge. These data similarly reflect extensive mining and drilling exposures. In the case of the Kalamazoo deposit, Prof. Einaudi's data are based on published descriptions by J. David Lowell (1968), and reflect a detailed three-dimensional model derived from surface, drilling, and mining exposures. The questions listed in the questionnaire for the PCDA model and Prof. Einaudi's answers regarding each of these three sites by Prof. Einaudi are given in Appendix B.

In the execution of Prospector, using the data for each of the three sites in turn, the following overall certainty scores were observed:

	Yerington	Bingham	Kalamazoo
Score:	.047	.444	.208

These data indicate that Version 1 of the PCDA model does not perform as Prof. Einaudi intended, since he considers these three sites to be exemplars of the PCDA model. Therefore, as in Section 4.3 for the KNSD model, our objective now is to account for the values that were observed. The overall score is determined from the scores of three spaces representing high-level conclusions in the model. These three spaces are called "Favorable Petrotectonic Setting" (FPTS), "Right Regional Environment" (RRE), and "Favorable PCDA Intrusive System" (FPCDAIS). For the three sites tested, the following certainty scores were observed:

	Yerington	Bingham	Kalamazoo
PCDA	.047	.444	.208
FPTS	4.091	2.273	2.273
RRE	.047	2.754	.208
FPCDAIS	.262	.440	.448

These data indicate deficiencies in each of the three major sections of the PCDA model, especially in the RRE and FPCDAIS sections. Accordingly, we shall expand each of these spaces into its constituents, as we did for the KNSD model in Section 4.3. Note, incidentally, that the top space of the PCDA model is defined as a logical conjunction of FPTS, RRE, and FPCDAIS, as is the case in the KNSD model. Hence, as in the KNSD model, the overall probability (and also, in this case the certainty score) of the PCDA space is limited by the minimum of the scores associated with the three major sections.

The topmost levels of the PCDA model's hierarchical structure are as follows:

	Total Number of Questions Defined in PCDA Model
Porphyry Copper deposit, Type A (PCDA)	88
Favorable petro-tectonic setting (FPTS)	7
Right regional environment (RRE)	10
Granitic intrusive system in the region (ISYS)	1
Favorable regional environment (FRE)	9
Favorable PCDA intrusive system (FPCDAIS)	71
Admissible PCDA intrusive system (APCDAIS)	6
Favorable intrusive system (FIS)	9
Favorable alteration and mineralization (FAMR)	56
Favorable zones present (FST)	56
Barren-core zone (RS)	15
Potassic zone (KS)	18
Propylitic zone (PS)	9
Sericitic zone (SS)	14
Favorable zone combination (FSC)	0

We now present in Table 11 the observed certainty scores for these various spaces--for each of the three sites in turn. (Table 11 is analogous to Table 7 for the KNSD model.)

Table 11

Prospector Scores for Several Levels of the PCDA Model
(Version 1)

Yerington Deposit

PCDA	.047			
FPTS		4.091		
RRE		.047		
ISYS			5.0	
FRE			.047	
FPCDAIS		.262		
APCDAIS			3.991	
FIS			4.392	
FAMR			.262	
FST				4.995
FSC				.262

Bingham Deposit

PCDA	.444			
FPTS		2.273		
RRE		2.754		
ISYS			5.0	
FRE			2.754	
FPCDAIS		.440		
APCDAIS			2.487	
FIS			4.281	
FAMR			.440	
FST				4.995
FSC				.440

Kalamazoo Deposit

PCDA	.208			
FPTS		2.273		
RRE		.208		
ISYS			5.0	
FRE			.208	
FPCDAIS		.448		
APCDAIS			5.0	
FIS			4.452	
FAMR			.448	
FST				4.995
FSC				.448

The data in Table 11 indicate that the FIS section of the model yields rather favorable scores, that the APCDAIS and FPTS sections might benefit from some revision, and that the FRE and FAMR sections may need significant revision. (Analogous results for the revised PCDA model are displayed in Table 15.)

4.4.1.2 Analysis of Zone Identification and Discrimination

We shall now undertake a more detailed analysis of the performance results for the FAMR section of the PCDA model, which concerns the interpretation of each of a number of spatially distinct zones distinguishable by the user at the prospect site. As discussed in Section 3.2, the PCDA model allows four possible interpretations of a given zone, namely, the barren-core zone, the potassic zone, the sericitic zone, and propylitic zone interpretations.

For each zone on the prospect, Prospector asks the user questions relevant to each of the four zone interpretations, ultimately assigning a certainty value to each. The observed zone-number/zone-interpretation matrix is given in Table 12 for each of the three sites tested. The bracketed abbreviation associated with each zone number in the table indicates the correct interpretation for that site, as determined by Prof. Einaudi. For ideal zone discrimination the diagonal entries from upper left to lower right in each matrix should have high certainty scores, and the other entries should have low or negative certainty. The entries in the table represented by dashes indicate that Prospector pursued the propylitic zone interpretation prior to the others, and that this interpretation was scored so high as to obviate the need to consider alternative interpretations.

Table 12

Zone Interpretations in the PCDA Model (Version 1)

Yerington Deposit

Zone number/ Interpretation	Barren-Core	Potassic	Sericitic	Propylitic
1 [B.C.]	3.567	-5	-4.98	-4.859
2 [Pot.]	-4.495	-5	.450	-4.999
3 [Ser.]	-5	-5	4.995	-4.999
4 [Prop.]	2.226	-5	.034	1.753

Bingham Deposit

Zone Number/ Interpretation	Barren-Core	Potassic	Sericitic	Propylitic
1 [B.C.]	3.375	-5	-4.999	-2.574
2 [Pot.]	-4.722	-5	.045	-4.915
3 [Ser.]	-5	-5	4.995	-4.928
4 [Prop.]	--	--	--	4.882

Kalamazoo Deposit

Zone Number/ Interpretation	Barren-Core	Potassic	Sericitic	Propylitic
1 [B.C.]	3.567	-5	-5	1.481
2 [Pot.]	-4.095	-5	-5	-4.915
3 [Ser.]	-5	-5	4.995	-5
4 [Prop.]	--	--	--	4.976

The data in Table 12 indicate that the propylitic zone is well discriminated from the other interpretations for each of the three sites, and the same is true for the sericitic zone. The barren-core zone was correctly identified in all three sites, but the propylitic interpretation also received a positive (but smaller) score in the case of the Yerington deposit. In contrast, the potassic zone interpretation

received a very negative score in all cases, even when that interpretation was the correct one. Consequently, this section of the model requires revision. The potassic zone deficiency is in fact responsible for the low scores reported for the FSC space (and therefore the FAMR space) in Table 11: FSC is established by a favorable combination of zones, for which the presence of a potassic zone is highly favorable. Since no zone was interpreted as potassic, FSC received a low score.

4.4.2 Sensitivity Analysis

Following the methodology outlined in Section 4.2, we performed a sensitivity analysis of the PCDA model analogous to that reported in Section 4.3.1 for the KNSD model. The performance measurements observed for the three sites tested are enumerated in Table 13.

Table 13

Prospector Scores for PCDA Model (Version1) Vary with degree of Certainty of Inputs ("Standard" Data from Table 11)

Yerington Deposit

Space Name	Less Certain	Standard	More Certain
PCDA	.016	.047	.088
FPTS	3.182	4.091	5.0
RRE	.016	.047	.088
ISYS	4.0	5.0	5.0
FRE	.016	.047	.088
FPCDAIS	.215	.262	.262
APCDAIS	2.982	3.991	5.0
FIS	3.993	4.392	4.452
FAMR	.215	.262	.262
FST	4.091	4.995	4.995
FSC	.215	.262	.262

Bingham Deposit

Space Name	Less Certain	Standard	More Certain
PCDA	.373	.444	.446
FPTS	1.364	2.273	3.182
RRE	2.139	2.754	2.754
ISYS	4.0	5.0	5.0
FRE	2.139	2.754	2.754
FPCDAIS	.369	.440	.441
APCDAIS	1.99	2.487	2.487
FIS	3.878	4.281	4.281
FAMR	.369	.440	.441
FST	4.495	4.995	4.995
FSC	.369	.440	.441

Kalamazoo Deposit

Space Name	Less Certain	Standard	More Certain
PCDA	.218	.208	.076
FPTS	1.364	2.273	3.182
RRE	.218	.208	.076
ISYS	4.0	5.0	5.0
FRE	.218	.208	.076
FPCDAIS	.369	.448	.449
APCDAIS	3.991	5.0	5.0
FIS	4.092	4.452	4.452
FAMR	.369	.448	.449
FST	4.495	4.995	4.995
FSC	.369	.448	.449

Note that for the Kalamazoo data the "less certain" score for the PCDA space exceeds the "standard" and "more certain" scores for that space. Table 13 shows these values to have propagated from the FRE space. The reason for this atypical ordering of the FRE scores is that negative as well as positive answers are made "less certain". For example, a certainty value of -5 in the "standard" case becomes -4 in the "less certain" case. Hence the impact of both negative and positive evidence is reduced by this transformation. Because of the particular likelihood ratio values of the rules in the FRE section of the model, negative evidence dominates positive evidence in this section of the model for the Kalamazoo deposit. Therefore, decreasing the certainty of both positive and negative evidence actually increases the favorability of the FRE hypothesis.

The relatively small discrepancies between the scores for "standard" runs and the corresponding scores for "more certain" runs reflects the fact that the questionnaire answers supplied by Prof. Einaudi consist mostly of "5"s and "-5"s, rather than less certain values. As a result, the values in the "more certain" questionnaire data set are identical in most instances to the corresponding values in the "standard" questionnaire data set, as was also the case for the evaluation of the KNSD model.

Inspection of Table 13 supports the conclusion that the PCDA model is in most cases rather insensitive to unit changes in the certainties of answers supplied by a user. To state this quantitatively and to facilitate comparing the sensitivity of the PCDA model with that of the KNSD model, we compute for the data in Table 13 the values of the sensitivity measure S(Y,Z) defined in section 4.3.2. Recall that Y identifies the known deposit and Z denotes the space in the model. The values of S(Y,Z) derived from the data of Table 13 are given in Table 14. The fourth column of Table 14 presents the average of the absolute values of entries listed in the first three columns.

Table 14

Sensitivity Measure (S) for the PCDA Model (Version 1)
(derived from data in Table 13)

	Yerington	Bingham	Kalamazoo	Average of Absolute Values
PCDA	66.0 %	16.0 %	-4.8 %	28.9 %
FPTS	22.2	40.0	40.0	34.1
RRE	66.0	22.3	-4.8	31.0
ISYS	20.0	20.0	20.0	20.0
FRE	66.0	22.3	-4.8	31.0
FPCDAIS	17.9	16.0	17.6	17.2
APCDAIS	25.3	20.1	20.2	21.9
FIS	9.1	9.4	8.1	8.9
FAMR	17.9	16.0	17.6	17.2
FST	18.1	10.0	10.0	12.7
FSC	17.9	16.0	17.6	17.2

The data in Table 14 are somewhat larger than the corresponding values for the KNSD model reported in Table 10, indicating a greater dependence by Prospector scores upon the degree of certainty of inputs for the PCDA model than for the KNSD model. Continuing the comparison, we observe that the largest value repoted in Table 14 is 66%, compared with 25.1% in Table 10. In Table 14 only 6 of the 33

values for the three sites are less than 10%, whereas 11 of the 27 values in Table 10 are less than 1%. The grand average for the 33 values in Table 14 is 21.7%, compared with 12.4% for the KNSD values in Table 10.

In general, as observed in Section 4.3.2, the less certain the user responses to questions posed by Prospector, the more questions Prospector will ask. In that section we compared for the KNSD model the number of questions asked by Prospector during the "less certain" run with the number asked during the "standard" run. We did the same for the PCDA model. With respect to the Yerington deposit, Prospector asked 122 questions during the "standard" run and 218 during the "less certain" run, an increase of 79%. For the Bingham deposit the corresponding numbers are 109 and 205, respectively, an increase of 97%. For the Kalamazoo deposit, they are 88 and 189, an increase of 115%. These large increases contrast sharply with the small increases observed for the KNSD model. Closer inspection reveals that almost all of this increase occurs in the FAMR section of the PCDA model, which identifies the zones present at the prospect site. The KNSD model, conversely, has no such representation of zones.

4.5 Performance Analysis of the PCDA Model (Version 2)

4.5.1 Analysis of Model Excluding Zones

We now evaluate the revised version of the PCDA model described in Section 3.2, and compare this version (which we refer to as Version 2) with the original one (Version 1). We executed Prospector with Version 2 of the model, using the same questionnaire data for the three sites tested in the evaluation of Version 1 of the model. The overall certainty scores are listed below and the analogous scores for Version 1 repeated for purposes of comparison.

	Yerington	Bingham	Kalamazoo
Version 1:	.047	.444	.208
Version 2:	4.769	4.721	4.756

These data indicate a dramatic improvement in the performance of the PCDA model. Incidentally, the revisions were made prior to the evaluation of Version 1 of the PCDA model.

To show performance in detail, we give below the hierarchical structure of the major sections of Version 2 of the model. At this coarse level of detail the two versions are nearly identical in the logical structure of the inference network. Included at the right in this enumeration is the total number of questions askable by Prospector for each of the major sections of the model, thus showing the relative distribution of these questions. (The number for FRE includes a question that is not strictly contained in this section, but serves to establish the correct context for inquiring about FRE. The same holds for FPCDAIS.)

	Total Number of Questions Defined in PCDA Model (Version 2)
Porphyry Copper deposit, type A (PCDA)	81
Favorable petrotectonic setting (FPTS)	4
Favorable regional environment (FRE)	9
Favorable PCDA intrusive system (FPCDAIS)	68
Favorable composition in differentiated sequence (FCDS)	4
Favorable intrusive system (FIS)	9
Favorable alteration and mineralization relations (FAMR)	53
Favorable zones present (FST)	53
Barren-core zone (RS)	17
Potassic zone (KS)	15
Propylitic zone (PS)	8
Sericitic zone (SS)	13
Favorable zone combination (FSC)	0

We performed a calibration exercise analogous to that reported in Section 4.3.1 for the KNSD model, i.e., we asked Prof. Einaudi to estimate the certainty score that should be assigned to each of the three deposits, as well as to each of the major components of the model for each site. His score is given on the same -5 to 5 scale as is used by Prospector, with the answers intended as target numbers for the purpose of future fine tuning of the PCDA model. The estimates are

given either in the form of a single number, or as two numbers establishing an upper and lower bound on a certainty interval. The estimates are listed in Table 15 on the left for each site in turn, with the scores as determined by execution of Prospector reproduced on the right. We informed Prof. Einaudi of the values on the right only after he had given us those on the left. The key NA in Table 15 signifies that Prof. Einaudi gave no estimate for that space. (Table 15 is analogous to Table 7 for the KNSD model.)

Table 15

Prospector Scores for Several Levels of the PCDA Model
(Version 2)

Yerington Deposit

Space Name	Einaudi's Estimate	Prospector Score
PCDA	4.5 - 5.0	4.769
FPTS	4.5 - 5.0	4.528
FRE	4.5	4.540
FPCDAIS	4.5 - 5.0	4.787
FCDS	5	4.524
FIS	5	4.744
FAMR	4.5 - 5.0	4.225
FST	NA	5.0
FSC	NA	4.225

Bingham Deposit

Space Name	Einaudi's Estimate	Prospector Score
PCDA	4.5	4.721
FPTS	3.5 - 4.0	4.449
FRE	4.0 - 4.5	4.829
FPCDAIS	4.5 - 5.0	4.729
FCDS	5	2.407
FIS	5	4.744
FAMR	4.0	4.225
FST	NA	5.0
FSC	NA	4.225

Kalamazoo Deposit

Space Name	Einaudi's Estimate	Prospector Score
PCDA	4.0 - 4.5	4.756
FPTS	4.0 - 4.5	4.449
FRE	3.5	1.784
FPCDAIS	4.5 - 5.0	4.791
FCDS	5	4.722
FIS	5	4.744
FAMR	4.0	4.225
FST	NA	5.0
FSC	NA	4.225

The data in Table 15 show that Prospector scores each of these sections of the model with high certainty for each site, with the exception that space FCDS for the Bingham deposit and space FRE for the Kalamazoo deposit are scored somewhat lower. This contrasts strongly with the analogous results for Version 1 of the model given in Table 11, in which a number of sections of that version are scored with rather low certainty values.

In most cases shown in Table 15 Prospector agrees very closely with Prof. Einaudi's estimate. To express this agreement quantitatively, we apply to the PCDA model the C(X, Y, Z) and E(Y, Z) notation defined in section 4.3.1. Hence, for example, the relative error of Prospector in predicting the score of FPCDAIS for Yerington is:

$$E(\text{Yerington, FPCDAIS}) = (4.75 - 4.787) / 4.75 = -.008$$

meaning that Prospector's prediction is accurate to within 0.8% in this case. Since Table 15 gives values for six spaces for each of three known deposits, we can compute the value of E for 21 different instances. For 5 of the 21 data points Prospector predicted Prof. Einaudi's estimate to within 1%, while 15 of the 21 data points show agreement to within 10%. The grand average over the 21 data points is 10.3% (compared to the analogous value of 25.0% for the KNSD model).

For convenience, we list these 16 values of E in Table 16, expressed as percentages. In particular, note that there is a more even distribution of positive and negative values than was the case for the KNSD model, which tends to underestimate Prof. Naldrett's expectations (Table 8 in Section 4.3.1).

Table 16

Relative Error (E) of Prospector Scores as Predictors of Einaudi's Estimates (derived from data in Table 15)

	Yerington	Bingham	Kalamazoo	Average of Absolute Values
PCDA	-.3 %	-4.9 %	-11.9 %	5.7 %
FPTS	4.7	-18.6	-4.7	9.3
FRE	-.9	-13.6	49.0	21.2
FPCDAIS	-.8	.4	-.9	.7
FCDS	9.5	51.9	5.6	22.3
FIS	5.1	5.1	5.1	5.1
FAMR	11.1	5.6	5.6	7.6

4.5.2 Analysis of Zone Identification and Discrimination

We shall give a more detailed analysis here of the performance results for the FAMR section of Version 2 of the PCDA model, which is concerned with the interpretation of each of a number of spatially distinct zones that the user can distinguish at the prospect site. This is a critical part of the model, because identification of certain combinations of these zones gives the best evidence of a match with the model. The results shown in Table 17 are analogous to those displayed in Table 12 for Version 2 of the model. (See the beginning of Section 4.4.1.2 for an explanation of what these numbers measure.)

Table 17

Zone Interpretations in the PCDA Model (Version 2)

Yerington Deposit

Zone	Barren-Core	Potassic	Sericitic	Propylitic
1 [B.C.]	3.457	-5.000	-4.980	-4.859
2 [Pot.]	-4.495	4.988	.450	-1.505
3 [Ser.]	-5.000	-5.000	5.000	-1.489
4 [Pro.]	-3.043	-5.000	0.360	1.753

Bingham Deposit

Zone	Barren-Core	Potassic	Sericitic	Propylitic
1 [B.C.]	3.257	-5.000	0	-2.574
2 [Pot.]	-4.722	4.988	.450	-4.915
3 [Ser.]	-5.000	-5.000	5.000	-4.700
4 [Pro.]	-3.043	-5.000	0	4.908

Kalamazoo Deposit

Zone	Barren-Core	Potassic	Sericitic	Propylitic
1 [B.C.]	3.457	-5	-5	1.481
2 [Pot.]	-4.095	4.988	-5	-4.915
3 [Ser.]	-5	-5	5	-5
4 [Pro.]	0.650	-5	-5	4.964

The data in Table 17 indicate that the deficiencies of Version 1 of the PCDA model with respect to zone identification and discrimination have been completely eliminated. In particular, the potassic zone is now scored correctly, and the barren-core zone is more distinctly differentiated from the propylitic zone than was the case in Version 1 of the model. Zone discrimination in Version 2 of the model approximates the ideal, in the sense that the diagonal entries from upper left to lower right in each matrix in the table have high certainty scores, whereas the other entries have very negative scores in most cases. In no case are these latter entries scored high enough to cause ambiguous interpretation of the zone.

It is a straightforward task to perform a sensitivity analysis of Version 2 of the PCDA model analogous to those for Version 1 of that model and for the KNSD model, but it has not yet been done.

4.6 Conclusions and Future Work in Performance Analysis

4.6.1 Summary of Experimental Results

We have developed an experimental methodology for measuring quantitatively the current performance of Prospector in some detail along several dimensions. Each run reported here uses input data supplied by the model designer about a known deposit with which he or she is familiar. The results span two models (the KNSD model and two versions of the PCDA model), three sites per model, and three variants of the degree of certainty of the input data--for a total of 21 distinct executions of Prospector (9 for the KNSD model, 9 for Version 1 of the PCDA model, and 3 for Version 2 of the PCDA model).

For each run we recorded not only the overall score Prospector assigned to that deposit, but also the score assigned to each major section of the model and its immediate subsections. In addition, for the PCDA runs we recorded the zone interpretation scores. The total number of these performance measurements recorded over all 21 distinct Prospector executions amount to 81 for the KNSD model, 147 for Version 1 of the PCDA model, and 75 for Version 2 of that model.

The results of this preliminary performance analysis can be summarized as follows.

(1) For both the KNSD model and Version 2 of the PCDA model, we measured Prospector's accuracy in predicting the judgment of the model designer about known deposits. To do this we obtained the model designer's expectations for each of various major sections of the model, asking him to score each of these for each site tested on the same scale as is used by Prospector. As a calibration exercise, we then compared these target scores with the Prospector scores actually observed during execution.

The results indicate that Version 2 of the PCDA model predicts very accurately the assessments that Prof. Einaudi, its designer, made for each of the known deposits tested (Tables 15 and 16). For overall favorability the differences between Prof. Einaudi's estimates and Prospector's scores range from 0.3% for the Yerington deposit to 11.9% for the Kalamazoo deposit, with an average of a 5.7% difference over the three sites. To obtain some detail in these results, we also made an analogous comparison for each of 6 major sections of the PCDA model, for a total of 21 comparisons over the three known deposits tested. Of these 21 data points 6 show a difference of less than 1%, and 13 a difference of less than 10%. The average for all 21 is 10.3%. For the KNSD model (Tables 7 and 8), the average for 27 analogous comparisons of Prospector's score with Prof. Naldrett's estimate shows a 25% difference.

Although covering a limited number of cases that include only exemplars of the model, these results are extremely encouraging. They demonstrate quantitatively Prospector's potential in accurately predicting the conclusions of exploration geologists who are authorities on particular types of ore bodies. They also confirm that careful and systematic revision of an existing model can improve its predictive ability dramatically.

(2) By means of sensitivity analysis experiments we measured quantitatively the robustness of the KNSD model and Version 1 of the PCDA model, as well as of various major sections thereof. The results (Tables 9, 10, 13, and 14) identify those sections of each model that are particularly sensitive to changes in the user's certainty. On the whole, these results indicate that neither the KNSD nor PCDA models, nor their principal subsections, are particularly sensitive to unit changes in the certainty of the user. On the average, the PCDA model is somewhat more sensitive (S = 21.7%) than the KNSD model (S = 12.4%).

(3) We have measured the capability of both versions of the PCDA model to identify and discriminate effectively the four zones represented in that model. The results reveal certain deficiencies in Version 1 of that model (Table 12), but also demonstrate that these observed deficiencies have been completely eliminated in Version 2 of that model (Table 17).

(4) We have compared two versions of the PCDA model in detail for the same known deposits, thereby measuring quantitatively the degree of improvement obtained by revising that model. These comparison results involve not only the zone-identification section mentioned above, but the entire model.

(5) By using several known deposits as test cases for each of the two models tested, we have measured quantitatively the consistency in each model's performance across several sites. For example, in Section 4.3.1 we observed that among the four constituents of the FCEM section of the KNSD model the space FMALE consistently receives a higher score than the others, followed in descending order by spaces FPTS, FOBU, and PTO. In terms of concurrence with Prof. Naldrett's estimates, and of the results of applying the sensitivity measure S, the same ordering (with but few exceptions) is observed among these spaces. By means of these and other measurements we have observed the KNSD model to show somewhat greater consistency across sites than does the PCDA model. It remains an open question whether these intermodel differences are accounted for by a greater variability among the three sites tested for the PCDA model than among their three counterparts tested for the KNSD model, or whether differences in the respective properties of the models themselves are the cause of the disparity.

Taken as a whole, these performance analysis results prove useful in several ways. First, they demonstrate the potential utility of performance analysis as a routine diagnostic tool for "fine tuning" of models. The results establish precise priorities for model revisions in that they identify the particular sections of a model that would most benefit from improvement, both as regards concurrence with the model designer's expectations and robustness of the model. Second, the results confirm the value of routinely measuring the performance of each version of a model as a means of measuring quantitatively the extent to which the revisions achieve the specific objectives that motivated them. Third, the results indicate that early versions of a model may not

perform as well as the model designer intended. This holds certain implications for the model design process, which by its very nature requires feedback. Performance analysis can accelerate this refinement process, and we should attempt to devise a variety of procedures and tools to increase the effectiveness of the initial model design process.

4.6.1.1 Future Work

The preliminary performance analysis of Prospector reported here has laid the foundation for more extensive evaluation efforts in the future. We enumerate several such possibilities below.

* Analogous Experiments on Additional Cases

It would be useful to duplicate the experiments reported here for additional cases of known deposits, including "near miss" cases. Presumably, the supplementary data would permit more reliable statements about Prospector's performance. Given a sufficiently large sample set of known deposits and prospect sites at advanced or early stages of exploration, conclusions about the performance of a model could be stated with statistical precision. Furthermore, with additional cases we could meaningfully compare Prospector's ranking of the test sites with a ranking obtained from the model designer. We see this extension of the present experiments to more cases as a major focus of future performance analysis efforts.

* Critical Factor Analysis

The present sensitivity analysis experiments measure the sensitivity of a given section of a model to changes in the user's answers. Since all the answers in the input data are changed, the results obscure the possible effect of changing only the response to any single particular question, leaving the answers to the others unchanged.

An alternative to the present method of changing all answers supplied by the user is to change just those that are rather uncertain (say, in the range -2 to 2). Each such reply would be changed to +5 or -5 in two ways: (1) so that the overall certainty for the model

is maximized, and (2) so that it is minimized. This would indicate which of the questions the user was uncertain about could have an important effect on the final outcome, had the user responded with greater certainty.

* Comparison of Data Supplied by Several Geologists About the Same Site

The sensitivity analysis experiments indicate the extent to which Prospector scores vary with the degree of certainty on the part of the user in answering Prospector's questions. To complement these measurements it is interesting to measure directly the degree to which two or more geologists familiar with the same known deposit concur in their answers about that site. If they always agreed perfectly with one another, there would be no need for sensitivity analysis experiments, because Prospector is a uniform mechanism that always yields the same results when given identical inputs. If the geologists disagree widely in response to certain questions, it may be necessary to extend the present sensitivity analysis experiments to greater changes in degree of certainty, e.g., to two units instead of one (as in the present experiments). Hence, measuring directly the range of variability among geologists' answers and calibrating the sensitivity analysis experiments accordingly may render the results of the latter more useful.

* Sensitivity Analysis of Rule Strength Values

In the sensitivity analysis experiments reported here we considered the models as fixed and we varied the inputs to the Prospector program. The "standard," "less certain," and "more certain" variations characterize in a simplified way the range of potential differences among field geologists in their observations about a given known deposit or prospect site.

However, the likelihood ratio values (i.e., LS and LN) of the rules supplied by the model designer and subsequently incorporated into a model are also subject to the designer's judgment. It may therefore be useful to perform experiments in which the likelihood ratio values vary while the inputs given during a run are held fixed, so as to

observe how sensitive the model is to changes in the likelihood ratio values.

Perhaps the simplest such experiment is to define three variations of a given model: a "standard rules" model, a "stronger rules" model, and a "weaker rules" model, analogous to the "standard," "more certain," and "less certain" versions defined for the sensitivity analysis experiments. The "standard rules" model is the one supplied by the model designer. In the "stronger rules" variation of a "standard rules" model, all LS (LN) values are increased (decreased) by a factor of two, say, whereas in the "weaker rules" variation these values are decreased (increased) by a factor of two.

It would then be possible to obtain for each site tested a 3-by-3 matrix of numbers reflecting all combinations of model variations and input variations. Given such data for a number of known deposits, some type of factor analysis or "credit assignment" techniques may permit us to determine the relative influence on Prospector's performance of the model designer's uncertainties, on the one hand, and of the user's responses, on the other. The analysis of variance (ANOVA) techniques of Kirk (1968) and Winer (1971) may prove to be particularly appropriate. Gillogly (1978) and Paxton (1977) have successfully applied these statistical techniques to certain artificial intelligence programs.

* Larger Uncertainties in Sensitivity Analysis Experiments

It would be interesting to extend the present sensitivity analysis experiments to two or three or four units of change in certainty. The results this obtained could enable Prospector's score to be plotted as a function of certainty in the inputs.

* Incorporation of Performance Analysis into the Explanation System

With some additional development effort several of the performance measurements reported or proposed above could be automated, making it possible for Prospector to execute these experiments during a

consultation session and report the results to the user immediately. For example, at the end of a session (at the user's option), Prospector might silently execute the "less certain" and the "more certain" runs, treating the answers just given by the user as constituting the "standard" run. In this scenario, Prospector would then present to the user tables analogous to Tables 9 and 10, or a summary thereof.

Similarly, Prospector might perform a critical factor analysis at the end of a consultation session, indicating which questions might have significantly changed the overall outcome, had the user answered them with greater positive or negative certainty. The effect would be to identify those observations about a prospect site (especially one at an early stage of exploration), the refining of which by additional investigation in the field would benefit a user.

The present experiments also offer the possibility of an interesting "fallout" effect, namely, to incorporate the questionnaire data for the known deposits already tested into a computer data base accessible to Prospector. In this scenario Prospector could compare the replies made by the user during the consultation session with those for the known deposits on record, and determine which of these the user's case matches most closely. This matching process could involve both the user's observations and the strengths of Prospector's several conclusion levels. For example, Prospector might report, "The answers you have given indicate that your prospect is very similar to the Yerington deposit, except that in your prospect the intrusive system is not as favorable as the one at Yerington, for the following reasons:..." Quite detailed quantitative comparisons could be performed in this manner.

5 RULES FOR DRILLING-SITE SELECTION

5.1 Introduction

The basic function of Prospector is to match geological field data to ore-deposit models. These data are always spatially dependent, and Prospector has several methods for taking spatial location into account. One of these is to partition the model into sections dealing with properties that are uniform over nested regions of varying size, such as the petrotectonic setting, the regional environment, and the local prospect. When there is significant variation within the prospect--as is the case with zones of alteration--Prospector can refer to homogeneous ares within the prospect symbolically (e.g., as Zone-1 or Zone-2) and use location variables to keep the information about each area separate.

Neither of these procedures is convenient when nontrivial spatial relationships among the data are important. It is often remarkably difficult to describe in words or in semantic networks something that can be expressed easily and naturally through a diagram or a map. Thus, it is clearly desirable for Prospector to be able to accept and utilize graphical input data.

We have begun developing a facility for using graphical input by applying Prospector's procedures to the problem of selecting a drilling site for a porphyry copper deposit--with the goal of finding the best hypogene grade of ore. This application was suggested by Victor F. Hollister, Manager for Canadian Exploration for the Duval International Corporation. He generously worked with us to develop an inference network that would combine various kinds of geochemical, geophysical, and geological evidence--such as proximity to faults or to boundaries of alteration zones--to determine the best drilling site.

The input used by the program is derived directly from certain contour maps. The user is asked whether he or she has certain kinds of map data, such as copper concentration contours derived from geochemical soil sampling. If the data are available, the user uses a digitizing tablet to enter the contours. The resulting digitized map is uniformly sampled at some convenient sampling interval. For any given sample point, the evidence for the inference network is typically a certainty measure based on how close that sample point is to a particular contour. The favorability of that particular point as a candidate drilling site is computed in the usual manner--by propagating the input certainties through the inference network. By repeating this process for every sample point, an array or sampled map of favorabilities is produced.

Mr. Hollister's initial set of rules treated the case in which there was stockwork, with a quartz-bearing intrusive normally present. This case is referred to as Model-1. An initial set of site-selection rules for Model-1 was described in Appendix B of our Second Semiannual Report (Hart et al., 1978). As a result of preliminary testing this inference network was revised, and rules were developed for two other models--Model-2 (the diorite model: stockwork, with a non-quartz-bearing intrusive normally present), and Model-3 (a breccia-pipe system). The revised network and the test results are described is the following sections.

5.2 Revision of Hollister's Rules

Version-1 of the inference network contained 81 spaces and 23 rules. Version-2--shown in Figure 6--contains 112 spaces and 53 rules. As with the porphyry copper model, some of the new rules were due to the replacement of conjunctive combinations by inference rules. However, most of the growth was due to inclusion of the two new models. In addition to these structural changes, all the parameter values (prior probabilities and likelihood ratios) have been systematically revised.

Of the 112 spaces in Version 2, 37 provide map input to the inference network. These spaces are explicitly listed in Table 18.

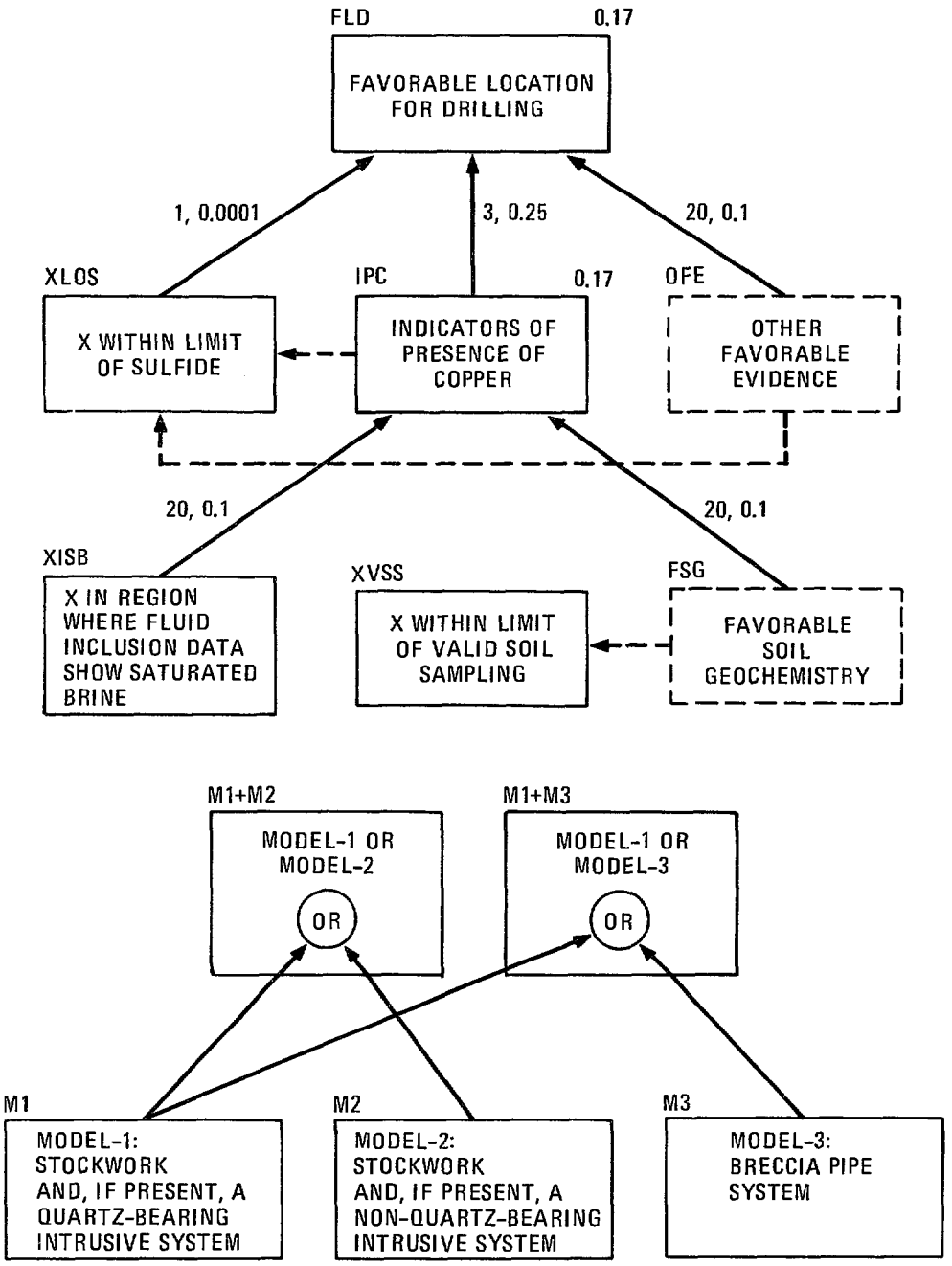


FIGURE 6 INFERENCE NETWORK FOR DRILLING SITE SELECTION

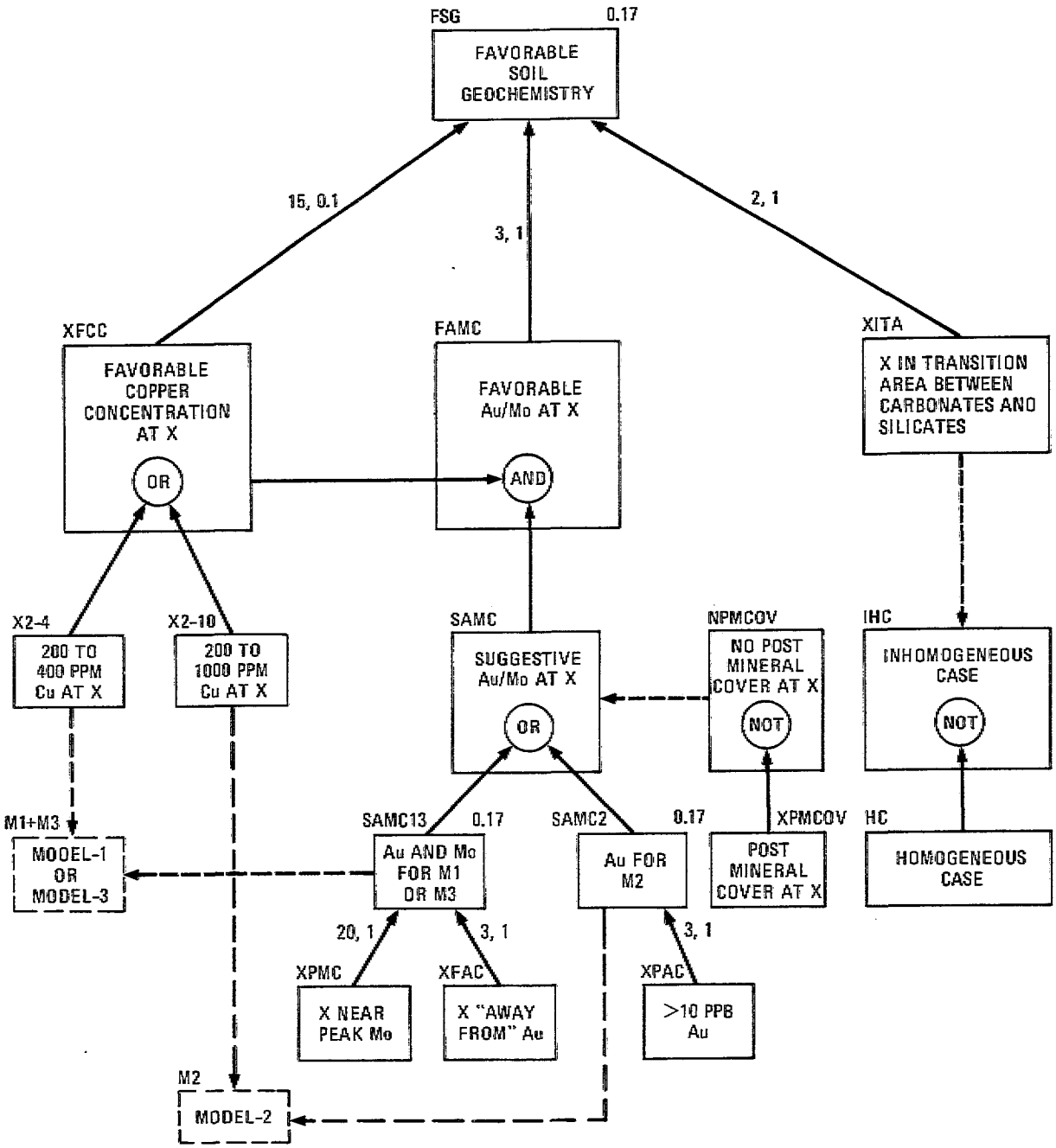


FIGURE 6 INFERENCE NETWORK FOR DRILLING SITE SELECTION (Continued)

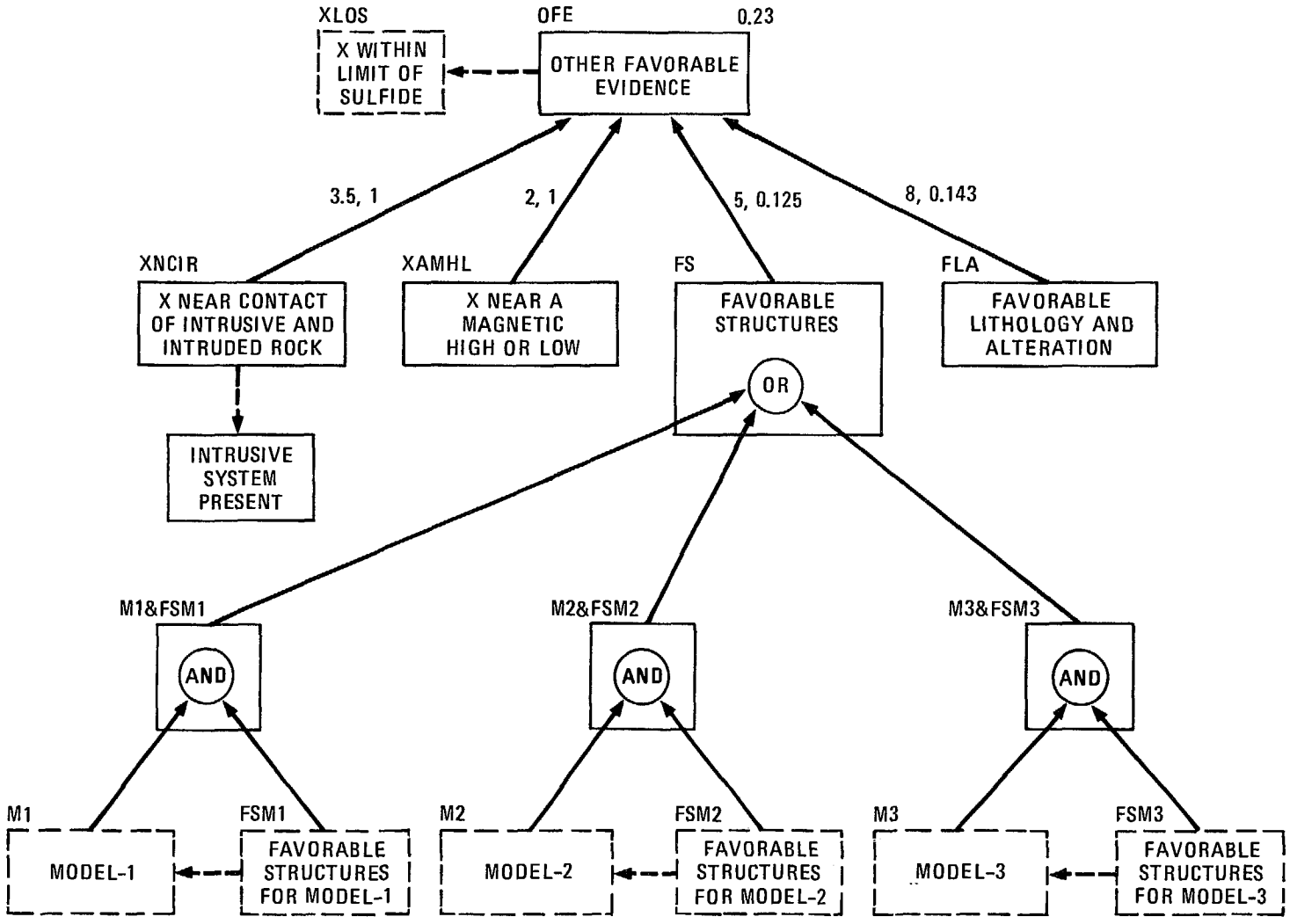
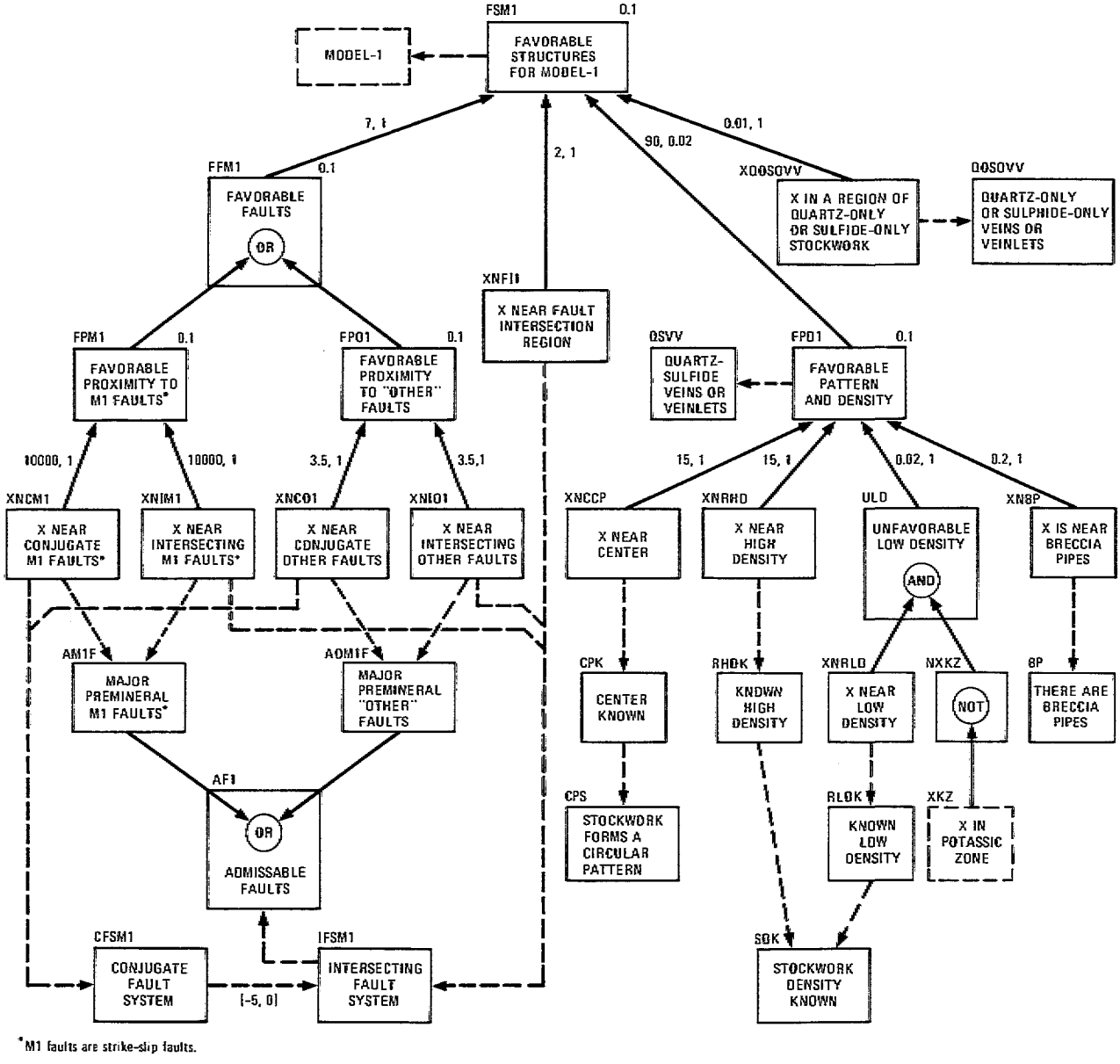
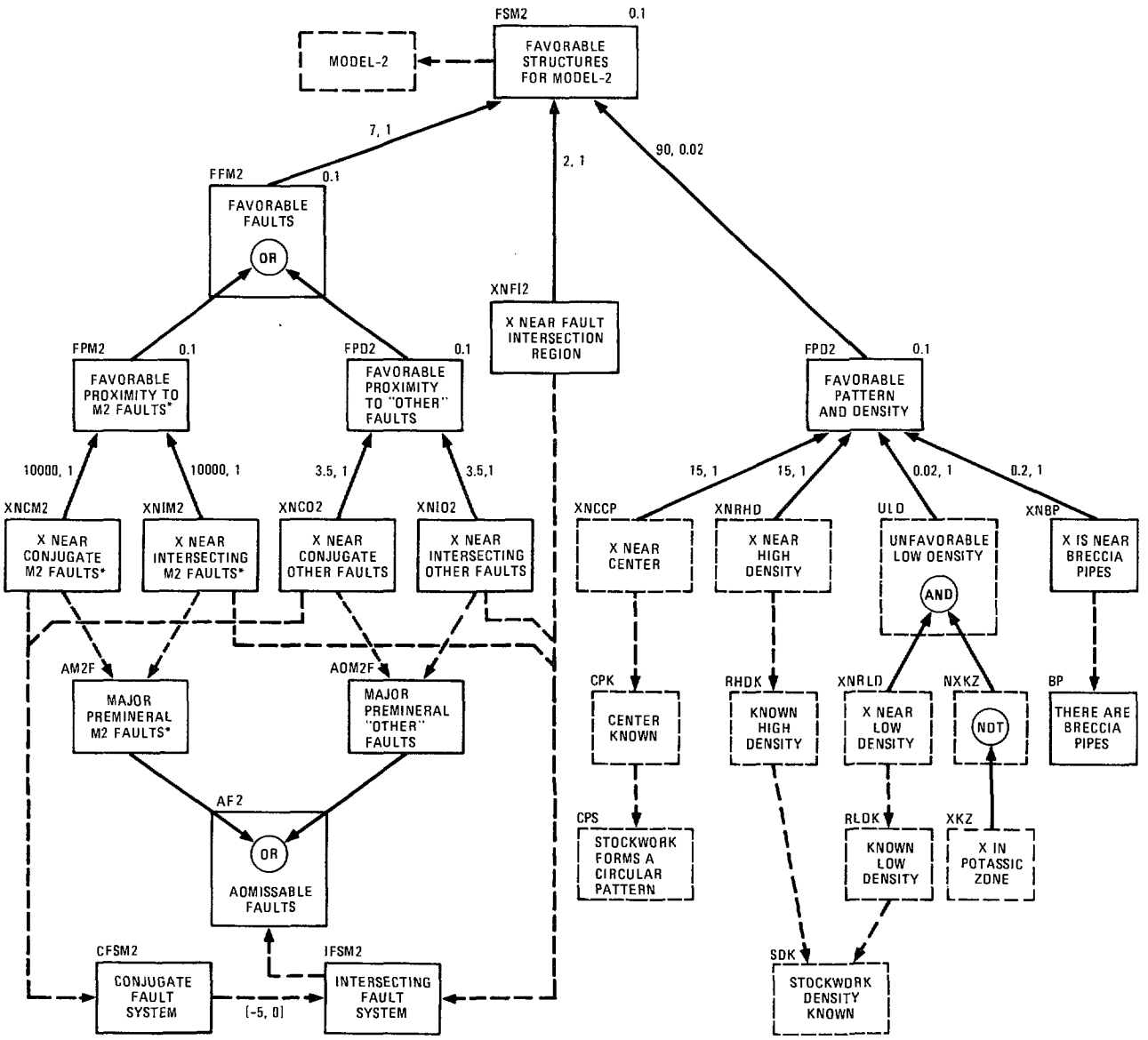


FIGURE 6 INFERENCE NETWORK FOR DRILLING SITE SELECTION (Continued)



*M1 faults are strike-slip faults.

FIGURE 6 INFERENCE NETWORK FOR DRILLING SITE SELECTION (Continued)



*M2 faults are strike-slip or dip-slip faults.

FIGURE 6 INFERENCE NETWORK FOR DRILLING SITE SELECTION (Continued)

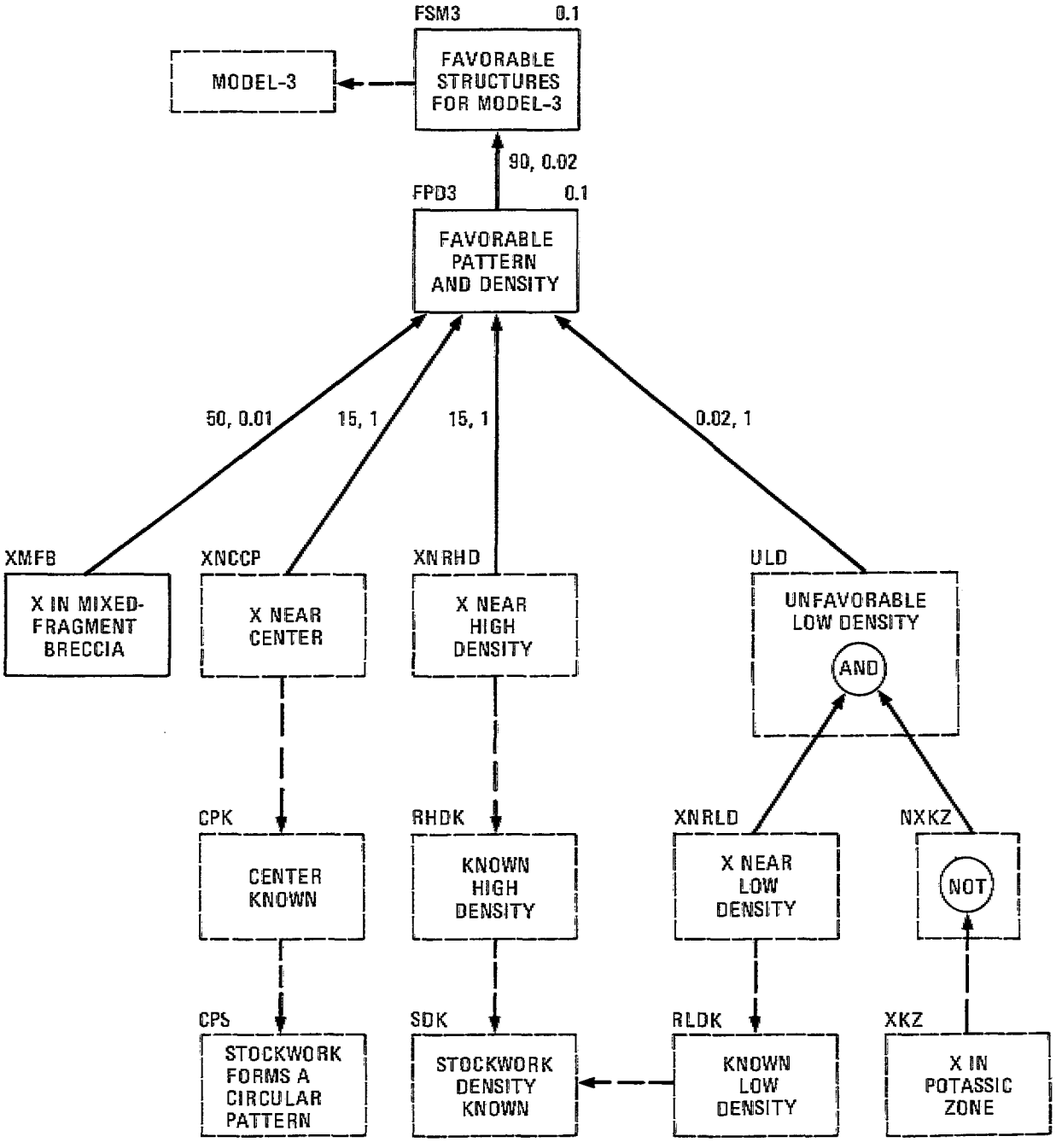


FIGURE 6 INFERENCE NETWORK FOR DRILLING SITE SELECTION (Continued)

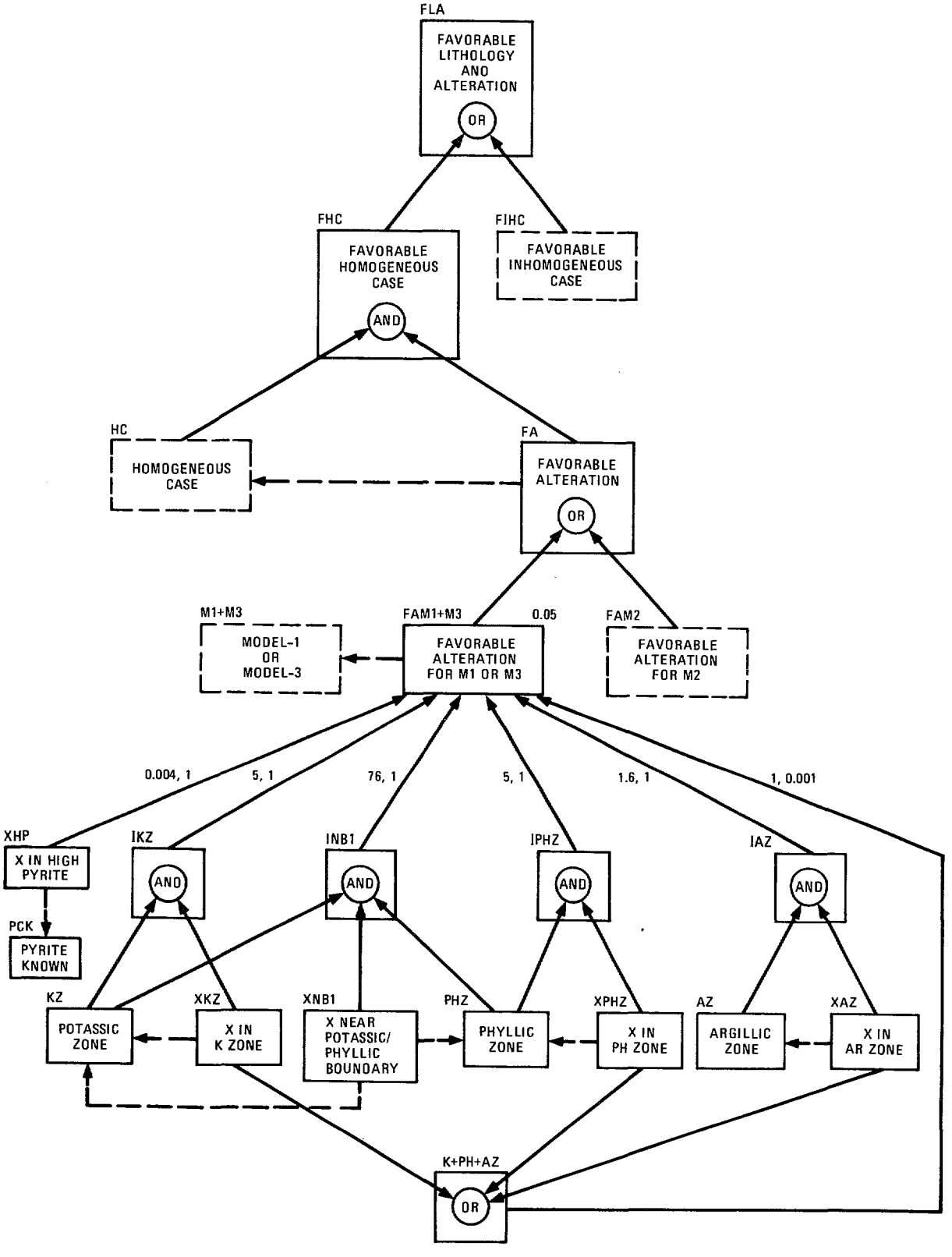


FIGURE 6 INFERENCE NETWORK FOR DRILLING SITE SELECTION (Continued)

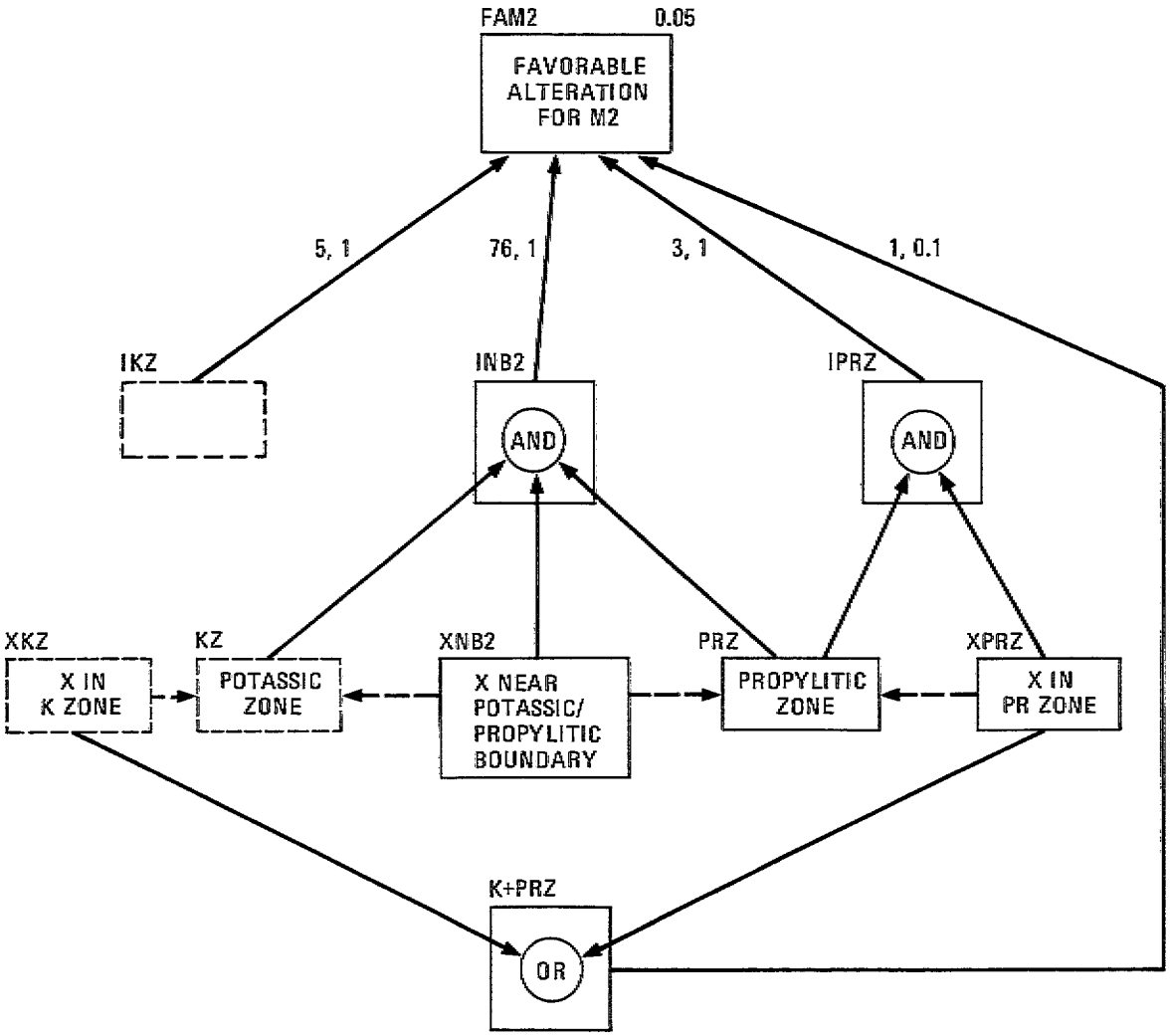


FIGURE 6 INFERENCE NETWORK FOR DRILLING SITE SELECTION (Continued)

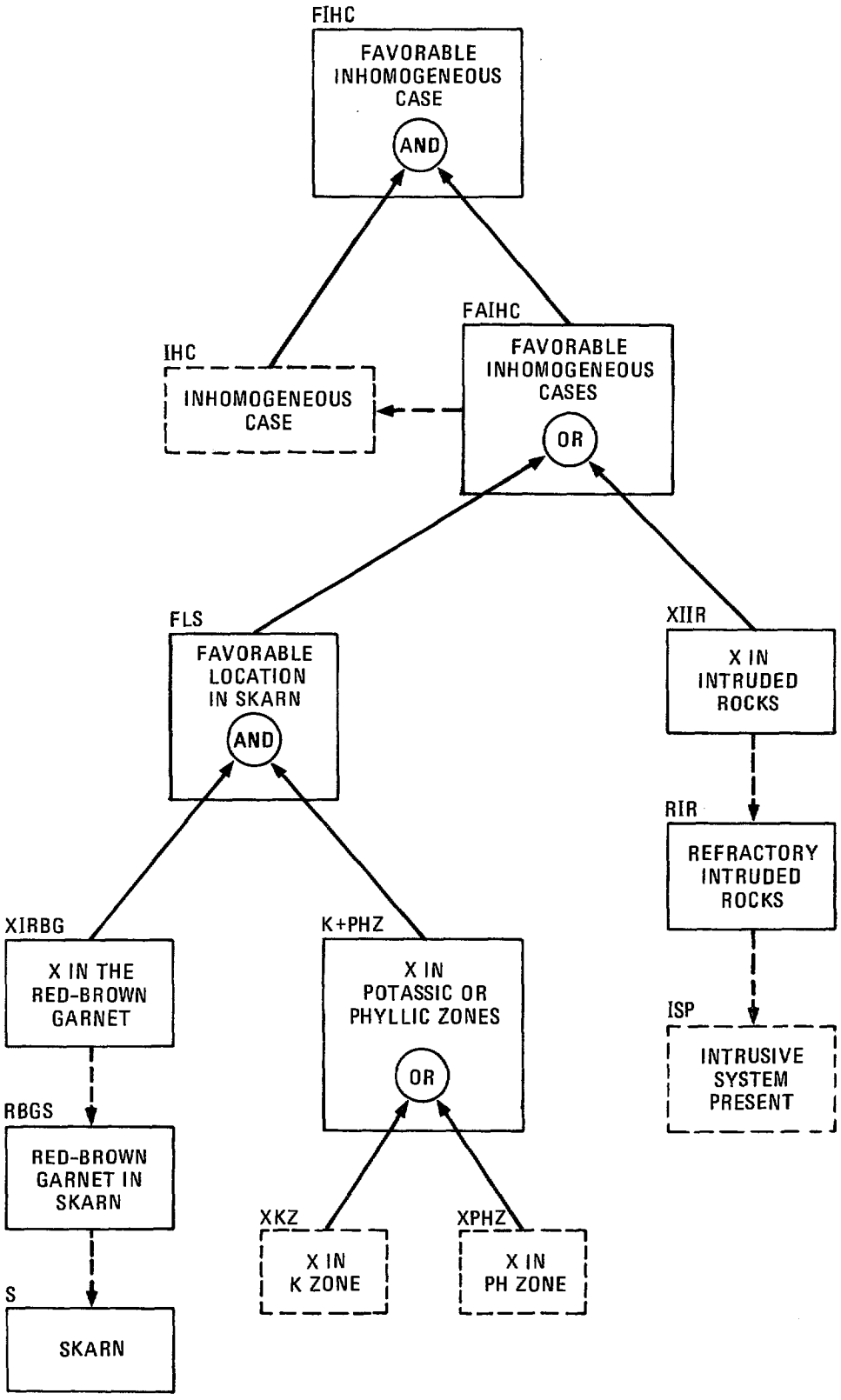


FIGURE 6 INFERENCE NETWORK FOR DRILLING SITE SELECTION (Concluded)

They describe one or the other of two basic situations--being within some region or near some contour. In both cases there are transition regions--places where it is not clear whether or not the situation holds. To effect the transition smoothly we employ a piecewise-linear "favorability function" that gradually reduces the probability from 1 to 0.

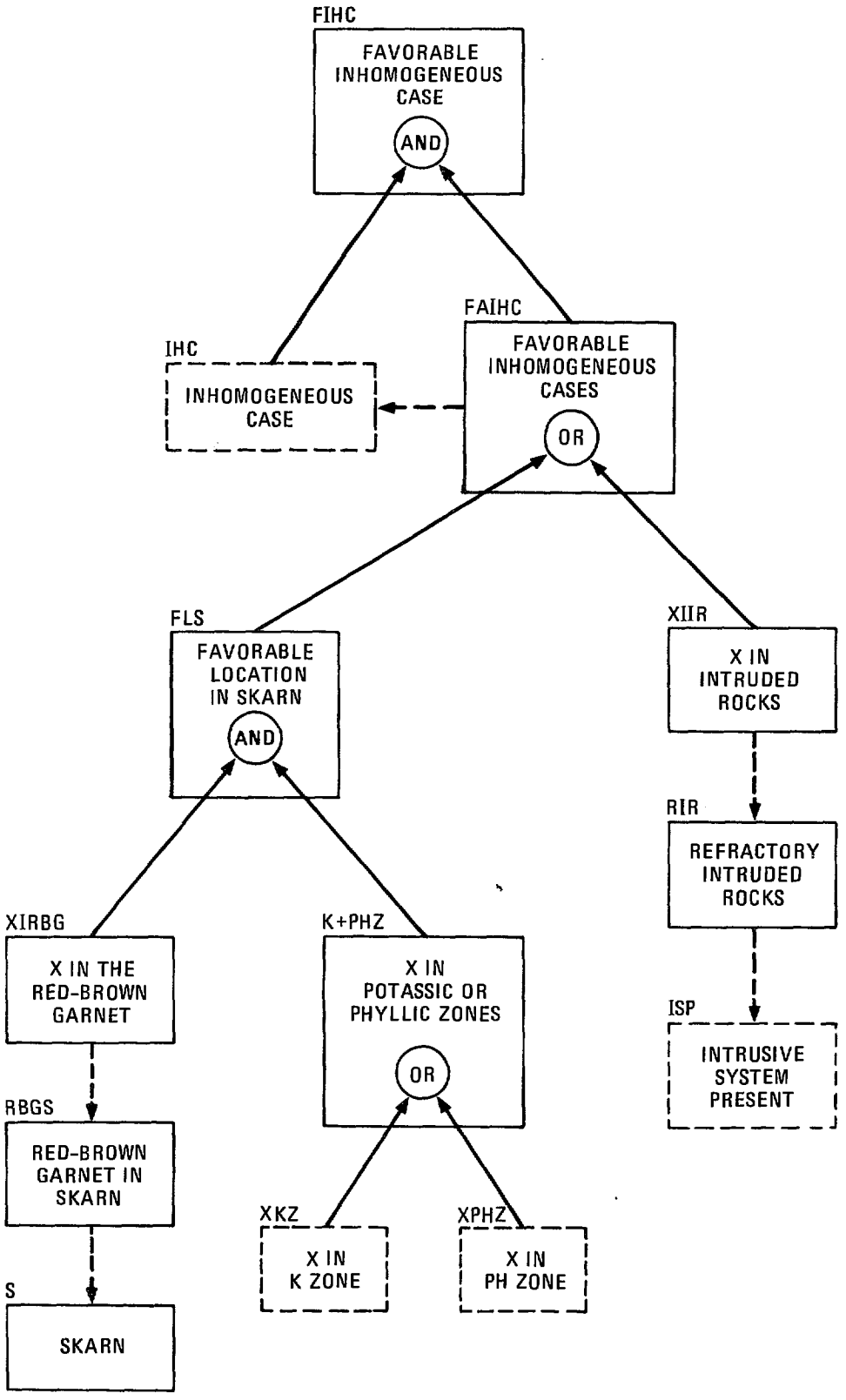


FIGURE 6 INFERENCE NETWORK FOR DRILLING SITE SELECTION (Concluded)

They describe one or the other of two basic situations--being within some region or near some contour. In both cases there are transition regions--places where it is not clear whether or not the situation holds. To effect the transition smoothly we employ a piecewise-linear "favorability function" that gradually reduces the probability from 1 to 0.

Table 18

Parameters for Favorability Functions

Space Name	Space Description	Favorability Parameters: a (m)	b (m)
XAMHL	in a magnetic anomaly region	0	100
XAZ	in an argillic zone	0	0
XFAC	"favorable" with respect to Au concentration	0	100
XHP	in a region of high pyrite	0	0
XIIR	in intruded rocks	0	0
XIRBG	in red-brown garnet	0	0
XISB	fluid inclusion data show saturated brine	0	0
XITA	in transition area, carbonates to silicates	0	100
XKZ	in a potassic zone	0	0
XLOS	within limit of sulfide	0	0
XFMB	in mixed-fragment breccia	0	0
XNBP	in a breccia-pipe region	0	50
XNB1	near potassic/phyllic boundary	0	0
XNB2	near potassic/propylitic boundary	0	0
XNCCP	near center of circular stockwork pattern	300	1000
XNCIR	near intrusive contact	20	100
XNCM1	near conjugate strike-slip faults	100	500
XNCM2	near conjugate strike-or-dip-slip faults	100	500
XNCO1	near conjugate non-strike-slip faults	100	500
XNCO2	near conjugate non-strike-or-dip-slip faults	100	500
XNFI1	in intersection region of strike-slip faults	0	0
XNFI2	in intersection region of strike-or-dip-slip faults	0	0
XNIM1	near intersecting strike-slip faults	100	500
XNIM2	near intersecting strike-or-dip-slip faults	100	500
XNIO1	near intersecting non-strike-slip faults	100	500
XNIO2	near intersecting non-strike-or-dip-slip faults	100	500
XNRHD	in a region of high stockwork density	0	0
XNRLD	in a region of low stockwork density	0	0
XPAC	in a region of more than 10 ppb Au	0	100
XPHZ	in a phyllic zone	0	0
XPMC	in a region of peak Mo concentration	0	100
XPMCOV	within postmineral cover	0	0
XPRZ	in a propylitic zone	0	0
XQOSOVV	in quartz-only or sulfide-only veins or veinlets	0	0
XVSS	within limit of valid soil sampling	0	0
X2-4	in 200 to 400 ppm Cu concentration	0	100
X2-10	in 200 to 1000 ppm Cu concentration	0	100

To define this function, let d be the distance from the sample point to the boundary. Then, for points outside a region or not exactly on a contour, we compute the probability of being within the region or near the contour by the function

$$F(d,a,b) = \begin{cases} 1 & \text{if} \quad 0 < d \leq a \\ \dfrac{b-d}{b-a} & \text{if} \quad a < d < b \\ 0 & \text{if} \quad b \leq d \,. \end{cases}$$

Thus, the transition zone extends from a to b, points closer than a considered as being inside the region (or near the contour), while points farther than b considered as being outside the region (or far from the contour). Values for the distance parameters a and b are given in Table 7.

5.3 Test Examples for Hollister's Rules

Three well-known porphyry copper deposits in British Columbia were used to evaluate these site-selection rules: Island Copper, Bell Copper, and Gibraltar. Since all three of these deposits have been mined, the actual locations of the ore bodies are known and could be used to evaluate the results. However, since all of these cases involve stockwork and a quartz-bearing intrusive (Model-1), they do not test either Model-2 or Model-3. The results obtained, therefore, must be viewed as preliminary tests of our approach.

5.3.1 Island Copper

A basic geologic map of Island Copper, showing the intrusive system and the outline of the ore body, appears in Figure 7 (Northcote, 1970). Additional information about premineralization faults, breccia pipes, geophysics and geochemistry were obtained from Young and Rugg (1970). (See also Pratt, 1970 and Island, 1972).

Digitized map contours prepared from these data were sampled at a 25-meter sampling interval to produce the following 128-by-128 data arrays:

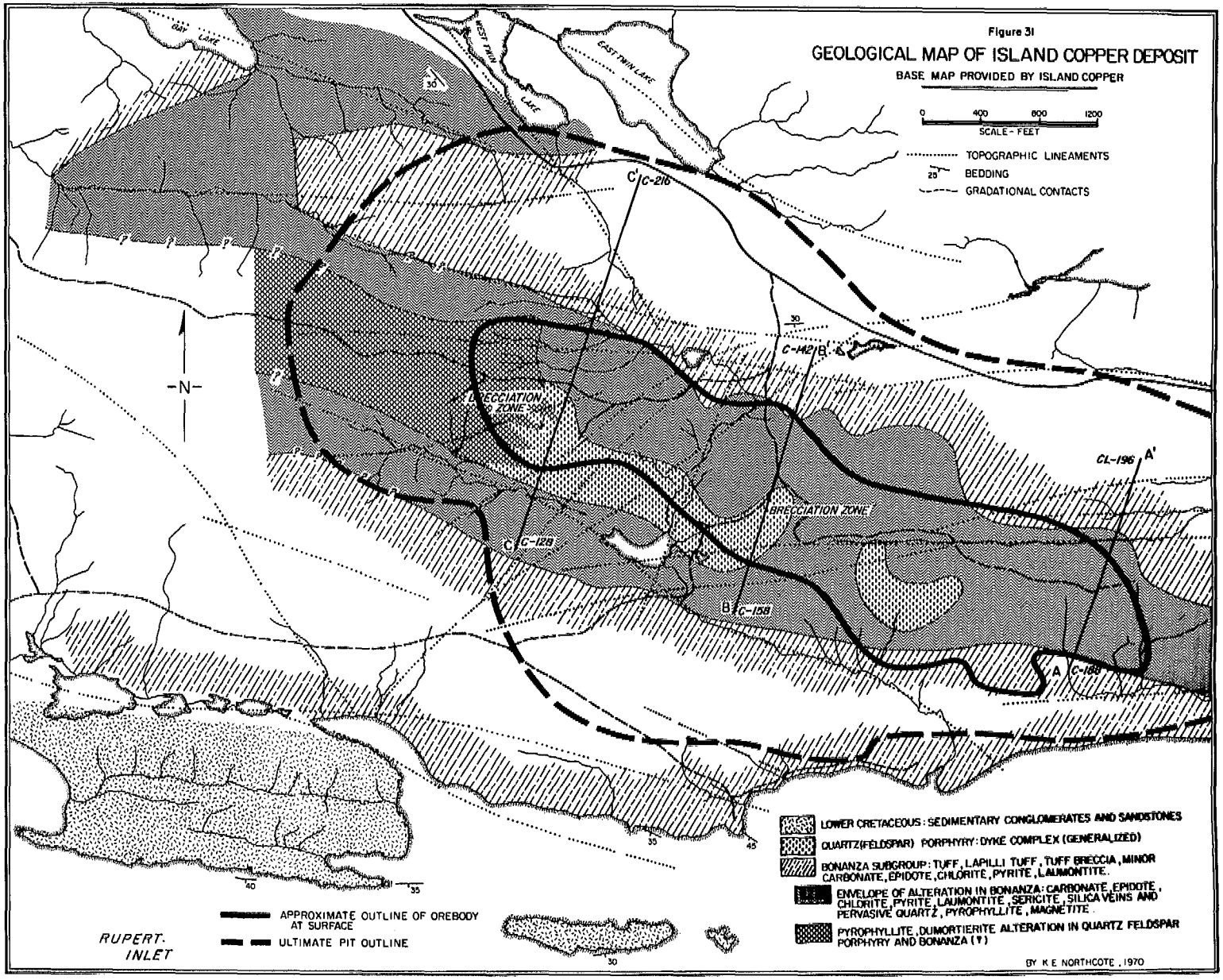
Structures:	the outline of the intrusive system, breccia pipes, a region of sulfide-only stockwork, and the major premineralization strike-slip faults (Figure 8a).
Alteration:	the zones of potassic, phyllic, argillic, and propylitic alteration (Figure 8b).
Geochemistry and Geophysics:	contours showing copper concentration obtained from induced polarization (IP), and contours for magnetic anomalies (Figure 8c).

(The limit of sulfide mineralization is shown in all the data maps, both as data and as a reference contour.)

It is interesting to see how the inference network combined these data to form the final results. For example, the data in Figure 8a were combined to produce a favorability map based on structural information alone (corresponding to Space FSM1 in Figure 6). This 128-by-128 array of certainty values is shown as an image in Figure 9a. In these favorability images black corresponds to a certainty of -5, white to a certainty of +5--with intermediate gray values corresponding to certainties as shown in the legend. Thus, based on structural information alone, the most favorable sites are seen to be near the fault intersections, the least favorable sites in the breccia pipes or in the region of sulfide-only stockwork.

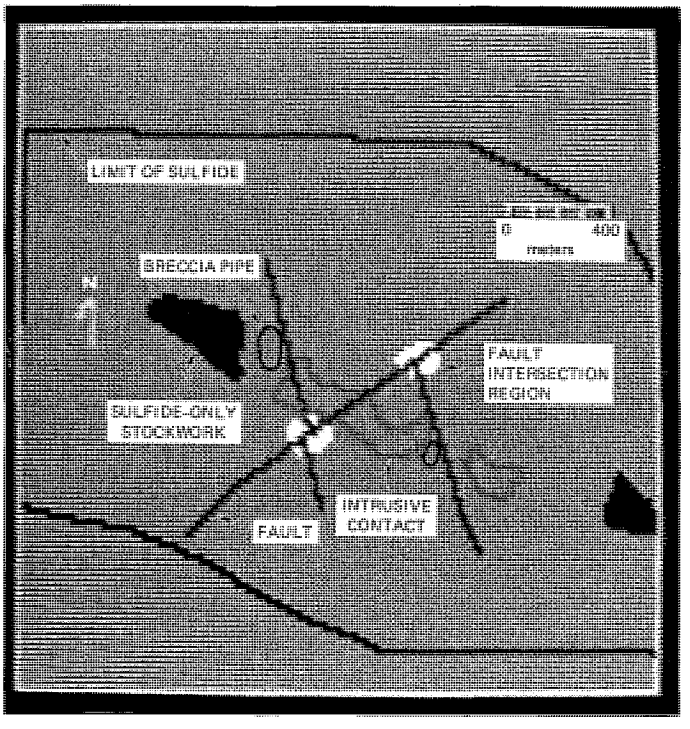
Figure 9b shows how the data on lithology and alteration were combined with the IP data (corresponding to Space FLA in Figure 6). Based on these data alone, the most favorable sites are near the boundary of the potassic and the phyllic zones, the least favorable in the regions of high pyrite.

These results on structures and alteration are combined with the data on magnetic anomalies and the intrusive contact to produce the results portrayed in Figure 9c (see Space OFE in Figure 6). The

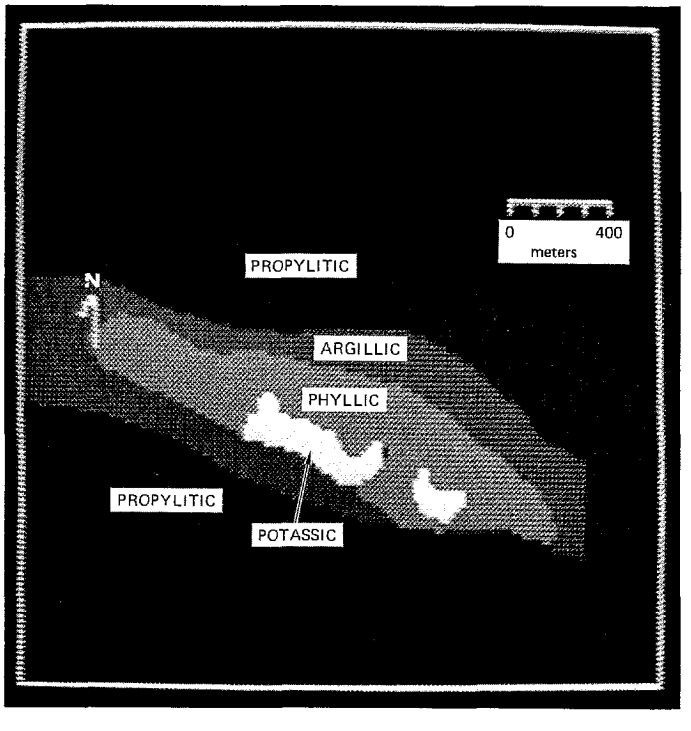


SOURCE: K. E. Northcote, "Rupert Inlet — Cape Scott Map Area," Geology, Exploration and Mining, British Columbia Department of Mines and Petroleum Resources, pp. 254–272, 1970.

FIGURE 7 GEOLOGICAL MAP OF ISLAND COPPER

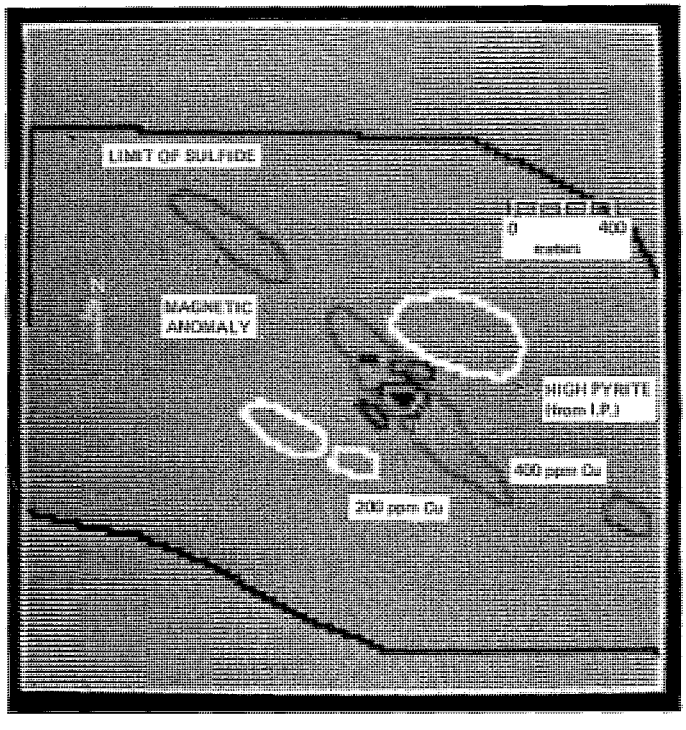


(a) STRUCTURES



(b) ALTERATION

FIGURE 8 DATA FOR ISLAND COPPER



(c) GEOCHEMISTRY AND GEOPHYSICS

FIGURE 8 DATA FOR ISLAND COPPER (Concluded)

addition of data from geochemical soil sampling produces the final favorability map, shown in Figure 10a. In addition to indicating the most favorable drilling sites, this map gives some idea of the expected size of the ore body. Figure 10b shows the outline of the ore body superimposed on the final map. Clearly, the size of the ore body has been underestimated, but the most favorable drilling sites are in high-grade ore prospects.

5.3.2 Bell Copper

A basic geologic map of Bell Copper, showing the intrusive system, major faults, and the outline of the ore body, is shown in Figure 11 (Carson et al., 1976; additional information on Bell Copper is given in Newman, 1969, and Wilson and Kesler, 1978). The data on Bell Copper are not as extensive as the data on Island Copper, but are still sufficient to allow use of the program. The available data were sampled at a 25-meter sampling interval to produce the following 128-by-128 data arrays:

Structures: the outline of the intrusive system, breccia pipes, and the major premineralization strike-slip faults (Figure 12a).

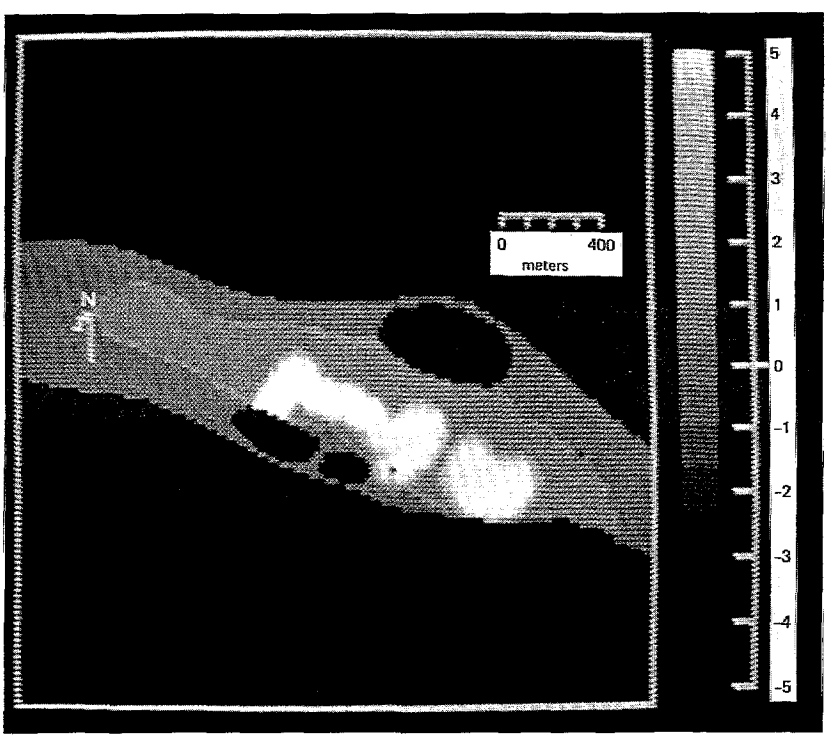
Alteration: the zones of potassic, phyllic, argillic and propylitic alteration (Figure 12b).

The resulting favorability map appears in Figure 13a, which is reproduced in Figure 13b with the ore body superimposed. As with Island Copper, the most favorable drilling sites are in high-grade ore.

5.3.3 Gibraltar

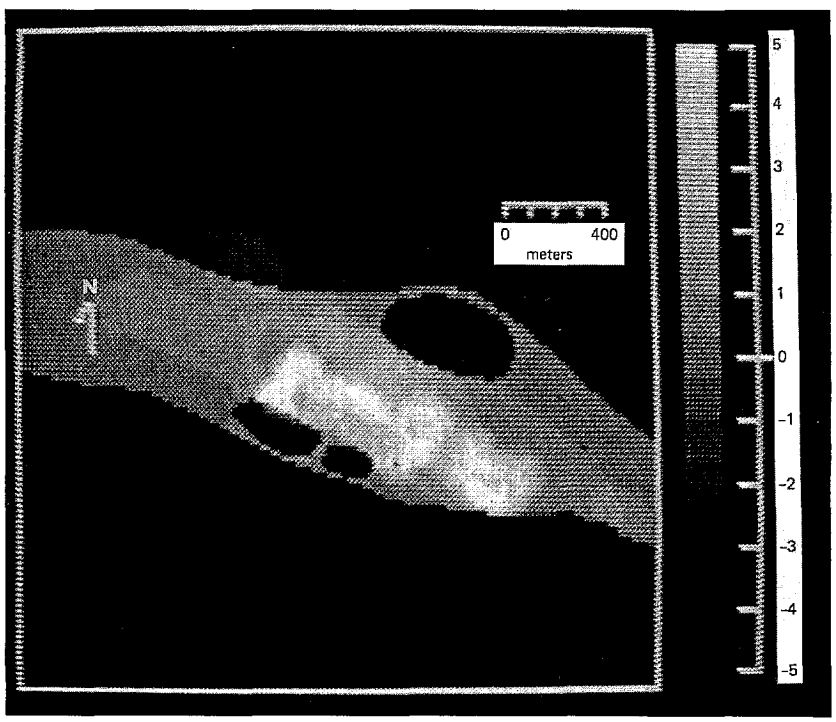
The Gibraltar Mines include four ore bodies known as Gibraltar East, Gibraltar West, Pollyanna, and Granite Lake. A basic geologic map of the area, showing the intrusive system and the four pit outlines, is depicted in Figure 14 (Drummond et al., 1973; additional information on Gibraltar is contained Brown, 1966, Simson, 1969, Rotherham et al.,

(a) FAVORABLE STRUCTURES



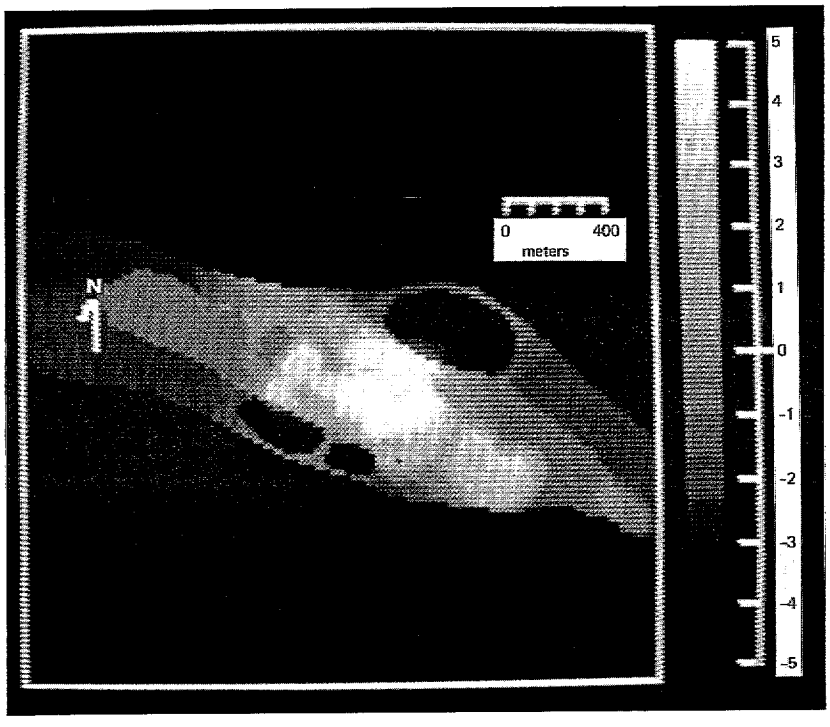
(b) FAVORABLE LITHOLOGY AND ALTERATION

FIGURE 9 INTERMEDIATE RESULTS FOR ISLAND COPPER

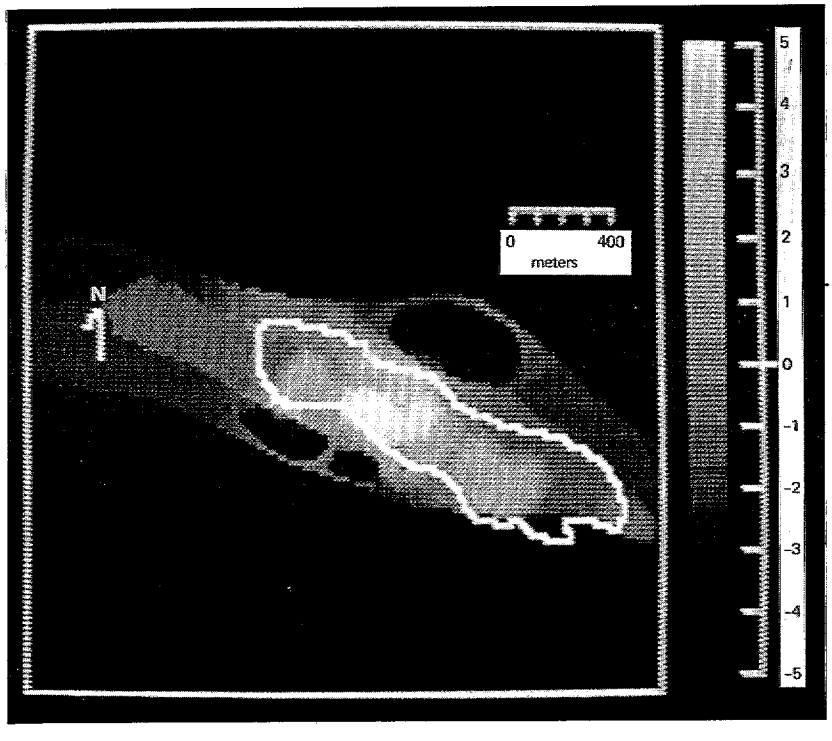


(c) OTHER FAVORABLE EVIDENCE

FIGURE 9 INTERMEDIATE RESULTS FOR ISLAND COPPER (Concluded)

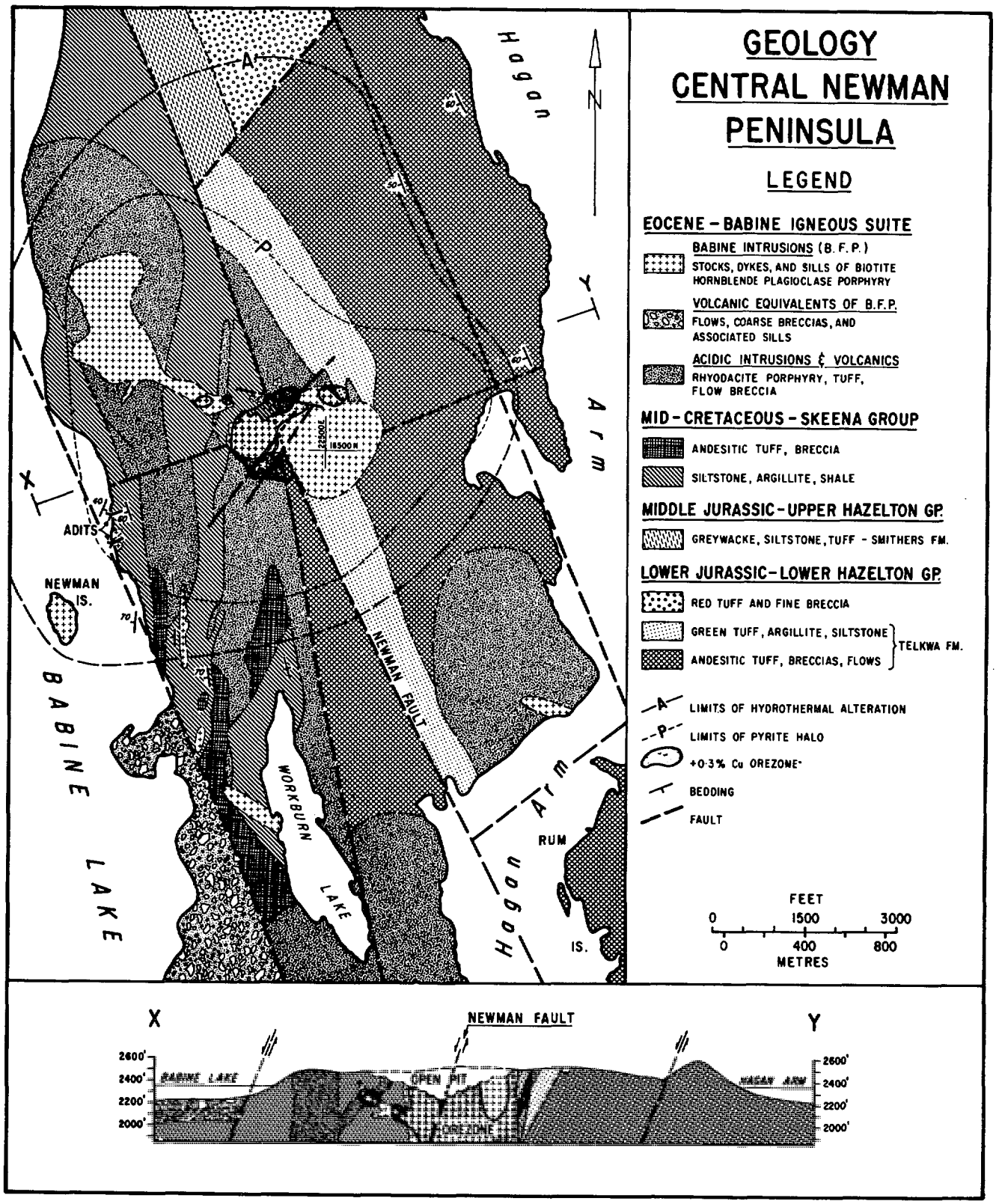


(a) FINAL FAVORABILITY MAP



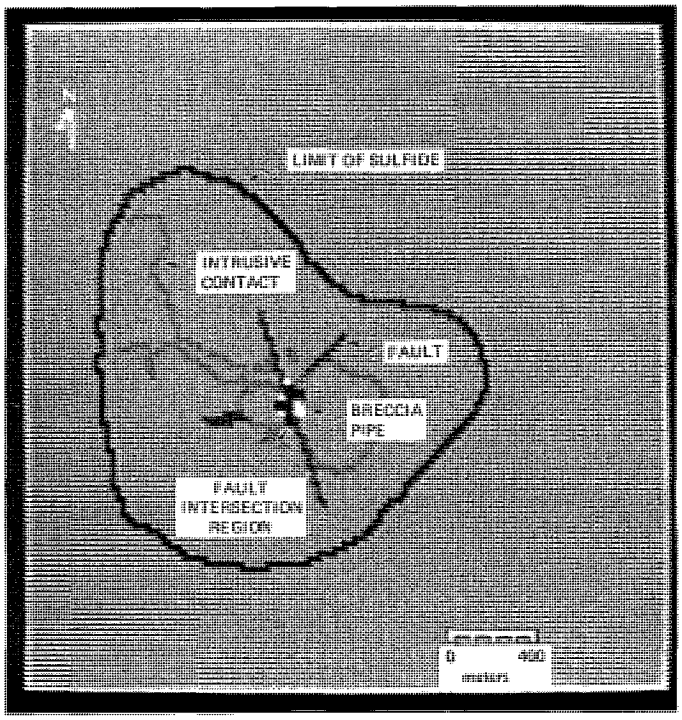
(b) OUTLINE OF ORE BODY

FIGURE 10 FINAL RESULTS FOR ISLAND COPPER

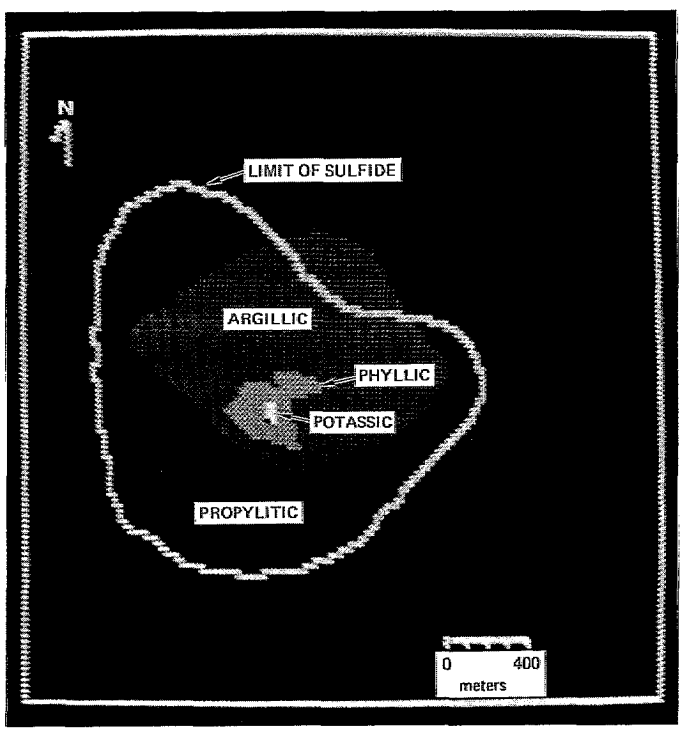


SOURCE: D. J. T. Carson et al., "Geology, Geochemistry and Genesis of a Supergene-Enriched, Biotized Porphyry Copper Deposit with a Superimposed Phyllic Zone," Porphyry Deposits of the Canadian Cordillera, Canadian Institute of Mining and Metallurgy, Special Volume 15, A. S. Brown, ed., pp. 245-263, 1976.

FIGURE 11 GEOLOGICAL MAP OF BELL COPPER

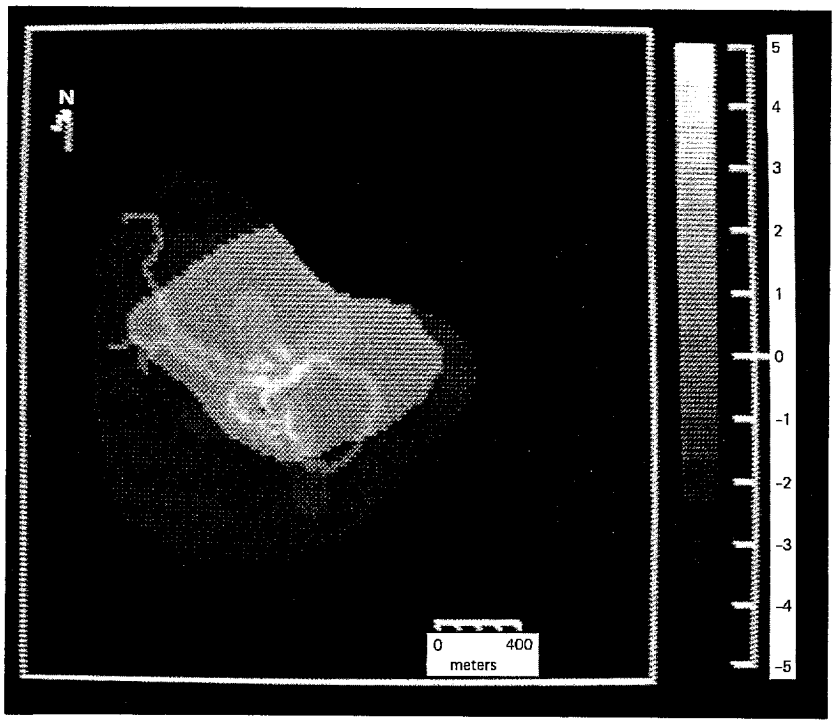


(a) STRUCTURES

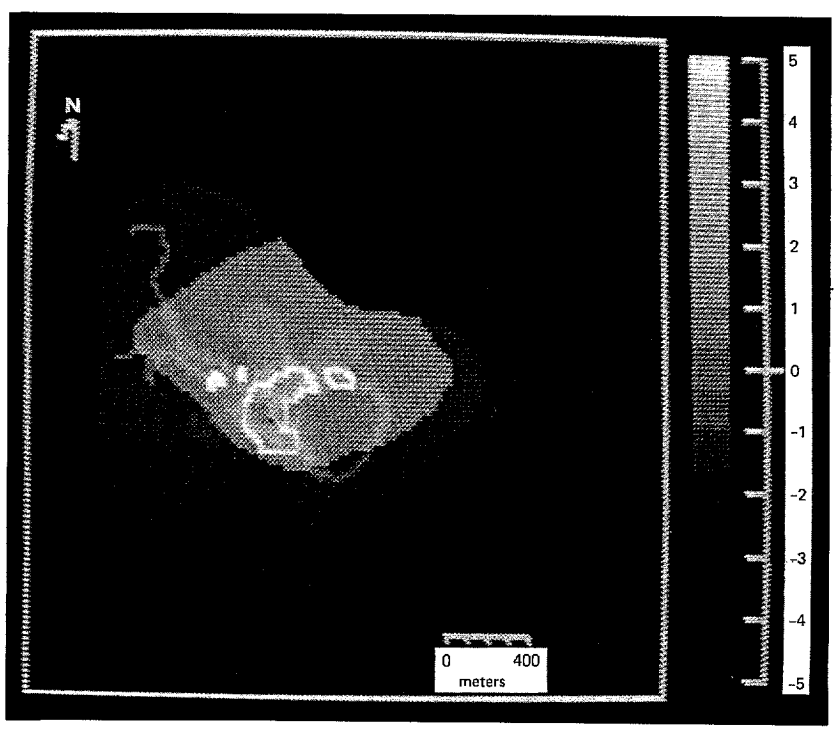


(b) ALTERATION

FIGURE 12 DATA FOR BELL COPPER



(a) FINAL FAVORABILITY MAP



(b) OUTLINE OF ORE BODY

FIGURE 13 FINAL RESULTS FOR BELL COPPER

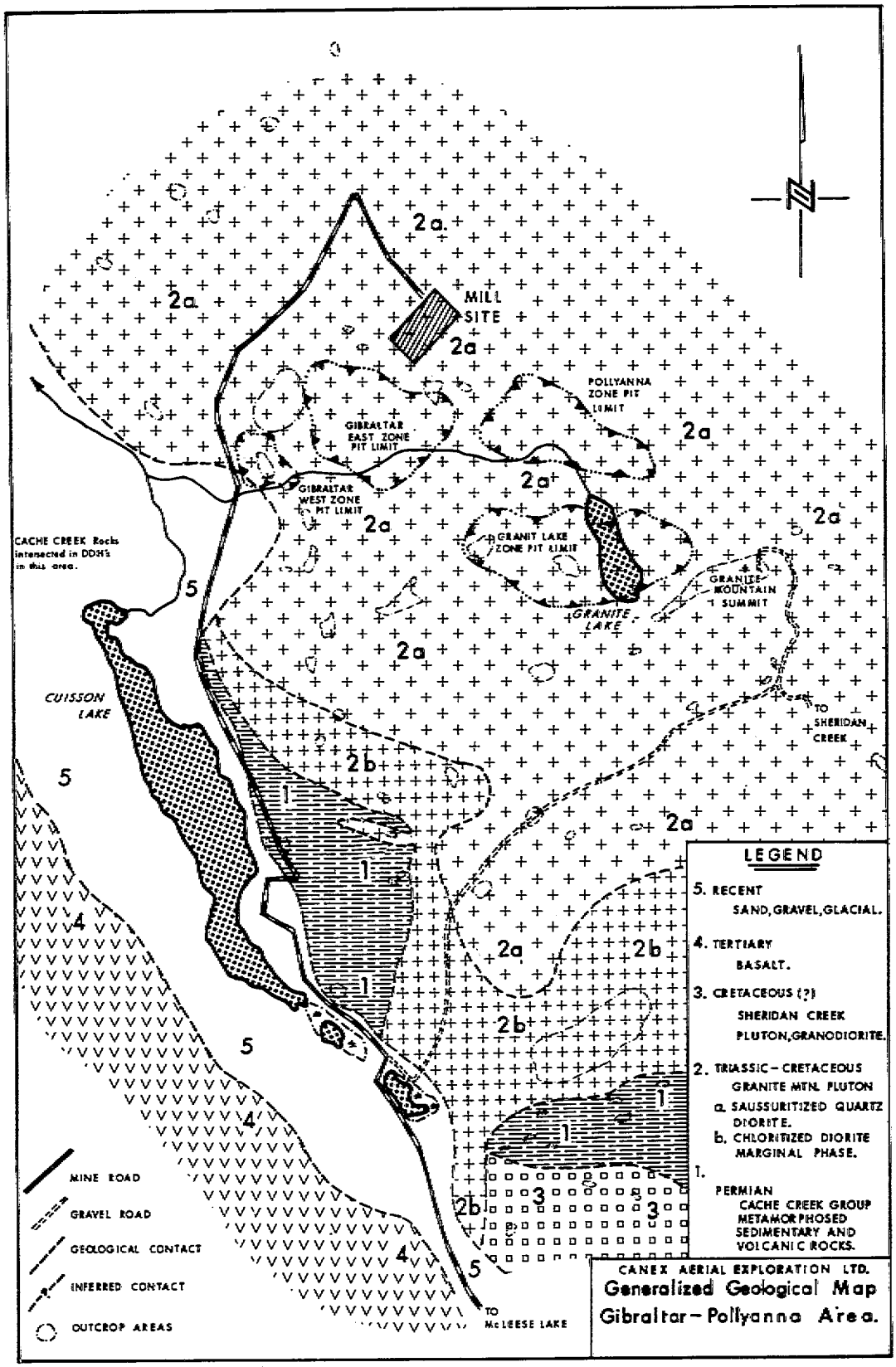
1972, and Cannon et al., 1972). The available data for Gibraltar were sampled at a 70-meter sampling interval to produce the following 128-by-128 data arrays:

Structures: the outline of the intrusive system, and the major premineralization strike-slip faults (Figure 15a).

Alteration: the zones of potassic, phyllic and propylitic alteration (Figure 15b).

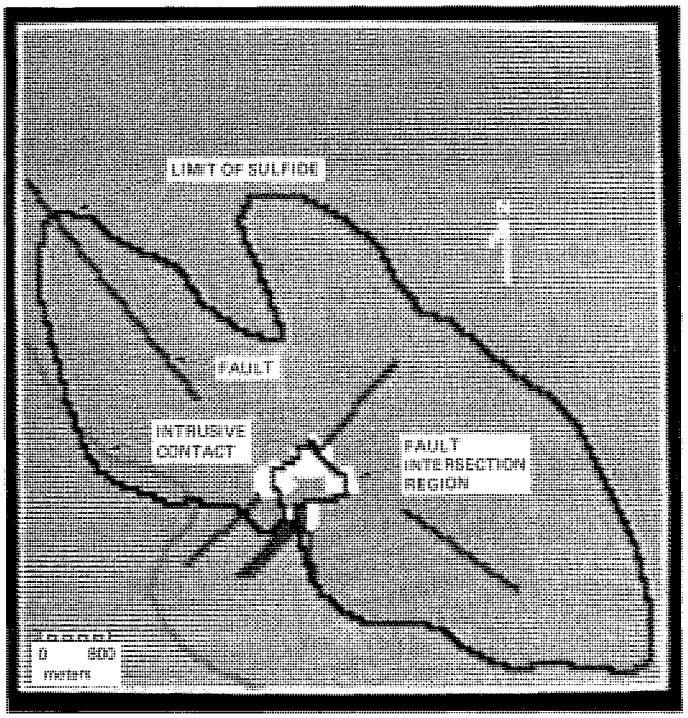
Geophysics: the high-pyrite zone, as obtained from induced polarization (Figure 15c).

The resulting favorability map is shown in Figure 16a, which is reproduced in Figure 16b with the ore body superimposed. The results are clearly not quite as favorable as those for Island Copper and Bell Copper. A significant portion of the most favorably scored points occurred in the potassic zone, which is actually barren. Based on the available information, however, a decision to drill at one of these points might well have been geologically sound, even though the ore body would have been missed. Thus, as with its other applications, Prospector should not be expected to be infallible--but its purpose will be achieved if it can indeed enable a significant increase in the probability of success.

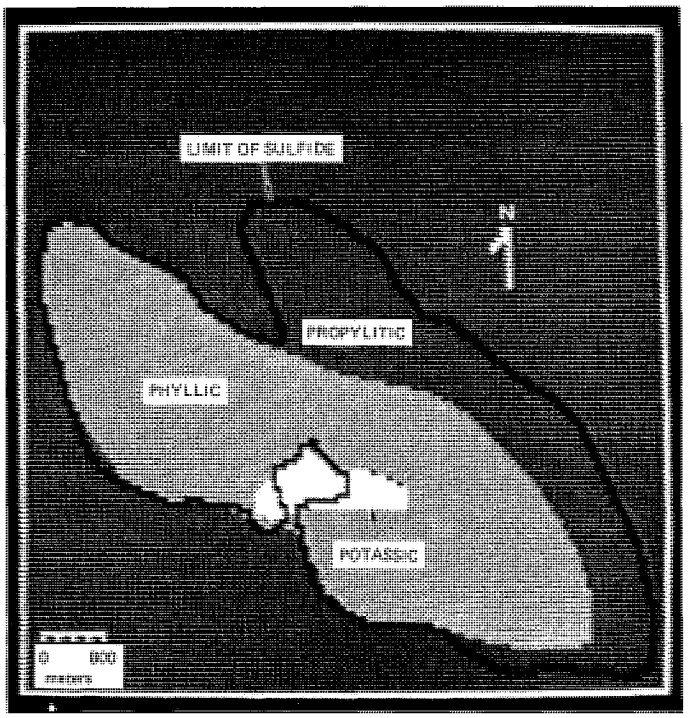


SOURCE: A. J. Drummond et al., "The Interrelationship of Regional Metamorphism, Hydrothermal Alteration and Mineralization at the Gibraltar Mines Copper Deposit in B. C.," Bulletin of the CIM, pp. 48-55, February 1973.

FIGURE 14 GEOLOGICAL MAP OF GIBRALTAR

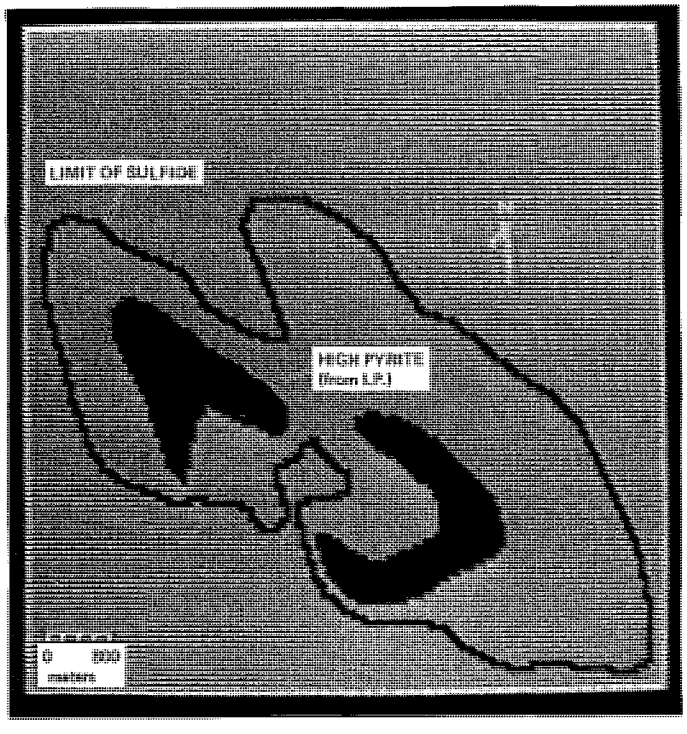


(a) STRUCTURES



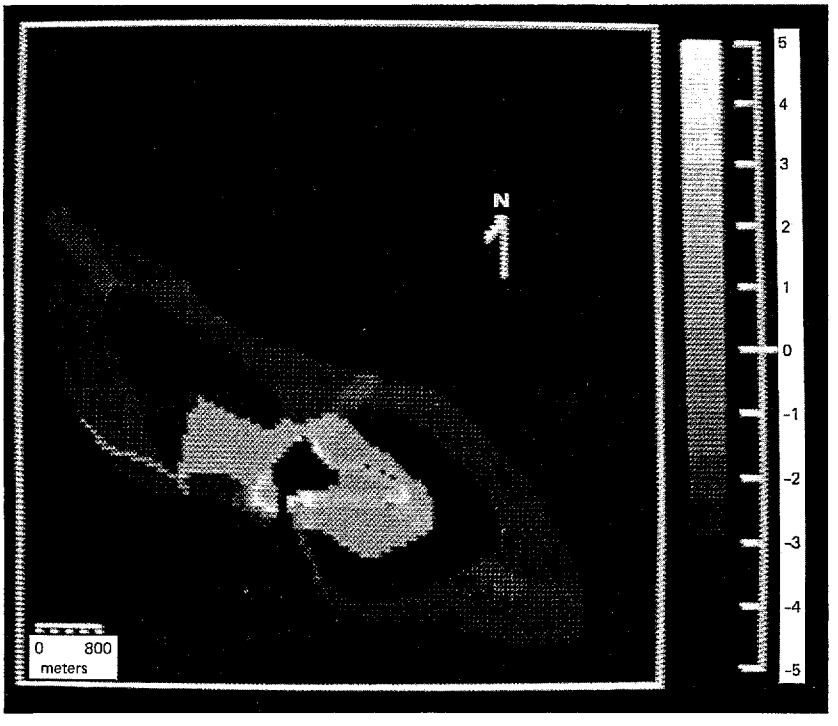
(b) ALTERATION

FIGURE 15 DATA FOR GIBRALTAR

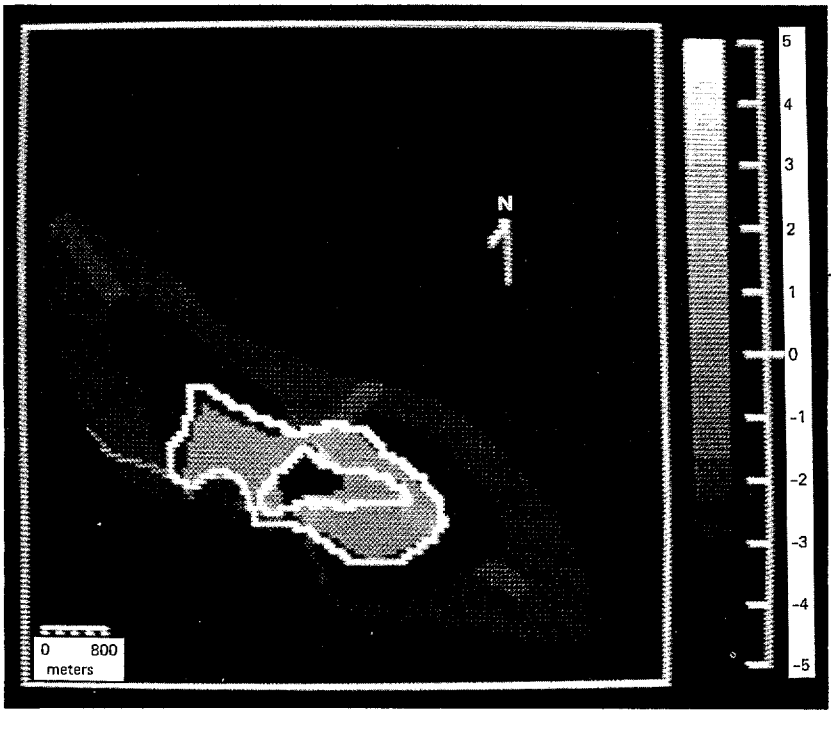


(c) GEOPHYSICS

FIGURE 15 DATA FOR GIBRALTAR (Concluded)



(a) FINAL FAVORABILITY MAP



(b) OUTLINE OF ORE BODY

FIGURE 16 FINAL RESULTS FOR GIBRALTAR

6 SYSTEM IMPROVEMENTS

During Prospector's development, much effort has been devoted to improving and extending the system code. In this section we describe the most significant of these improvements.

6.1 Quantitative Rules

Many of the rules in Prospector involve the value of some quantity, such as the size of the intrusive, the age of the host rock, or the percentage of some constituent. While many of the relevant questions can be formulated as true/false propositions, such as "The age of the host rock is post-Paleozoic," it is often convenient to know the actual value (or a reasonably certain interval of values) for several reasons:

(1) If the strength of the rule varies with the value x, it is simpler to leave one rule with a single likelihood ratio L(x) than to split the domain of x into intervals and to have separate spaces and rules for each interval.

(2) Even if several rules must be used because they go to different consequent spaces, all these rules can be used as soon as the user supplies the value of x.

(3) Use of a quantitative rule allows a clear distinction to be made between the numerical value of a quantity and the user's certainty regarding that value.

In our last annual report (Duda et al., 1977) we derived some basic equations for such quantitative rules. They can be summarized as follows. Given values for x and L(x), one can compute the posterior odds O(H|x) from the prior odds O(H) by

$$O(H|x) = L(x) * O(H).$$

From the odds one can immediately compute the posterior probability P(H|x). The problem arises when the value of x is not known exactly. Let E' denote the observed evidence and let p(x|E') be the resulting

probability density function for x. It follows from the law of total probability that

$$P(H|E') = \int_{-\infty}^{\infty} P(H,x|E')\, dx$$
$$= \int_{-\infty}^{\infty} P(H|x,E')\, p(x|E')\, dx \ .$$

If E' did not add any information about H that is not supplied by x, we could simplify this equation by writing $P(H|x,E') = P(H|x)$. With this in mind, we define the theoretical posterior probability $P_t(H|E')$ by

$$P_t(H|E') = \int_{-\infty}^{\infty} P(H|x)\, p(x|E')\, dx$$
$$= \int_{-\infty}^{\infty} \frac{L(x)\ O(H)}{1 + L(x)\ O(H)}\ p(x|E')\, dx \ .$$

This is essentially the main result given in our last report. Unfortunately, two problems limit its immediate usefulness:

(1) Consistency. If nothing at all is known about x, $p(x|E')$ must be the expert-supplied prior density $p(x)$. However, there is no guarantee that substitution of $p(x)$ for $p(x|E')$ will yield the expert-supplied prior probability $P_t(H|E') = P(H)$.

(2) Certainty. If the user has partial knowledge of x, he or she is asked to provide it in the form of a function, $p(x|E')$. This is even more difficult than asking for a probability $P(E|E')$ in the propositional case.

In the propositional case these problems were solved by replacing the linear relation between $P(H|E')$ and $P(E|E')$ with a piecewise linear relation, and by communicating in terms of certainties rather than probabilities. We investigated several generalizations of these procedures without finding an exactly analogous extension. However, the following procedure does solve the problems mentioned above, and has consequently been implemented in the system.

Let L_{min} denote the minimum value of $L(x)$ and let L_{max} denote the maximum value. Corresponding to these extreme values are minimum and maximum values of $P(H|x)$, which we denote by P_{min} and P_{max}, respectively. Let P_0 be the value of $P_t(H|E')$ obtained when the prior density $p(x)$ is substituted for $p(x|E')$. (It is easy to show that P_0 and $P(H)$ lie between P_{min} and P_{max}.) Then the consistency problem is solved by arbitrarily defining $P(H|E')$ as the following piecewise linear function of $P_t = P_t(H|E')$:

$$P(H|E') = \begin{cases} P_{min} + \dfrac{P(H) - P_{min}}{P_0 - P_{min}} (P_t - P_{min}) & \text{if } P_t < P_0 \\ P_0 + \dfrac{P_{max} - P(H)}{P_{max} - P_0} (P_t - P_0) & \text{otherwise.} \end{cases}$$

This is analogous to the piecewise linear interpolation used in the propositional case. It yields the correct values for $P(H|E')$ at the three points $P_t = P_{min}$, P_0, and P_{max}, and elsewhere interpolates between those values. The main defect in this procedure is that if the user knows the value of x, so that $P_t(H|E') = P(H|x)$, and if the expert is not consistent, so that $P(H) \neq P_0$, then, in general, $P(H|E')$ will not be equal to $P(H|x)$. However, if the expert is consistent, then $P(H|E') = P_t(H|E')$, and no approximation is involved.

The problem of communicating in terms of certainties is more difficult to solve, and stems from the fact that certainties are not cumulative. To see this, let $F(x|E')$ be the cumulative distribution function for x,

$$F(x|E') = Pr\{X<x|E'\} = \int_{-\infty}^{x} p(u|E')\, du \ .$$

If $F(x)$ is the prior distribution function, the corresponding certainty can be computed by the formula in Section 2.4. For convenience, we eliminate the factor of 5 and define the certainty factor CF (on a -1 to 1 scale) by

$$CF(x|E') = \begin{cases} \dfrac{F(x|E') - F(x)}{1 - F(x)} & \text{if } F(x|E') > F(x) \\[2ex] \dfrac{F(x|E') - F(x)}{F(x)} & \text{otherwise.} \end{cases}$$

While both $F(x)$ and $F(x|E')$ are monotonically nondecreasing functions of x, this is not true of $CF(x|E')$. In particular, CF is zero whenever $F(x|E') = F(x)$, and changes algebraic sign each time these two curves cross one another. Thus, certainties do not have the additive properties of probabilities. In particular, while the probability $F(a,b|E')$ that x is between a and b is given by $F(b|E') - F(a|E')$, there is no necessary relation between the certainty that x is between a and b and the two certainties $CF(a|E')$ and $CF(b|E')$.

Our basic problem is to give the user a way of saying something about $p(x|E')$ that will be consistent with the expert's prior density $p(x)$. In particular, when the user has no opinion about x we want to obtain $p(x|E') = p(x)$. We have "solved" this problem by limiting the user to specifying a single interval (a,b) in which he or she thinks x lies, and by allowing him or her to assign a certainty value to that estimate. Now, by definition, the certainty that x lies in (a,b) is given by

$$CF(a,b|E') = \begin{cases} \dfrac{F(a,b|E') - F(a,b)}{1 - F(a,b)} & \text{if } F(a,b|E') > F(a,b) \\[2ex] \dfrac{F(a,b|E') - F(a,b)}{F(a,b)} & \text{otherwise.} \end{cases}$$

Thus, given the user's certainty and the expert's prior density, we can solve for the posterior probability and obtain

$$F(a,b|E') = \begin{cases} F(a,b) + (1-F(a,b))CF & \text{if } CF > 0 \\ F(a,b)\ (1 + CF) & \text{otherwise.} \end{cases}$$

Of course, $F(a,b|E')$ does not give us the density function $p(x|E')$; it merely constrains the area from a to b to be $F(a,b|E')$. To obtain the function $p(x|E')$, we assume that it is proportional to the prior density $p(x)$, and we select the proportionality constants so that the area constraint is satisfied. To be more specific, we assume that

$$p(x|E') = \begin{cases} A\ p(x) & a < x < b \\ B\ p(x) & \text{otherwise.} \end{cases}$$

The resulting values for A and B are given by

$$A = \frac{F(a,b|E')}{F(b) - F(a)}$$

and

$$B = \frac{1 - F(a,b|E')}{1 - [F(b) - F(a)]} .$$

This procedure is admittedly ad hoc. It produces a posterior density having peculiar discontinuities at a and b. Still, it does offer several advantages. Its implementation is straightforward, requiring nothing more of the user than an interval and a certainty value. For the propositional case in which x can assume only one of two values, it reduces to our standard procedure. Finally, if the interval is small, $p(x|E')$ will approach an impulse function, as it should, and, if the interval is large, $p(x|E')$ will approach $p(x)$, as desired.

6.2 Spaces with Variables

In computer science terms, Prospector can be viewed as an instance of a general class of systems known as production systems (Newell and Simon, 1972; Hayes-Roth et al., 1978). A classical production system consists of a data base of assertions, a set of production rules of the form <antecedent> --> <consequent>, and an interpreter that matches the antecedents in the rules against the assertions in the data base to determine which rules are applicable. When a rule is applied, it is typical for the consequent part of the rule to modify the assertions in the data base, which in turn may cause other rules to be applicable.

In a general production system the antecedent part of a rule contains variables that can be matched against more than one assertion in the data base. Thus, the same rule can be used many times, with different bindings for its variables in each application. This is the source of a production system's computational power, which is equivalent to that of a Turing machine. The price for this generality, however, is that much time must be spent matching antecedents against the data base searching for possibly applicable rules.

In Prospector, the problem of search for a match arises only when the user volunteers information. The linking of rules is done by the Network Creator when the Inference Network is built, and the consequences of "applying" a rule can be propagated immediately through this existing network. But the price for this efficiency is that different instances of what might be essentially the same general rule must be explicitly present in the Inference Net, which, under some circumstances, can lead to replication and large networks.

The major place where this problem appears in Prospector is in the rules for identifying zones (see Section 3.2). Since an indefinite number of zones can occur in a given prospect, and since a large number of rules are devoted to identifying a zone, it would be very inefficient (as well as inelegant) to replicate these rules for each possible zone in advance.

As a first step toward providing Prospector with a more general mechanism for using rules that contain variables, we have implemented a limited procedure for using the rules repeatedly for zone identification. Each zone location on the prospect is assigned a name by the system, such as Zone-1, Zone-2, etc. The values for state variables (such as the posterior probabilities) for all spaces associated with the zone identification rules are stored on a property list. Thus, the values for Zone-i are stored as the i-th entries on the lists. If a space is an antecedent for rules going to other parts of the network, the posterior probability used is the maximum of the values on that list. Thus, each space is interpreted externally as affirming that the situation described exists in at least one zone of the prospect.

When the Questioning System first encounters a space in the set of zone identification spaces, it must take certain actions to preserve the distinctions in information about different zones. Our current strategy is to ask the user if there are any zones at all, and then to consider each zone in turn. Thus, a whole series of questions concerning Zone-1 is followed by a series of questions concerning Zone-2, and this process is continued until no zones remain.

This is a rather inflexible procedure of limited generality. However, it does allow repeated use of a large set of rules and it supports such standard features as the ability to change answers and to obtain explanations. We anticipate that future versions of Prospector will incorporate more general procedures for the inclusion of rules containing variables.

6.3 Graphic Input

In Section 5 we described an application of Prospector that required map input. This necessitated the development of simple but useful facilities for acquiring, editing, storing, processing and displaying graphical data. Many of the basic procedures had already been programmed for various image-processing projects at SRI, and merely

had to be called from Prospector. However, these procedures were written in SAIL, an ALGOL-like language particularly suited to numerical processing, whereas Prospector is written in INTERLISP, which is a very different, interactive list-processing language. Thus, special methods had to be developed to interface these procedures with Prospector.

Figure 17 shows the final program organization. When the program first starts, the user is talking to the INTERLISP side of the system. If a new map is to be acquired, that command is passed to the SAIL programs, which then assume control. Working through a simple executive program on the SAIL side, the user can enter registration data, digitize contours, display the results, and write the results on disk files (the map data base). The results can be stored either as lists of coordinate points or as sampled 128-by-128 arrays, either of which can be subsequently read and displayed.

The SAIL routines also compute the favorability arrays from the raw map data (see Section 5.2). This is done in response to a command from the LISP side that includes the name of the source map and the values for the distance parameters a and b. The resulting favorability array is automatically stored in the map data base, and is accessed from there by LISP.

Thus, when the LISP program is used to perform the inference network calculations, all input is obtained from the map data base. The final output, which is a set of digitized maps for the various spaces of interest, is stored back in the map data base, where it is available for subsequent inspection.

6.4 The Network Compiler

6.4.1 The Need for a Compiler.

The Questioning System and the Inference Procedures together can be thought of as a single subsystem that we shall call the Net Interpreter (see Figure 1). In its normal interactive mode of operation the Net Interpreter has many tasks to perform. It will search portions

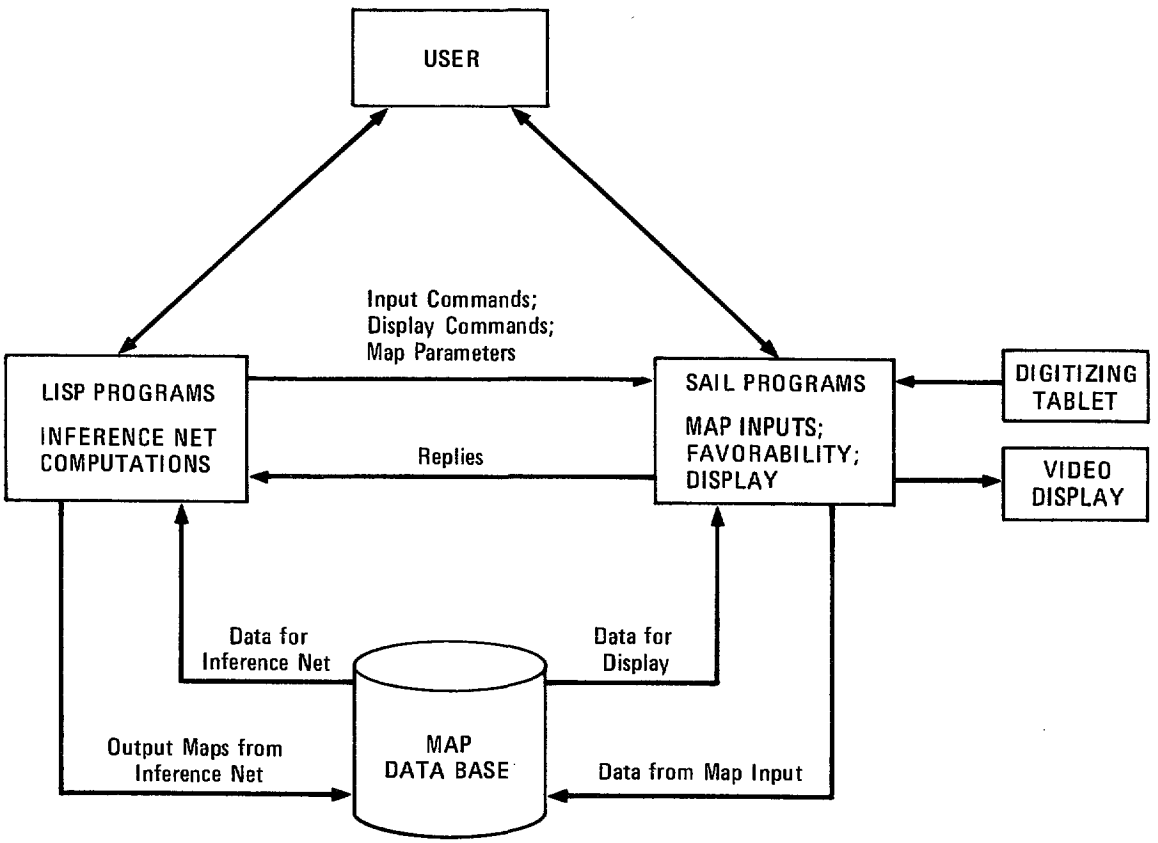


FIGURE 17 PROGRAM ORGANIZATION FOR GRAPHICAL DATA PROCESSING

of the net to find an appropriate question to ask next. Given an answer from the user, it will check that answer for consistency, then propagate the results through the net until there are no more probability changes. It will then repeat the process until all relevant questions have been asked.

Efficiency requirements in this interactive mode are not very severe. Each cycle of finding a question and propagating results could take several CP (central processor) seconds without causing undue frustration on the part of the user (at least on an unloaded system). A complete session could thus consume several minutes of CP time and still be acceptable.

In contrast to the interactive mode, there are several applications of the inference net which would involve many runs of the same inference net over an incrementally changing set of answers. In dealing with data from geologic maps, we typically divide the maps into a grid of 128-by-128 sample points and propagate probabilities through the same inference net for each sample point. In performing sensitivity analysis, starting from an initial set of values for the askable spaces, we could plot the change in favorability of a hypothesis as a result of changing certainties for many combinations of evidence spaces. Both these applications involve a great many runs of the inference net. If each run were to take several minutes of CP time for each data point,* computational costs would be prohibitive.

Thus the question of efficient evaluation--in terms of CP time--becomes important. We have developed a technique for compiling the action of the Net Interpreter upon a given model into a sequence of machine instructions that runs several orders of magnitude faster than the Net Interpreter. In this section we discuss the inference net compilation procedure, and give some results from a practical compiler.

---

* A data point is a set of answers to all askable spaces.

6.4.2 Using the Compiled Net

A compiler for inference nets converts the action of the Net Interpreter, for a given model, into a sequence of machine instructions. The resulting code, called a compiled inference net, does not mimic the action of the Net Interpreter in every way. A key assumption is that the compiled inference net is run noninteractively, that is, the answer to every askable space is made available at the beginning of the run.

Since we will want to run the compiled inference net many times on different data points, these data points must be made available to the compiled net at the start of the run. The particular problem to which the compiled net is applied will dictate how these data points are generated. For problems involving geologic maps, the digitized maps will be used to supply an array of points for some of the askable spaces. For sensitivity analysis, a user-defined function may generate a set of incrementally varying answers for some spaces.

The output of a single run of the compiled inference net is a set of values giving the certainties for the unaskable spaces in the net. Running the compiled net many times for a set of input data points will yield a set of values for each unaskable space, one for each input data point. If generated from map data, these results could be displayed on a graphics device to show a favorability map for any hypothesis in the model. For sensitivity analysis, the results for any hypothesis could be graphed to show the variation of the certainty in the hypothesis as a function of the certainty in the evidence. The compiled net is indifferent to the source of the input or the destination of the output. It simply takes a set of input data points for askable spaces and produces a corresponding set of output data points for unaskable spaces; the problem is to do it very quickly.

Because it is run in a noninteractive mode, the compiled inference net does not do run-time consistency checking on the input data it received. Thus, inconsistent answers to askable spaces, which would have been caught by the Net Interpreter, will not be noticed by the compiled net.

6.4.3 Compilation Procedure

The overhead associated with the Net Interpreter comes from the control strategy, from accessing required parameters, and from using LISP as the interpreter language. For each of these areas we describe how the compiler bypasses the efficiency bottlenecks of the Net Interpreter.

6.4.3.1 Control Strategy

When the Net Interpreter is considering a space in the inference net, there is a variety of tasks it performs before it calculates the posterior probability of that space. Consider the partial inference net below, in which the goal of the Interpreter is to establish H1:

```
                    S
                    ^
                    |
              --------------- . . . .
              |       |
C <- - - - H1         H2  . . . .
           ^
           |
      ---------------
      |     |     |
      E1    E2    E3
```

First, any context space C must be found, checked, and established if no previous attempt to establish it had been made. Then, H1 must be checked for askability. All rules leading to H1 must be fetched and an order for asking them computed.* Finally, the results are collected and used to update H1; then any consequents of H1 must be found and probabilities propagated up through the net. This procedure may be repeated many times and H1 updated each time, until all askable spaces below H1 have been considered.

---

* The algorithm used to compute an order for asking antecedent rules is the J* algorithm described in Section 4.5 of our last annual report (Duda et al., 1977).

We are going to consider compiling the control strategy for Space H1 under the following condition: All present probabilities for its antecedent spaces are available at the start of the calculation. This condition essentially means that the compiled code must calculate E1, E2, and E3 before it calculates H1. If this condition is satisfied, then the compiled code need calculate the posterior probability of H1 only once. The code generated by the compiler looks like this:

```
        set P(H1|E') to P(H1)                    * initialize P(H1|E')
        if NOTCONTEXT(C) goto ENDH1              * check for valid context
        { code to calculate L' of E1,
            save in temporary storage }

        { code to calculate L' of E2,
            multiply to L' by E1 and resave }

        { code to calculate L' of E3,
            multiply by L's of E1 and E2 and resave }

        compute P(H1|E') from L's and save in temporary storage

ENDH1

        { code to calculate P(H2|E') }
                .
                .
                .

        compute P(S|E') and save
```

All control strategy decisions which can be made at compile time are encoded by the sequence of function and control statements. With a simple control strategy the only run-time decision is for valid contexts.

When one considers the inference net as a whole, several additional constraints on compilation arise.

(1) If context space C itself is not directly askable, it should have been calculated before H1. In general, context arcs induce a partial ordering on the sequence of probability calculations. The compiler follows this partial ordering, always producing compiled code for a space after its context spaces.

(2) An inference net will not always be a tree, i.e., some space may appear in more than one path to the top-level node. The compiler produces code that calculates the posterior probability of such a space only once and saves the result in temporary storage to be used again.

The net result of compiling the inference net control strategy into a sequence of operations is that almost all the overhead associated with the control strategy is removed. The sequence of operations is such that the posterior probability of each space is calculated only once, after the probabilities of all its antecedents are computed. A single propagation sweep through the net, starting at the tip spaces and ending at the root space S, suffices to calculate all hypothesis probabilities.

6.4.3.2 Accessing Parameters

Currently all parameters are stored associatively in an appropriate place; e.g., the prior probability for a space is stored on the property list of that space under the label PRIOR. A routine can be invoked to find and return these parameters.

One can consider PRIOR as a variable associated with a space. Typically, compilers set aside specific storage cells for variable bindings so that compiled code can access the variable bindings directly rather than associatively. This results, for example, in a large efficiency increase in compiled LISP.

The inference net compiler sets aside storage locations for parameters needed in its calculations. Two storage cells are needed for each space--for prior and posterior probabilities. For each rule $P(H|E)$ and $P(H|\sim E)$ are stored in two cells. In addition, temporary storage for various intermediate results is needed. The compiled code knows the location of the parameter cells, and uses simple storage fetches and saves to access them.

One problem with compiling the network is that if the network is changed it must be recompiled. However, merely changing priors or likelihood ratios does not entail recompilation. An

initialization routine is used to preset storage cells to desired values before the compiled code is run.

6.4.3.3 LISP Overhead

Performing the various calculations needed to update a single space can consume considerable time if it is done by LISP code. Function calls and arithmetic operations are notoriously slow, even if the LISP code is compiled.* For this reason, all probability calculations in the compiled net are done in-line by appropriately generated machine code. This results in a significant saving in time needed to calculate the posterior value for a single hypothesis space.

All techniques discussed so far are combined into a single compile pass over an inference net, generating assembly code for the LISP assembler. The resultant procedure is callable from LISP and can be saved as an object module for loading by another language.

6.4.4 Timing Results

The compiler has been tested on Mr. Hollister's drilling-site-selection model for porphyry copper deposits. The version of this model which was tested contained 32 rules and 84 spaces; 25 of these spaces could take map input data. This is a medium-size model, similar in size to KNSD.

The resulting compiled code consisted of about 5000 machine words on a DEC PDP KL-10, broken down as follows:

initialization routine:	1675 words
inference net code:	2585 words
storage cells:	711 words
total:	4971 words

Running the compiled net for a single data point takes approximately 3

---

* Prospector is written in INTERLISP, which does not have a compiler that optimizes arithmetic operation calls.

milliseconds on the KL-10.* More than half of this time is spent in floating-point operations necessary to calculate the values of the effective likelihood ratios for rules; most of the remaining time is devoted to calculating posterior probabilities for logical nodes, and fetching and storing parameters and results. The context checking takes a negligible amount of time. This is in contrast to the inference net interpreter, where most of the overhead time is spent within the control strategy.

The compiled inference net is about four orders of magnitude faster than the net interpreter operating on the same model. A complete favorability map for the top-level hypothesis is produced in just under one minute of CP time (16,384 data points). This makes the use of geologic map input data eminently practical with currently available computers.

6.5 Savemode and Batchmode

There are various reasons for wanting to be able to keep a computer-readable record of the answers supplied by a user during a normal interactive run. One reason is that unexpected things can happen when a new model or a new system facility is being debugged, and it is convenient to be able to restore the system to its exact state before the difficulty was encountered. Another reason is that sometimes the user wants to terminate the run temporarily and then resume it later. A third reason is that sensitivity investigations can be made by systematically varying the answers from one saved standard set of answers. Three new commands (SAVEMODE, SAVE and NOSAVE) have been added to allow the user to save his or her answers on a disk file.

When Prospector starts it asks whether or not the user wants to read answers from a file. If so, the program leaves the interactive mode and runs like a batch program, returning to the interactive mode

---

*The actual time can vary with the number of askable spaces which actually do have map data, as well as the number of hypotheses for which results are outputted. Typically, for 15 inputted maps and 7 outputted hypotheses, a 10 to 20 percent increase in CP time can be expected.

only if the last answer in the file does not terminate the run. For sensitivity studies, the program can automatically make certain systematic changes to the answers it reads, such as decreasing the certainty for all answers having a positive certainty. The user invokes this mode by defining and supplying the name of a function (LISP subroutine) that specifies the desired modification. These simple features have proved quite valuable to us in reducing the chance of human error during repeated Prospector runs.

6.6 Model Implementation Aids

The process of transition from a geologist's verbal description of an ore deposit model to full implementation of the model in Prospector is described in Section 7. We have carefully designed Prospector to keep the encoded model and the general geological knowledge on files that are distinct from the program itself. This minimizes the need to make changes in Prospector per se when a new model is developed. Despite this modularity, the encoding of a new model is a time-consuming task, and so we have developed additional software tools to make this process simpler and less error prone.

Our chief tool is a simple keyword-based language for specifying the spaces and rules that comprise a model. A description of the syntax for that language is given in Section 7.2. The language has the advantages of being easy to understand, easy to edit with a conventional text editor, and sufficiently constrained to allow it to be read by the Network Builder and printed in a uniform format. In addition, a number of consistency checks are automatically made when a model file is read. These include (a) noting any spaces that are referenced but not defined, (b) noting syntactic errors in semantic network representations, (c) calculating prior probabilities for spaces that are logical combinations of other spaces and (d) noting any omissions or inconsistencies in probability assignments.

While this version of the Network Builder has proved to be quite useful, we are aware that even more powerful tools are needed for model

implementation. A major problem with using a text editor to make changes that affect network structure is that unintended changes can introduce major structural errors that may not be discovered until much later. Furthermore, during editing it is virtually mandatory to draw reference diagrams to keep a clear picture of the structural changes being made. Finally, one encounters errors during debugging that one would like to correct at the time, rather than leave the system, edit the files, load the edited files, and rerun to get to the point at which the problem was detected. We are currently developing a network editor that should make on-line structural editing subsequently more straightforward and efficient.

7 THE MODEL-BUILDING PROCESS: TECHNIQUES AND TOOLS

7.1 Introduction

In the development of models during this two-year contract period, one of the most significant results has been the evolution of a methodology for encoding exploration models in the network form used by Prospector. This methodology involves the following elements: interviewing techniques, principles for determining the overall structure of a model, programs for the interactive input and modification of models in the computer, and the use of performance analysis procedures as a diagnostic tool for revising a model. The methodology has been applied in various stages to the encoding of the PCDA and the KNSD models with encouraging results. Accordingly, there is good reason to believe that it can be usefully applied to the construction of any model for the Prospector system.

In this section we describe the encoding of a new model as a step-by-step procedure. At appropriate points below, for the purposes of illustration, we cite examples from the development of the KNSD and PCDA models.

It should be emphasized that the encoding of a new model, although now easier and more systematic than it was at the onset of the Prospector project, is still not a routine matter. It requires the collaboration of two types of people: one or more exploration geologists who are authorities on the type of ore deposit to be represented in the model, and one or more computer scientists who can translate the geologist's intensions into the computer representation used in the Prospector system. For brevity, we will refer to these two agents as DS, the Domain Specialist, and MI, the Model Implementor. In addition to interviewing DS and providing guidance in developing an acceptable overall structure for the model, MI must also be aware of potential

defects and/or deficiencies in representations of models, must detect their existence when present, and must cooperate with DS in correcting them.

To date the roles of DS and MI have been played, respectively, by exploration geologists having extensive field experience with the type of deposit to be modeled, and computer scientists knowledgeable about the use of Prospector and the model construction methodology. Both of these people currently have creative roles to play. In time, however, both roles may be fulfilled by a single person. It is precisely toward this end that our efforts to develop an effective model-encoding methodology are directed. Thus, we have developed procedures and tools to facilitate the more routine tasks of the MI, and to make more routine those which seemed previously to be more creative.

Although model encoding is not a tightly structured process, it does progress through several distinct phases. The rest of this section is devoted to explaining the activities that take place in the following six phases:

A. Initial preparation
   1. Familiarization of DS and MI with each other's terminologies
   2. Listing of known deposits to be covered by the model

B. Initial encoding of the model
   1. Development of inference network structure
   2. Choosing numeric values for parameters
   3. Writing of the elaboration/explanation/announce text
   4. Creation of the semantic-network representation for each space in the model

C. Installation of the model in Prospector
   1. Creation of computer file containing model definition
   2. First loadup
   3. Creation of model questionnaire

D. Initial debugging of the model
   1. Initial runs
   2. Implementation of initial modifications and corrections

E. Preliminary performance analysis
   1. Completion of the model questionnaire for known deposits
   2. Running of Prospector using questionnaire data
   3. Sensitivity analysis experiments

4. Analysis of experimental data for indications of model deficiencies

F. Subsequent revision of the model based on results of preliminary performance analysis
   1. Revision of the model
   2. Performance analysis of the revised model

7.2 Procedure for Model Construction

PHASE A: Initial preparation

Step A-1. Familiarization of DS and MI with each other's terminologies

Before the actual design of a model can commence, DS must become familiar with the the Prospector system and with the Prospector language for describing models, e.g., by reading Section 2 of this report. Similarly, it is useful for MI to become familiar with the general characteristics of the type of ore deposit to be modeled. In our experience to date, the initial meeting between DS and MI has commenced with an hour or two of informal discussion to establish a common language. At this time, DS becomes familiar with the following terms, among others:

1. Inference network
2. Space
3. Certainty of a space vs. probability of a space
4. Logical combination of spaces (i.e., AND, OR, and NOT)
5. Rules from one space to another
6. Prior probability of a space
7. Sufficiency strength of a rule (i.e, LS value)
8. Necessity strength of a rule (i.e., LN value)
9. Spaces requiring numerical input (i.e., L(x) spaces)
10. Context spaces for a given space
11. Zones
12. Semantic-network representation of a space

Step A-2. Listing of known deposits to be covered by the model

Typically, DS has in mind a number of particular, known deposits that serve as examples for the model to be developed. These may be of three types: known deposits that DS intends the model to match very well; known deposits that DS intends the model to match, but perhaps not as well; known deposits that should serve as "near miss" cases, i.e., cases that are partially favorable but lack certain characteristics that are critical to establishing overall favorability for the model.

Such a concrete list of known deposits facilitates the identification of critical factors for favorability that hold generally for the listed sites. It also makes more apparent the intersite differences that must be accounted for in the model.

PHASE B: Initial design of the model

Step B-1. Development of inference network structure

In this, the most creative and intellectually demanding step of the model-encoding process, DS enunciates his knowledge about the ore deposit type, and organizes it, with the aid of MI, into a hierarchical inference network structure of the sort shown in Figures 4 and 5. The inference network is roughly tree-shaped, with the "leaf" or "tip" nodes of the tree representing field observations and the other nodes representing interpretations and conclusions drawn therefrom. Hence, DS must identify the particular field observations that will serve as the foundation of the model (i.e., the questions that are asked of a user by the resulting Prospector implementation), and must specify how these factual inputs are combined into several levels of interpretations and conclusions.

We have found that a "top-down" development of the inference network gives quite satisfactory results. Using this approach, DS first identifies the principal general factors required to establish overall favorability. If it makes sense to ask a user directly about any of these factors, then they represent "askable" concepts, and serve

directly as evidence spaces. However, many of these factors are general hypotheses, such as the hypothesis that the petrotectonic setting is favorable for a particular model. These factors are "unaskable" concepts and serve as hypotheses to be established by evidence. Thus, DS must next identify the principal general factors needed to establish the favorability of each of the unaskable hypotheses. This process of progressive elaboration is continued until everything is related to field-observable evidence. Hence, each level of the inference network represents a more detailed expansion of the level immediately above it.

For example, the first-level spaces in KNSD are called "Favorable Conditions, Excluding Mineralization" and "Favorable Mineralization." Each of these first-level spaces itself represents a conclusion based on more detailed considerations. For example, the space "Favorable Conditions, Excluding Mineralization" is itself determined by four factors, represented by spaces called "Favorable Petrotectonic Setting," "Favorable Magmatic Association and Local Environment," "Favorable Ore-bearing Unit," and "Proximity to Ore."

In addition to these "top-to-bottom" considerations, our experience has suggested an effective "left-to-right" development as well in which the several first-level spaces represent a hierarchy of spatial settings. Using this approach, a model has several major sections concerning, in turn: overall characteristics or geological setting; regional (or possibly local) environment; and target area characteristics. This spatial hierarchy is clearly evident in the first four factors that establish "Favorable Conditions, Excluding Mineralization" in the KNSD model, and also runs through the PCDA model.

DS must make numerous other choices as well in constructing the inference network. Among these are the following:

1. Type of connection between spaces

As we pointed out in Section 2.2, evidence can be combined through a combination of inference rules and logical expressions. The basic idea is to use disjunctions (ORs) when any one piece of evidence is

sufficient, conjunctions (ANDs) when all pieces of evidence are necessary, and rules when each piece of evidence contributes votes. In practice none of these canonical methods is exactly right, and DS must be willing either to settle for an approximation or to work with MI in constructing a satisfactory network from these primitive elements.

2. Numerical quantity question vs. simple certainty question

Many parts of a model concern favorable intervals for a parameter value, such as the age of the host rock or the concentration of constituents. If this can be expressed as a simple proposition, such as "The age of the host rock is Archean," then the user need only indicate the degree to which he or she thinks that the proposition is true. However, if different ranges are favorable in different degrees, then the quantitative-rule formulation should be used (see Section 6.1), and the user must supply both the interval and a certainty for that interval.

3. Question sequencing

If DS has no preference about the order in which questions are asked, the Questioning System will select them in accordance with its control algorithm (see Section 4.3 of our last annual report (Duda et al., 1977)). Sometimes a particular order is mandatory because a question may make sense only in a particular context, one that must first be established. Sometimes it is merely conventional to ask questions in a particular order. Finally, a particular question about a weak inference should sometimes be asked only after an attempt to draw a stronger inference was made and produced inconclusive results. DS can use the context mechanism to handle all these situations.

Step B-2. Choosing numeric values for parameters

The inference network is not complete until DS has chosen a value for the prior probability of each space and two numeric values for each rule. The two numeric values associated with a rule represent the degree to which the evidence space is necessary and sufficient in establishing the hypothesis of the rule (see Section 2.3).

A critically important principle in choosing values for the rule strengths is that of "independence of spaces." The objective is to specify the favorability of a single hypothesis space based solely on its evidence spaces, independently of the rest of the model. This simplifies the task of DS in that he or she need take into account only a few factors at a time. Otherwise it is hopelessly complex to predict how the dozens of numeric values will interact when the complete model is executed in Prospector, with the result that the actual interaction may not reflect DS's intentions.

We have experimented with different ways of eliciting the numerical parameter values. For both the PCDA and the KNSD model we have found it effective to work top-down through the model, assigning prior probabilities to hypothesis spaces and computing rule strengths to give the desired posterior probabilities. The assignment of a prior probability to a space that represents a rare situation S is always difficult. In the absence of any constraining contexts, we usually resort to estimating the ratio of the area in which S occurs to the area that might reasonably be explored for ore deposits. As contexts are established, prior probability values can be estimated with greater confidence. Since rules usually concern evidence that can have a significant effect in establishing S, posterior probabilities are usually easier to estimate. However, interactions among rules often disclose that combinations of evidence only crudely approximate the desired behavior. A satisfactory approximation can often be obtained by adjusting the values of the prior probabilities and the rule strengths, but sometimes this is inadequate and we must go back to Step B-1 to revise the structure of the inference network. Simple programs that allow DS to see the effects of choosing different parameter values quickly have proved helpful in this process. In addition to satisfying DS, the numeric values must also satisfy certain consistency constraints, principally the following:

(1) Since the categories represented by a space that is the antecedent for a quantitative rule are mutually exclusive and exhaustive, the prior probabilities assigned to the categories must sum to one.

(2) Spaces that are defined as a logical conjunction (i.e., AND) of other spaces must have a prior probability equal to the minimum of those of these evidence spaces. Similarly, a space defined as the logical disjunction (i.e., OR) of its evidence spaces must have a prior probability equal to the maximum of those of its evidence spaces. Finally, a space A defined as the logical negation (i.e., NOT) of another space B must have a prior probability that, when added to that of B, equals one.

Step B-3. Writing of the elaboration/explanation/announce text

Before an implementation is complete DS must provide three different kinds of text called the elaboration text, the explanation text and the announce text. The elaboration text is a detailed, rephrased version of the question asked of the user. It is output in response to a "?" command from the user. The explanation text is a brief paragraph or two explaining the reason the question was asked. It is output in response to a "WHY" command and often contains a citation of references that provide more information. The announce text is output preceding exploration of a major new section of the model; it indicates to the user the purpose of the subsequent questions. While ultimately necessary, the task of creating such text is not on the critical path in developing the logical structure of the model and can be done whenever it is convenient.

Step B-4. Creation of the semantic network representation for each space in the model

To allow the system to recognize certain logical connections between the assertions that spaces represent and to allow connections to be made between volunteered statements and spaces in the models, each space must be articulated as a semantic network (see Section 2.5). The chief problem in creating these networks is to decide on a representation that is sufficiently detailed to discriminate among assertions, but not so detailed that it contains fine distinctions that are virtually never used. In particular, spaces that represent very abstract situations that are unlikely either to link logically to other spaces or to be volunteered may receive no semantic-network representation at all.

Since initial implementation and debugging can be done without volunteering information, considerable development work is accomplishable before the semantic networks are needed. Thus, the latter are usually created only after the model has been tested sufficiently to ensure that relatively minor future revisions of the spaces in the inference network may be anticipated. However, since the semantic networks do form links between the rules, final logical testing cannot be done prior to completion of the semantic networks.

PHASE C: Installation of the model in Prospector

Step C-1. Creation of computer file containing model definition

Part of the Network Creator is a special program called PARSEFILE that loads a fully specified model into Prospector. PARSEFILE reads in the description of a model contained in a computer disk file and creates from this information the internal representation used by the Prospector system. This instantiation of the Prospector system with models is then saved on a disk file and can be run by users.

To simplify its task PARSEFILE requires that the model be specified in a uniform format. In effect, a model is defined using a highly constrained formal language. In conventional computer science terminology, the format is the grammar for that formal language and the PARSEFILE program is an interpreter for sentences written in that formal language.

A detailed description of the grammar is given below. The notation employed uses uppercase words and lowercase words enclosed in angle brackets to indicate different types of entries in the file. An uppercase entry (e.g., SPACE) is a literal indicator, i.e., a particular keyword that tells PARSEFILE how to interpret what immediately follows. A lowercase entry enclosed in angle brackets (e.g., <space-name>) is a generic indicator, meaning that a certain type of entry is expected, but the actual value used varies from space to space and from model to model. The requirements of each generic indicator are elaborated below as essential for clarity.

Note that unless otherwise specified, all entries shown are required to appear. Also, while the entries may be given in any order, a model definition must start with the "MODEL <model-name>" entry, a <space-entry> with "SPACE <space-name>".

The format to define a model is:

```
MODEL <model-name>
TOPSPACE <space-name>
VERSION <version-number>                    (optional)
<space-entry>
<space-entry>
    .
    .
    .
<space-entry>
STOP
```

The format for a <space-entry> is:

```
SPACE <space-name>
DESC <description-text>     (Note: DESC is short for DESCRIPTION)
ANNOUNCE <announce-text>                    (optional)
? <elaboration-text>                        (optional but desirable)
WHY <explanation-text>                      (optional but desirable)
PRIOR <prior-probability-value>
<network-connection-entry>
<semantics-entry>                           (optional)
<context-entry>                             (optional)
ASKABLE                          (omitted if space is not askable)
<properties-entry>
```

A <network-connection-entry> takes one of two alternative forms:

(1) LOGICAL DEFINITION <logical-type> <space-name> ... <space-name>, where <logical-type> is either AND, OR, NOT, or *PROC. (*PROC indicates that the following <space-name> is the top of the section of the model dealing with zones.)

(2) RULES <direction> <rule-strength-entry> where <direction> is either TO or FROM, and <rule-strength-entry> specifies the likelihood ratio values (see example for KNSD model below).

Other generic indicators used above (e.g., <semantics-entry>) are clarified in the example below. The properties-entry is mostly for the benefit of certain routine tasks of MI and is not discussed here.

We illustrate the PARSEFILE format below with excerpts from the file for the KNSD model. Annotations and comments regarding this example follow, and are keyed to the numbers appearing to the right of the entries in the example. Note the variation in format used below for Space AGB, which is the antecedent for a quantitative rule.

```
                                                        Key for comments:
MODEL KNSD
TOPSPACE KNSD

SPACE KNSD
DESC /* Komatiitic Nickel Sulfide Deposit */
PRIOR .001
LOGICAL DEFINITION AND FCEM FM                                          1

SPACE FCEM
DESC /* Favorable Conditions, Excluding Mineralization */
LOGICAL DEFINITION AND FPTS FMALE FOBU PTO

SPACE FPTS
DESC /* Favorable PetroTectonic Setting */
PRIOR .001
RULES FROM GB LH 5 1.0E-4                                               2
          BPGB LH 2 .0001                                               3
          FETH LH 1.5 1
          UHAME LH 3 .001
          SUMR LH 300 .0001
          LMC LH 2 .5

SPACE GB
DESC /* The target area lies in a greenstone belt */
PRIOR .05
SEMANTICS COMP-OF E1 GREENSTONE /                                       4
          FORM-OF E1 BELT       /
          LOC     E1 REGION     /
ASKABLE                                                                 5

SPACE AGB
DESC LX /* What is the age of the greenstone belt (in millions
of years) */
                                                                        6
PRIOR IN 0 1500 USE .7                                                  7
      IN 1500 2500 .1
      IN 2500 3000 .15
      IN 3000 3500 .04
      IN 3500 4000  .01
RULES TO FPTS LHS .2 4.0 10 4 1                                         8
         RAGB LHS .1 1 4 10 100
ASKABLE
CONTEXT OF GB                                                           9
```

```
SPACE RAGB
DESC /* The age of the greenstone belt is Archean or Proterozoic */
PRIOR .01
SEMANTICS AGE-OF E1<GB> PRECAMBRIAN
CONTEXT OF GB
        FOR BPGB                                                    10

SPACE DVC
DESC YES/NO /* Can you identify more than one discrete volcanic
cycle within the greenstone belt */                                 11
ASKABLE
CONTEXT FOR DSS
        OF GB
```

The following comments are keyed to the above example.

(1) Space KNSD is the top space for the KNSD model. Its entry has only four items: a space name, a description text, a prior probability value, and a network connection entry of the logical-combination type (defining the space KNSD to be the logical conjunction of spaces FCEM and FM). Note that by default this space is not askable, because the keyword ASKABLE has been omitted. An explicit indication can be given by using the keyword UNASKABLE.

(2) Space FPTS is established by rules from spaces GB, BPGB, etc. The entry for the rule from GB says that the LS value for this rule is 5, and the LN value is 0.0001 .

(3) Note that the identifier "RULES FROM" need not be repeated if several rules having a common hypothesis space are defined contiguously.

(4) The keyword SEMANTICS identifies this as the semantic-network description for the space. Following the keyword can be any number of expressions of the form <relation-name> <argument> ... <argument> /. An argument of the form E<n>, where <n> is an integer, indicates a entity that is local to that space. An argument of the form E<n> followed by a <space-name> enclosed in angle brackets indicates an entity in the named space. All other arguments are global and must be in the taxonomy.

(5) Space GB is an askable space, since it has the entry ASKABLE.

(6) The keyword LX appearing immediately after the keyword DESC indicates that space AGB is of the numerical-input type (i.e., L(x)). Note that the description text forms a complete sentence (a question), whereas for the standard type of space Prospector precedes the

description text with the phrase "To what degree do you believe...".

(7) DS divides the entire range of the numeric quantity into several disjoint and exhaustive intervals according to the relative favorability of that interval in establishing the hypothesis. In the case of space AGB, DS has split the entire range of geologic age into six intervals, the first of which is from 0 to 1500 million years. A prior probability value must be given for each interval (e.g., .7 for the first interval), and these must sum to one. The keyword USE is optional. Note that the keyword PRIOR need not be repeated if the intervals are given contiguously.

(8) Rule strengths for a L(x) type rule differ from those of simple rules in that a single likelihood ratio value is given for each interval. In this case, the value .2 corresponds to the interval 0 to 1500 MY, 4.0 corresponds to the interval 1500 to 2500 MY, and so on. Note also that a rule may be entered in the entry of the evidence space, as here, in which case the keywords used are RULES TO (and then the hypothesis space is named, e.g., FPTS). Alternatively, a rule may be entered in the entry of the hypothesis space (as is the case for space FPTS), in which case the keywords RULES FROM are used.

(9) It would be nonsensical to ask the question of AGB if space GB has not yet been established. This context entry in the definition of space ABG will cause Prospector to ask the question of GB before that of AGB, and the latter only if the user responds positively to the former.

(10) Like RULES TO and RULES FROM, one can specify either CONTEXT OF or CONTEXT FOR (or both).

(11) The keyword YES/NO immediately following the keyword DESC indicates that this is a "yes or no" type of question.

Step C-2. First loadup

When the model definition file is complete, MI executes PARSEFILE and creates a version of Prospector containing the new model. PARSEFILE automatically checks for syntactic consistency of the model definition and prints a message if any of the following syntactic errors are detected:

(1) A space has no prior probability given.

(2) The given prior probability of one space is inconsistent with those of others (e.g., if Space A is the AND of Spaces B, C, and D, then the prior of A must equal the minimum of those given for B, C, and D).

(3) A particular space is referred to but not defined (usually indicating a typographical error in the model definition file).

Typically, several inconsistencies of this type are detected during the first loadup. In this event, MI corrects the model definition. This is done either by editing the model definition file and re-executing PARSEFILE, or by using PARSEFILE in an interactive mode to redefine the erroneous entries in the model.

Step C-3. Creation of model questionnaire

We have developed a computer program, called QUESTIONNAIRE, that uses the internal representation of a model to create a questionnaire listing all the questions asked by a model. After a successful loadup such a questionnaire can be produced. The principal function of the questionnaire is as a data sheet for DS to complete for each of the known deposits he or she listed in Step A-2 for the performance analysis in Phase E. However, it has also proved useful for such tasks as refining the wording of questions and identifying appropriate places for inserting "announce text."

PHASE D: Initial debugging of the model

Step D-1. Initial runs

At this point, MI (or DS, if he or she is present) makes several runs of the new model in Prospector to gain some initial impressions concerning its behavior. Certain deficiencies of the model may thereby become apparent. For example, grossly unexpected behavior may be traced to typographical errors in the model definition file; in addition, general program tendencies toward overconfidence or underconfidence may be noted for subsequent rectification.

As a rule, DS is not physically present at this point, and MI has not yet received completed questionnaires from DS. This step, therefore, is typically undertaken by MI acting alone. We have observed that an experienced MI can detect and correct fairly routine problems without DS's intercession, but must yield responsibility to DS for all graphically validated changes. If DS is not present, MI communicates the results to DS by mailing transcripts of these runs to DS, perhaps including his own inspection-based comments and annotations.

Step D-2. Implementation of initial modifications and corrections

If the results of Step D-1 indicate the need for any modifications, the next step is to implement these by editing the model file and then repeating Steps D-1 and D-2.

PHASE E: Preliminary performance analysis

Step E-1. Completion of the model questionnaire for known deposits

The completed questionnaires are used by MI as input for the preliminary performance analysis experiments. This is discussed fully in Section 4 of this report.

Step E-2. Running of Prospector using questionnaire data

MI inputs the questionnaire data supplied by DS into machine-readable form (i.e., a computer disk file), using for this purpose the "data input" mode of the QUESTIONNAIRE program. Then MI executes Prospector for each of these cases, typically using its "batchmode" facility (see Section 5.5) to obviate the need for locating each individual question asked by Prospector on the questionnaire and typing the answers one by one. Hence the mechanics of executing Prospector for the purposes of this step have been largely automated.

MI uses the tracing and/or summarizing facilities of Prospector for extracting performance data from these runs. MI then performs an

analysis using the methodology described in Section 4. As described in that section, the results of this part of the preliminary performance analysis can indicate specific sections of the model that need to be revised.

Step E-3. Sensitivity analysis experiments

MI continues with the sensitivity analysis portion of the preliminary performance analysis, as discussed in Section 4. The results of this analysis can indicate those specific sections of the model that are especially sensitive to the user's degree of certainty.

Step E-4. Analysis of experimental data for indications of model deficiencies

This step is discussed in detail in Section 4 of this report.

PHASE F: Subsequent revision of the model based on results of preliminary performance analysis

Step F-1. Revision of the model

The results of Phase E will typically indicate sections of the model that may need revision; the analysis may suggest particular types of revision. However, DS retains responsibility for the geological validity of any proposed changes to the model. Since the first preliminary performance analyses have only recently been completed, we do not as yet have sufficient experience with this step to warrant further discussion of it in this report.

Step F-2. Performance analysis of the revised model

Once a model has been revised, MI repeats Phase E for the new version. Of particular interest are comparisons between several versions of a model (see Section 4). The results of performance analysis provide an objective measure of the amount of improvement in model performance subsequent to the revisions.

7.3 General Observations

This description of the model construction process supports certain general conclusions. First, the creation of new models for Prospector at present is far from a simple matter; it requires the services of highly trained specialists, it employs many specialized technical procedures, and it requires several man-months to accomplish. However, the methodology has matured to the point at which it should be feasible to train non-computer scientists to play the role of Model Implementor (MI). Finally, our experience has demonstrated that models definitely vary in their accuracy in matching known deposits. The incorporation of performance analysis into our model construction methodology shows promise as an objective measure of model quality.

In evolving this methodology for model construction, we have pursued a strategy of devising routine procedures and building computer tools wherever such aids showed promise of saving time and/or yielding more uniform results. The results to date encourage us to continue this development. In particular, it seems desirable to write additional computer programs to extend the automatic consistency-checking facilities of the PARSEFILE program, for example:

(1) The probability of each unaskable consequent space can only be established by the incident rules from antecedent spaces. DS should be informed if no combination of states of knowledge about the antecedent spaces can make the consequent space very certain. We plan to develop a program to perform this check, to be used in conjunction with the PARSEFILE program.

(2) It would be desirable to devise additional computer programs for automating the analysis of a model based on executions concerning known deposits. Some such tools are suggested at the end of Section 4.

8 CONCLUSIONS AND RECOMMENDATIONS

The work on developing Prospector has shown that it is possible to provide valuable computer-based consultation services for mineral exploration at a low computational cost. We have developed a technology for encoding exploration models that possesses evident capability of representing information about a wide range of different types of models. Our methods for using that information to address particular exploration problems produce systematic investigations that identify the extent to which the data match the models. Although our tests of the validity of the conclusions are still in a preliminary phase, the results obtained to date have confirmed that the known deposits on which the models are based match the models well.

Several tasks remain to bring this technology into practical use. The first category of tasks is the improvement of existing models and the construction of new ones. The Kuroko-type massive sulfide model, the first one we developed, needs extensive revision. The Mississippi-Valley-type-deposit model, although structured more like our later models and probably less in need of revision, should nevertheless be revised. Under a separate contract we are currently developing models of sandstone and vein-type uranium deposits. Other candidates for future models include podiform chromite, Archean massive sulfide, porphyry molybdenum, porphyry skarn, and bulk low-grade silver deposits.

As more models are created, the interrelations among them will become more evident. Certain models, such as those for porphyry copper and porphyry molybdenum, or the Kuroko and Archean massive sulfide models, will undoubtedly share common submodels. Because different experts will someday be involved in developing similar models, systematic methods should be developed for keeping shared submodels compatible.

This report has focused on the use of these models for interactive evaluation of a particular prospect. However, as was mentioned in the Introduction, the knowledge encoded in these models can be applied to different kinds of tasks at different scales. The way this knowledge is used is determined primarily by the procedures of the Questioning System. More effort should be devoted to generalizing the Questioning System, so that Prospector can be used in a greater variety of applications.

There are several reasons for devoting considerable effort to rigorous testing of the Prospector system. The most obvious is to provide objective evaluation of the system. In addition, well-designed tests aid the model construction and revision procedure. Moreover, they promise ultimately to provide an objective method for comparing different models that purport to represent the same class of ore deposits, and thus to advance the science of economic geology. This process will begin when the consulting geologists who have contributed their valuable efforts to developing models in this formalism publish their models and the results of such tests in the literature of their discipline.

Besides the addition of special knowledge about ore deposit models, some effort should be devoted to augmenting the general geological knowledge contained in the taxonomy. The greatest need here is to augment the information about rocks, so that, rather than having to depend on rock names, the system can make use of such measurements as grain size and mineral concentrations. This should be done in such a way that rock types can be recognized that are similar to or easily confused with specific types mentioned in a particular model.

A second category of tasks is the augmentation of existing system facilities and the addition of new ones. Since the essence of a consultation system is to provide explanations and advice, improvements in the Explanation System are especially important. Specifically, more attention should be devoted to sensitivity analysis, so that the value of additional information can be more accurately determined.

Furthermore, the Explanation System should be able to answer questions about the system itself--such as the rules it contains and the contents of its taxonomy.

Information currently volunteered to Prospector is limited to naming the types, ages, and forms of rocks and minerals that are present in the target area. These descriptions must be given in independent, simple sentences. There are several ways in which these restrictions can be relaxed and the power of the English Analyzer increased. An ability to localize observations--for example, by referring to what was found in a particular outcrop--is definitely necessary. Ultimately the user should be enabled to communicate by means of a natural combination of text and maps or diagrams.

Finally, more work should be devoted to developing aids to model acquisition. The ultimate goal would be a system that would interact with an expert geologist in English to implement a new model. In the near term model implementation will have to be a cooperative enterprise between a geologist who understands the model and a computer scientist who understands the system. The development of tools that will aid both partners in the enterprise will simultaneously contribute to achievement of the ultimate goal.

REFERENCES

"Island Copper Project," Mining Magazine, pp. 344-349 (October 1972).

Brown, A. S., "Geology of the Granite Mountain Cusine Lake Area," British Columbia Department of Mines and Petroleum Resources, Geology Exploration and Mining, pp. 121-124 (1966).

Cannon, R. W., J. M. Thornton and D. C. Rotherham, "I.P. and Resitivity Prove Invaluable in Finding Gibraltar Ore," Mining Engineering, pp. 43-46 (September 1972).

Carson, D. J. T., J. L. Jambar, P. L. Orgyzlo and T. A. Richards, "Geology, Geochemistry and Genesis of a Supergene-enriched, Biotized Porphyry Copper Deposit with a Superimposed Phyllic Zone," in "Porphyry Deposits of the Canadian Cordillera," Canadian Institute of Mining and Metallurgy, Special Volume 15, A. S. Brown, ed., pp. 245-263 (1976).

D'allesandro, M., et al.,"Cogeodata Inquiry on Rock Names and Descriptions," Neues Jahrbuch fur Mineralogie, Vol. 129, No. 1, p. 43 (1977). Abstracted in Cogeodata Newsletter, published by International Union of Geological Sciences, Vol. 4, No. 1, p. 3 (January 1978).

Drummond, A. J., S. J. Tennant and R. J. Young, "The Interrelationship of Regional Metamorphism, Hydrothermal Alteration and Mineralization at the Gibraltar Mines Copper Deposit in B.C.," Bulletin of the CIM, pp. 48-55 (February 1973).

Duda, R. O., P. E. Hart, N. J. Nilsson, R. Reboh, J. Slocum and G. L. Sutherland, "Development of a Computer-Based Consultant for Mineral Exploration," Annual Report, SRI Projects 5821 and 6415, SRI International, Menlo Park, California (October 1977).

Eastwood, G. E. P., "Geology of Granite Mountain Stock," British Columbia Department of Mines and Petroleum Resources, Geology Exploration and Mining, pp. 162-172 (1966).

Ewers, W. E. and D. R. Hudson, "An Interpretive Study of a Nickel-Iron Sulfide Ore Intersection, Lunnon Shoot, Kambalda, Western Australia," Economic Geology, Vol. 67, pp. 1075-1092 (1972).

Gillogly, J. J., "Performance Analysis of the Technology Chess Program," Report CMU-CS-78-109, Dept. of Computer Science, Carnegie-Mellon University, Pittsburgh, Pennsylvania (March 1978).

Hart, P. E., R. O. Duda and K. Konolige, "A Computer-Based Consultant for Mineral Exploration," Second Semiannual Report, SRI Project 6415, SRI International, Menlo Park, California (April 1978).

Hayes-Roth, F., D. A. Waterman and D. B. Lenat, "Principles of Pattern-Directed Inference Systems," in Pattern-Directed Inference Systems, D. A. Waterman and F. Hayes-Roth, eds., pp. 577-601, Academic Press, New York (1978).

Hendrix, G. G., "Expanding the Utility of Semantic Networks Through Partitioning," Proc. 4th International Joint Conference on Artificial Intelligence, pp. 115-121, Tbilisi, Georgia, USSR (1975).

Hendrix, G. G., "Human Engineering for Applied Natural Language Processing," Proc. 5th International Joint Conference on Artificial Intelligence, pp. 183-191, Cambridge, Massachusetts (1977).

Kirk, R. E., Experimental Design: Procedures for the Behavioral Sciences, Brooks/Cole, Belmont, California (1968).

Lowell, J. D., "Geology of the Kalamazoo Orebody, San Manuel District, Arizona," Economic Geology, Vol. 63, pp. 645-654 (1968).

Lowell, J. D., and Guilbert, J. M., "Lateral and Vertical Alteration-Mineralization Zoning in Porphyry Copper Deposits," Economic Geology, Vol. 65, pp. 373-408 (1970).

Newell, A. and H. A. Simon, Human Problem Solving, Prentice-Hall, Englewood Cliffs, New Jersey (1972).

Newman, British Columbia Department of Mines and Petroleum Resources, Geology Exploration and Mining, p. 114 (1969).

Northcote, K. E., "Rupert Inlet-Cape Scott Map Area," British Columbia Department of Mines and Petroleum Resources, Geology Exploration and Mining, pp. 254-272 (1970).

Paxton, W. H., "A Framework for Speech Understanding,", Technical Note 142, Artificial Intelligence Center, SRI International, Menlo Park, California (June 1972).

Pratt, M. E., "The Island Copper Project," Western Miner, pp. 72-80 (April 1970).

Raiffa, H., Decision Analysis, Addison-Wesley Publishing Co., New York (1968).

Raphael, B., The Thinking Computer, W. H. Freeman and Co., San Francisco (1976).

Ross, J. R. and G. M. F. Hopkins, "The Nickel Sulfide Deposits of Kambalda, Western Australia," in Economic Geology of Australia and Papua-New Guinea., C. Knight, ed., Australian Inst. Min. Met., Monograph 5, pp. 100-121 (1975).

Rotherham, D. C., A. D. Drummond and S. J. Tennant, "Exploration of Gibraltar," Western Miner, pp. 25-28 (February 1972).

Shortliffe, E. H., Computer-Based Medical Consultations: MYCIN, Elsevier, New York (1976).

Simson, Y., "Geology of the Gibraltar-Polyanna Copper Deposit," B.S. Thesis, University of British Columbia, Vancouver, British Columbia (1969).

Waterman, D. A. and F. Hayes-Roth, eds., Pattern-Directed Inference Systems, Academic Press, New York (1978).

Wilson, J., and S. E. Kesler, "Fluid Inclusion Geochemistry of the Bell and Granisle Porphyry Copper Deposits, B.C.," paper No. 133, 80th Annual General Meeting, The Canadian Institute of Mining and Metallurgy, Vancouver, B.C. (April 23-27, 1978).

Winer, B. J., Statistical Principles in Experimental Design, Second Edition, McGraw-Hill, New York (1971).

Woodall, R. and G. A. Travis, "The Kambalda Nickel Deposits, Western Australia," Proc. 9th Commonwealth Mining and Metallurgical Congress, Vol. 2, pp. 517-533 (1969).

Young, M. Y., and E. S. Rugg, "Island Copper Deposit," Western Miner, pp. 31-40 (April 1970).

Appendix A

INPUT DATA FOR THREE KNOWN DEPOSITS FOR THE KNSD MODEL

Appendix A

INPUT DATA FOR THREE KNOWN DEPOSITS FOR THE KNSD MODEL

This appendix lists the input data to Prospector that Prof. Anthony Naldrett supplied concerning three known nickel sulfide deposits: the Langmuir deposit (in Ontario), the Alexo deposit (in Ontario), and the Lunnon deposit (in Kambalda, Western Australia). For each deposit, Prof. Naldrett completed a copy of the questionnaire reproduced below. To facilitate comparison of these data, the answers for the three deposits are listed here in three columns, keyed as follows: La = Langmuir, A = Alexo, Lu = Lunnon. The key NR indicates that the question is not relevant, because of the answer given to a preceding question.

SRI PROSPECTOR SYSTEM

DATA SHEET FOR SITE EVALUATION

Prospect site name and location: _____

_____

Maturity of field observations: (indicate one)

Preliminary or reconnaissance (a week to a month)? _____

Detailed surface map (one or more field seasons, surface observations only)? _____

Preliminary three-dimensional model (several field seasons, surface and drilling observations)? _____

Detailed 3-D model (several field seasons, surface + drilling + significant mining exposures)? _____

Your name and organization: _____

_____

Date: _____

This is a listing of the questions askable by PROSPECTOR for the KNSD model.

With the possible exception of a few yes/no questions and a few numeric interval questions, each question below takes the form,

To what extent do you believe S?,

where S is a statement, as listed below. The questions should be answered by a number from -5 to 5, where 5 indicates certainty that that the statement is true (i.e., the evidence is definitely present), -5 indicates certainty that the statement is false, and numbers in between indicate lesser certainty. In particular, the answer zero (0) indicates a complete absence of information about the statement.

		La	A	Lu
1) The target area lies in a Greenstone Belt	ANS:	5	5	5
2) (If positive answer to question 1) What is the age of the Greenstone Belt (in millions of years) INTERVAL ANS: 2500-3000 (all) What is your degree of belief in this interval?		5	4	2
3) (If positive answer to question 1) Can you identify more than one discreet volcanic cycle within the greenstone belt	ANS:	yes	yes	yes
4) (If positive answer to question 3) Can you determine the stratigraphic sequence in which these cycles occur	ANS:	yes	yes	yes
5) (If positive answer to question 4) The target area lies within the lowermost cycle of the series.	ANS:	-5	-3	5
6) (If positive answer to question 1) The target area is in a region containing significant volume of ultramafic rocks	ANS:	5	5	5
7) (If negative answer to question 6) The target area lies in a basaltic province in the greenstone belt	ANS:	NR	NR	NR
8) (If positive answer to question 7) The target area is in a portion of the basaltic province with an unusually high regional aeromagnetic expression	ANS:	NR	NR	NR

9) (If positive answer to question 7)
The target area is in a portion of the basaltic province containing significant volume of iron-rich rocks, with >15% Fe and >1% TiO_2 ANS: NR NR NR

So far the questioning has been designed to determine whether the prospect lies within a favorable petro-tectonic setting. The next few questions are designed to determine whether a favorable magmatic association and local environment exist. (covered by questions 10 through 24)

10) The target area is within 10 miles of known concentrations of nickel-bearing sulfides associated with komatiites ANS: 5 5 5

11) The prospective ore body lies in a sequence of mafic to ultra-mafic igneous rocks ANS: 5 5 5

12) (If positive answer to question 11)
This sequence of mafic to ultra-mafic rocks forms a komatiitic rock suite ANS: 5 5 5

13) Has a systematic chemical study of this sequence of mafic to ultra-mafic igneous rocks been done ANS: yes yes yes

14) (If positive answer to question 13)
The concentration of TiO_2 and MgO obey the equation:
$TiO_2 < 1.2 - (MgO / 54)$ ANS: 5 5 5

15) (If positive answer to question 13)
The concentration of FeO*, MgO, and Al_2O_3 obey the equation:
$FeO^* / (FeO^* + MgO) < .15 + .035 \times Al_2O_3$ ANS: 5 5 5

16) This sequence of mafic to ultra-mafic igneous rocks contains non-cumulate textured peridotite ANS: 5 5 3

17) (If positive answer to question 16)
This peridotite has 20-35% MgO ANS: 5 5 3

18) Some olivine-rich rocks are present as part of this sequence of mafic to ultra-mafic igneous rocks ANS: 5 5 5

19) This sequence of mafic to ultramafic igneous rocks forms a sequence of shallow-seated intrusives or flows ANS: 5 5 5

20) (If positive answer to question 18) These olivine-rich rocks have spinifex texture	ANS:	5	5	5
21) This sequence of mafic to ultramafic rocks contains pyroxene rich rocks	ANS:	0	5	5
22) (If positive answer to question 21) These pyroxene-rich rocks have spinifex texture	ANS:	NR	5	0
23) This sequence of mafic to ultra-mafic igneous rocks contains non-cumulate textured pyroxenite	ANS:	0	5	0
24) (If positive answer to question 23) This pyroxenite has 12-20% MgO	ANS:	NR	5	NR

Having established that we are dealing with a sequence of komatiite textured rocks, we now wish to determine whether favorable structural indicators are present. (covered by questions 25 through 25)

25) The target area lies in or close to a zone of major paleo faulting	ANS:	5	5	5
26) The target area is close to domes with granitic intrusive cores	ANS:	5	-4	5
27) The prospective ore body lies in a peridotite sub-volcanic feeder or flow	ANS:	5	3	5

The focus of the next few questions will now be to determine whether we are within a favorable ore-bearing unit. (covered by questions 28 through 35)

28) There are known concentration of nickel-sulfides within the unit	ANS:	5	5	5
29) (If uncertain answer to question 28) The concentration of sulfur in the central portion of the unit is greater than 1000 parts per million	ANS:	NR	NR	NR
30) (If uncertain answers to questions 28 and 29) The target area is surrounded by sulfur-bearing exhalite	ANS:	NR	NR	NR
31) The concentration of MgO in the central portion of the unit is greater than 40%	ANS:	0	5	5
32) The target area lies in a flow	ANS:	5	0	5

33) (If uncertain answer to question 32) The thickness of the unit is greater than 10 meters	ANS:	5	NR	5
34) (If positive answer to question 32) The thickness of the flow is greater than 10 meters	ANS:	5	NR	5
35) The target area is the feeder to or in the lowermost flow of a sequence of peridotite flows	ANS:	5	0	5
36) (If negative answer to question 35) The target area is in the second or third lowermost flow of a sequence of peridotite flows	ANS:	NR	0	NR

Now that we have established the presence of a favorable ore-bearing unit, the following few questions are designed to determine the chances of being proximal to ore within the unit. (covered by questions 37 through 43)

37) (If positive answer to question 30) There is evidence for the local absence of sulfur-bearing iron formation	ANS:	NR	NR	NR
38) The prospective ore body lies within or below an area of the flow or feeder in which the MgO content is unusually high	ANS:	0	0	0
39) (If positive answer to question 32) The prospective ore body lies in a zone of increased spinifex texture in the flow	ANS:	-2	0	0
40) (If positive answer to question 32) The prospective ore body is within 500 meters of a feeder	ANS:	5	NR	0
41) The prospective ore body lies in an irregularity at the base of the flow or feeder	ANS:	5	5	5
42) (If uncertain answer to question 41) The prospective ore body lies in a zone of minor paleo faulting	ANS:	NR	NR	NR
43) (If positive answer to question 32) The prospective ore body lies in a thicker zone along the flow	ANS:	2	0	3

The following questions are designed to determine whether the sequence

of ore types and the minerals within the ore body are typical of an ore deposit of the type in question. (covered by questions 44 through 50)

44) The prospective ore body consists of a sulfide assemblage	ANS:	5	5	5
45) The ratio of Cu and Ni in the sulfide assemblage conforms to the equation: Cu / [Cu + Ni] < .1	ANS:	5	5	5
46) This sulfide assemblage contains high nickel content minerals such as pentlandite	ANS:	5	5	5
47) (If positive answer to question 46) These nickel minerals are 20-40% of the total sulfides in the sulfide assemblage	ANS:	5	-5	5
48) This sulfide assemblage contains pyrrhotite	ANS:	5	5	5
49) (If positive answer to question 48) Pyrrhotite is 55-80% of the total sulfides in the sulfide assemblage	ANS:	5	5	5
50) Pyrite is less than 20% of the total sulfides in the sulfide assemblage	ANS:	5	0	5
51) Within the sulfide assemblage the original stratigraphic sequence - from the base to the top - consisted of massive sulfides followed by net textured sulfides followed by disseminated sulfides	ANS:	5	5	5

FOLLOW UP

Are your answers based primarily on: (check all that apply)

Extensive first-hand knowledge? _____

Minor first-hand knowledge? _____

Published references? (please list)

_____

_____

_____

_____

Unpublished reports? (please identify)

_____

_____

_____

How do you rate your understanding of this prospect?

Fair? _____
Good? _____
Excellent? _____

Your time spent completing this site evaluation data sheet is much appreciated. The information you have provided will aid in evaluating the effectiveness of the Prospector system and of the KNSD model.

Appendix B

INPUT DATA FOR THREE KNOWN DEPOSITS FOR THE PCDA MODEL

Appendix B

INPUT DATA FOR THREE KNOWN DEPOSITS FOR THE PCDA MODEL

This appendix lists the input data to Prospector that Prof. Marco Einaudi supplied concerning three known porphyry copper deposits: the Yerington deposit (located in Nevada), the Bingham deposit (located in Utah), and the Kalamazoo deposit (located in Arizona). For each deposit, Prof. Einaudi completed a copy of the questionnaire reproduced below. To facilitate comparison of these data, the answers for the three sites are listed here in three columns, keyed as follows: Y = Yerington, B = Bingham, K = Kalamazoo. The introductory and follow-up portions of the questionnaire are omitted here, but are shown in Appendix A.

This is a listing of the questions askable by Prospector for Version 1 of the PCDA model. (Most of the questions asked in Version 2 of the model are identical to those given below. Answers for other questions in Version 2 are derived from the answers given below.)

		Y	B	K
1) The target area is in a continental margin mobile belt	ANS:	5	5	5
2) The target area is subject to tectonic and magmatic activity related to subduction	ANS:	4	2	2
3) The continental margin mobile belt is post-Paleozoic	ANS:	5	5	5
4) The continental margin mobile belt contains granitic intrusives	ANS:	5	5	5
5) The belt contains metamorphosed Precambrian rocks	ANS:	-5	5	5
6) The belt contains metamorphosed late Paleozoic to early Mesozoic eugeosynclinal volcanic and sedimentary rocks	ANS:	5	-5	-5

7) The belt contains folded and faulted Paleozoic to early Mesozoic miogeosynclinal sedimentary rocks	ANS:	-5	5	5
8) There is a granitic intrusive system in the region	ANS:	5	5	5
9) There is a throughgoing fault system in the region	ANS:	1	5	5
10) The throughgoing fault system is older than the intrusive system	ANS:	1	5	-4
11) There are volcanic rocks in the region contemporaneous with the intrusive system (i.e., coeval volcanic rocks)	ANS:	3	5	-5
12) Igneous rocks with fine to medium grain size abound in the region	ANS:	5	5	5
13) Igneous rocks in the region have porphyritic texture	ANS:	5	5	5
14) The region contains an abundance of small stocks	ANS:	-5	5	5
15) The region contains an abundance of dikes	ANS:	5	-5	5
16) The region contains intrusive breccias	ANS:	-5	5	-5
17) The region contains volcanic plugs	ANS:	-5	5	-5
18) There is an intrusive system in the target area	ANS:	5	5	5
19) The host rock is intermediate to silicic igneous or metamorphic	ANS:	5	5	5
20) There is quartz monzonite in the intrusive system	ANS:	5	5	5
21) There is granodiorite in the intrusive system	ANS:	5	-5	5
22) There is quartz diorite in the intrusive system	ANS:	1	-5	0
23) The phases in the intrusive system have the same age	ANS:	4	5	5
24) The youngest major stock contains significant porphyry rock	ANS:	5	5	5
25) The grain size of the phenocrysts is .5 - 3.0 mm	ANS:	5	5	5
26) The phenocrysts are subhedral or euhedral	ANS:	5	5	5

27) The phenocrysts are composed of quartz, feldspar, and biotite or hornblende		ANS:	5	5	5
28) The groundmass texture is aplitic		ANS:	5	5	5
29) The grain size of the groundmass is less than .25 mm		ANS:	5	5	5
30) The groundmass is composed of quartz and feldspar		ANS:	5	5	5
31) The diameter of the youngest major stock in the intrusive system is less than 5000 feet		ANS:	5	5	5
32) The youngest major stock is a multiple plug or multiple dike		ANS:	5	5	5
33) There are fresh-looking feldspars	ZONE 1	ANS:	5	5	5
	ZONE 2	ANS:	5	5	5
	ZONE 3	ANS:	-5	-5	-5
	ZONE 4	ANS:	5	5	5
34) There is hornblende altered to actinolite	ZONE 1	ANS:	5	5	-4
	ZONE 2	ANS:	-5	-5	-4
	ZONE 3	ANS:	-5	-5	-5
	ZONE 4	ANS:	5	5	0
35) Hornblende has been altered to minor chlorite	ZONE 1	ANS:	5	5	5
	ZONE 2	ANS:	-5	-5	-5
	ZONE 3	ANS:	-5	-5	5
	ZONE 4	ANS:	5	5	2
36) Hornblende has been altered to major epidote	ZONE 1	ANS:	-5	-5	-5
	ZONE 2	ANS:	-5	-5	-5
	ZONE 3	ANS:	-5	-5	-5
	ZONE 4	ANS:	5	5	5
37) Plagioclase has been altered to albite	ZONE 1	ANS:	5	3	0
	ZONE 2	ANS:	5	5	-2
	ZONE 3	ANS:	5	3	-5
	ZONE 4	ANS:	4	3	0
38) Plagioclase has been altered to minor sericite	ZONE 1	ANS:	5	4	5
	ZONE 2	ANS:	5	5	5
	ZONE 3	ANS:	5	5	-5
	ZONE 4	ANS:	5	-3	-4

39) Plagioclase has been altered to major epidote	ZONE 1 ANS:	-5	-5	-5
	ZONE 2 ANS:	-5	-5	-5
	ZONE 3 ANS:	-5	-5	-5
	ZONE 4 ANS:	5	5	-3
40) There is hornblende with little or no secondary biotite	ZONE 1 ANS:	5	5	5
	ZONE 2 ANS:	-5	-5	-5
	ZONE 3 ANS:	-5	-5	-5
	ZONE 4 ANS:	5	5	5
41) There is chalcopyrite	ZONE 1 ANS:	5	5	5
	ZONE 2 ANS:	5	5	5
	ZONE 3 ANS:	5	5	5
	ZONE 4 ANS:	5	2	-3
42) There is bornite-chalcopyrite	ZONE 1 ANS:	5	-5	-5
	ZONE 2 ANS:	5	5	5
	ZONE 3 ANS:	5	-5	-5
	ZONE-4 ANS:	-5	-4	-5
43) There is bornite	ZONE 1 ANS:	5	-5	-5
	ZONE 2 ANS:	5	5	5
	ZONE 3 ANS:	-5	-5	-5
	ZONE 4 ANS:	-5	-5	-5
44) There is pyrite-molybdenite	ZONE 1 ANS:	-5	5	5
	ZONE 2 ANS:	-5	5	5
	ZONE 3 ANS:	-5	5	-3
	ZONE 4 ANS:	-5	-2	-5
45) The total amount of chalcopyrite, bornite-chalcopyrite, bornite and/or pyrite-molybdenite is less than one percent	ZONE 1 ANS:	5	5	5
	ZONE 2 ANS:	-5	-5	-5
	ZONE 3 ANS:	-5	-5	-5
	ZONE 4 ANS:	5	5	5
46) There is magnetite or pyrite in disseminated form	ZONE 1 ANS:	5	5	5
	ZONE 2 ANS:	5	-5	-4
	ZONE 3 ANS:	5	-5	-5
	ZONE 4 ANS:	5	5	3
47) There is magnetite or pyrite in veinlet form	ZONE 1 ANS:	5	5	5
	ZONE 2 ANS:	5	-5	-5
	ZONE 3 ANS:	5	-5	-5
	ZONE 4 ANS:	-2	-5	-3

48) Secondary K-feldspar is replacing plagioclase in the vein walls	ZONE 1 ANS:	-3	5	-5
	ZONE 2 ANS:	5	5	5
	ZONE 3 ANS:	5	5	-3
	ZONE 4 ANS:	-3	-4	-5
49) There is secondary K-feldspar in quartz veinlets	ZONE 1 ANS:	-5	5	5
	ZONE 2 ANS:	5	5	5
	ZONE 3 ANS:	5	5	2
	ZONE 4 ANS:	-4	-4	-5
50) There are abundant quartz-sulfide veinlets	ZONE 1 ANS:	-5	-5	-5
	ZONE 2 ANS:	5	5	5
	ZONE 3 ANS:	5	5	0
	ZONE 4 ANS:	-5	-5	-5
51) An alteration halo is apparent	ZONE 1 ANS:	-5	-5	-5
	ZONE 2 ANS:	5	5	5
	ZONE 3 ANS:	5	5	0
	ZONE 4 ANS:	-5	-5	-5
52) There is pervasively biotized hornblende	ZONE 1 ANS:	-5	-5	-5
	ZONE 2 ANS:	5	5	5
	ZONE 3 ANS:	5	5	-5
	ZONE 4 ANS:	-5	-5	-5
53) There is partially-biotized hornblende	ZONE 1 ANS:	5	5	-5
	ZONE 2 ANS:	-5	-5	-5
	ZONE 3 ANS:	-5	-5	-5
	ZONE 4 ANS:	2	-4	-5
54) There is pyrite-chalcopyrite	ZONE 1 ANS:	5	5	5
	ZONE 2 ANS:	5	5	5
	ZONE 3 ANS:	5	5	5
	ZONE 4 ANS:	5	5	-3
55) There is pyrite	ZONE 1 ANS:	5	5	5
	ZONE 2 ANS:	5	5	0
	ZONE 3 ANS:	5	5	5
	ZONE 4 ANS:	5	5	5
56) There are finely disseminated sulfides	ZONE 1 ANS:	5	5	5
	ZONE 2 ANS:	5	5	5
	ZONE 3 ANS:	5	5	-2
	ZONE 4 ANS:	5	5	0
57) There are sulfides in quartz veinlets	ZONE 1 ANS:	5	5	-5
	ZONE 2 ANS:	5	5	5
	ZONE 3 ANS:	5	5	5
	ZONE 4 ANS:	5	5	-4

58) There is glassy limonite	ZONE 1 ANS:	5	0	0
	ZONE 2 ANS:	5	0	0
	ZONE 3 ANS:	-5	0	0
	ZONE 4 ANS:	5	0	0
59) There is tenorite	ZONE 1 ANS:	5	0	0
	ZONE 2 ANS:	5	0	0
	ZONE 3 ANS:	-2	0	0
	ZONE 4 ANS:	-4	0	0
60) There is chrysocolla	ZONE 1 ANS:	5	0	0
	ZONE 2 ANS:	5	0	0
	ZONE 3 ANS:	-4	0	0
	ZONE 4 ANS:	-4	0	0
61) There is malachite	ZONE 1 ANS:	5	0	0
	ZONE 2 ANS:	5	0	0
	ZONE 3 ANS:	-4	0	0
	ZONE 4 ANS:	-4	0	0
62) There are limonite pseudomorphs	ZONE 1 ANS:	5	0	0
	ZONE 2 ANS:	5	0	0
	ZONE 3 ANS:	5	0	0
	ZONE 4 ANS:	5	0	0
63) There is an abundance of leached cavities	ZONE 1 ANS:	5	0	0
	ZONE 2 ANS:	-5	0	0
	ZONE 3 ANS:	-5	0	0
	ZONE 4 ANS:	-5	0	0
64) There are finely disseminated oxidation products	ZONE 1 ANS:	5	0	0
	ZONE 2 ANS:	5	0	0
	ZONE 3 ANS:	5	0	0
	ZONE 4 ANS:	-3	0	0
65) There are oxidation products in quartz veinlets	ZONE 1 ANS:	5	0	0
	ZONE 2 ANS:	5	0	0
	ZONE 3 ANS:	5	0	0
	ZONE 4 ANS:	5	0	0
66) There are partially epidotized feldspars	ZONE 1 ANS:	5	-4	-5
	ZONE 2 ANS:	5	-2	-4
	ZONE 3 ANS:	5	4	-5
	ZONE 4 ANS:	5	5	5
67) There are feldspars altered to sericite and montmorillonite	ZONE 1 ANS:	5	-1	5
	ZONE 2 ANS:	5	5	5
	ZONE 3 ANS:	5	5	5
	ZONE 4 ANS:	4	-4	4

68) Plagioclase has been altered to albite and calcite	ZONE 1 ANS:	-4	-4	0
	ZONE 2 ANS:	-4	2	2
	ZONE 3 ANS:	-4	3	-4
	ZONE 4 ANS:	3	3	0
69) Plagioclase has been altered to epidote and calcite	ZONE 1 ANS:	-5	-4	-3
	ZONE 2 ANS:	-5	-3	-5
	ZONE 3 ANS:	-5	-5	-5
	ZONE 4 ANS:	4	2	5
70) There is moderate to major chlorite	ZONE 1 ANS:	-5	0	4
	ZONE 2 ANS:	5	-4	-4
	ZONE 3 ANS:	-5	-5	-5
	ZONE 4 ANS:	-5	4	-2
71) There is moderate to major epidote	ZONE 1 ANS:	-5	-4	-4
	ZONE 2 ANS:	-5	-5	-5
	ZONE 3 ANS:	-5	-5	-5
	ZONE 4 ANS:	4	4	4
72) There is fresh hornblende	ZONE 1 ANS:	5	5	5
	ZONE 2 ANS:	-5	-5	-5
	ZONE 3 ANS:	-5	-5	-5
	ZONE 4 ANS:	5	5	5
73) There is very minor to absent secondary biotite	ZONE 1 ANS:	5	5	5
	ZONE 2 ANS:	-5	-5	-5
	ZONE 3 ANS:	5	-5	5
	ZONE 4 ANS:	5	5	5
74) There are veins of galena, sphalerite, pyrite or tennantite	ZONE 1 ANS:	-5	-5	-5
	ZONE 2 ANS:	-5	4	-5
	ZONE 3 ANS:	-5	5	-5
	ZONE 4 ANS:	-4	5	5
75) There are altered looking rocks	ZONE 1 ANS:	5	-4	-5
	ZONE 2 ANS:	5	5	-5
	ZONE 3 ANS:	5	5	5
	ZONE 4 ANS:	5	5	-5
76) The pyrite content exceeds 5 percent	ZONE 1 ANS:	-5	-5	-5
	ZONE 2 ANS:	-5	-5	-5
	ZONE 3 ANS:	5	5	5
	ZONE 4 ANS:	-5	-4	-5

77) There are pyrite veins or veinlets in pervasive quartz-sericite-pyrite	ZONE 1 ANS:	-5	-5	-5
	ZONE 2 ANS:	-5	-5	-5
	ZONE 3 ANS:	5	5	5
	ZONE 4 ANS:	-5	-5	-5
78) There are pyrite veins or veinlets with quartz-sericite-pyrite alteration halos in fresh-looking rocks	ZONE 1 ANS:	-3	-4	5
	ZONE 2 ANS:	5	5	-5
	ZONE 3 ANS:	5	5	5
	ZONE 4 ANS:	4	-3	-5
79) There are limonite veins or veinlets in pervasive quartz-sericite-limonite	ZONE 1 ANS:	-5	0	0
	ZONE 2 ANS:	-5	0	0
	ZONE 3 ANS:	5	0	0
	ZONE 4 ANS:	-5	0	0
80) There are limonite veins or veinlets with quartz-sericite-limonite alteration halos in clay altered rocks	ZONE 1 ANS:	-4	0	0
	ZONE 2 ANS:	5	0	0
	ZONE 3 ANS:	5	0	0
	ZONE 4 ANS:	-4	0	0
81) There is porous quartz-sericite rock	ZONE 1 ANS:	-5	0	0
	ZONE 2 ANS:	-5	0	0
	ZONE 3 ANS:	5	0	0
	ZONE 4 ANS:	-5	0	0
82) There are abundant leached cavities	ZONE 1 ANS:	-5	0	0
	ZONE 2 ANS:	-5	0	0
	ZONE 3 ANS:	5	0	0
	ZONE 4 ANS:	-5	0	0
83) There is quartz-sericite alteration in clay altered rocks	ZONE 1 ANS:	-5	0	0
	ZONE 2 ANS:	5	0	0
	ZONE 3 ANS:	5	0	0
	ZONE 4 ANS:	-5	0	0
84) There is alteration on the fractures	ZONE 1 ANS:	-5	0	0
	ZONE 2 ANS:	5	0	0
	ZONE 3 ANS:	5	0	0
	ZONE 4 ANS:	-5	0	0
85) There are leached cavities	ZONE 1 ANS:	-5	0	0
	ZONE 2 ANS:	5	0	0
	ZONE 3 ANS:	5	0	0
	ZONE 4 ANS:	-5	0	0

86) There is intense leaching	ZONE 1 ANS:	-5	0	0
	ZONE 2 ANS:	-5	0	0
	ZONE 3 ANS:	5	0	0
	ZONE 4 ANS:	-5	0	0
87) Chalcocite-covellite is replacing pyrite in pervasive quartz-sericite-pyrite	ZONE 1 ANS:	-5	0	0
	ZONE 2 ANS:	-5	0	0
	ZONE 3 ANS:	5	0	0
	ZONE 4 ANS:	-5	0	0
88) Chalcocite-covellite is replacing pyrite	ZONE 1 ANS:	-5	0	0
	ZONE 2 ANS:	-5	0	0
	ZONE 3 ANS:	5	0	0
	ZONE 4 ANS:	-5	0	0

WLK - I got a copy. I have seen the entire package and a big report they wrote. Also chatted w. CF. Park about it. I'm not interested. They are now trying to assemble a bulk low grade Ag model
FTG.
5/30/78

ASARCO

RECEIVED
MAY 22 1978
EXPLORATION DEPARTMENT

Northwestern Exploration Division
John C. Balla
Manager

May 16, 1978

Mr. W.L. Kurtz, Manager
Western USA
ASARCO Incorporated
P.O. Box 5747
Tucson, Arizona 85703

RCL - Consultants

PROSPECTOR - A Computer-Based Consultation System for Mineral Exploration

Dear Sir:

While at Stanford University recently, I paid a visit to the Stanford Research Institute (no longer affiliated with Stanford University) to view their "PROSPECTOR" Program. Mr. Peter E. Hart, Director of the Artificial Intelligence Center, and Richard O. Duda, associated with the Artificial Intelligence Center, explained the program, and demonstrated its utilization. For an explanation of the program, the Introduction and Overview from the October 1977 Annual Report "DEVELOPMENT OF A COMPUTER BASED CONSULTANT FOR MINERAL EXPLORATION, prepared by the Stanford Research Institute for the U.S. Geological Survey is an excellent summary, and is presented, attached.

With the above as background knowledge, we began a demonstration of the program. I was asked to imagine that I was on an outcrop some place, and was describing the rock found in outcrop to the computer. I had been mapping in the general area of the outcrop.

In order to keep things simple, I assumed that I was standing on an outcrop that represented the transition zone between the barren core and the potassic zone of the San Manuel Porphyry copper deposit. Oxidation and leaching were minimal (I mentally "moved" the deposit to Alaska).

The computer asked me a series of questions which were displayed on a TV screen. In answering the questions, I had a choice of answering from +5 (positive certainty that the answer was yes or present) through 0 (no knowledge, or can't say) through -5 (negative certainty that the answer was no or not present). In the attached

computer print out, the answers immediately follow the questions asked by the computer.

The computer initially compared the answers I gave it with the three ore deposit models it has in its memory (porphyry copper, Kuroko-type massive sulfide, Mississippi Valley-type carbonate lead-zinc deposit) then began comparing the various alteration zones with the answers I was providing it. By making these comparisons, the computer was deciding the probability of what alteration zone I was describing (The highest possible probability is 5 in the computer).

In the attached computer printout sheet, the letter "N" after a question means no. A question mark after a question (such as in question 26) is the same as asking the computer why it is asking the questions. It then provided an explanation. One can also ask the computer "why", as in question 27, and obtain an answer.

Periodically the computer, when questioned as to why it asked a particular question, would answer, as in question 43, "Presently I cannot provide any explanation to this question." What this means is that the computer does not "know" the explanation yet.

After 74 questions, the computer was asked to summarize its position. As shown in the summary, its "current hypothesis" was that I was looking at a type A porphyry (San Manuel type) copper deposit (as opposed to a type B). Also, its certainties as to which alteration zone I was in were: Barren core zone, 3.654 (out of a possible 5.0), potassic zone, 4.697 (out of a possible 5), which, of course, was correct.

The whole exercise took about 2 hours. The actual computer time used was only a couple of minutes, and the total computer cost was about $10.

As shown on the attached computer print-out, the program is designed to ask fairly detailed, geologically sophisticated, questions. I was favorably impressed with this aspect of the program.

Since this program is funded by the U.S. Geological Survey, the question may be asked: How useful is this program for mineral exploration in the U.S.A.? Frankly, I doubt that it will be very useful to the mineral exploration community. I believe that most mineral exploration geologists are sufficiently astute to recognize a porphyry copper deposit environment, or a Mississippi Valley-type environment when they see it. Those that are not that astute will probably not know about the "PROSPECTOR" program, or won't be able to answer the questions.

Who, then will be the primary beneficiaries of the program? Primarily, I would guess, the unknowledgeable (in mineral deposits) geologists conducting mineral exploration, or general mapping. These geologists would probably be U.S. G.S. type geologists, or governmental geologists working for foreign countries.

In addition, it may also be used in mineral resource appraisal of a given area, such as a proposed wilderness, or an existing wilderness or withdrawn (from mineral exploration) area. Under such a scenario, a U.S.G.S. geologist, after conducting a superficial mineral resource evaluation, would then "discuss" with the PROSPECTOR program, the results of his field work. The PROSPECTOR program would then give to the geologist the "probability" rating that a certain kind of deposit might exist in the proposed withdrawn area. Given two or more such areas (such as the RARE II areas or the BLM land areas), the PROSPECTOR program could "rate" the various areas in terms of their probability for a given type of mineral deposit to occur. The end result could be that the U.S.G.S. would be able, at some point in the future, to know where to go and explore for, say a porphyry copper deposit, by simply asking the computer where the highest probability is for the discovery of a porphyry copper deposit.

By the same token, this program could assist a company in mineral exploration. Given the geological data in our files, this data could be encoded into a computer. The ASARCO computer data would then be hooked up to the PROSPECTOR program, which would then, based upon the data provided, "rate" the probability that each prospect has for being associated with a particular type of mineral deposit. We could then examine those prospects that had the highest probability of being associated with a certain kind of mineral deposit.

The above scenarios are all predicated on the data base stored in the computer. If the data base does not contain the data on a particular type of deposit, then quite obviously the computer will not be able to evaluate prospects vis-a-vis that type of deposit.

Attached is a copy, out of the 1977 Annual Report, of the chapter on "Status and Future Work". In addition to the professors and consulting geologists noted who are providing data to the PROSPECTOR program, Professor Einaudi indicated that V.F. Hollister is also providing data, mainly on porphyry copper deposits. As indicated, they are attempting to expand their models of ore deposits. I was asked if I or we (ASARCO Incorporated) would assist in broadening their data base. I indicated that I didn't really know if we could or would.

You will notice section 5.2.1 that there is a plan to interface the PROSPECTOR program with the CHARAN program. Here, as an example, regional geochemical data will be rated by the PROSPECTOR program to aid in mineral resource appraisal. This could end up being used, as the geological data could be used, in evaluating numerous areas, such as the RARE II areas, for their mineral potential.

Mr. Hart indicated that one ultimate utilization of the PROSPECTOR program was that one day a geologist might be on an outcrop in, say, Alaska, and talk, via a radio, to a satellite, which would relay the message to the computer, which would verbally communicate with him. Thus he would know, essentially instantaneously, what kind of deposit he was looking at.

I indicated to Mr. Hart that I didn't think the PROSPECTOR program was going to be too useful in the southwestern U.S., due to the advanced state of exploration in the southwest. Also, I wasn't too enthused about its usefulness in the Northwestern U.S. I felt that its principal usefulness would probably be in assisting geological work in underdeveloped countries, where the governmental geologists might not be too knowledgeable about mineral deposits (The U.N. geologists work in Mexico at La Caridad might be one such example).

The underlining permise of my reasoning is that exploration in the U.S. is at an advanced state, and that most mineral exploration geologists should already have sufficient knowledge about mineral deposits not to have need of the PROSPECTOR program. By the time the program has all of the available data on all of the various types of mineral deposits that one might encounter, there probably will be little need for the program, except for the unknowledgeable.

One positive aspect of the program is that it makes for an excellent teaching tool. It really forces a person to think about the rock or outcrop, and the geologic characteristics that are present, but not perhaps noted, by the average geologist. The availability of this program to Stanford geology student will significantly increase their knowledge of mineral deposits, I believe.

I have the complete Annual Report in the Spokane Office. Enclosed is a copy of the table of contents of the report. Also I have a copy of a paper presented by

Messrs. Hart and Duda on PROSPECTOR, given at the Taita Hills Conference on Standards for Computer Applications in Resource Studies (Taita Hills, Kenya, November 8-15, 1977). Both are available to anybody who wishes to read them.

Yours very truly,

John C Balla

John C. Balla

JCB/mc
Enclosures

cc: TCOsborne w/enc.
SAAnzalone w/enc.
DPCadwell w/enc.
FTGraybeal w/enc.
DMSmith w/enc.

CONTENTS

ABSTRACT ii

LIST OF ILLUSTRATIONS iv

LIST OF TABLES v

ACKNOWLEDGMENTS vi

I INTRODUCTION AND OVERVIEW 1

II INFERENCE NETWORKS 33

III SEMANTIC NETWORKS 57

IV CONTROL 63

V STATUS AND FUTURE WORK 74

REFERENCES 79

APPENDICES

A Porphyry Copper Deposits: A Summary of Present Models . . . 82

B The Rule of Theories of Ore Deposition in Mineral Exploration 101

C Semantic Network Encoding of Charles F. Park, Jr.'s Kuroko-Type Massive Sulfide Deposit Model . . . 111

D Semantic Network Encoding of Neil Campbell and Alan N. Campbell's Model of Mississippi-Valley-Type Carbonate Lead/Zinc Deposits 121

E Semantic Network Encoding of Marco T. Einaudi's Model of Type-A Near-Continental-Margin Porphyry-Copper Deposits 132

F The Taxonomy 168

G Index of Rocks, Minerals and Alteration Products in the Taxonomy 196

I INTRODUCTION AND OVERVIEW

1.1 Introduction

This report describes the results of the first year of a two-year research and development program funded jointly by the Office of Resource Analysis of the U.S. Geological Survey, the RANN Division of the National Science Foundation, and the Exxon Corporation. The goal of this project is the creation of a computer-based system called PROSPECTOR that can serve as a consultant to aid exploration geologists in their search for ore deposits.

Three primary considerations motivated us to develop PROSPECTOR: the need for continuing exploration for mineral resources, the difficulty of staying technically abreast of an expanding technical discipline, and the desirability of bringing the knowledge of several specialists to a given resource problem. The success of related efforts to develop computer-based consultation systems for problems in medical diagnosis provided the basis for believing that our goal was achievable.

We envision two different modes of use for such a system. In the first mode, an exploration geologist starts by telling the program the characteristics of a particular prospect of interest -- the geologic setting, structural controls, and kinds of rocks, minerals, and alteration products present or suspected. The program compares these observations with models of various kinds of ore deposits, noting the similarities, differences, and missing information. The program then engages the geologist in a dialog to obtain additional relevant information and uses that information to make an assessment of the mineral potential of the prospect. Our goal here is to provide the geologist with a service comparable to giving him telephone access to

authorities on many different kinds of ore deposits. In many cases, the main benefit of the consultation may be to alert the geologist to unsuspected possibilities, and to establish which additional observations would be most valuable for further exploration.

In the second mode, the program would "talk" not to a person but to a large data base of mineralized properties. Here the goal would be to screen the data base, either to select for particular commodities or to make regional assessments of mineral resources. In general, this is a more difficult mode of operation, primarily because the facts recorded in data bases often require interpretation; simple mechanical accessing may fail to provide an answer, even though a trained geologist would recognize that relevant information was present. Thus, although we recognize the value of this mode of use, we have begun by addressing the problem of providing interactive consultation to an exploration geologist.

The ability of PROSPECTOR to provide expert consultation rests on a base of knowledge about economic geology. This knowledge base has several components, the most important of which are the models that contain information relevant to exploration for various classes of ore deposits. The program currently contains three different models, one for Kuroko-type massive sulfide deposits, one for Mississippi-Valley-type carbonate lead/zinc deposits, and one for a major class of near-continental-margin porphyry copper deposits. The models are stored in the computer in a special way, to be described later, that enables PROSPECTOR to use them to reason about geologic data. In addition, the overall system has been designed in such a way that its competence can be continually improved by the incorporation of additional models.

Mineral exploration is perhaps as much art as science, and the state of this art does not admit the construction of models as rigorous and complete as, say, those of Newtonian mechanics. This state of affairs has two important affects on the design of PROSPECTOR. First, the system must accommodate plausible or probabilistic styles of reasoning. Second, the models often reflect the subjective judgments of

expert economic geologists more than objectively derivable facts. Of course, the use of subjective judgments and probabilities to make technical evaluations is not unique to PROSPECTOR. Subjective probabilities have been used in resource evaluation (Harris, 1970; see Harris, 1977 for a comprehensive treatment), while panels of experts have been frequently used in Delphi studies (Linstone and Turoff, 1971; Ellis et al., 1975) to forecast technological events. What is unique is PROSPECTOR's combination of plausible and logical reasoning using a knowledge base supplied by experts to provide a computer-based consultation service.

As the foregoing description suggests, three different groups of people are involved in the design and use of PROSPECTOR: computer scientists who design the computer programs that provide the framework of the system, expert economic geologists who provide the knowledge base for the system, and the end users of the system who are seeking consultation about particular prospects of interest to them. We should emphasize here that the end user is expected to be a geologist himself, and not a layman. Because PROSPECTOR cannot make direct observations, the quality of its conclusions must depend upon the skill of the user.

To the best of our knowledge, PROSPECTOR represents the first attempt to build a computer system able to consult actively on problems of mineral exploration. The general notion of a computer-based consultation system, however, has been explored before. Procedures for performing the required plausible and logical reasoning have been developed through computer science research on artificial intelligence (Nilsson, 1971; Raphael, 1976). These procedures have been applied in several fields, with the most advanced development being in the field of medicine. In particular, computer-based consultation systems have been developed for infectious diseases (Shortliffe 1976), for glaucoma (Weiss et al., 1977), for kidney disease (Pauker et al., 1976), and for a substantial part of internal medicine (Pople, 1977). While these systems are rather recent and are still experimental, some have already attained levels of competence that in their specialties rival the

performance of expert physicians. These successes provided an additional motivation for us to try to achieve similar performance in the area of mineral exploration.

The remainder of this report is organized as follows. Section 1 is an overview of the system. It includes an example of an interactive computer session with PROSPECTOR as it currently exists, and a general discussion of both the form in which geological knowledge is stored and the mechanisms by which plausible inferences are drawn from specific observations. Sections 2 through 4 provide material that is more technical than that in Section 1. Section 2 is concerned with the mechanisms for drawing plausible inferences, Section 3 with the structuring of the geological knowledge, and Section 4 with the strategies used by PROSPECTOR to pursue a line of reasoning. Finally, Section 5 summarizes our present status and future plans.

1.2 An Example

The following example illustrates the operation of the present system. Although the case used is believed to be plausible, it is a hypothetical case chosen to illustrate various aspects of the program, and does not correspond to an actual exploration problem. For clarity, all information supplied by the user has been printed in upper-case and has been underlined. [Editorial comments are enclosed by brackets.]

```
PROSPECTOR ...... Mark II ...... Version 2
          16-Oct-77 16:53:40

Do you want instructions? YES

    PROSPECTOR, an experimental computer based consulting system, is
designed for use both in searching for ore deposits and in evaluating
the mineral potential of large geographic areas. Being in the early
research stages, PROSPECTOR is based on only three models, C. F. Park's
model for Kuroko-type massive sulfide deposits, N. Campbell's model for
a class of Mississippi Valley type lead/zinc deposits, and M. Einaudi's
model for a class of near-continental-margin porphyry-copper deposits.
    In using PROSPECTOR, you will be asked to answer questions about
```

V STATUS AND FUTURE WORK

The present PROSPECTOR system, while demonstrable and occasionally even impressive in its performance, is still nonetheless in a strictly experimental state. Here we describe that state, and outline our development plans for the immediate future. For simplicity, we discuss separately the geological and computer science aspects of the project although in reality the two are closely connected.

5.1 Current Status

5.1.1 Geological Status

The knowledge base of the system is composed of three different models of ore deposits, together with an associated taxonomy of rocks and minerals. Encodings of these models are given in Appendices C, D, and E, and a listing of the taxonomy is given in Appendix F. There are two simple measures of the size of the knowledge base -- the number of rules and the number of spaces included in the semantic network encoding of these models. The current version of PROSPECTOR contains a total of 118 rules and 269 spaces.

Our first model, of Kuroko-type massive sulfide deposits, was obtained from Prof. Charles F. Park, Jr. of Stanford University. Its encoding contains 34 rules and 39 spaces. The second model, of Mississippi-Valley-type lead/zinc deposits, was provided by Dr. Neil Campbell and Dr. Alan N. Campbell, both independent consulting economic geologists. Its encoding contains 20 rules and 28 spaces. The third model, of near-continental-margin-type porphyry copper deposits, was obtained from Prof. Marco T. Einaudi of Stanford University. This model is more than twice the size of the other two models combined; 64 rules and 202 spaces are needed to represent it.

The greater size of the porphyry model is due in part to the large amount of work that has been done on porphyry deposits, but we believe that an equally important factor is the experience we have gained in structuring models of deposits. Indeed, it appears clear at this point that over the past year we have evolved a systematic methodology for structuring and encoding models of ore deposits. We plan to structure our subsequent models along the same general lines as the porphyry model, and in fact already have evidence that this approach will continue to be successful. As an indication of the effort required to encode such models, about 50 hours of interviewing time were needed to define the porphyry model in terms of inference rules.

The taxonomy that supports these models currently contains 182 rock types and 231 minerals. In addition, there are 506 synonyms for these rocks and minerals. As one might expect, it is rather straightforward to establish a mineral taxonomy, but the rock taxonomy is more difficult. The task is eased somewhat for us because we do not seek a strictly hierarchical taxonomy, but can allow one rock name to appear in more than one taxonomic location. Our major requirement of the taxonomy is that it should permit appropriate inferences to be drawn by the system.

5.1.2 Computer Science Status

The PROSPECTOR system is implemented in the INTERLISP language on a DEC PDP-10 computer running under the TOPS-20 operating system. INTERLISP is a special developmental language designed for the rapid implementation of interactive programs for list processing. A consultation session on this time-shared computer, using our existing models, would generally cost no more than a few dollars in computer time. The revolution in the field of microelectronics continues to reduce hardware costs by about a factor of ten every 5 to 7 years, so computation costs can be expected to remain low even though the size of the knowledge base will grow.

5.2 Future Plans

Development of the system to date has proceeded quite closely along the lines laid out in our initial plans well over a year ago, and we do not foresee a major departure from those plans in the next year. The following specific tasks are high on our agenda.

5.2.1 Geological Plans

The major geological task is the acquisition of additional models of ore deposits, the installation of those models in PROSPECTOR, and the testing and refinement of both the individual models and the interactions among them. As the subject of our next efforts, we have selected layered chromite deposits and a family of nickel sulfide deposits. Layered chromite models will be furnished by Prof. Eugene Cameron of the University of Wisconsin, in collaboration with Prof. Einaudi. The nickel models will be furnished by Prof. Anthony Naldrett of the University of Toronto, again in collaboration with Prof. Einaudi. Some initial work on these models has already begun, and we expect the methodology successfully employed on the porphyry model to be used here. Along with these models, we anticipate that some additional work in augmenting and refining the existing taxonomy of rocks and minerals will be necessary.

A second planned geological extension is the incorporation of an ability to do statistical analysis of geochemical data. Specifically, we plan to interface PROSPECTOR to the CHARAN program written by Dr. Joseph Moses Botbol of the Office of Resource Analysis of the U.S. Geological Survey (Botbol et al., 1976). This program accepts from a user the geochemical characteristics of a prospect, and compares them statistically with the "signature" of other known deposits or districts. We believe that a conceptually straightforward method of interfacing the two programs is possible, and will result in substantially increased power of the combined system.

5.2.2 Computer Science Plans

Our tasks in system design and implementation are motivated by several factors--the need to accommodate an increasing number of models of deposits, the need to make more effective use of existing models, the need to provide a richer vocabulary for the specification and definition of models, and the need to provide a more convenient and flexible interface to the user of the system.

An immediate need is to relax restrictions on memory space. The current version of PROSPECTOR exhausts primary computer memory, and the knowledge base will have to be partitioned so that major portions can be stored in secondary memory and swapped into primary memory as needed. Another immediate need is to relax restrictions on the use of a single inference rule several times within a given consultation session. For technical reasons, the current system allows a rule to be used only once, a temporary expedient that allowed us to make rapid progress in an earlier experimental stage.

Many geologists who have observed experimental sessions with PROSPECTOR have commented on the fact that its principal value lies not in the estimation of the probability occurrence of a deposit, but rather in the analysis and interpretation of the data on the prospect. These comments have reinforced our belief that the system must continue to improve in this regard, and we plan to devote effort to this task during the coming year.

The geologists who will use PROSPECTOR are not expected to be experts in computer science, so the motivation to simplify access to the system is strong. We plan to improve the ability of the system to accept English language input; some attention may also be given to the problem of utilizing graphical input and output.

5.3 Other Possibilities

We have focussed on the use of PROSPECTOR as a tool for the working

field geologist exploring a particular prospect. It is clear, however, that the same general mechanisms can also be employed as a tool in problems of regional resource assessment. There are at least two ways in which this might be done. One way would be to use the system to screen a database representing properties with mineral potential. Each entry in a typical database of such properties describes some of the geological characteristics of the property. The system would "talk" to the database, rather than to the geologist, to obtain the data. Developing this possibility would require a considerable amount of effort, because PROSPECTOR would not be able to ask questions as it would of a human user.

A second possibility would be to use PROSPECTOR to screen each of a suitably sized collection of cells covering the region of interest. In this mode, the available data on each of the cells would be the primary input to the system. Of course, it would be possible to provide PROSPECTOR with the general characteristics of the entire region, so that manual entry would be needed only to specify those cell characteristics that differed from the general description.

Following either of these two possibilities, the principal result of running the program would be an estimate of the probability of occurrence of each type of deposit for which PROSPECTOR has a model. In addition to possibly large savings in manpower, both approaches have the attractive property that the evaluation would be objective in the sense that it would be repeatable and uniform over the region.

```
          PROSPECTOR ...... Mark II ...... Version 4

                    24-Apr-78 10:07:58

Do you want instructions? N


Do you want to volunteer any information? Y

Pause for loading the language module...

When you are finished volunteering, please type DONE
Do you want to see the instructions? N

--------------------
1: QUARTZ-MONZONITE

   - QUARTZ-MONZONITE. (5)
--------------------
2: PRECAMBRIAN GRANITE

   - PRECAMBRIAN GRANITE. (5)
--------------------
3: TRACE
OK
--------------------
3: STOCK

--------------------
3: QUARTZ VEINLETS

(QUARTZ VEINLET)

    E-9008      COMP-OF: QUARTZ
                FORM-OF: VEINLET
                CERTAINTY: 5
SPACE-9011
Links:  to      SPACE-8L.MSD
        from    SPACE-OPCV.PCDA
        from    SPACE-SQV.PCDA
        from    SPACE-AQSV.PCDA
        from    SPACE-SKFV.PCDA
```

```
 Changing the certainty of (QUARTZ VEINLET)
        from 0.0 to 5.0

 Changing the certainty of (* Massive sulfide deposit mineralization)
        from 0.0 to 5.0

 Changing the certainty of (* Massive Sulfide Deposit)
        from 0.0 to .01992
   - QUARTZ VEINLET. (5)
--------------------
4: PYRITE

(PYRITE)

    E-9012      COMP-OF: PYRITE
                CERTAINTY: 5
SPACE-9014
Links:  to      SPACE-22R.MSD
        to      SPACE-8L.MSD
        from    SPACE-PYCHAL.PCDA
        from    SPACE-PYMOLY.PCDA
        =       SPACE-21R.MSD
        from    SPACE-PACC.PCDA
        from    SPACE-CCPPQSP.PCDA
        from    SPACE-PYGT5.PCDA
        from    SPACE-PVVQSPH.PCDA
        from    SPACE-PVVPQSP.PCDA
        from    SPACE-PY.PCDA


 Changing the certainty of (Pyrite)
        from 0.0 to 5.0

 Changing the certainty of (Sulfides)
        from 0.0 to 5.0
   - Pyrite. (5)
--------------------
5: CHALCOPYRITE

(CHALCOPYRITE)

    E-9015      COMP-OF: CHALCOPYRITE
                CERTAINTY: 5
SPACE-9017
Links:  to      SPACE-57.MVTD
        to      SPACE-22R.MSD
        to      SPACE-8L.MSD
        from    SPACE-PYCHAL.PCDA
        from    SPACE-BOCHAL.PCDA
        from    SPACE-CHAL.PCDA
```

Changing the certainty of (CHALCOPYRITE)
from 0.0 to 5.0

Changing the certainty of (* Prospect contains sulfide minerals other than those of Fe, Pb and Zn)
from 0.0 to 5.0

Changing the certainty of (* Prospect contains no sulfide minerals other than those of Fe, Pb and Zn)
from 0.0 to -5.0

Changing the certainty of (* Prospect contains no sulfide minerals other than those of Fe, Pb and Zn, or there is a satisfactory explanation for their presence)
from 0.0 to -4.0

Changing the certainty of (* Suggestive evidence for MVTD)
from 0.0 to -4.0

Changing the certainty of (MVTD)
from 0.0 to -4.0
- CHALCOPYRITE. (5)

6: NOTRACE
OK

6: MOLYBDENITE

- MOLYBDENITE. (5)

7: SERICITE

- Sericite. (5)

8: FELDSPAR

- FELDSPAR. (5)

9: DONE

I am considering the possibility of a
(* Type-A porphyry copper deposit -- PCDA)
in the target area.

Presently, the possible deposits are:

1 - Type-A porphyry copper deposit -- PCDA. (0.0)
                                    Current hypothesis
2 - Kuroko-type massive sulfide deposit -- MSD. (.01992)
3 - Mississippi Valley type Lead/Zinc deposit -- MVTD. (-4.0)
Which of the above would you like to rule out?
    (type ? for available options).....? 2

The following questions are intended to establish the nature of the petro-tectonic setting.

You may skip over this part of the model if you wish, in which case I will assume that the corresponding conditions are established.
Do you want me to skip? Y

The following questions are intended to determine the nature of the regional environment.

You may skip over this part of the model if you wish, in which case I will assume that the corresponding conditions are established.
Do you want me to skip? N

9--To what degree do you believe that:
(* There is a granitic intrusive system in the region) ? 5

10--To what degree do you believe that:
(* Igneous rocks with fine to medium grain size abound in the region) ? 5

11--To what degree do you believe that:
(* Igneous rocks in the region have porphyritic texture) ? 5

12--To what degree do you believe that:
(* The region contains an abundance of small stocks) ? 5

13--To what degree do you believe that:
(* The region contains an abundance of dikes) ? 5
[Pause for memory reclaiming.....Continuing]

My certainty in (* Suggestive morphology of igneous rocks) is now: 4.847
Do you want to pursue this further? N

14--To what degree do you believe that:
(* There are volcanic rocks in the region contemporaneous with the intrusive system (i.e., coeval volcanic rocks)) ? ?

*
Explosive venting of a porphyry intrusive system leads to the generation of coeval volcanic deposits; coeval volcanic rocks, either as massive flows, ash flow tuffs, or stratified volcanoclastic deposits, may be intruded by, or grade into, various intrusive bodies such as dikes, plugs, or breccia pipes. In some cases where contact relations are absent or non-definitive, the relative ages of intrusive and volcanic rocks can only be inferred from regional geological relations or determined by radiometric age dating. To what degree do you believe that volcanic rocks in your area are coeval with the intrusive system? -5

15--To what degree do you believe that:
(* There is a throughgoing fault system in the region) ? 5

16--To what degree do you believe that:
(* The throughgoing fault system is older than the intrusive system) ? 5

The following questions are intended to determine the existence of the right kind of an intrusive system in the target area.

You may skip over this part of the model if you wish, in which case I will assume that the corresponding conditions are established.
Do you want me to skip? N

17--To what degree do you believe that:
(* There is an intrusive system in the target area) ? 5

18--To what degree do you believe that:
(* The host rock is intermediate to silicic igneous or metamorphic) ? 5

You told me about:
- QUARTZ-MONZONITE. (5.0)

19--To what degree do you believe that:
(* There is quartz monzonite in the intrusive system) ? 5

20--To what degree do you believe that:
(* There is granodiorite in the intrusive system) ? 0

21--To what degree do you believe that:
(* There is quartz diorite in the intrusive system) ? -3

22--To what degree do you believe that:
(* The youngest major stock contains significant porphyry rock) ? 5

23--To what degree do you believe that:
(* The grain size of the phenocrysts is .5 - 3.0 mm) ? WHY

*
I am trying to establish the presence of the right porphyry texture. The data on porphyritic rocks associated with porphyry copper deposits indicate that there is a characteristic and relatively narrow range of sizes for phenocrysts. Typical sizes for feldspar phenocrysts are:

0.5 to 4 mm, rarely up to 10 mm (K-spar) at Yerington, Nevada
1.7 mm average for quartz monzonite prohpyry and 2.3 mm average for latite porphyry at Bingham, Utah
3 mm at Ajo, Arizona
5 mm at San Manuel, Arizona
2-3 mm at Island Copper, British Columbia
1-5 mm at Gibralter, British Columbia

23--To what degree do you believe that:
(* The grain size of the phenocrysts is .5 - 3.0 mm) ? 3

24--To what degree do you believe that:
(* The phenocrysts are subhedral or euhedral) ? ?

Presently I cannot elaborate any further on this question.

24--To what degree do you believe that:
(* The phenocrysts are subhedral or euhedral) ? 3

You told me about:
- FELDSPAR. (5.0)

25--To what degree do you believe that:
(* The phenocrysts are composed of quartz, feldspar, and biotite or hornblende) ? 5

26--To what degree do you believe that:
(* The groundmass texture is aplitic) ? ?

*
The finer mineral grains in porphyritic rocks constitute the groundmass. I am trying to establish the texture and grain size of the groundmass. It may be either glassy or crystalline; in the latter case it may be very fine grained and appear aphanitic to the unaided eye, but in many cases the grain size and texture can be discerned with the aid of a hand lens. If the individual grains in the groundmass are equant and anhedral, the texture is said to be xenomorphic-granular, aplitic, or sugary (saccharoidal). I am also trying to establish the mineral composition of the groundmass, and especially to determine whether it is relatively free of mafic minerals. To what degree do you believe that:
(* The groundmass texture is aplitic) ? 4

27--To what degree do you believe that:
(* The grain size of the groundmass is less than .25 mm) ? WHY

I am trying to establish (* Right groundmass)
*
A sugary-textured groundmass of quartz and feldspar, usually with a high alkali-feldspar/plagioclase-feldspar ratio, is a characteristic feature of porphyries associated with porphyry copper deposits. The texture and abundance of K-feldspar may be due in part to alteration and recrystallization, but more likely is due to pressure quenching.
27--To what degree do you believe that:
(* The grain size of the groundmass is less than .25 mm) ? 1

You told me about:
- FELDSPAR. (5.0)

28--To what degree do you believe that:
(* The groundmass is composed of quartz and feldspar) ? 5

29--To what degree do you believe that:
(* The diameter of the youngest major stock in the intrusive system is less than 5000 feet) ? 3

30--To what degree do you believe that:
(* The youngest major stock is a multiple plug or multiple dike) ? 5

Porphyry copper deposits are often composed of characteristic zones of alteration and mineralization. (In some cases these can be looked upon as different stages superimposed upon one another.) I shall call the most prominent such zone (or stage) Zone 1.

Even though you may have already identified this zone as, say, the propylitic zone or the sericitic zone, I need to consider other possibilities as well. Thus, the following questions are intended to determine the possibilities for Zone 1.

Hypothesis: Zone 1 is the (Barren core zone)

31--To what degree do you believe that:
(* There are fresh-looking feldspars) in Zone 1 ? 5

32--To what degree do you believe that:
(* There is hornblende with little or no secondary biotite) in Zone 1 ? 4

33--To what degree do you believe that:
(* There is hornblende altered to actinolite) in Zone 1 ? 0

34--To what degree do you believe that:
(* Hornblende has been altered to minor chlorite) in Zone 1 ? 0

35--To what degree do you believe that:
(* Hornblende has been altered to major epidote) in Zone 1 ? 0

36--To what degree do you believe that:
(* Plagioclase has been altered to albite) in Zone 1 ? 1

You told me about:
- Sericite. (5.0)

37--To what degree do you believe that:
(* Plagioclase has been altered to minor sericite) in Zone 1 ? 4

38--To what degree do you believe that:
(* Plagioclase has been altered to major epidote) in Zone 1 ? -3

39--To what degree do you believe that:
(* There is magnetite or pyrite in disseminated form) in Zone 1 ? 5

You told me about:
- CHALCOPYRITE. (5.0)

Your statements imply:
- Prospect contains sulfide minerals other than those of Fe, Pb and Zn. (5.0)
- Sulfides. (5.0)

40--To what degree do you believe that:
(* There is chalcopyrite) in Zone 1 ? 5

41--To what degree do you believe that:
(* The total amount of chalcopyrite, bornite-chalcopyrite, bornite and/or pyrite-molybdonite is less than one percent) in Zone 1 ? 4

Conclusion: my certainty in (Barren core zone) is now: 3.654

Hypothesis: Zone 1 is the (Propylitic zone)

42--To what degree do you believe that:
(* There is moderate to major epidote) in Zone 1 ? -4

43--To what degree do you believe that:
(* Plagioclase has been altered to albite and calcite) in Zone 1 ? WHY

Presently I cannot provide any explanation for this question.

43--To what degree do you believe that:
(* Plagioclase has been altered to albite and calcite) in Zone 1 ? -3

44--To what degree do you believe that:
(* Plagioclase has been altered to epidote and calcite) in Zone 1 ? 1

45--To what degree do you believe that:
(* There is fresh hornblende) in Zone 1 ? ?

46--To what degree do you believe that:
(* There is very minor to absent secondary biotite) in Zone 1 ? 5

You told me about:
- FELDSPAR. (5.0)
- Sericite. (5.0)

47--To what degree do you believe that:
(* There are feldspars altered to sericite and montmorillonite) in Zone 1 ? 2

48--To what degree do you believe that:
(* There is moderate to major chlorite) in Zone 1 ? 5

49--To what degree do you believe that:
(* There are veins of galena, sphalerite, pyrite or tennantite) in Zone 1 ? 4

Conclusion: my certainty in (Propylitic zone) is now: 3.682

Hypothesis: Zone 1 is the (Potassic zone)

50--To what degree do you believe that:
(* There is pervasively biotized hornblende) in Zone 1 ? SUMMARIZE

On a scale from -5 to 5,
my certainty in the type of possible deposits are:

1 - Type-A porphyry copper deposit -- PCDA. (.04505)
Current hypothesis
2 - Kuroko-type massive sulfide deposit -- MSD. (.01992)
3 - Mississippi Valley type Lead/Zinc deposit -- MVTD. (-4.0)

My certainties in the possible identities of Zone 1 are:
4 - Barren core zone. (3.654)
5 - Potassic zone. (0.0) **Current hypothesis**
6 - Propylitic zone. (3.682)
7 - Sericitic zone. (0.0)
For which of the above do you wish to see additional information?
(type ? for available options).....? NONE

50--To what degree do you believe that:
(* There is pervasively biotized hornblende) in Zone 1 ? 1

51--To what degree do you believe that:
(* There is partially-biotized hornblende) in Zone 1 ? CHANGE

Change Question number : SHOW

Show Question number : 45
45 - There is fresh hornblende. (3)
Change Question number : SHOW

Show Question number : 43-49

Not a legal answer, I am expecting primarily:
[a question number, DONE] or one of:
(? WHY WAIT SUMMARIZE HELP QUIT COMMENT CHANGE TRACE NOTRACE SHOW)

Show Question number : SHOW 43

```
   Show Question number :  43
    43   - Plagioclase has been altered to albite and calcite. (-3)
   Show Question number :  42
    42   - There is moderate to major epidote. (-4)
Change Question number :  SHOW

   Show Question number :  46
    46   - There is very minor to absent secondary biotite. (5)
Change Question number :  SHOW

   Show Question number :  47
    47   - There are feldspars altered to sericite and montmorillonite. (
2)
Change Question number :  SHOW

   Show Question number :  48
    48   - There is moderate to major chlorite. (5)
Change Question number :  SHOW

   Show Question number :  49
    49   - There are veins of galena, sphalerite, pyrite or tennantite. (
4)
Change Question number :  49
New answer: -5
More changes? NO
u.b.a.
UNASK

8_BT
u.b.a.
BT

9_BREAK1(T T)

(broken)
10:BT
```

TOP

11:OK
NIL
T
12_DRIBBLE()
NIL
17#OK
Returning to regular interaction

51--To what degree do you believe that:
(* There is partially-biotized hornblende) in Zone 1 ? 3
[Pause for memory reclaiming.....Continuing]

You told me about:
- QUARTZ VEINLET. (5.0)

Your statements imply:
- Sulfides. (5.0)

52--To what degree do you believe that:
(* There are abundant quartz-sulfide veinlets) in Zone 1 ? -3

53--To what degree do you believe that:
(* An alteration halo is apparent) in Zone 1 ? 0

54--To what degree do you believe that:
(* Secondary K-feldspar is replacing plagioclase in the vein walls) in Zone 1 ? 5

You told me about:
- FELDSPAR. (5.0)
- QUARTZ VEINLET. (5.0)

55--To what degree do you believe that:
(* There is secondary K-feldspar in quartz veinlets) in Zone 1 ? 5

Your statements imply:
- Sulfides. (5.0)

56--To what degree do you believe that:
(* There are finely disseminated sulfides) in Zone 1 ? 2

You told me about:
- QUARTZ VEINLET. (5.0)

Your statements imply:
- Sulfides. (5.0)

57--To what degree do you believe that:
(* There are sulfides in quartz veinlets) in Zone 1 ? 2

58--To what degree do you believe that:
(* There is glassy limonite) in Zone 1 ? -5

59--To what degree do you believe that:
(* There is tenorite) in Zone 1 ? 0

60--To what degree do you believe that:
(* There is chrysocolla) in Zone 1 ? 0

61--To what degree do you believe that:
(* There is malachite) in Zone 1 ? 0

62--To what degree do you believe that:
(* There are limonite pseudomorphs) in Zone 1 ? 0

63--To what degree do you believe that:
(* There is an abundance of leached cavities) in Zone 1 ? 0

64--To what degree do you believe that:
(* There are finely disseminated oxidation products) in Zone 1 ? 0

You told me about:
- QUARTZ VEINLET. (5.0)

65--To what degree do you believe that:
(* There are oxidation products in quartz veinlets) in Zone 1 ? -5

Conclusion: my certainty in (Potassic zone) is now: 4.697

Hypothesis: Zone 1 is the (Sericitic zone)

66--To what degree do you believe that:
(* There are altered looking rocks) in Zone 1 ? 1

You told me about:
- Sericite. (5.0)
- Pyrite. (5.0)

Your statements imply:
- Sulfides. (5.0)

67--To what degree do you believe that:
(* There are pyrite veins or veinlets in pervasive
quartz-sericite-pyrite) in Zone 1 ? -4

You told me about:
- Sericite. (5.0)
- Pyrite. (5.0)

Your statements imply:
- Sulfides. (5.0)

68--To what degree do you believe that:
(* There are pyrite veins or veinlets with quartz-sericite-pyrite
alteration halos in fresh-looking rocks) in Zone 1 ? -4

You told me about:
- Pyrite. (5.0)

Your statements imply:
- Sulfides. (5.0)

69--To what degree do you believe that:
(* The pyrite content exceeds 5 percent) in Zone 1 ? -5

You told me about:
- Sericite. (5.0)

70--To what degree do you believe that:
(* There are limonite veins or veinlets in pervasive
quartz-sericite-limonite) in Zone 1 ? -5

You told me about:
- Sericite. (5.0)

71--To what degree do you believe that:
(* There are limonite veins or veinlets with quartz-sericite-limonite
alteration halos in clay altered rocks) in Zone 1 ? -5

You told me about:
- Sericite. (5.0)

72--To what degree do you believe that:
(* There is porous quartz-sericite rock) in Zone 1 ? -5

You told me about:
- Sericite. (5.0)

73--To what degree do you believe that:
(* There is quartz-sericite alteration in clay altered rocks) in Zone 1
? -5

You told me about:
- Sericite. (5.0)
- Pyrite. (5.0)

Your statements imply:
- Sulfides. (5.0)

74--To what degree do you believe that:
(* Chalcocite-covellite is replacing pyrite in pervasive
quartz-sericite-pyrite) in Zone 1 ? -4
That is all I have to ask about Zone 1
My best guess for this Zone is:
- Potassic zone. (4.697)

Are there any other zones or stages present in the target area? SUMMARIZE

On a scale from -5 to 5,
my certainty in the type of possible deposits are:

1 - Type-A porphyry copper deposit -- PCDA. (.04505)
Current hypothesis
2 - Kuroko-type massive sulfide deposit -- MSD. (.01992)
3 - Mississippi Valley type Lead/Zinc deposit -- MVTD. (-4.0)

My certainties in the possible identities of Zone 1 are:
4 - Barren core zone. (3.654)
5 - Potassic zone. (4.697)
6 - Propylitic zone. (2.917)
7 - Sericitic zone. (-4.0)
For which of the above do you wish to see additional information?
(type ? for available options).....? ?
The following options are currently available:

-a number (or a list of numbers) referring to the above hypothesis.
-?? retypes the hypotheses you may summarize.
-ALL summarizes all of the hypotheses.
-NONE resumes normal operation.

For which of the above do you wish to see additional information?
(type ? for available options).....? NONE
Are there any other zones or stages present in the target area? YES

Hypothesis: Zone 2 is the (Barren core zone)

75--I SHOULD NOT HAVE ASKED ABOUT THIS SPACE
(* There are fresh-looking feldspars) ? in Zone 2 ? SUMMARIZE

On a scale from -5 to 5,
my certainty in the type of possible deposits are:

1 - Type-A porphyry copper deposit -- PCDA. (.04505)
Current hypothesis
2 - Kuroko-type massive sulfide deposit -- MSD. (.01992)
3 - Mississippi Valley type Lead/Zinc deposit -- MVTD. (-4.0)

My certainties in the possible identities of Zone 2 are:
4 - Barren core zone. (0.0) **Current hypothesis**
5 - Potassic zone. (0.0)
6 - Propylitic zone. (0.0)
7 - Sericitic zone. (0.0)
For which of the above do you wish to see additional information?
(type ? for available options).....? ?
The following options are currently available:

-a number (or a list of numbers) referring to the above hypothesis.
-?? retypes the hypotheses you may summarize.
-ALL summarizes all of the hypotheses.
-NONE resumes normal operation.

For which of the above do you wish to see additional information?
(type ? for available options).....? 1

On a scale from -5 to 5 ,
my certainty in (PCDA) is now: .04505

In summary, the following have been observed:

The major favorable conclusions were:
- Right petro-tectonic setting. (5.0)
- Right regional environment. (3.962)

The following important factors have not yet been determined:
- Right PCDA intrusive system. (0.0)

The encouraging observations were:
- There is a granitic intrusive system in the region. (5.0)
- The throughgoing fault system is older than the intrusive system. (5.0)
- Igneous rocks in the region have porphyritic texture. (5.0)
- Igneous rocks with fine to medium grain size abound in the region. (5.0)
- There is a throughgoing fault system in the region. (5.0)
- The region contains an abundance of small stocks. (5.0)
- The region contains an abundance of dikes. (5.0)

I have insufficient information on the following observations:
- There is granodiorite in the intrusive system. (0.0)
- The intrusive system has the right age. (0.0)
- The phases in the intrusive system have the same age. (0.0)
- Sericitic zone. (0.0)

There are 4 more observations,
How many additional observations do you wish to see? 4
- Potassic zone. (0.0)
- Barren core zone. (0.0)
- Propylitic zone. (0.0)
- Right spatial relations. (0.0)

75--I SHOULD NOT HAVE ASKED ABOUT THIS SPACE
(* There are fresh-looking feldspars) ? in Zone 2 ? -5

Conclusion: my certainty in (Barren core zone) is now: -5.0

Hypothesis: Zone 2 is the (Propylitic zone)

You told me about:
- FELDSPAR. (5.0)

76--To what degree do you believe that:
(* There are partially epidotized feldspars) in Zone 2 ? -5

Conclusion: my certainty in (Propylitic zone) is now: -5.0

Hypothesis: Zone 2 is the (Sericitic zone)

77--To what degree do you believe that:
(* There are altered looking rocks) in Zone 2 ? -5
That is all I have to ask about Zone 2
My best guess for this Zone is:
- Sericitic zone. (-5.0)

Are there any other zones or stages present in the target area? NO

On a scale from -5 to 5 ,
my certainty in (PCDA) is now: .714

In summary, the following have been observed:

The major favorable conclusions were:
- Right petro-tectonic setting. (5.0)
- Right regional environment. (3.962)
- Right PCDA intrusive system. (.675)

The encouraging observations were:
- There is a granitic intrusive system in the region. (5.0)
- There is an intrusive system in the target area. (5.0)
- The throughgoing fault system is older than the intrusive system. (5.0)
- Igneous rocks in the region have porphyritic texture. (5.0)

There are 9 more observations,
How many additional observations do you wish to see? 9
- Igneous rocks with fine to medium grain size abound in the region. (5.0)
- There is quartz monzonite in the intrusive system. (5.0)
- The youngest major stock is a multiple plug or multiple dike. (5.0)
- The host rock is intermediate to silicic igneous or metamorphic. (5.0)
- There is a throughgoing fault system in the region. (5.0)
- The region contains an abundance of small stocks. (5.0)
- The region contains an abundance of dikes. (5.0)
- Potassic zone. (4.697)
- The diameter of the youngest major stock in the intrusive system is less than 5000 feet. (3.0)