ARCHER: A Prototype Expert System for Identifying Some Meteorological Phenomena

W. R. MONINGER

NOAA/Environmental Sciences Group, Boulder, Colorado

21 January 1987 and 16 June 1987

ABSTRACT

ARCHER is a computer expert system that attempts to identify meteorological phenomena from their signatures on Doppler radar. Probabilities that phenomena under study could be any one of a number of possible meteorological archetypes (i.e., stereotyped phenomena such as convective cells or gust fronts) are determined by comparing current evidence with stored meteorological knowledge.

1. Introduction

Over the next several years, the amount of data available to operational weather forecasters will increase by several orders of magnitude. Doppler radars, atmospheric profilers, and other new instruments will provide temporally and spatially dense datasets, revealing unprecedented detail to the forecaster. Research meteorologists have used these dense datasets for several years, have learned a great deal about atmospheric behavior from them, and typically take more than two years to analyze the data from a single storm. When such datasets become operationally available, a forecaster may have only a few minutes to assimilate enough of the data to produce a valid and accurate forecast. If more than a tiny fraction of the available data is to affect the forecast, computer aid must be provided.

The computer aid that has been available to forecasters in the past has been algorithmic, in the form of numerical models, meteorological indices, and statistical guidance products. Computers have not been able to provide help in the more subjective, heuristic aspects of weather forecasting; they have been unable to recognize patterns upon which expert forecasters often key their forecasts. However, if examples from other fields can serve as a guide, this limitation need no longer apply. In fields as diverse and nonalgorithmic as medicine and mineral exploration, computer expert systems have provided guidance showing substantial skill (e.g., Buchanan and Shortliffe, 1984; Duda et al., 1979).

Expert systems are nontraditional computer programs in that their program structure comprises three distinct parts:

- A "knowledge base," which is a database of factual knowledge having to do with the subject domain (for instance, meteorology).
- An "inference engine," which is a computer program that uses rules of logical inference to draw conclusions from the knowledge base.
- A "working memory" of current facts, either observed or inferred, about the specific situation under study.

(For a good basic introduction to expert systems, also called knowledge-based systems, see Winston, 1984.)

The ARCHER computer program represents an attempt to apply the technology of expert systems to the needs of weather forecasters. ARCHER is not designed as an operational tool; it was written as a prototype with which to evaluate various aspects of the technology of expert systems. The program is the result of approximately five person-months of effort. Two months were spent learning the LISP computer language, one month was spent on planning and initial coding, and two months were spent refining and evaluating.

2. Goals of ARCHER

The original and underlying goal of our work with ARCHER was to uncover the problems and advantages inherent in developing and using an expert system in the field of meteorology in general, and in weather forecasting in particular. As the design effort took shape these subgoals evolved:

- Investigate frame structures (Winston, 1984, Chapter 8) as a method of encoding meteorological knowledge. A frame structure is a database in which an attempt is made to take all knowledge relevant to a particular kind of object (such as an archetypal convective cell) and store it in one place (a "frame"). The frame has "slots" in which are stored knowledge about particular attributes of the object.

Corresponding author address: Dr. William R. Moninger, NOAA Mail Code R/E2, 325 Broadway, Boulder, CO 80303.

• Investigate the utility of Bayesian probability theory as a vehicle to translate the kind of knowledge typically gained from statistical studies into operational forecast probabilities.

• Investigate the LISP language as a programming environment, and see if it can be used effectively by a scientist accustomed to programming in Fortran.

As a computer program, the goal of ARCHER is to identify a meteorological phenomenon correctly by comparing statements about evidence observed on a radar screen with stored knowledge about meteorological archetypes. The program requests information from the user, and accepts information volunteered by the user, to make its deductions. For instance, on the basis of user input, the program might distinguish a gust front from a microburst.

3. ARCHER's knowledge base

The frame structure of ARCHER's knowledge developed from an attempt to encode the meteorological knowledge relevant to translating Doppler radar displays into meteorological archetypes, such as warm front, cold front, microburst, and supercell. Operational forecasters seem to forecast, at least in part, by comparing current data with their individual images of classically defined, or archetypal, examples of various phenomena. The archetype structure also provides a good vehicle for encoding research results, which are often presented as case studies of (presumed) archetypal examples of meteorological phenomena.

At present, ARCHER has knowledge of only about 15 archetypes, and only a limited number of attributes have been included for each one. Rather than expanding the knowledge at this stage (a process that may well take years), we have placed emphasis on developing a structure for meteorological knowledge that will be sufficiently flexible and general to be useful.

The knowledge associated with each meteorological archetype in ARCHER is stored in individual frames. Each frame (see Table 1) currently consists of six parts: 1) the name of the archetype (e.g., microburst, gust front); 2) a series of assertions having to do with the radar signature of the particular archetype (discussed in more detail below); 3) the initial priority of the archetype, indicating how potentially damaging this archetype is (microbursts have a higher initial priority than warm fronts, for instance); 4) generalizations of the archetype (e.g., an air mass boundary is a generalization of a cold front); 5) specializations of the archetype (e.g., a newgust front is a specialization of gust front—it is a gustfront with an associated convective cell); 6) the name of the author or authors of the frame for the archetype.

Each assertion in the frame refers to a particular property of the archetype, such as size or reflectivity, observable on a radar screen. Each assertion consists of five parts: 1) The name of the observable quantity to which the assertion refers, such as size, reflectivity, weak echo region. 2) A range of acceptable values for that quantity. 3) Possibly, a location to which the assertion refers, such as "at surface," "at wind shift line." 4) A set of evidence strengths, $p1$ and $p2$: $p1$ is the probability that this piece of evidence will occur given that the phenomenon is truly an example of the archetype; $p2$ is the probability that the evidence will occur if the phenomenon is *not* an example of the archetype. 5) The name of the author of this particular assertion.

As an example, the frame for the archetype "cycles cold front" is shown in Table 1.

As an example of the knowledge-gathering stage of developing ARCHER (called knowledge engineering in expert system circles), the evidence in the frame for the archetype Cycles Cold Front was generated during discussions with one of the authors of Hobbs et al. (1980). Initially, we had generated a frame based on a thorough reading of Hobbs et al., but discussion revealed that the article stressed those aspects of middle-latitude cyclones that are most interesting from a physical point of view, rather than those that are most useful

TABLE 1. Knowledge frame for archetype Cycles Cold Front.

Name: Cycles Cold Front

Assertions:

Name (Location)	Acceptable range	Evidence strength p1	Evidence strength p2	Author
Reflectivity-max	>40 dBZ	0.80	0.30	John
Altitude	linear 1 3*	1.00	0.82	Bill
Surface-front	exists	1.00	0.82	John
Wind shift line (at the surface-front)	exists	0.95	0.20	John
Low-level jet (ahead of the surface-front)	exists	0.50	0.40	John
Pbl winds (ahead of the surface-front)	veering	0.95	0.95	John
Pbl winds (behind the surface-front)	backing	1.00	0.50	John
Warm-sector rainband (ahead of the surface-front)	exists	0.50	0.10	John
Wide cold frontal rainbands (near the surface-front)	exist	0.80	0.10	John
Narrow cold frontal rainband (at the surface-front)	exists	0.95	0.30	John

Initial Priority: 3
Generalizations: Air mass boundary
Specializations: None
Author: John Doe and Bill Moninger

* An acceptable range is designated by, for example, linear 1 3. This indicates that altitudes less than 1 km are considered certain evidence against the archetype (see section 3b), altitudes greater than 3 km are considered certain evidence for the archetype, and altitudes between 1 and 3 km are considered equivocal.

in distinguishing middle-latitude cyclones from other atmospheric phenomena, such as warm fronts. The scientist who provided the information for the knowledge frame was uneasy at being asked to provide the weights $p1$ and $p2$ for individual pieces of evidence. He found it difficult to treat each piece of evidence in isolation from the others. Nevertheless, he was willing to provide a first guess at weights, under the assurance of anonymity and the assumption that the weights would be refined as the program is tested against real data.

Several psychological studies (as cited in, e.g., Stewart and McMillan, 1987) indicate that in many cases, a simple linear combination of cues, each based on an individual piece of evidence, is sufficient to describe the decision behavior of experts, even when the experts believe they are weighing interactions among several pieces of evidence. Thus, ARCHER's knowledge representation scheme, which although not linear comprises individual pieces of evidence taken in isolation from each other, may in fact be adequate to describe meteorological knowledge. However, meteorologists are not comfortable expressing their knowledge in the required form.

4. Inference engine: Probability generation

The evidence strengths and acceptable values are used to establish the probabilities for the various archetypes, using the method of Bayesian probability theory, with extensions initially developed for the Prospector expert system for mineral exploration. We provide a brief overview of that method here; for a more detailed discussion, see Duda et al. (1979) or any standard Bayesian probability text (e.g., Walpole and Myers, 1978).

Bayes' theorem provides a way of finding how the probability of a hypothesis H changes depending on the truth value of a particular piece of evidence. In what follows, we will find it easier to speak about the odds of hypothesis H, rather than the probability of H. Odds and probability are related simply:

$$O = P/(1 - P). \tag{1}$$

Thus, if the probability of hypothesis H is 0.67, the odds on H are 2 to 1, or simply 2.

Now a piece of evidence E in support of hypothesis H will not be strong evidence for H if E is nearly always true whether or not H is true. For example, "size greater than 0.2 km" would not be a very strong piece of evidence for a convective cell, because most convective cells (at least those that produce radar echoes) have sizes greater than 0.2 km. Bayes' theorem accounts for this effect by considering two weights for a piece of evidence and using them both in deducing the new odds on the hypothesis in question. The weights are $p1$, the probability that E will be true, given that H is true, and $p2$, the probability that E will be true even though H is false. Then the prior odds, $O1$, and the conditional posterior odds, $O2$, on H (conditional on E) are related by an odds factor, OF, given by

$$\mathrm{OF} = \begin{cases} p1/p2, & \text{if } E \text{ is true} \quad (2)\\ (1-p1)/(1-p2), & \text{if } E \text{ is false.} \quad (3)\end{cases}$$

Then,

$$O2 = \mathrm{OF} * O1. \tag{4}$$

This is simply Bayes' theorem written in terms of odds rather than probabilities.

In meteorology, evidence is usually not known with certainty. So in ARCHER, as in Prospector, the certainty with which evidence is known is expressed on a scale of -5 to $+5$; $+5$ indicates that the evidence is certainly true, 0 indicates that the evidence is unknown, and -5 indicates that the evidence is certainly false. This is used to scale the odds factor. The scaled odds factor, OFS, is given by

$$\mathrm{OFS} = 1 + (\mathrm{OF} - 1) * |C|/5 \tag{5}$$

where C is the certainty level and the OF is calculated from (2) or (3), depending on whether C is positive or negative, respectively. The prior and posterior odds on H are then related by

$$O2 = \mathrm{OFS} * O1. \tag{6}$$

In some cases the certainty level is supplied directly by the user. In other cases, where a physically meaningful numerical value is called for, as for size or reflectivity, the physical value that is provided by the user is translated into the -5 to $+5$ certainty scale using information stored in the knowledge frame of the archetype.

5. Inference engine: Procedures

ARCHER can query the user for information and accept volunteered information. The program accepts volunteered information because things often change rapidly in the atmosphere; if a dialog had to be restarted each time the evidence changed, the system would not be useful. (Accepting volunteered information also appears to be necessary for user acceptance. Buchanan and Shortliffe, 1984, report that physicians using the MYCIN program were annoyed that they could not volunteer information they knew would disprove a particular hypothesis.)

In answering questions or volunteering information, the user initially is describing any chosen phenomenon that is visible on the radar display screen. ARCHER considers this phenomenon to be object 0. Later in the dialog, in order to identify object 0, ARCHER may have to shift its attention to other objects at different (related) locations. ARCHER will then ask questions in an effort to identify these other, subsidiary, objects. For instance, for the archetype Cycles Cold Front, when ARCHER encounters the evidence assertion "warm-sector rainband" at the location "ahead of the surface-front," ARCHER will start asking questions in an effort to prove the identity (and existence) of a phenomenon, labeled object 1, at the location "ahead of the surface-front."

For each object, ARCHER has four stages of inference: 1) search, 2) learn, 3) process new information, and 4) conclude. These are discussed below.

a. Search

1) Find the archetype with the highest priority. The initial priority is a part of the knowledge frame of the archetype, and is an approximate measure of the a priori importance of the weather event. As facts become known, the priority is given by

$$\mathrm{PRI} = (\text{initial priority}) + 100 * (\text{current probability}) + \text{hysteresis}, \quad (7)$$

where hysteresis is set to 1 if a question was just asked about this archetype, and 0 otherwise. The hysteresis term gives the program a focus of attention so that it does not question randomly about equally likely candidate archetypes. For instance, if at some point in the dialog, both warm front and cold front have equal probabilities, ARCHER will continue to ask questions about only one of the archetypes until some other archetype becomes more probable. This focus, not necessary from a computational point of view, seems to be useful in building confidence among users (Davis and Lenat, 1982). In fact, the priority concept itself is not necessary; ARCHER could simply ask about the archetype with the highest probability. However, the priority concept provides added freedom to make ARCHER'S questions appear more intelligent (and hence more trustworthy) to the user.

2) Find the first missing piece of evidence for the archetype (if any) that can be requested from the user. Evidence cannot be asked for if it has already been asked for (even if the answer was "unknown"), or if its location refers to an object (such as a frontal surface) that is not currently known to exist.

3a) If the missing piece is not another archetype at a new location, ask the user for the evidence.

3b) If the missing piece is another archetype at a new location, shift the focus of attention of ARCHER to this new object and start over with step 1.

b. Learn

1) Add the evidence to the working memory of facts.

2) Update probabilities for all archetypes for which the evidence is relevant and update priorities as discussed in section 3b.

3) Report to the user, if requested, whenever the probabilities for the archetypes change.

4) Remove from further consideration archetypes whose probabilities drop below 0.5% (but see "resurrection" below). Declare those archetypes "impossible," and add them to the list of impossibles.

c. Process new information

The user can volunteer evidence at any time instead of answering the question that ARCHER is asking simply by typing "new," and then telling ARCHER what kind of information is to be provided. ARCHER then asks for this new information. When volunteered information is provided:

1) Perform the entire learn and process stages on the volunteered evidence.

2) If the volunteered evidence has to do with new information about a fact that previously caused an archetype to be declared impossible, resurrect that archetype by recalculating its probability on the basis of all that is relevant in the fact base. Then perform the learning and thinking stages on the resurrected archetype.

d. Conclude

Conclude the search for the identity of a given object when an archetype reaches a probability greater than or equal to 0.9, or when there are no more questions to ask. If the conclusion refers to a subsidiary object, ARCHER follows up the implications of the conclusion. For example, say that ARCHER has just concluded that an example of archetype warm-sector rainband exists at the location ahead of the surface-front, with a probability 0.7. ARCHER will convert the probability into a certainty ($C = +2$ in this case), and feed the fact

warm-sector rainband exists,
ahead of the surface-front, $C = +2$

to itself; ARCHER treats that fact just as it would had the user entered it.

6. Technical details

ARCHER was developed in Common LISP (Steele, 1984), and runs on both a Data General MV4000 computer and a MicroVax AI workstation. When the program is run interpretively, the slowest program response to an entered piece of evidence is about 7 seconds. (This usually occurs early in a program run, when few archetypes have been rejected.) In compiled form, the program runs substantially faster.

7. Summary and initial conclusions

ARCHER is a program that attempts to identify meteorological phenomena from Doppler radar signatures. Initial results from building and using the program indicate the following:

1) A knowledge base centered around frames that store knowledge about archetypal radar signatures for various phenomena seems to provide a useful structure for encoding meteorological knowledge.

2) The kind of statistical information needed to relate case study information regarding various meteorological phenomena to forecast probabilities is often not well known, but when it is known, the frame structure of ARCHER accommodates the data well.

3) It is often difficult for experts to think in terms of, and establish weights for, individual pieces of evidence for an archetype. The experts prefer to think in terms of interacting pieces of evidence.

4) The LISP language, with the kinds of programming tools that are typically available, is a productive environment, even taking into account the time it takes for a Fortran programmer to learn a new and very different language.

In developing ARCHER further, we plan to 1) make the knowledge base both broader and more thorough, 2) increase the reasoning capabilities of the program, 3) evaluate the program, or its offspring, as a training aid and as an aid to research and operational forecasters.

REFERENCES

Buchanan, B. G., and E. H. Shortliffe (Eds.), 1984: *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project,* Addison-Wesley, 748 pp.

Davis, R., and D. B. Lenat, 1982: *Knowledge-Based Systems in Artificial Intelligence,* McGraw-Hill, 490 pp.

Duda, R., J. Gaschnig and P. Hart, 1979: Model design in the Prospector consultant system for mineral exploration, *Expert Systems in the Microelectronic Age,* D. Michie, Ed., Edinburgh University Press, 334–348.

Hobbs, P. V., T. J. Metejka, P. H. Herzegh, J. D. Locatelli and R. A. Houze, Jr., 1980: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. Part I: A case study of a cold front. *J. Atmos. Sci.,* **37,** 568–596.

Steele, G. L., Jr., 1984: *Common LISP Reference Manual,* Digital Press, 465 pp.

Stewart, T. R., and C. McMillan, Jr., 1987: Descriptive and prescriptive models for judgment and decision making: Implications for knowledge engineering. *Expert Judgment and Expert Systems,* J. Mumpower, L. Phillips, O. Renn and V. R. R. Uppuluri, Eds., Springer-Verlag, (in press).

Walpole, R. E., and R. H. Myers, 1978: *Probability and Statistics for Engineers and Scientists,* Macmillan, 580 pp.

Winston, P. H., 1984: *Artificial Intelligence.* Addison-Wesley, 524 pp.