CE213 Artificial Intelligence – Lecture 9

MYCIN and Reasoning with Uncertainty

(https://en.wikipedia.org/wiki/Mycin)

MYCIN is an expert system for identifying bacterial infection and appropriate antibiotic treatment, developed at Stanford University.

How many rules are there?
How to organise them?
How to search and interpret them?
If rule matching is not certain, how to handle uncertainty?

MYCIN

A classic backward chaining (mainly) expert system developed in early 1970s by a research group led by ***Edward Feigenbaum*** at Stanford University.

The first large expert system to perform at the level of a human expert.

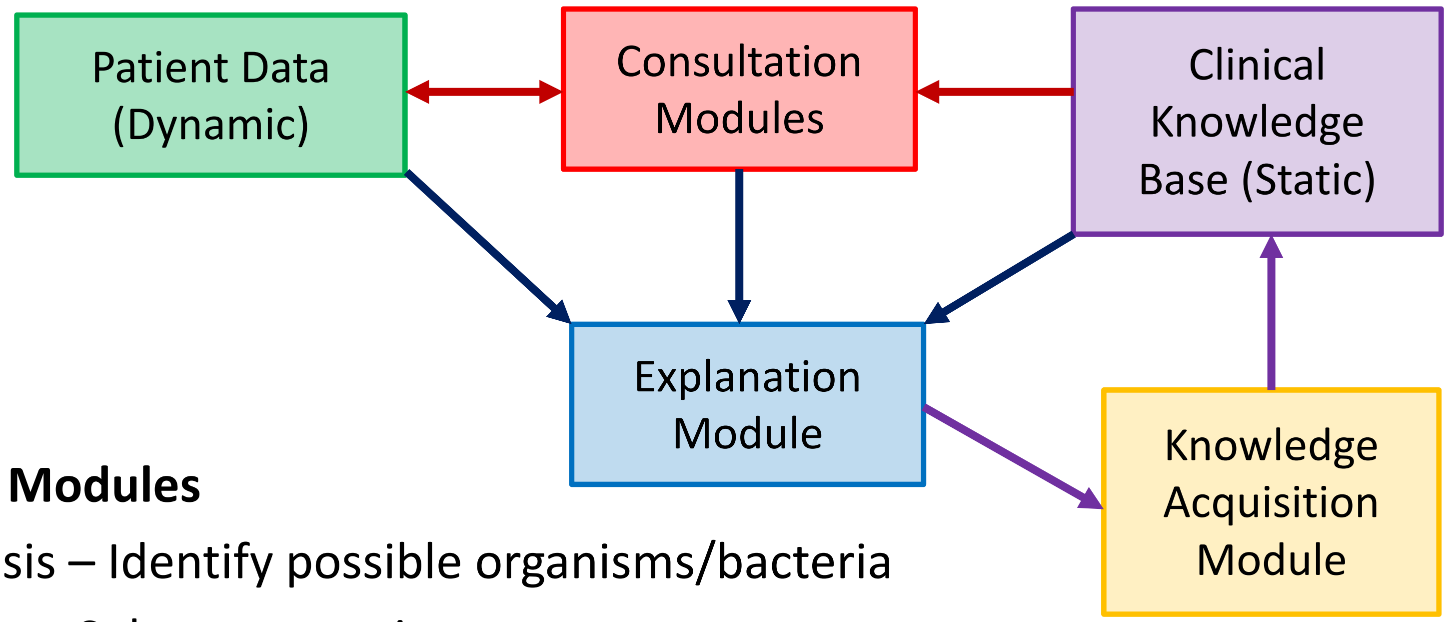
A widely adopted architecture for other expert systems.

MYCIN’s Task

Determine the most appropriate antibiotic treatment (therapy) for a patient suffering from a bacterial infection (diagnosis).

Although not a purely **diagnostic task**, identifying the bacteria responsible for the infection is the major part of MYCIN’s task.

MYCIN's Basic Architecture



Consultation Modules

1. Diagnosis – Identify possible organisms/bacteria
2. Therapy – Select appropriate treatment

Explanation Module

Provides an explanation for both the conclusions reached and the questions asked during a consultation.

Knowledge Acquisition Module

Is used to build and modify rule base.

MYCIN's Knowledge/Facts Representation

MYCIN represents most of the entities it reasons about as triplets:

(Object, Attribute, Value)

Each triplet, or an OAV, comprises an object, one of its attributes, and the current value of that attribute,

e.g., The fact that John is 50 could be represented by a triplet

(John, Age, 50)

This type of representation of facts has proved remarkably effective for most programming languages to handle.

Attributes Used in MYCIN

Original version of MYCIN used **65 attributes**, which were divided into **6 groups** in terms of the type of objects they could be applied to:

Attribute Groups	***Example Attributes***
Cultures	Site where collected
Drugs	Duration of administration
Surgical Procedures	Body cavity opened
Organisms	Morphology (form and structure)
Patients	Age
Therapies	Dosage

Attributes Used in MYCIN (2)

As part of its **static knowledge**, MYCIN stores various facts about each attribute, including:

Range of possible values.

How to ask the user for its value.

Whether the value can be obtained from a lab test.

Which rules refer to the attribute in their condition part.

Which rules may draw conclusions about the attribute.

As knowledge representation, most production rules in MYCIN establish relationships among objects, attributes, and values.

Most of MYCIN's reasoning through rule interpretation attempts to **obtain or deduce the values of unknown attributes.**

Production Rules in MYCIN

Condition part: in the form of ‘attribute of object is value’

Some relational predicates (e.g., =, >, etc.) can be used in condition expressions.

Compound conditions (e.g., conjunctions and disjunctions) can be formed using AND and OR.

Consequent/action Part: also in the form of ‘attribute of object is value’

It typically draws conclusions about the **value** of the **attribute** of an **object**. These conclusions usually take the form of evidence for or against a particular fact being true, with some **certainty**.

Production Rules in MYCIN (2)

There are about 600 production rules in MYCIN.

An Example Rule:

	(attribute)	(object)	(value)
IF	the stain of the current organism is gramneg		
AND	the morphology of the current organism is rod		
AND	the current patient is a compromised host		
THEN CONCLUDE	there is evidence with certainty 0.6 that the identity of the current organism is pseudomonas (a bacteria).		

What are the objects, attributes, and values in this rule?

MYCIN’s Rule Interpretation

The diagnosis module of MYCIN uses ***exhaustive backward chaining***.

By ‘exhaustive’ we mean that ***evidence from all matched rules,*** for and against every possible hypothesis, will be accumulated before drawing any conclusions.
(no need of conflict resolution in exhaustive backward chaining)

The therapy module of MYCIN uses ***forward chaining***.

There may be many rules whose RHSs match a hypothesis about an antibiotic treatment. However, there may be one rule only whose LHS matches a bacteria infection. Therefore, forward chaining could be more efficient for the therapy module.

Forward Chaining or Backward Chaining?

	Forward Chaining	Backward Chaining
Procedure	iterative	recursive
Need of conflict resolution	yes	no
Suitable problems or applications	where it is hard to make hypothesis or there may be too many possible hypotheses; or where there may be too many rules that match a specific hypothesis. e.g., prediction of the value of a continuous variable	where there are obvious hypotheses to make and the number of possible hypotheses is limited; or where there may be too many rules that match a specific situation but with different actions/conclusions. e.g., medical diagnosis, fault detection

Uncertainty – A Key Issue in MYCIN’s Rule Interpretation

In the medical domain, in which MYCIN operates, much of the information is uncertain.

- Facts may be believed with some degree of confidence rather than known for certain.
- Rules may lend support in favour of a hypothesis rather than confirm it definitely.

Ideally, manipulation of such uncertainties would be based on some rigorous foundations, such as **Bayesian methods** derived from probability theory.

In practice, it is not feasible to use such methods because:

- Information from medical experts is usually in vague terms like *fairly strong evidence.*
- Even if the experts could be persuaded to provide precise estimates of **conditional probabilities**, an enormous number of these would be needed to build a system using Bayesian methods.

MYCIN's Uncertainty Representation

The creators of MYCIN (researchers at Stanford University) developed their own system for **representing and manipulating uncertainty**.

It is hard to justify, but its chief merits are that it is easy to use and it works well. (Compared to Bayesian methods, it is simple and effective.)

Representing uncertain facts:

Degrees of belief in a fact or conclusion are represented by ***Certainty factors*** *(CF), r*anging from 0 (certainly false) to 1 (certainly true).

[N.B. In some applications, the range of certainty factors could be from –1 (certainly false) through 0 (no belief either way) to +1 (certainly true).**]**

MYCIN's Uncertainty Representation (2)

Drawing a conclusion through one rule with uncertainty:

To determine the certainty of a conclusion to be drawn from a rule, the certainty of the condition is multiplied by the certainty stated in the consequent part of the rule:

$$\mathbf{CF_{conclusion} = CF_{condition} \times CF_{consequent}}$$

- The certainty of a *simple condition* is just the one stated in the condition part.
- The certainty of the ***conjunction*** (AND) of two or more conditions is the ***lowest*** certainty of the individual conditions: $\mathbf{CF_{condition}} = \min(\mathbf{CF_{condition1}}, \mathbf{CF_{condition2}})$
- The certainty of the ***disjunction*** (OR) of two or more conditions is the ***highest*** certainty of the individual conditions: $\mathbf{CF_{condition}} = \max(\mathbf{CF_{condition1}}, \mathbf{CF_{condition2}})$

An Example of Rule Interpretation with Uncertainty

Consider the rule given earlier:

IF the stain of the current organism is gramneg
AND the morphology of the current organism is rod
AND the current patient is a compromised host
THEN CONCLUDE there is evidence with certainty 0.6 that the identity of the current organism is pseudomonas (a type of bacteria).

Suppose the following facts had been established:

Stain of the current organism is gramneg with certainty 0.4
Morphology of the current organism is rod with certainty 0.7
Current patient is a compromised host with certainty 0.9

Then MYCIN will conclude the organism is pseudomonas with certainty

$$0.4 \times 0.6 = 0.24 \qquad (CF_{condition}=0.4,\ CF_{consequent}=0.6)$$

In other words there is fairly weak evidence to support the hypothesis that the organism is pseudomonas.

Combining Conclusions from Different Rules

If two or more rules draw conclusions about the same fact, then in exhaustive backward chaining they must be combined.

MYCIN uses the following formula to combine the certainties of conclusions drawn from two rules:

$$\mathbf{CF_{combined} = CF_{r1} + (1 - CF_{r1}) \times CF_{r2} = CF_{r2} + (1 - CF_{r2}) \times CF_{r1}}$$

where $0 \leq \mathbf{CF_{r1}} \leq 1$ and $0 \leq \mathbf{CF_{r2}} \leq 1$.

Note that the result of combination is greater than each of the individually concluded certainties but less than their sum, i.e.,

$$\mathbf{CF_{r1} \text{ or } CF_{r2} \leq CF_{combined} \leq CF_{r1} + CF_{r2}.}$$

Also note also that no certainty produced by combining certainties, which are less than or equal to 1, could exceed 1, i.e., $0 \leq \mathbf{CF_{combined}} \leq 1$.

Combining Conclusions from Different Rules (2)

If CF_{r1} and CF_{r2} are negative or they have different signs, different formula should be used.

If $-1 \leq CF_{r1} \leq 0$ and $-1 \leq CF_{r2} \leq 0$,

$$CF_{combined} = CF_{r1} + (1 + CF_{r1}) \times CF_{r2}$$

If CF_{r1} and CF_{r2} have different signs,

$$CF_{combined} = (CF_{r1} + CF_{r2})/[1-\min(|CF_{r1}|, |CF_{r2}|)]$$

In this module, we mainly consider situations where $0 \leq CF_{r1} \leq 1$ and $0 \leq CF_{r2} \leq 1$.

(This slide is for information only. It is optional for this module.)

Another Example of Rule Interpretation with Uncertainty

Suppose we have two rules:

R1: If A AND B
Then Conclude X with certainty 0.6

R2: If C
Then Conclude X with certainty 0.5

Suppose the system has already discovered
A is satisfied with certainty 0.8
B is satisfied with certainty 0.7
C is satisfied with certainty 0.6

In attempting to establish X, the system will try both rules.

Note: A, B, C are conditions and X represents a conclusion. They are all in the form of '**attribute of object is value**'.

Another Example of Rule Interpretation with Uncertainty (2)

Starting with R1:

Certainty of B is less than that of A.

Using the rule for conjunctive conditions, A AND B has a certainty of 0.7.

Hence X is concluded with certainty 0.7 x 0.6 = 0.42

Then using R2:

Certainty of C is 0.6.

Hence X is concluded with certainty 0.6 x 0.5 = 0.3

Thus we have two different values for the certainty of X and these must be combined.

Another Example of Rule Interpretation with Uncertainty (3)

The formula is

$$CF_{combined} = CF_{r1} + (1 – CF_{r1}) \text{ x } CF_{r2}$$

Using the conclusion of R1 gives us $CF_{r1} = 0.42$

The conclusion of R2 gives us $CF_{r2} = 0.3$

So the certainty of X derived from both rules is

$$CF_{combined} = 0.42 + (1 – 0.42) \text{ x } 0.3 = 0.42 + 0.58 \text{ x } 0.3 = 0.42 + 0.174 = \mathbf{0.594}$$

Thus as a result of using both rules X now has a certainty of 0.594.

Exactly the same result would have been obtained if we use

$$CF_{combined} = CF_{r2} + (1 – CF_{r2}) \text{ x } CF_{r1} = \mathbf{0.594}$$

In the next class, there will be an exercise of reasoning with uncertainty using exhaustive backward chaining, with a whole set of production rules.

Explanation Generation (mainly for self study)

MYCIN is able to answer two types of question:

WHY: Why the system asked the user a particular question.

HOW: How the system reached a particular conclusion.

As the consultation proceeds, MYCIN maintains a record of all the rule activations, indicating which rules were fired in an attempt to satisfy the conditions of which other rules (rule chaining).

This record will form **a tree of hypotheses and subsidiary hypotheses**.

- To answer a *WHY* question, the system must **look up** the tree from the current position to identify the hypotheses that led to the current question.
- To answer a *HOW* question, the system must **look down** the tree from the current position to identify the evidence that led to the current conclusion.

(only possible with backward chaining)

(This slide is for information only. It is optional for this module.)

Explanation Generation (mainly for self study) (2)

Such explanation facilities are easy to implement and produce two benefits:

- They increase the user’s confidence in the correctness of the conclusions.
- They are invaluable in debugging a system that draws erroneous conclusions.

(This slide is for information only. It is optional for this module.)

Summary

MYCIN's basic architecture

Knowledge/fact representation in MYCIN – OAV triplets

MYCIN's production rule format

MYCIN's production rule interpreters

MYCIN's system for representing and manipulating uncertainty

Reasoning with uncertainty with the MYCIN's system

Explanation generation in MYCIN