Rule-Based Systems

6.871-- Lecture 6

Outline

- MYCIN
  - Introduction
    » Task
    » Sample Interaction
    » Reasons for Success
  - How It Worked
    » Knowledge Representation
    » Architecture and Control Structure
    » Inexact Inference
    » Explanation
    » Interaction Help
  - Reasons for Success Revisited
  - Contributions
- Prospector
- "Rules Glasses"

MYCIN

- What was the task?
  - Medical diagnosis and therapy selection for bacterial infections of the blood

- Why this domain?
  - Overuse of antibiotics

Why This Domain

- Overuse of antibiotics

- Irrational use of antibiotics

- Maldistribution of expertise

- Domain is "small" and "isolated"

Sample Interaction: Clinical Data

```
---------PATIENT-248----------
1) Patient's name: (first-last)
**   CR
2) Sex:
**    MALE
3) Age:
**    52
4) Have you been able to obtain positive microbiological information
     about a possible infection of C.R.?
**   YES

---------INFECTION-------------

5) What is the infection?

**   ENDARTERITIS
```

Sample Interaction (cont'd)

```
The first significant organism from this blood culture (CULTURE-1)
  will be referred to as:
--------------ORGANISM-1-----------------
6) Enter the identity of ORGANISM-1
**    UNKNOWN
7) Is ORGANISM-1 a rod or coccus (etc.):
**    ROD
8) The gram stain of ORGANISM-1
**    GRAMNEG
9) HAVE THERE BEEN positive cultures yielding organisms
     about which you will NOT be seeking advice?
** YES
```

Sample Interaction (cont'd)

9) Enter the identity of ORGANISM-2:

** SALMONELLA (.4)   E.COLI (.3)

Sample Interaction: Explanation

Is the patient's illness with ORGANISM-1 a hospital-acquired infection?
** WHY
   [i.e. WHY is it important to determine whether or not the infection with
   ORGANISM-1 was acquired while the patient was hospitalized]
[2.0] This will aid in determining the identity of ORGANISM-1.
   It has already been established that
   [2.1] the morphology of ORGANISM-1 is rod
   [2.2] the gram stain of ORGANISM-1 is gramneg
   [2.3] the aerobicity of ORGANISM-1 is facul
Therefore, if
   [2.4] the infection with ORGANISM-1 was acquired while
          the patient was hospitalized
   then
          there is weakly suggestive evidence (.2) that the identity
          of ORGANISM-1 is pseudomonas [RULE050]

Sample Interaction: Diagnosis

INFECTION-1 is ENDARTERITIS with BACTEREMIA
<Item 1>   E. COLI
<Item 2>   SALMONELLA (species unknown)
<Item 3>   KLEBSIELLA-PNEUMONIAE
<Item 4>   PSEUDOMONAS-AERUGINOSA
<Item 5>   ENTEROBACTER
<Item 6>   PROTEUS-NON-MIRABILUS

Sample Interaction: Therapy

[REC 1] My preferred therapy recommendation is as follows:
   In order to cover for Items <1 3 4 5 6 >:
       Give: GENTAMICIN
       Dose:128 mg (1.7. Mg/kg) q8h IV (or IM) for 10 days
       Comments: Modify dose in renal failure
   In order to cover for Item <2>:
       Give: CHLORAMPHENICOL
       Dose:563 mg (7.5 mg.kg) q6h for 14 days
       Comments: Monitor patient's white count
Do you wish to see the next choice therapy?
**   NO

How It Worked: Representation

Rule 27

If:
  1)the gram stain of the organism is gram
  negative, and
  2)the morphology of the organism is rod, and
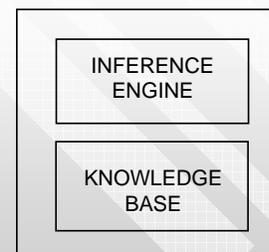 **3)the aerobicity of the organism is**
  anaerobic,

Then:
  There is suggestive evidence (.7) that the
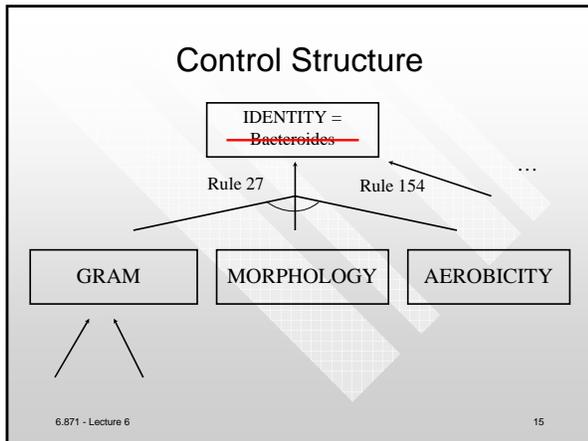 **identity of the organism is Bacteriodes.**

Architecture

INFERENCE
ENGINE

KNOWLEDGE
BASE

Control Structure

IDENTITY =
~~Bacteroides~~

Rule 27   Rule 154   …

| GRAM | MORPHOLOGY | AEROBICITY |

Control Structure

- What kind of search?
  - Backward-chaining, depth-first

- With what variations?
  - Generalized subgoals
    E.g. "determine identity", not "is identity = E.coli"
  - Every rule relevant to goal is explored
  - Inexact inference

Inexact Inference

R1: A and B ➡ X (.4)
R2: C and D ➡ X (.3)

A and B and C and D ➡ ?

Desiderata:
  - Commutative
  - Asymptotic
  - Balances positive and negative evidence

Inexact Inference

R1: A and B ➡ X (.4)
R2: C and D ➡ X (.3)

A and B and C and D ➡ ?

A and B ➡ X (.4)
C and D:

"increase certainty .3 more from current value"

$X(.4 + .3(1 - .4)) = X(.58)$

current value   new value   remaining uncertainty

Inexact Inference

R1: A (.5) and B (.6) ➡ X (.4)
R2: C (.4) and D (.7) ➡ X (.3)

A and B and C and D ➡ ?

AND: min (A and B) = .5
        min (C and D) = .4
R1: .5 * .4 = .2
R2: .4 * .3 = .12
Combining: .2 + .12(1 - .2) = .30

current value   new value   remaining uncertainty

Explanation

** HOW DID YOU DECIDE THAT ORGANISM-1 WAS AN E.COLI?

I used rule084 to conclude that the identity of ORGANISM-1 is e.coli. This gave a cumulative c.f. of (.51).

I used rule003 to conclude that the identity of ORGANISM-1 is e.coli. This gave a cumulative c.f. of (.43).

Explanation

**   WHY DIDN'T YOU THINK THAT ORGANISM-1 WAS NEISSERIA?

The following rules could have been used to determine that the identity of ORGANISM-1 is neisseria: 9.

However, none of these succeeded in the context of ORGANISM-1. If you would like an explanation for why any of these rules failed, please enter their numbers:

** 9

Clause 2 of rule009 ["the morphology of the organism is coccus"] was already known to be false for ORGANISM-1, so the rule was never tried.

6.871 - Lecture 6   21

Successful?

- Appropriate task: Narrow domain, heuristic knowledge, articulate experts, etc.
- The match of knowledge and representation:
  - Knowledge about of how the body works, vs.
  - Knowledge about how to diagnose it
- Answer was a ranked list, not a single result
- Perhaps most important:
  Having the right abstractions, the right language

6.871 - Lecture 6   22

The Right Abstraction

```
If …
  the organism is gram-negative, and
  the portal of entry is skin-wound,
Then the organism is likely to be …
```

- It's an abstraction
- It's the right abstraction for this task
- Where did it come from?

6.871 - Lecture 6   23

What ought a rule to be?

Answer:
An *independent*
single-step
inferential
empirical association.

6.871 - Lecture 6   24

What ought a rule to be?

- Independent
  - Does it make sense in isolation?
- Single step: mental hygiene
- Inferential: focus on the knowledge level
- Empirical association
  - Association: symptom & disease
  - Empirical: noticed but not understood
  - Logic-like, but:
    » Inexact
    » Not truth-preserving

6.871 - Lecture 6   25

What Ought a Rule to Be?

IF   There is pressure on project leader,
         Respondent is R&D manager,
         No increase in growth stage of life cycle,
         Increased probability of commercial success,
         Project champion did not appear at end,
         Product not in infancy stage of life cycle,
         Top management support,
         Association between commercial and technological aspects,
         R&D perceives project mgmt commitment as high,
         There is a project champion,
         Don't know about newly enacted favorable international regulations,
         Project champion appeared in the middle,
         Respondent is not the VP,
         Project Champion didn't appear at beginning,
         Respondent is not marketing manager,
THEN   the Project is likely to succeed.

From Gallant and Balachandra, Using automated techniques to generate an expert system for R&D project monitoring, *First IFAC Intl Symposium on Economics and AI*, Sept. 1986.

6.871 - Lecture 6   26

What Ought a Rule to Be?

IF   the stock-price is determined, and
         the stock-earnings is determined,
THEN   the price-earnings-ratio is stock-price/stock-earnings.

IF   the patient is between 16 and 21, and
         the chief complaint is fatigue, and the
         patient has a mild fever
THEN   the disease is likely (.8) to be mono.

Contributions

- Existence proof of adequacy of rule-based systems

- Knowledge can be captured as a set of mostly independent rules

- Knowledge can obviate search: looking through the right space

- Experts can be debriefed

Contributions

- The adequacy of a simple model of uncertainty.

- Explanation as a core property of KBS

- Preview of multiple uses of knowledge

- Meta-rules and meta-knowledge

Prospector

- Like Mycin:
  - System creation as a forcing function for making knowledge explicit (hence accessible, testable, communicable, etc.)

- Unlike Mycin:
  - Network view of a knowledge base (Lecture 7)

  - An alternate model for inexact reasoning: Bayesian probability theory (Lecture 10)

What are the "Rules Glasses?"

- An individual inference rule is the appropriate unit of modularity.
- Information can be captured as single independent statements, single independent inferences.
- Expertise is the accumulation of good guesses derived from experience.
- Expertise is can be captured as a set of informal, heuristic inferences, each of which is a simple if/then rule.

What are the "Rules Glasses?"

- Capture the expertise by dissecting it into individual inferences.
- Impose your ontology by choosing your predicates carefully.
- The choice of language is an important source of power: the right choice of language vastly reduces search.

6.871 - Lecture 6   38

687.1 Lecture 6   6