The USL/DBMS NASA/RECON Working Paper Series contains a collection of reports representing results of activities being conducted by the Computer Science Department of the University of Southwestern Louisiana pursuant to the specifications of National Aeronautics and Space Administration Contract Number NASW-3846. The work on this contract is being performed jointly by the University of Southwestern Louisiana and Southern University.

For more information, contact:

Wayne D. Dominick

Editor
USL/DBMS NASA/RECON Working Paper Series
Computer Science Department
University of Southwestern Louisiana
P. O. Box 44330
Lafayette, Louisiana 70504
(318) 231-6308
KNOWLEDGE BASED SYSTEMS:
A CRITICAL SURVEY OF MAJOR CONCEPTS,
ISSUES AND TECHNIQUES:
VISUALS

Srinu Kavi

Center for Advanced Computer Studies
University of Southwestern Louisiana
P. O. Box 44330
Lafayette, Louisiana 70504

December 11, 1984
KNOWLEDGE BASED SYSTEMS:
A CRITICAL SURVEY OF MAJOR CONCEPTS,
ISSUES, AND TECHNIQUES

SRINU KAVI
FALL 1984
OBJECTIVES

- To examine various issues and concepts involved in KBSs

- To examine various techniques used to build KBSs

- To examine (at least one) KBS in detail (i.e., case study)

- To list and identify limitations and problems with KBSs

- To suggest future areas of research

- To provide extensive references
TABLE OF CONTENTS

1. INTRODUCTION

2. KNOWLEDGE-BASED SYSTEMS (KBSs)
   - Introduction
   - Hypothetical KBS
   - KBS Components

3. TECHNIQUES USED TO CONSTRUCT KBSs
   - Introduction
   - Methods of Representing Knowledge
   - Inference Engine
   - Workspace Representation
   - The Interface

4. KBS BUILDING TOOLS AND LANGUAGES
   - General Purpose Programming Languages
   - Skeletal Systems
   - General Purpose Representation Languages
   - Computer-Aided Design Tools
   - Case Studies
5. APPLICATION CONSIDERATIONS

Initial Considerations
Technology Considerations
Environmental Considerations

6. CONCLUSIONS

7. POTENTIAL FUTURE RESEARCH AREAS

APPENDICIES

A. CASE STUDY - MYCIN
B. LIST OF KBSs
C. FIFTH GENERATION PROJECT

REFERENCES
CHARACTERISTICS OF KBSs

- Organization of Knowledge
- Performance
- Utility (or Usefulness)
- Understandability (or Explainability)
- Heuristics
- Uncertainty
- Flexibility
- Modularity
FIGURE 2-1. KBS ELEMENTS AND THEIR RELATIONSHIP
BASED ON [HAYES-ROTH, ET AL, '83] AND [BARNETT & BERSTEIN, '77]
TECHNIQUES USED TO CONSTRUCT KBSs

Introduction

Origins of KBS Techniques
Choices and Restrictions
Knowledge Representation Problems
Knowledge Representation Forms
Knowledge Representation Unit
Credibility Factors
Procedural Versus Declarative Representation

Methods of Representing KB

Finite-State Machine Programs
Predicate Calculus
Production Rules
Semantic Networks
Frames
Inference Engine (IE)

Primary Functions of IE
Some Definitions
IE Strategies
Methods of Implementing the IE
Measures of Performance

Workspace Representation

Introduction
HEARSAY-Blackboard
AND/OR Graph
Blackboard Versus AND/OR Graph

The Interface

Functions of the Interface
User Interface
Expert Interface
Knowledge Acquisition (KA) Process
Table 3-1 ORIGINS OF KBS TECHNIQUES
(Based on [Barnett & Bernstein, 77])

ARTIFICIAL INTELLIGENCE (AI)

Heuristic Search
Inference and Deduction
Pattern Matching
Knowledge Representation and Acquisition
System Organization

LANGUAGE PROCESSING

Parsing and Understanding
Question and Response Generation
Knowledge Representation and Acquisition

THEORY OF PROGRAMMING LANGUAGES

Formal Theory of Computational Power
Control Structures
Data Structures
System Organization
Parsing
MODELING AND SIMULATION

Representation of Knowledge
Control Structures
Calculation of Approximations

DATA BASE MANAGEMENT

Information Retrieval
Updating
File Organization

SOFTWARE ENGINEERING

System Organization
Documentation
Iterative System Development

APPLICATION AREAS

Domain-Specific Algorithms
Human Engineering
Figure 3-1. Restrictions on choices of KBS methodologies based on [Barnett & Berstein, '77].
FIGURE 3-2. KNOWLEDGE REPRESENTATION FORMS
BASED ON [BARNETT & BERNSTEIN, '77]
KNOWLEDGE REPRESENTATION METHODS

- Finite state machines
- Programs
- Predicate calculus
- Production rules
- Semantic networks
- Frames

Representation = Knowledge + Access

[Newell, 82]
PRODUCTION SYSTEM COMPONENTS

Three parts [Barr & Feigenbaum, 81]:

- A Workspace: A buffer like data structure.
- An Interpreter: Which controls the system activity.
INTERPRETER TASKS

- Matching or Building a Conflict-Set

- Conflict-Resolution

- Action or Execution

CONFLICT RESOLUTION STRATEGIES

- Rule Order

- Rule Precedence

- Generality Order

- Data Order

- Regency Order

- Non-Deterministic
AN EXAMPLE

Automotive Repair Agency

The System Contains

- Knowledge Base of production rules
  (Performance characteristics and Measurable attributes)

- A Database
  (Past problems, Repairs, and Service performed)
R1  IF fan belt tension is low
    THEN alternator output will be low [.5]
    and engine will overheat [.2]

R2  IF alternator output is low
    THEN battery charge will be low [.7]

R3  IF battery is low
    THEN car will be difficult to start [.5]

R4  IF automatic choke malfunctions OR
    automatic choke needs adjustment
    THEN car will be difficult to start [.8]

R5  IF battery is out of warranty
    THEN battery charge may be low [.9]

Figure 3-9  PRODUCTION RULES FOR AUTOMOTIVE SYSTEMS KS
R6  IF coolant is lost OR coolant system pressure cannot be maintained
    THEN engine will overheat [.7]

R7  IF there is a high resistance short
    AND fuse is not blown
    THEN battery charge will be low [.8]

R8  IF battery fluid is low
    THEN battery will boil off fluid [.3]

R9  IF battery fluid is low
    THEN battery charge will be low [.4]

Figure 3-9. PRODUCTION RULES FOR AUTOMOTIVE SYSTEMS KS (CONT'D)
<table>
<thead>
<tr>
<th>OBSERVATIONS</th>
<th>AGENT</th>
<th>DIFFICULTY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternate Output Level</td>
<td>Mech</td>
<td>4</td>
</tr>
<tr>
<td>Battery Charge Level</td>
<td>Mech</td>
<td>3</td>
</tr>
<tr>
<td>Battery Fluid Level</td>
<td>SrvR</td>
<td>2</td>
</tr>
<tr>
<td>Choke Adjustment</td>
<td>Mech</td>
<td>5</td>
</tr>
<tr>
<td>Choke Function</td>
<td>Mech</td>
<td>5</td>
</tr>
<tr>
<td>Coolant Level</td>
<td>SrvR</td>
<td>2</td>
</tr>
<tr>
<td>Coolant System Pressure</td>
<td>Mech</td>
<td>5</td>
</tr>
<tr>
<td>Difficulty to Start</td>
<td>Cust</td>
<td>1</td>
</tr>
<tr>
<td>Engine Temperature</td>
<td>Cust</td>
<td>1</td>
</tr>
<tr>
<td>Fan Belt Tension</td>
<td>Mech</td>
<td>3</td>
</tr>
<tr>
<td>Fuse Condition</td>
<td>SrvR</td>
<td>2</td>
</tr>
<tr>
<td>Short in Electric System</td>
<td>Mech</td>
<td>8</td>
</tr>
<tr>
<td>Voltage Regulator Level</td>
<td>Mech</td>
<td>4</td>
</tr>
<tr>
<td>Warranties</td>
<td>Data Base</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 3-10. DATA GATHERING PROCEDURE
FACT FILE
DIFFICULT TO START

R4

CHOKE MALFUNCTION

R4

CHOKE OUT OF ADJUSTMENT

R7 &

FUSES NOT BLOWN

R7

SHORT HIGH RESISTANCE

R3

BATTERY CHARGE LOW

R5

BATTERY OUT OF WARRANTY

R9

BATTERY FLUID LOW

R8

VOLTAGE REGULATOR OUTPUT HIGH

R1

ALTERNATOR OUTPUT LOW

R1

LOW FAN BELT TENSION

R1

HIGH ENGINE TEMPERATURE

AGENTS

MECH(5)

MECH(5)

SRVR(2)

MECH(8)

MECH(3)

DB(0)

SRVR(2)

MECH(4)

MECH(4)

MECH(3)

CUST(1)

FIGURE 3.11 EXAMPLE FLOW IN AUTO DIAGNOSTIC SYSTEM
FIGURE 3-13. FRAGMENT OF GRAPH STRUCTURE
FIGURE 3-14. CHARACTERISTICS OF PRODUCTION SYSTEMS
BASED ON [BARNETT & BERNSTEIN, '77]
Semantic networks are used in

- Psychological modeling of human memory
- Programming languages
- Natural language understanding
- Database management systems

A SEMANTIC NETWORK (or NET) consists of nodes and links.
RELATIONS

TEMP(WARM-BLOODED MAMMAL)
ISA(DOG, MAMMAL) ISA(CAT, MAMMAL)
ISA(FIDO, DOG) ISA(BOWSER, DOG) ISA(PUFF, CAT)
LOC(MARY'S, FIDO) LOC(FIREHOUSE, BOWSER) LOC(BOB'S, PUFF)
COLOR(TAN, FIDO) COLOR(TAN, BOWSER) COLOR(BLACK, PUFF)
SIZE(40LB, FIDO) SIZE(14LB, BOWSER) SIZE(4LB, PUFF)
BETWEEN(MARY'S, FIREHOUSE, BOB'S)

SEMANTIC NETWORK

RULES OF INference

ISA(X, Y) ^ ISA(Y, Z) => ISA(X, Z)
SIZE(X, Y) ^ SIZE(U, V) ^ X<U => SMALLER(Y, V)
ISA(X, Y) ^ R(U, Y) => R(U, X)

FIGURE 3.15 EXAMPLE SEMANTIC NETWORK
INFERENCES

First Rule

PUFF is a cat and CAT is a MAMMAL; therefore, PUFF is a MAMMAL.

Second Rule

SIZE(4, PUFF) & SIZE(14, BOWSER) & 4 < 14
=> SMALLER(PUFF, BOWSER)

Third Rule

ISA(FIDO, DOG) & ISA(DOG, MAMMAL)
=> ISA(FIDO, MAMMAL)

ISA(FIDO, MAMMAL) &
TEMP(WARM_BLOODED, MAMMAL) =>
TEMP(FIDO, WARM_BLOODED)
MEANINGLESS INFERENCES

\[ \text{ISA(DOG, MAMMAL)} \ & \ \text{ISA(CAT, MAMMAL)} \]
\[ \Rightarrow \ \text{ISA(DOG, CAT)} \]

\[ \text{INHERITABLE(TEMP)} \]

\[ \text{ISA(x, y)} \ & \ \text{r(u, y)} \ & \]
\[ \text{INHERITABLE(r)} \Rightarrow \ r(u, x) \]
CURRENT RESEARCH

- What does a node (object) really mean?

- Is there a unique way to represent an idea?

- How is the passage of time to be represented?

- How does one represent things that are not facts about the world but rather ideas or beliefs?

- What are the rules about inheritance of properties in networks?
FRAMES

Frame Characteristics

- Description
- Instantiation
- Prediction or Expectation
- Justification
- Variation
- Correction
- Perturbation
- Transformation
dog FRAME ISA mammal
kind breed
color SUBSET.OF {tan brown black white rust}

FROM color OF kind
leggedness 0...4
weight '0, FROM size OF kind
state adult OR puppy if age < 1
age '0, now birthday
birthday date
name string

END dog

(a)

Figure 3-16. EXAMPLE FRAME DEFINITIONS
[Barnett & Bernstein, 77]
boxer FRAME ISA breed OF dog
\[2\]
color ONE. OF \{tan brown brindle\}
\[3\]
size \quad 40...60
\[4\]
tail \quad bobbed OR long
\[5\]
ears \quad bobbed OR floppy
\[6\]
temperment \quad playful
\[7\]
COMPLAINTS IF weight > 100 THEN ASSUME (great dane)
\[8\]
END boxer
\(b)\)

Figure 3-16. EXAMPLE FRAME DEFINITIONS (CONT'D) [Barnett & Bernstein, 77]
LOW-LEVEL INFORMATION

OBJECT . 654

color = tan
ears = bobbed
leggedness = 4
size = 40 - 45
temperment = mean

TRIAL IDENTIFICATION

[OBJECT 654 ISA dog

kind boxer WITH [color tan
size 40 - 45
tail ASSUMED bobbed
ears bobbed
temperment EXCEPTIONAL
mean]

color tan
leggedness 4
weight 40 - 45
state ASSUMED adult]

Figure 3-17. INEXACT MATCH BY A FRAME SYSTEM
[Barnett & Bernstein, 77]
INFERENCE ENGINE CONTROL STRATEGIES

- Forward chaining
- Backward chaining
- Chain both ways
- Middle term chaining
- Fixed directionality
- Variable directionality
- Hybrid strategy
- Breadth-first
- Depth-first
FIGURE 3.18 CHAINING EXAMPLES
An Example: 8-Puzzle

![Diagram of 8-Puzzle states a and b]
Figure 3-21. The tree produced by a breadth-first search based on [Nilsson, '71]
FIGURE 3-22. DEPTH-FIRST BACK CHAINING
Figure 3-23. The tree produced by a depth-first search based on [Nilsen, '71]
METHODS OF IMPLEMENTING THE IE

- Search Methods
- Simulation Methods
- Pattern Matching

SEARCH SYSTEM COMPONENTS

Five major components

- Select
- Expand
- Evaluate
- Prune
- Terminate
EVALUATION FUNCTION

"The purpose of an evaluation function is to provide a means for ranking those nodes (activities) that are candidates for expansion to determine which one is most likely to be on the best path to the goal" [Nilsson, 71].
A* - AN OPTIMAL SEARCH ALGORITHM

In A*, the evaluation function, $f'(x)$ is the cost of a solution path constrained to go through node $x$; $f'$ should be minimized.

$$f'(n(i)) = \sum_{j=1}^{m-1} K(n(j), n(j+1)) \quad 1 \leq i \leq m$$

$$f'(n) = f'(\text{start}, n) + f'(n, \text{goal})$$

$$f'(n) = g(n) + h(n)$$

$$f(n) = g'(n) + h'(n).$$
MEASURES OF PERFORMANCE

Search Techniques

Penetrance

\[ P = \frac{L}{T} \]

\( L \) length of the derived path from initial to goal
\( T \) total number of nodes

Branching Factor
FIGURE 3-27. EXAMPLE MOVE GRAPH AND BALANCED TREE

\[ T = 15 \]
\[ L = 3 \]
\[ P = 1/5 \]
\[ B = 2 \]
WORKSPACE REPRESENTATION

- Plan
- Agenda
- History
- Solution Set

TWO METHODS

- HEARSAY Blackboard
- AND/OR Graph
HEARSAY BLACKBOARD

A data structure

- Hypotheses and support criteria stored

- Intermediary between KSs and IE
FIGURE 3-28. HEARSAY II LEVELS OF REPRESENTATION* AND KNOWLEDGE SOURCES BASED ON [ERMAN, ET AL., '80]
USER INPUT

Parsing Strategies

- Backtracking Versus Parallel Processing
- Top Down Versus Bottom Up Processing
- Choosing How to Expand or Combine
- Multiple Knowledge Sources
PARSING SYSTEMS

- Template matching
- Transition networks
- Semantic grammar parsers
TEMPLATE MATCHING

E.g., ELIZA, SIR, STUDENT

$1 x(i) \{ IS/ARE \} NOT$ 2

WHAT IF $x(i)$ WERE $2$ ?

"Today's temperature is not hot"

"What if temperature were hot?"
RECURSIVE TRANSITION NETWORKS

E.g., "The little boy in the swimsuit kicked the red ball"

NP: The little boy in the swimsuit

PP: in the swimsuit

NP: the swimsuit

Verb: kicked

NP: the red ball
FIGURE 3-32. A RECURSIVE TRANSITION NETWORK
BASED ON [BARR & FEIGENBAUN, '81]
AUGMENTED TRANSITION NETWORKS

ATN --> RTN extended in three ways

- Registers
- Tests
- Actions
DIFFICULTIES IN EXPLANATIONS

Explanations

- Must be in terms of
  Knowledge chunks
  Problem parameters
  Inference rules

- Must be translated to human understanding

METHODS OF PROVIDING EXPLANATIONS

- Workspace Representation
- Using Knowledge Source(s)
- Re-solve the Problem
Figure 3-33. Stages of Knowledge Acquisition
Based on [Hayes-Roth, et al., '83]
DIFFICULTIES IN KA

- Representational mismatch
- Verbalization by the expert (Protocol study)
- Limitations on current technology

KA bottleneck
KBS BUILDING TOOLS AND LANGUAGES

- General purpose programming languages
- Skeletal systems
- General purpose representation languages
- Computer-aided design tools for KBSs

Case Studies

- EMYCIN
- HEARSAY-III
- AGE
INITIAL CONSIDERATIONS

- Task suitability
- Availability of expert
- KA process
- Agreement with the domain theory
- Expert's model
- Expert's principles of reasoning
- Intermediate levels of abstraction
- General versus domain specific knowledge
- End users
- Unanticipated support
- Cost versus benefits
TECHNOLOGY CONSIDERATIONS

- Building the prototype system
- Chunk size
- Representation of knowledge
- Inference engine
- Meta knowledge
- Procedural knowledge
- Addition of knowledge by the users
- Extensibility
- Knowledge representation tools
- Design of tools
KNOWLEDGE REPRESENTATION TOOLS

- Generality
- Appropriateness
- Accessibility
- Explanation/Interaction

- Problem characteristics versus Tool features
DESIGN OF TOOLS FOR BUILDING KBSs

- Generality
- Completeness
- High-level representation language
- Explanation/interaction facilities
- Data representation
- Control structure
ENVIRONMENTAL CONSIDERATIONS

- Interactive KBSs
- Interactive development
- Local operating environment
SUMMARY
CONCLUSIONS

- Wide spectrum of application areas

- Highly successful

- Some systems are being used routinely (DENDRAL, MYCIN, R1, PROSPECTOR)

- Not yet commonly understood (Few "data points")

- Major motivations

  Replication/Distribution expertise

  Union of expertise

  Documentation
- Building ESs expensive and time-consuming ($1-2 million; 5 person-years with tools)

- General level of accomplishment high

- Number of unresolved issues

- Difficulties and potential risk

- High expectations/misunderstandings
POTENTIAL FUTURE RESEARCH AREAS

- Knowledge acquisition
- Representation theory
- Comparison of techniques
- KBS building tools
- Explanation
- Evaluation
- Parallel processing
- Learning from experience
- Management of knowledge
- Abstraction and hierarchies
- Technological innovations
- Uniform terminology

This report represents one of the 72 attachment reports to the University of Southwestern Louisiana's Final Report on NASA Grant NGT-19-010-900. Accordingly, appropriate care should be taken in using this report out of the context of the full Final Report.