U S L / D B M S   N A S A / R E C O N

W O R K I N G   P A P E R   S E R I E S

Report Number

DBMS.NASA/RECON-10

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

| N A S A |

| N A S A |

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

| DBMS.NASA/RECON-10 |

| WORKING PAPER SERIES |

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 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

CHARACTERISTICS OF KBSs

- Organization of Knowledge
- Performance
- Utility
  (or Usefulness)
- Understandability
  (or Explainability)
- Heuristics
- Uncertainity
- Flexibility
- Modularity

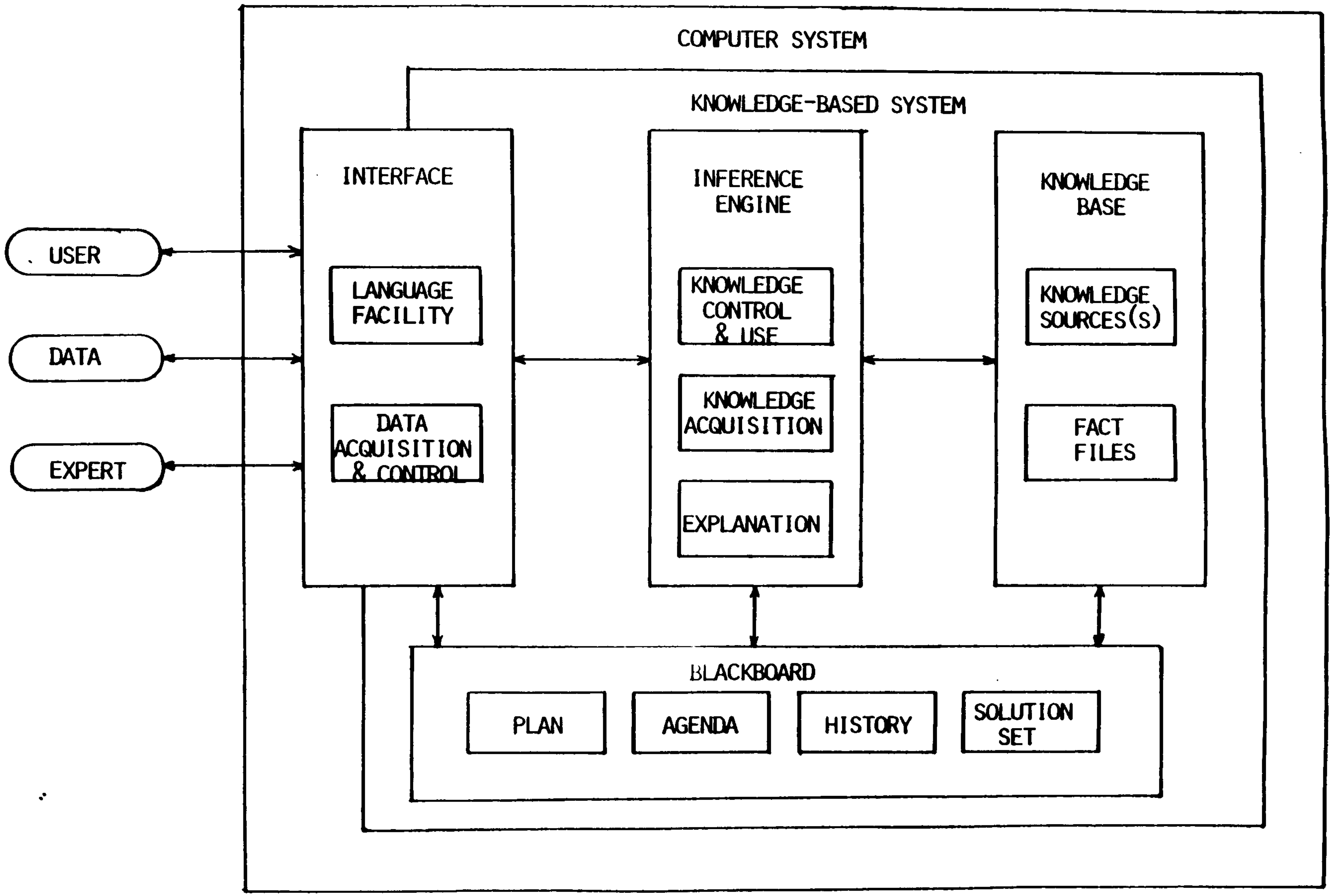


FIGURE 2-1. KBS ELEMENTS AND THEIR RELATIONSHIP
BASED ON [HAYES-ROTH, ET AL, '83] AND [BARNETT & BERSTEIN, '77]

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

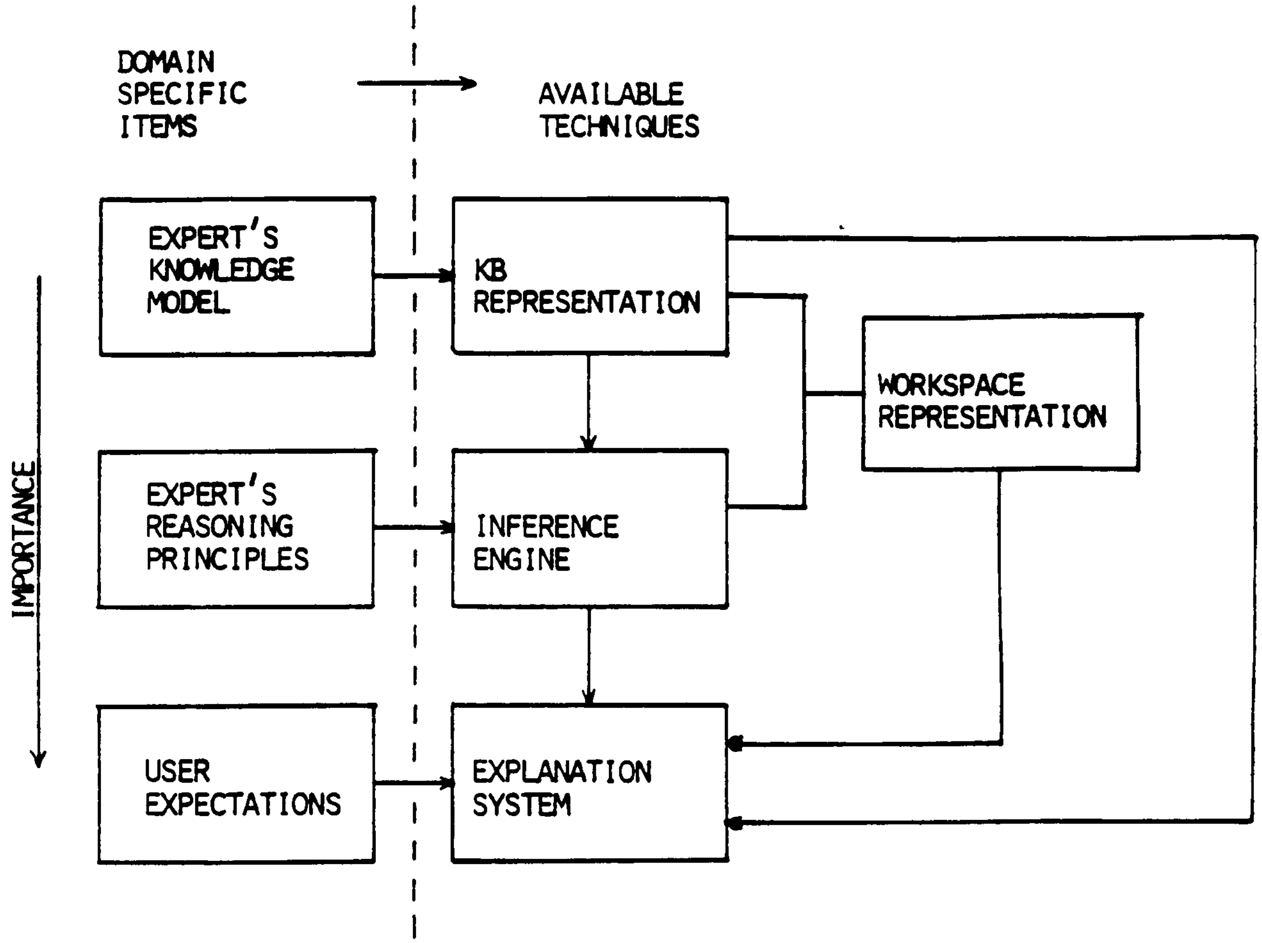


FIGURE 3-1. <u>RESTRICTIONS ON CHOICES OF KBS METHODOLOGIES</u>
BASED ON [BARNETT & BERSTEIN, '77]

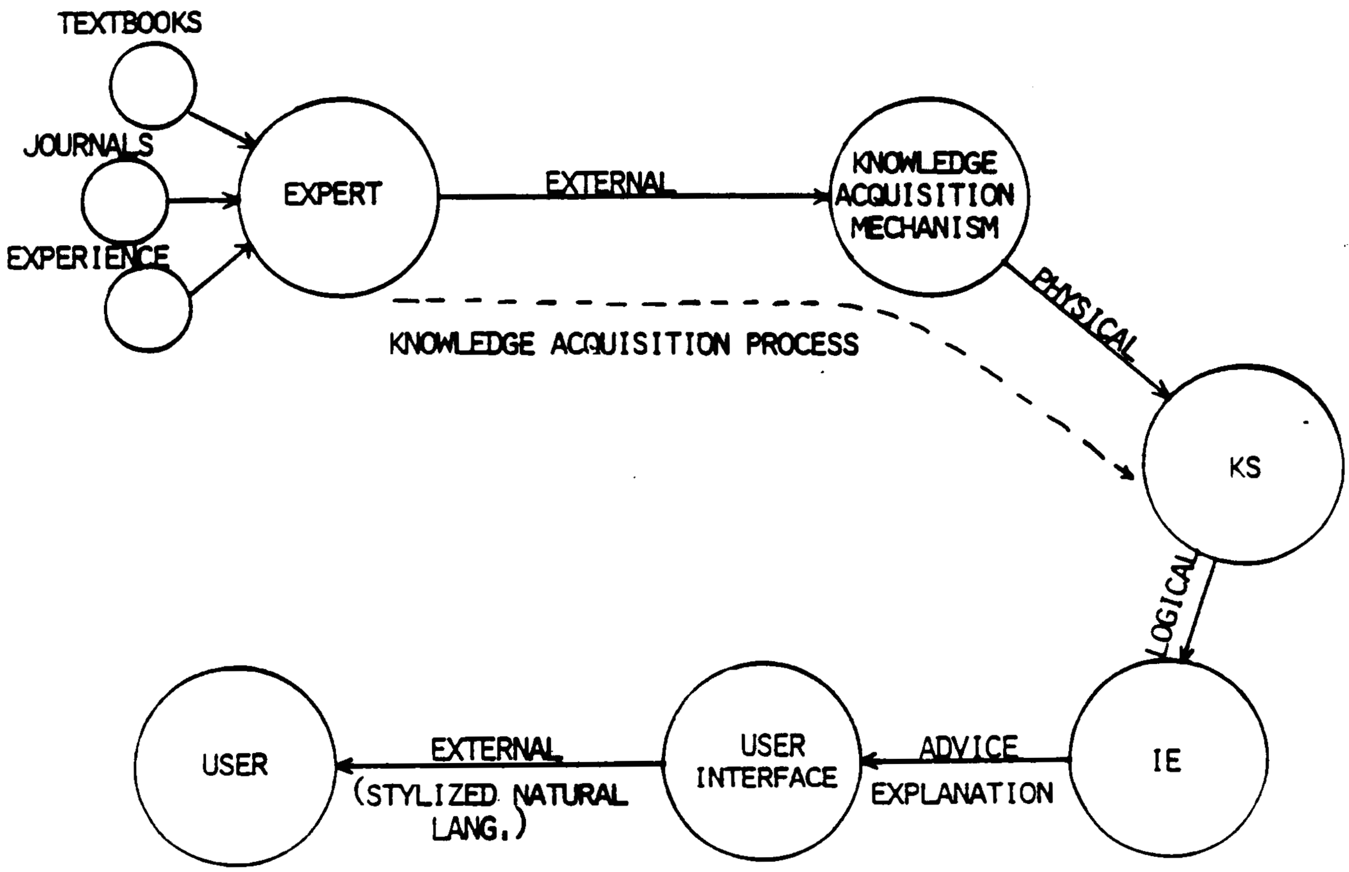


FIGURE 3-2. KNOWLEDGE REPRESENTATION FORMS
BASED ON [BARNETT & BERNSTEIN, '77]

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 Rule Base: A collection of production rules.
- 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)

OBSERVATIONS	AGENT	DIFFICULTY
Alternate Output Level	Mech	4
Battery Charge Level	Mech	3
Battery Fluid Level	SrvR	2
Choke Adjustment	Mech	5
Choke Function	Mech	5
Coolant Level	SrvR	2
Coolant System Pressure	Mech	5
Difficulty to Start	Cust	1
Engine Temperature	Cust	1
Fan Belt Tension	Mech	3
Fuse Condition	SrvR	2
Short in Electric Systm	Mech	8
Voltage Regulator Level	Mech	4
Warranties	Data Base	0

Figure 3-10. DATA GATHERING PROCEDURE FACT FILE

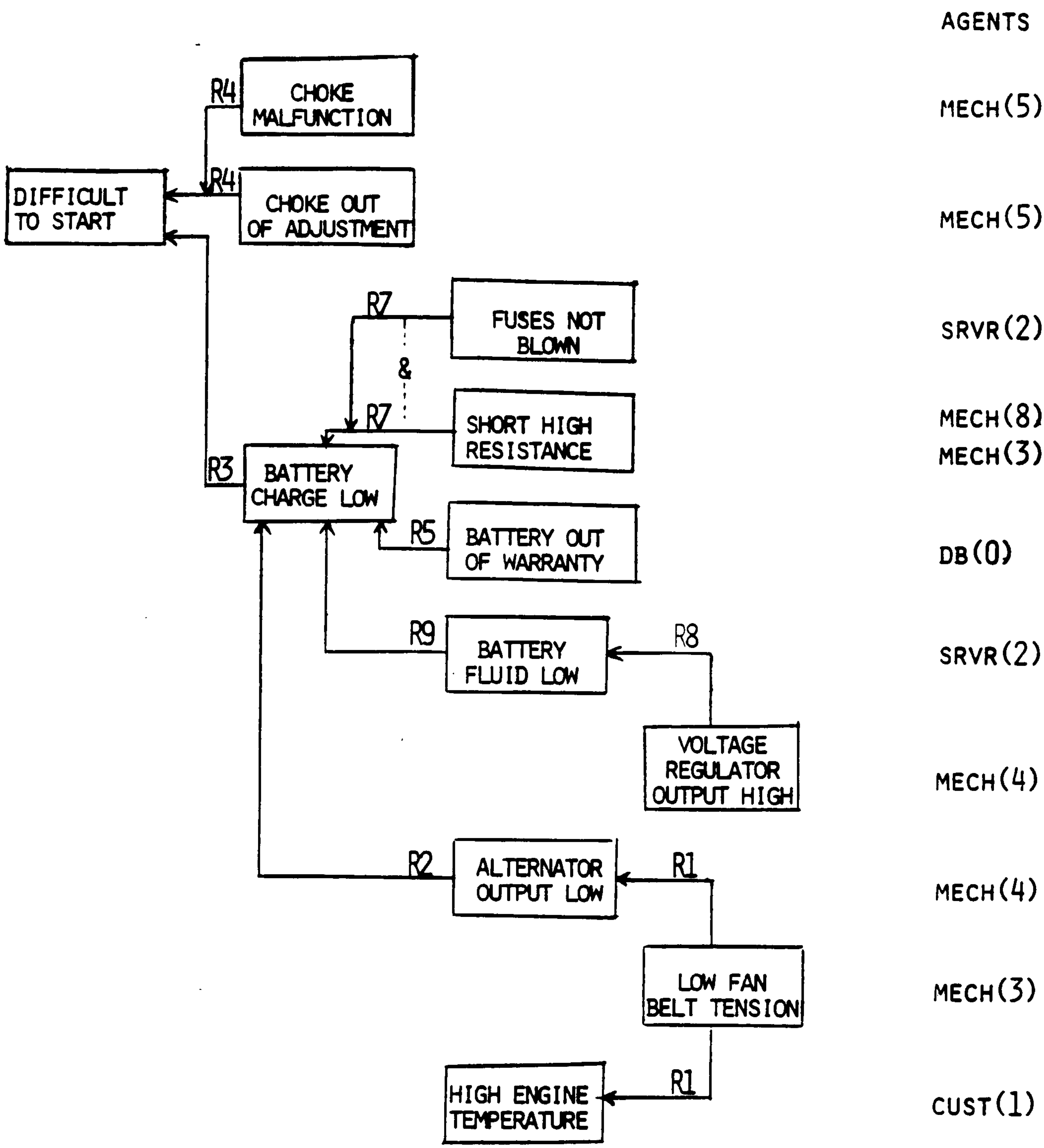


FIGURE 3.11 EXAMPLE FLOW IN AUTO DIAGNOSTIC SYSTEM

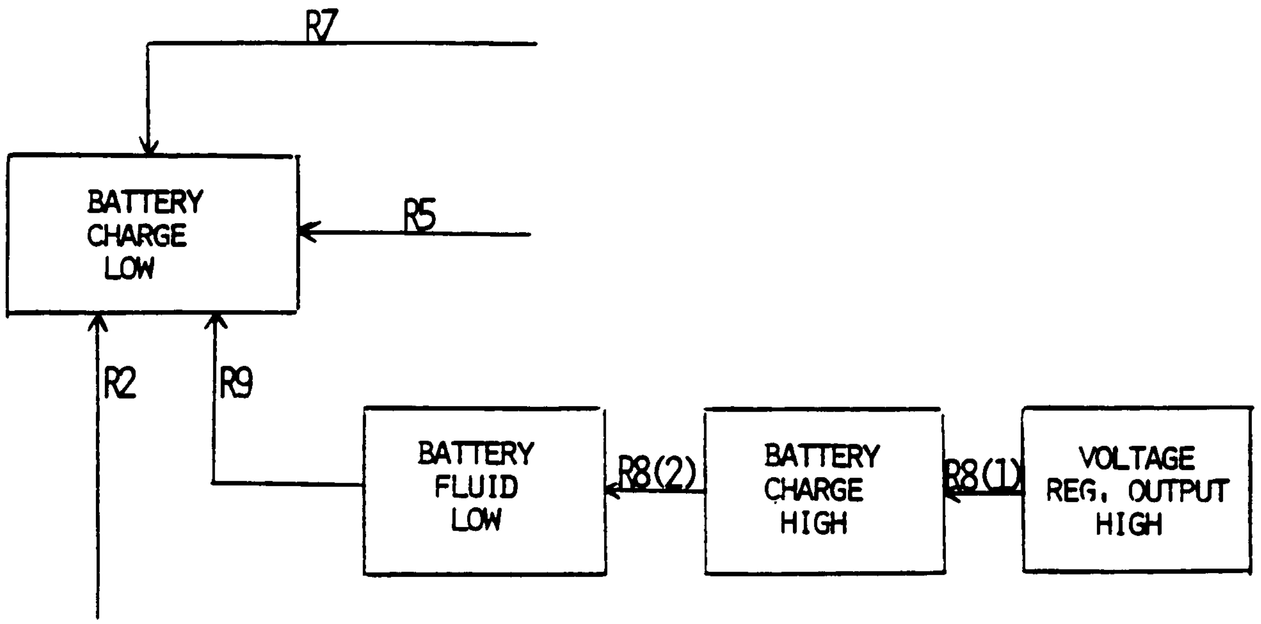


FIGURE 3-13. FRAGMENT OF GRAPH STRUCTURE

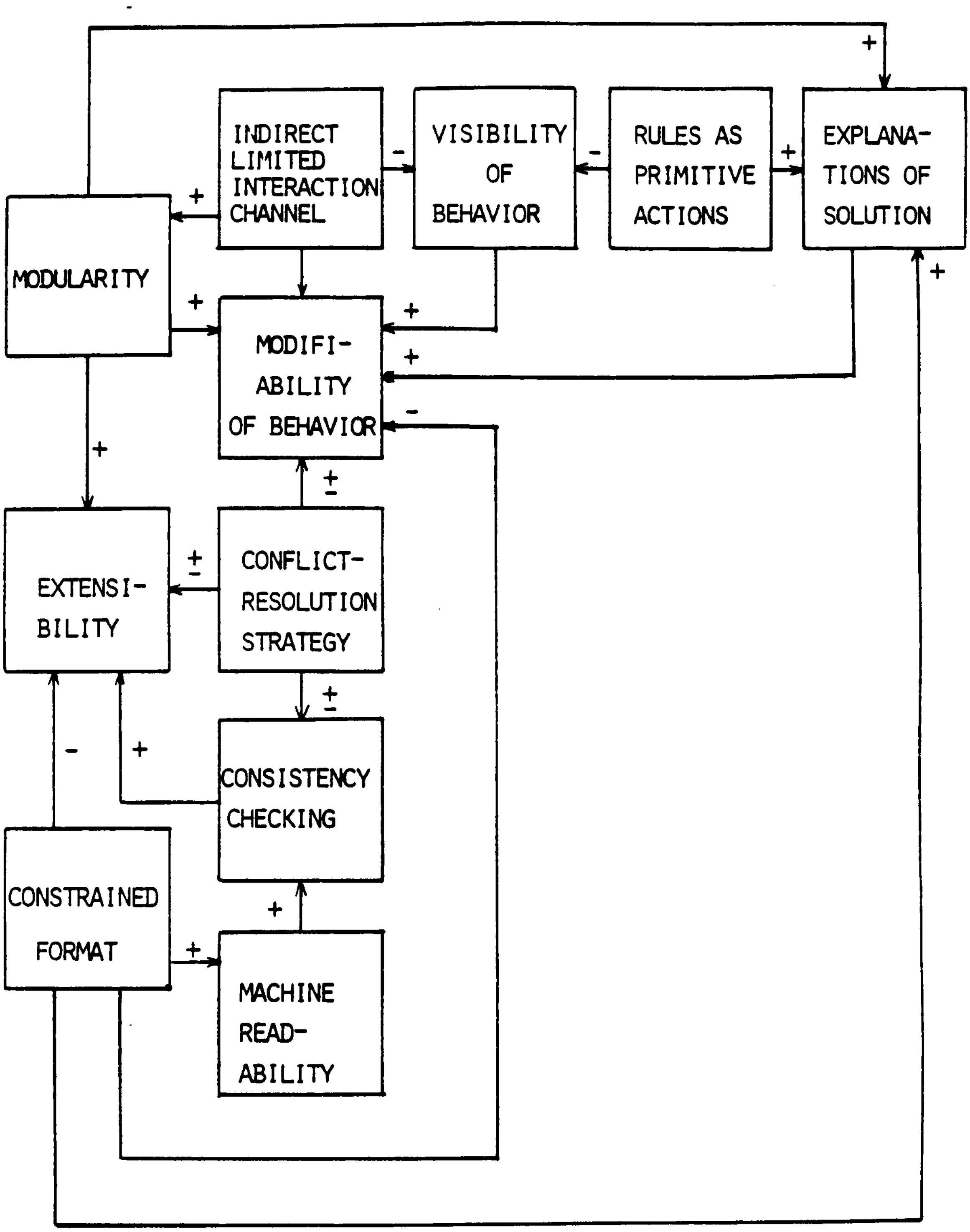


FIGURE 3-14. CHARACTERISTICS OF PRODUCTION SYSTEMS
BASED ON [BARNETT & BERNSTEIN, '77]

SEMANTIC NETWORKS

Semantic networks are used in

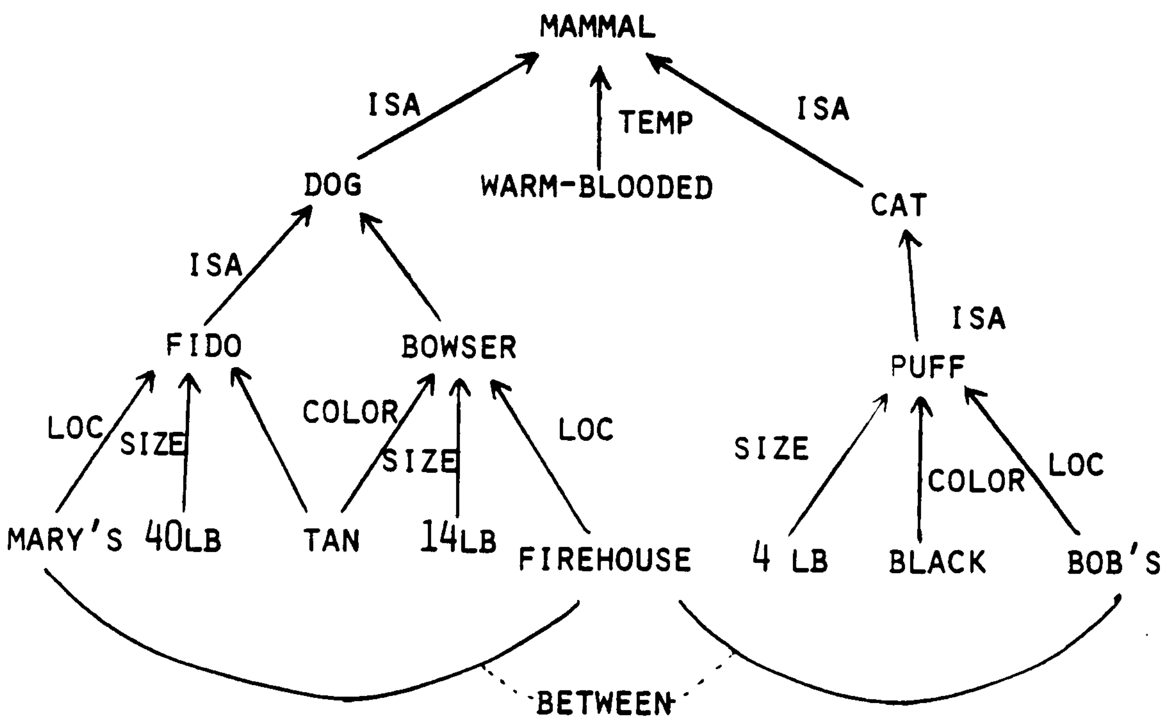
- Psychological modeling of human memory
- Programming languages
- Natural language understanding
- Data base management systems

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 INFERENCE

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

INHERITABLE(TEMP)

ISA(x, y) & r(u, y) &
INHERITABLE(r) => 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

```
1   dog   FRAME   ISA   mammal
2         kind    breed
3         color   SUBSET.OF {tan brown black
                             white rust}
4         FROM color OF kind
5         leggedness   0...4
6         weight       >0,  FROM size OF kind
7         state        adult  OR
                         puppy if age < 1
8         age          >0,  now birthday
9         birthday     date
10        name         string
11        END          dog
```

(a)

Figure 3-16. EXAMPLE FRAME DEFINITIONS
[Barnett & Bernstein, 77]

```
1  boxer FRAME  ISA   breed OF dog
2        color        ONE.OF {tan
                             brown brindle}
3        size         40...60
4        tail         bobbed OR long
5        ears         bobbed OR floppy
6        temperment   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

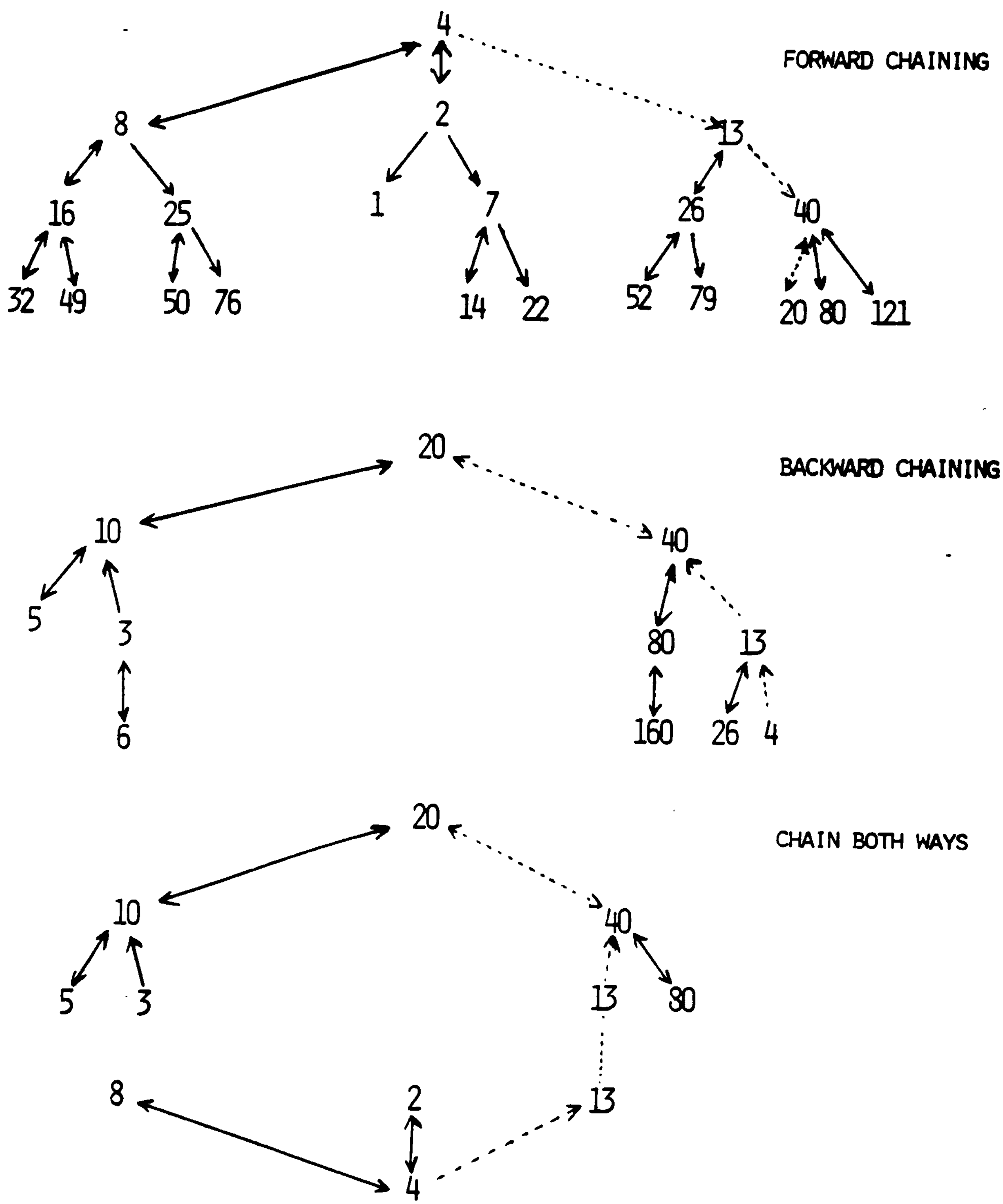


FIGURE 3.18 CHAINING EXAMPLES

BREADTH-FIRST CONTROL STRATEGY

An Example: 8-Puzzle

```
+---+---+---+          +---+---+---+
| 2 | 8 | 3 |          | 1 | 2 | 3 |
+---+---+---+          +---+---+---+
| 1 | 6 | 4 |          | 8 | - | 4 |
+---+---+---+          +---+---+---+
| 7 | - | 5 |          | 7 | 6 | 5 |
+---+---+---+          +---+---+---+
      a.                     b.
```

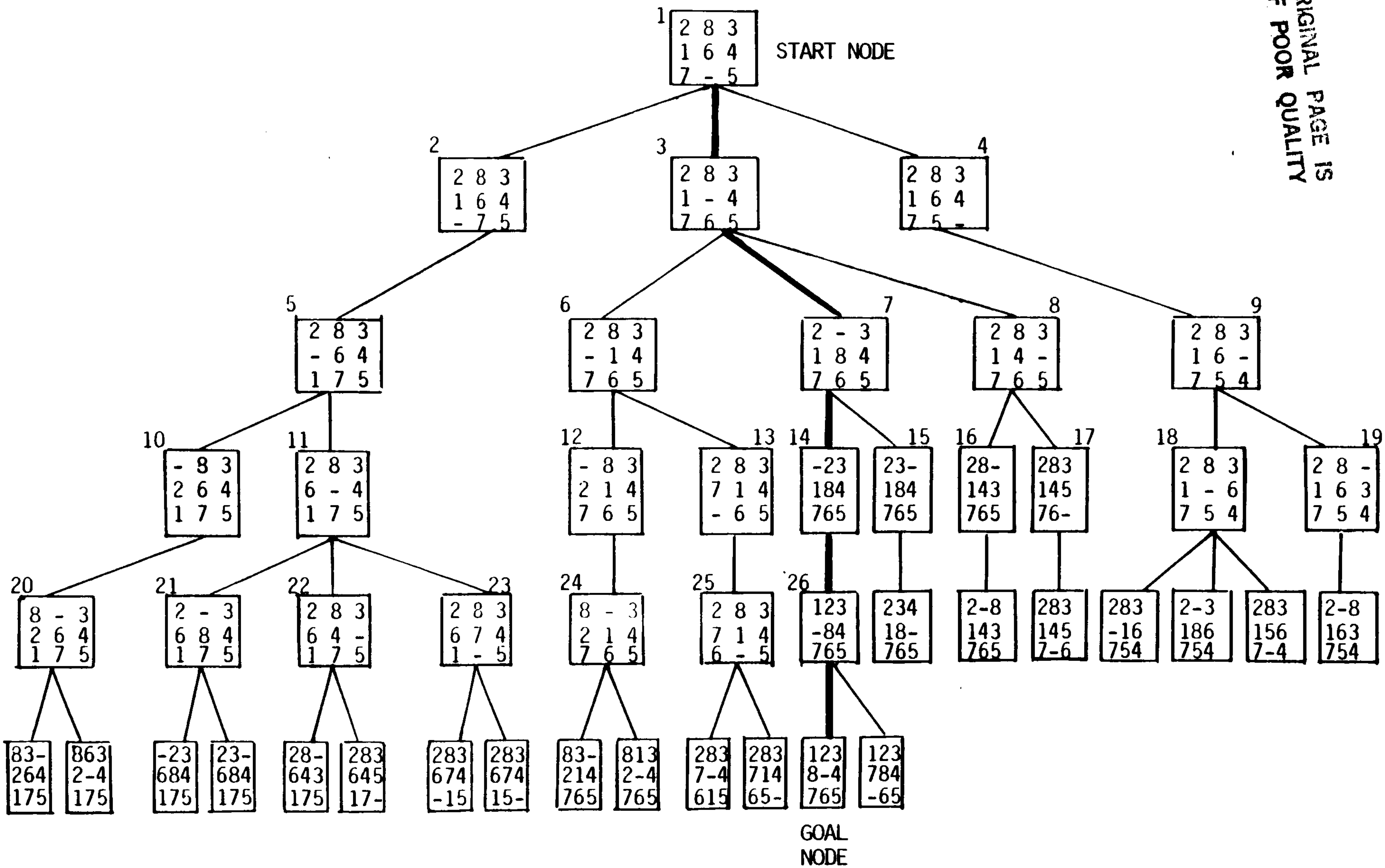


FIGURE 3-21. THE TREE PRODUCED BY A BREADTH-FIRST SEARCH
BASED ON [NILSSON, '71]

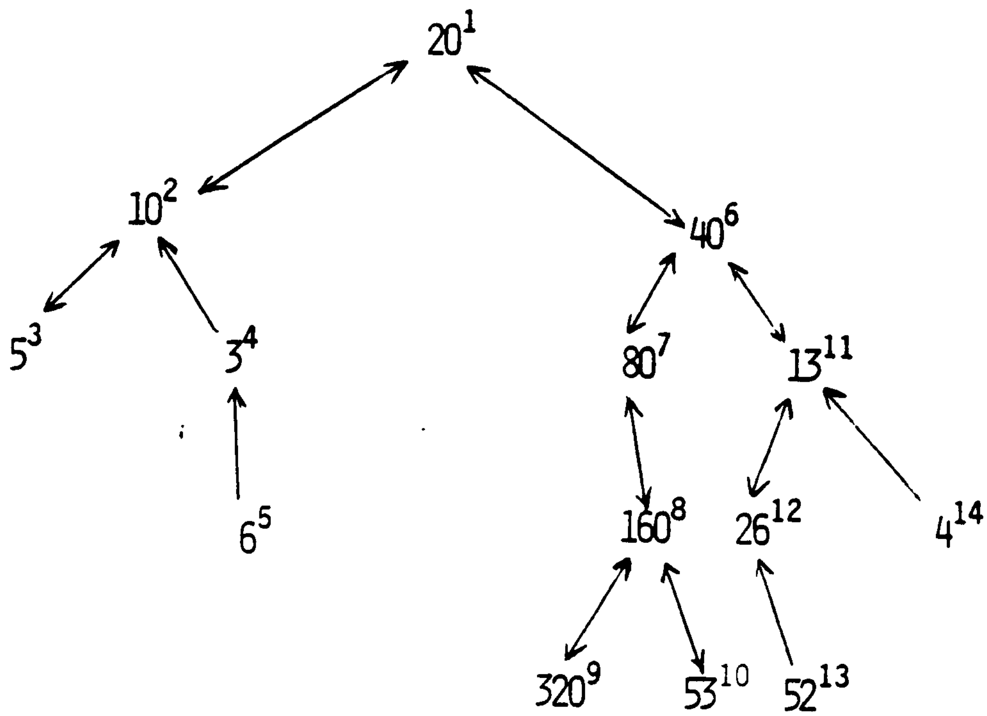


FIGURE 3-22. DEPTH-FIRST BACK CHAINING

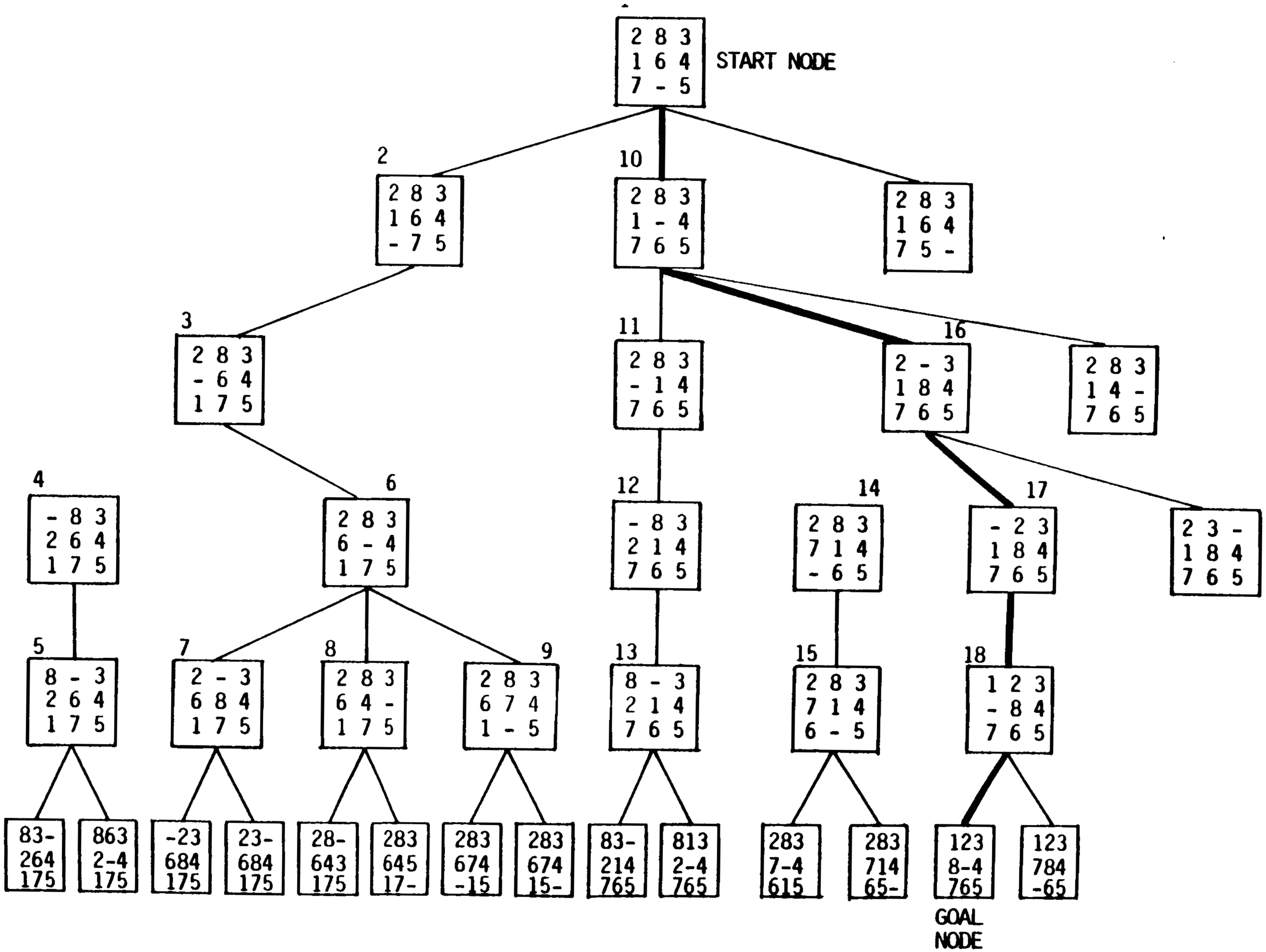


FIGURE 3-23. <u>THE TREE PRODUCED BY A DEPTH-FIRST SEARCH</u>
BASED ON [NILSSON, '71]

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 <u>minimized</u>.

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

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

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

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

MEASURES OF PERFORMANCE

Search Techniques

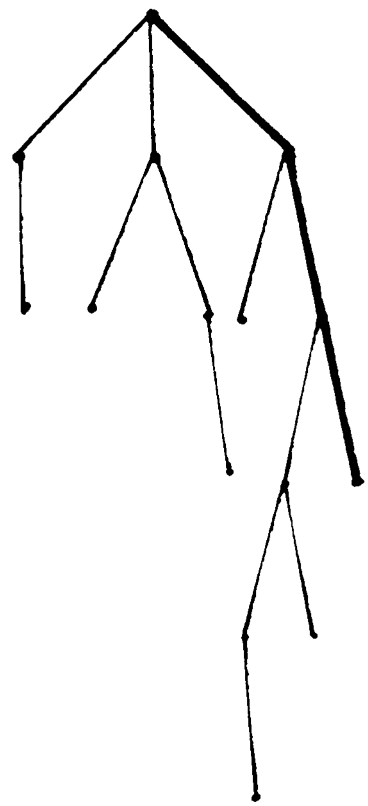
Penetrance

$P = L/T$

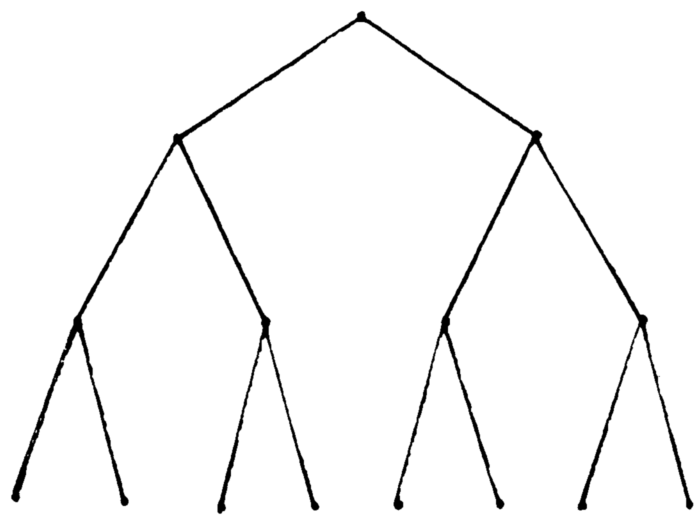
L length of the derived path from initial to goal

T total number of nodes

Branching Factor



a.



b.

T = 15

L = 3

P = 1/5

B = 2

FIGURE 3-27. EXAMPLE MOVE GRAPH AND BALANCED TREE

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

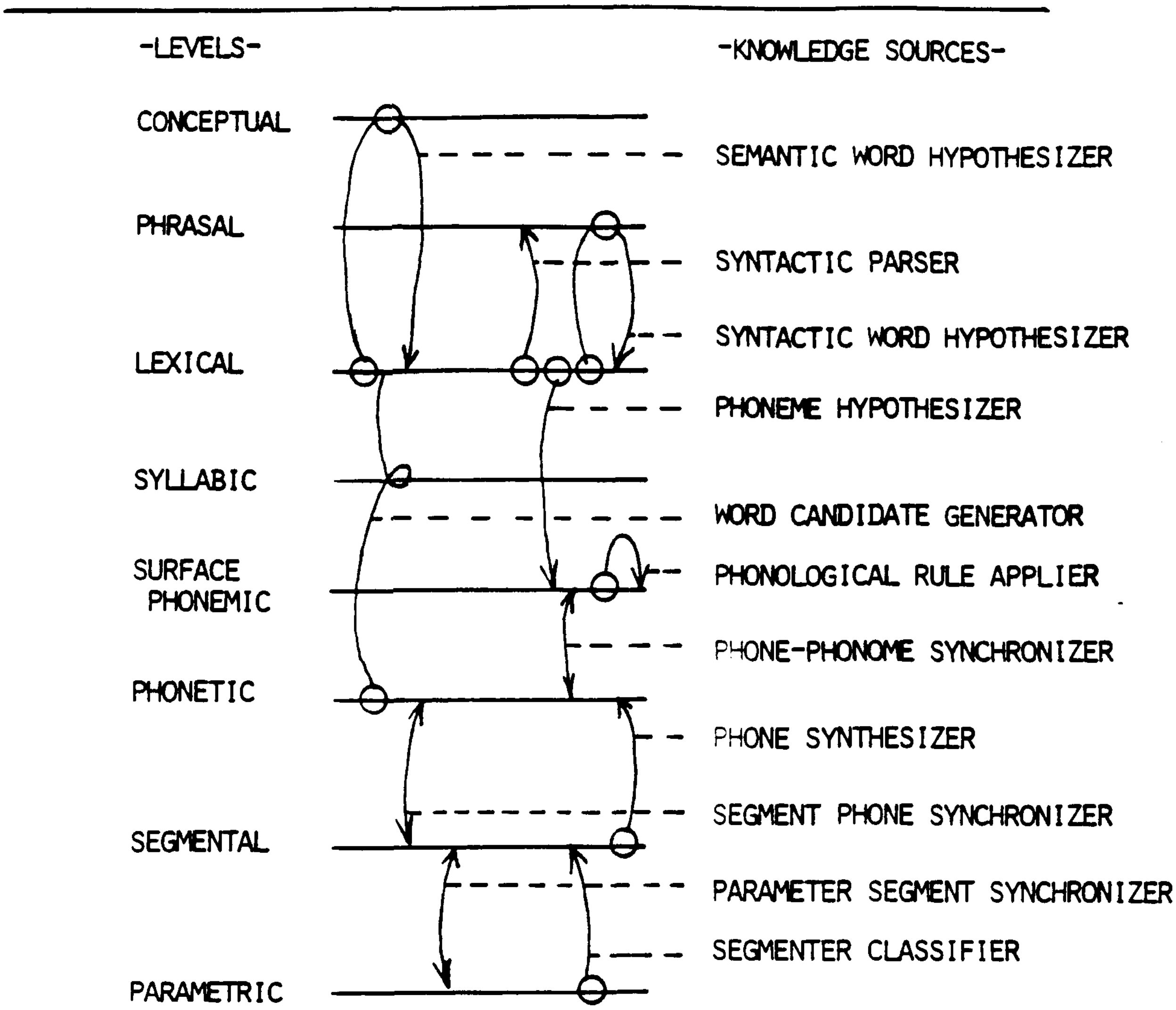


FIGURE 3-28. HEARSAY II LEVELS OF REPRESENTATION AND KNOWLEDGE SOURCES BASED ON [ERMAN, ET AL, '80]

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

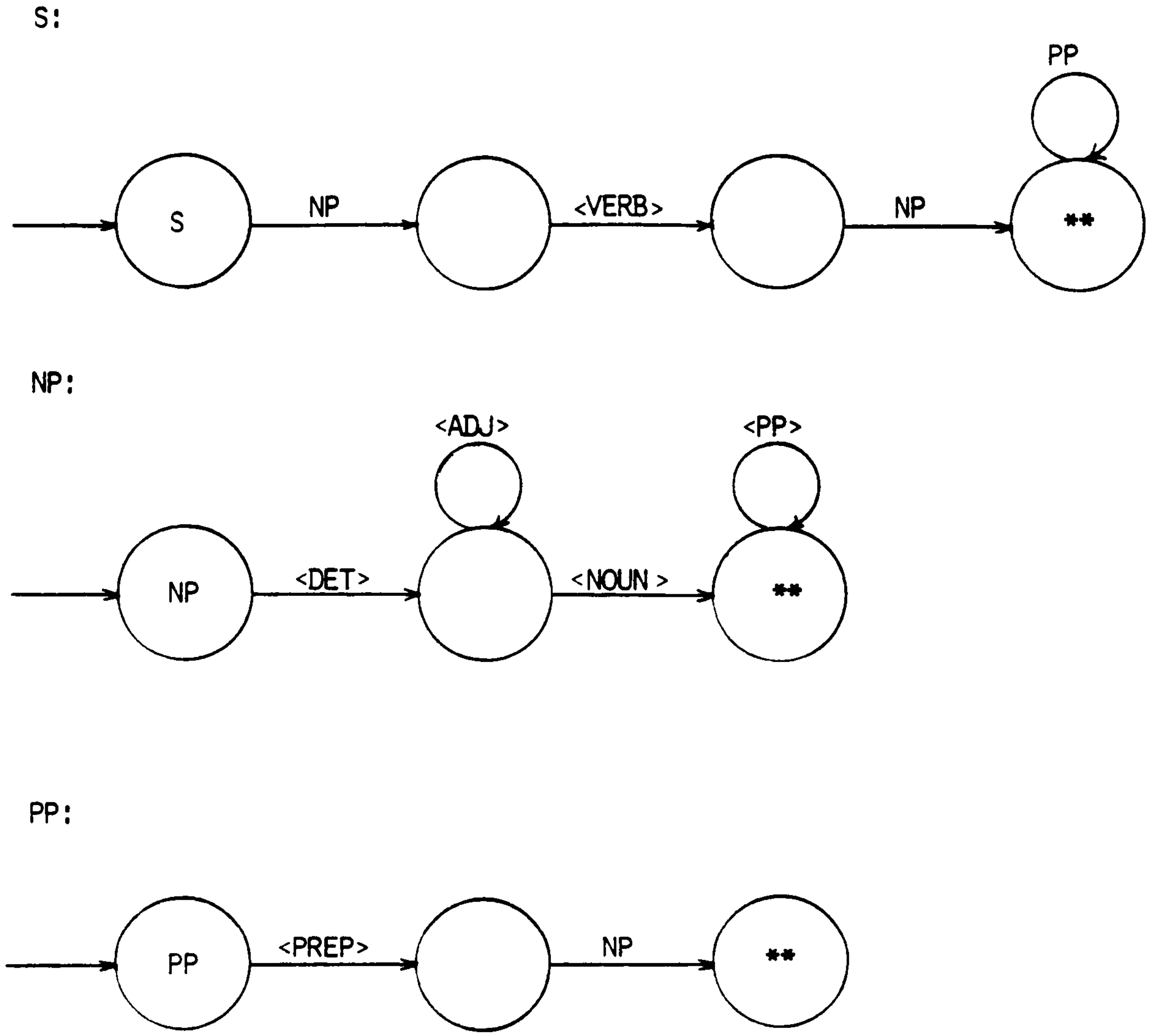


FIGURE 3-32. A RECURSIVE TRANSITION NETWORK
BASED ON [BARR & FEIGENBAUN, '81]

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

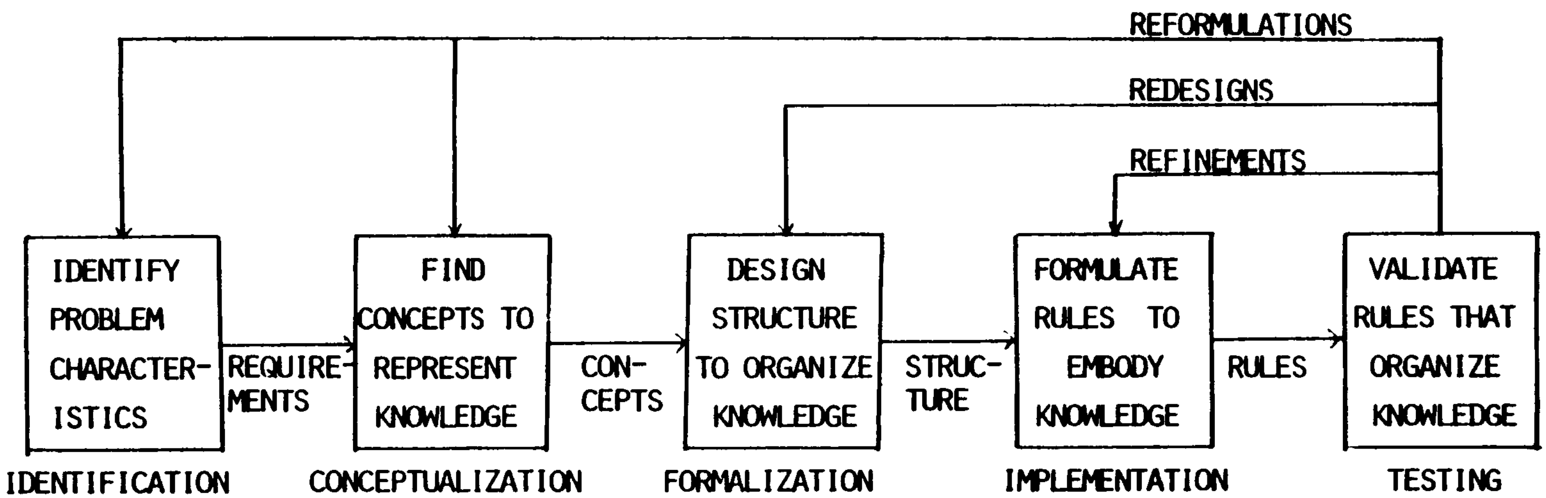


FIGURE 3-33. STAGES OF KNOWLEDGE ACQUISITION
BASED ON [HAYES-ROTH, ET AL, '83]

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
- Comparision of techniques
- KBS building tools
- Explanation
- Evaluation
- Parallel processing
- Learning from experience
- Management of knowledge
- Abstraction and hierarchies
- Technological innovations
- Uniform terminology

4.11

1. Report No. IN-82	2. Government Accession No. 183556	3. Recipient's Catalog No.
4. Title and Subtitle USL/NGT-19-010-900: KNOWLEDGE BASED SYSTEMS: A CRITICAL SURVEY OF MAJOR CONCEPTS, ISSUES, AND TECHNIQUES: VISUALS		5. Report Date December 11, 1984 6. Performing Organization Code
7. Author(s) SRINU KAVI		8. Performing Organization Report No.
9. Performing Organization Name and Address University of Southwestern Louisiana The Center for Advanced Computer Studies P.O. Box 44330 Lafayette, LA 70504-4330		10. Work Unit No. 11. Contract or Grant No. NGT-19-010-900
12. Sponsoring Agency Name and Address		13. Type of Report and Period Covered FINAL; 07/01/85 - 12/31/87 14. Sponsoring Agency Code
15. Supplementary Notes		

16. Abstract

This Working Paper Series entry represents a collection of presentation visuals associated with the companion report entitled "Knowledge Based Systems: A Critical Survey of Major Concepts, Issues, and Techniques", USL/DBMS NASA/RECON Working Paper Series report number DBMS.NASA/RECON-9.

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.

17. Key Words (Suggested by Author(s)) Knowledge-Based Systems, Information Storage and Retrieval Systems		18. Distribution Statement	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 64	22. Price*

*For sale by the National Technical Information Service, Springfield, Virginia 22161