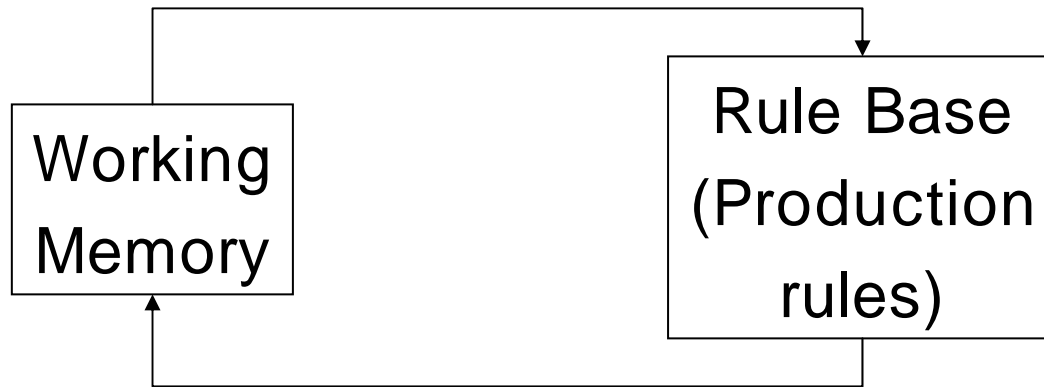


Problem Solving and Reasoning

Rule - Based System (Production System)



C_i : condition

A_i : Action

R_i : rule or production

- Loop continues until working patterns no longer matches the condition of any rules

- PS

- rule base (production memory, long term memory)
- working memory (shorter term memory)
- interpreter (controller)

< rule >

If condition 1, then action 1

If condition 1, then action 1

...

“ rule working memory
.”

- PS

1. Separation of knowledge & control

2. Modularity

가, , 가 가

→

*

가 가

Modularity

3. Uniformity

4. Naturalness

()

“

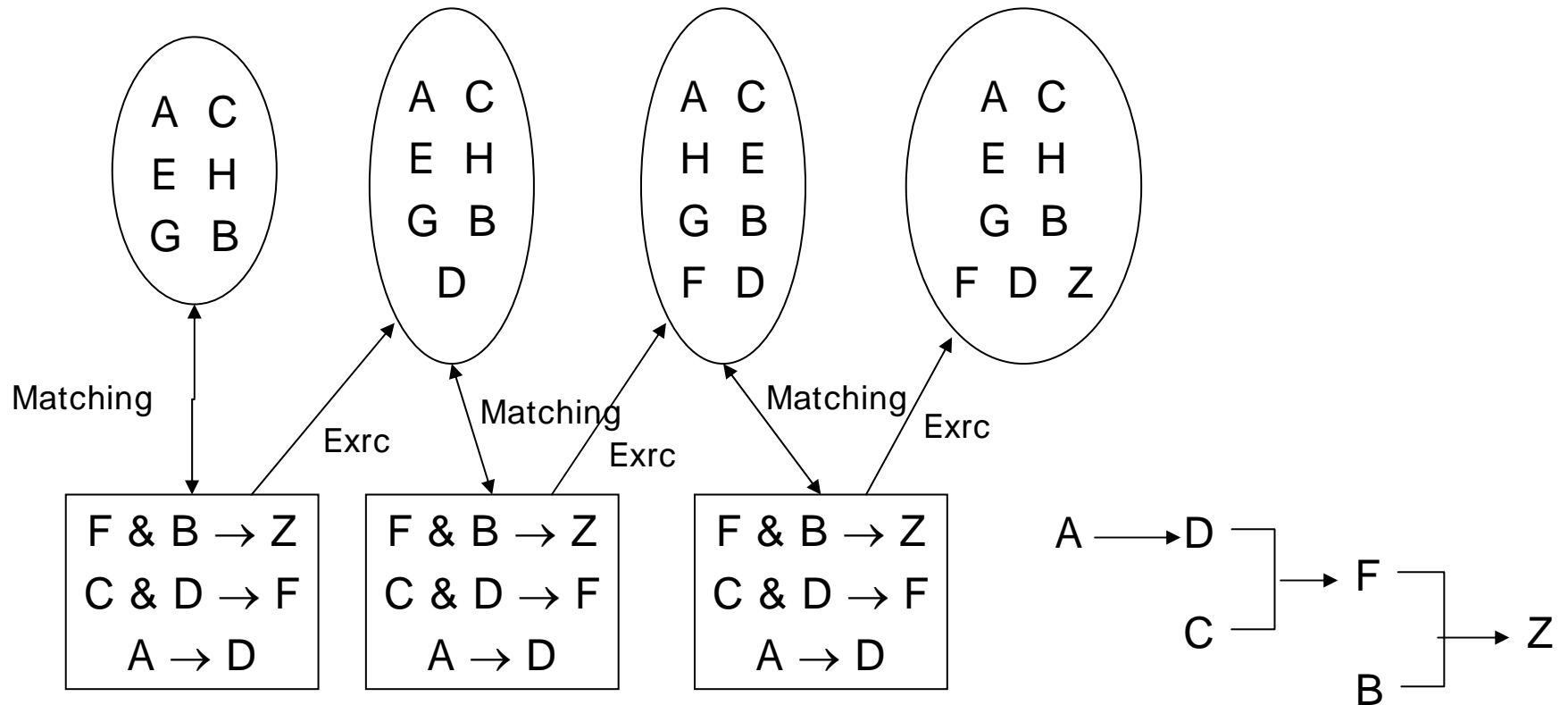
situation

.

가”

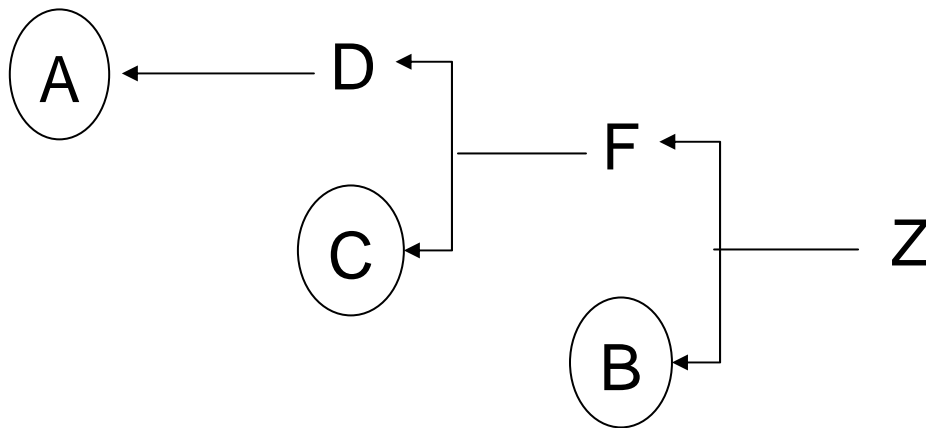
PS

- forward chaining
- backward chaining



< Forward chaining example >

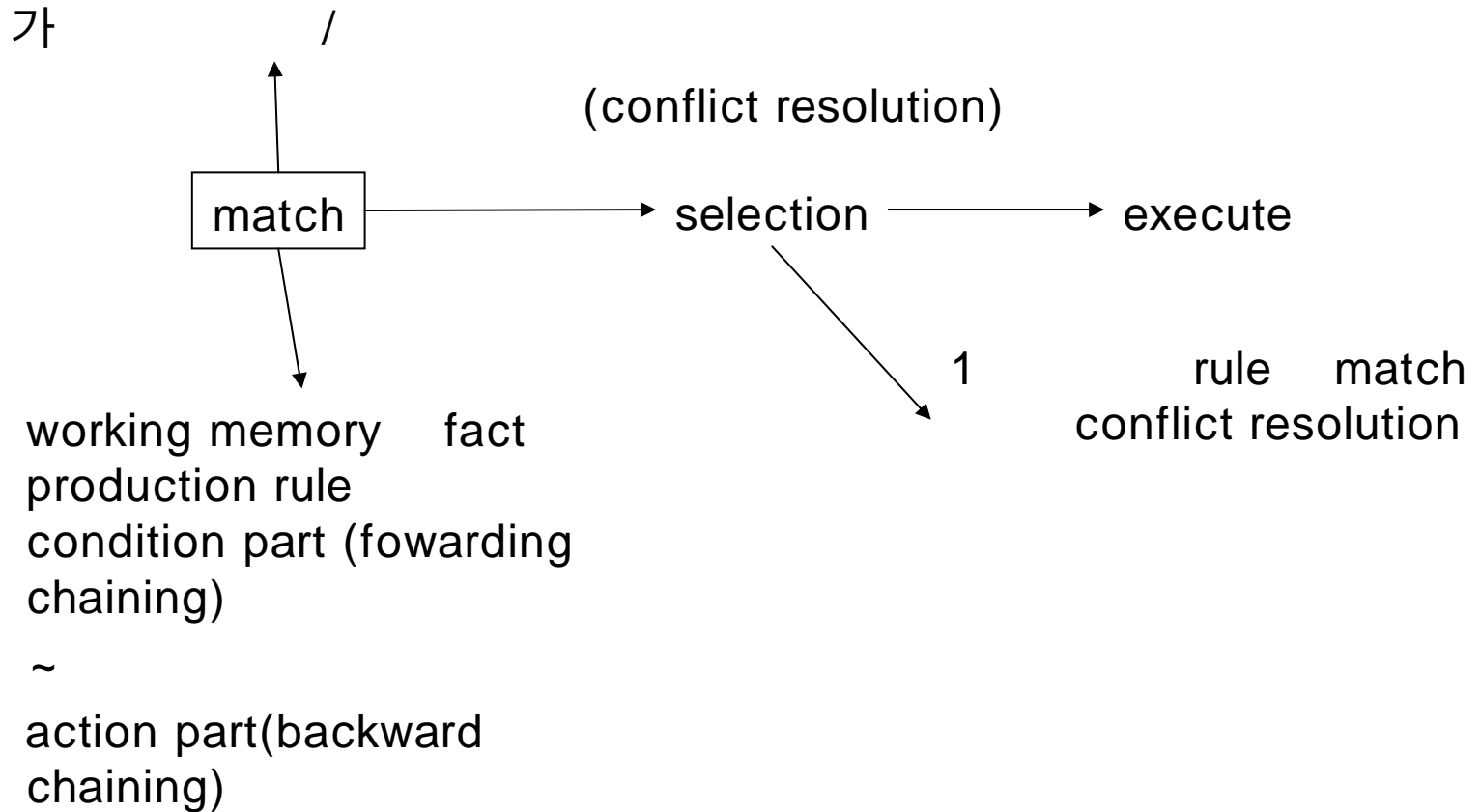
- If problem is to see “Z” is true or not,
 ← Backward chaining is better.
- Z가
 . ← more efficient! Z



? Backward chaining vs. Abduction

- Backward chaining abduction
 - backward chaining $A \rightarrow B$ $B \nVdash \text{true}$
 $\quad \quad \quad , A \nVdash \text{true} \quad \nVdash$
 $(A \rightarrow B \ \& \ A) \rightarrow B$
 - abduction : $(A \rightarrow B \ \& \ B) \rightarrow A$
- Working memory size \nVdash memory
 - $>$ MYCIN context tree
 prospector semantic net
 - production rule grouping

< Production system >



- Conflict Resolution

- 가

- Specify ordering : 가

- Rule ordering : rule base

- Data ordering : priority

- Recency ordering : 가
가

- Context limiting :

(rule) ()
grouping

- (ex)
production rule :
1. $ba \rightarrow ab$
2. $ca \rightarrow ac$
3. $cb \rightarrow bc$

| iteration | WM | conflict set | rule fired |
|-----------|----------------|--------------|------------|
| 0 | c <u>b</u> aca | 1,2,3 | 1 |
| 1 | <u>c</u> abca | 2 | 2 |
| 2 | ac <u>b</u> ca | 3,2 | 2 |
| 3 | ac <u>b</u> ac | 1,3 | 1 |
| 4 | ac <u>a</u> bc | 2 | 2 |
| 5 | aa <u>c</u> bc | 3 | 3 |
| 6 | aabcc | 0 | stop |

Matching \rightarrow Conflict Resolution \rightarrow execution
(interpreter)

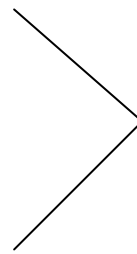
- Matching complexity

- working memory : w elements
- production rules : r
- # of elements in condition part : n
- # of cycles for solving problem : c
- total # of match (unification) = $wrnc$
- simpler expert system : $w = 100$

$$r = 200$$

$$n = 5$$

$$c = 1000$$



100M

- rule-based system (,)
 rule (structuring)
 (ordering) .

– p & q & r → S

(check) confirm
 fail 가 .

fail rule AND
 check 가

→ easier check first
 more likely fail first

Rule 1 : if

the engine is getting gas, and
the engine will turn over,
then
the problem is spark plugs.

Rule 2 : if

the engine does not turn over,
and
the lights do not come on,
then
the problem is battery.

Rule 3 : if

the engine does not turn over,
and
the lights do come on
then
the problem is the starter motor.

Rule 4 : if

there is gas in the fuel tank,
and
there is gas in the carburetor,
then
the engine is getting gas.

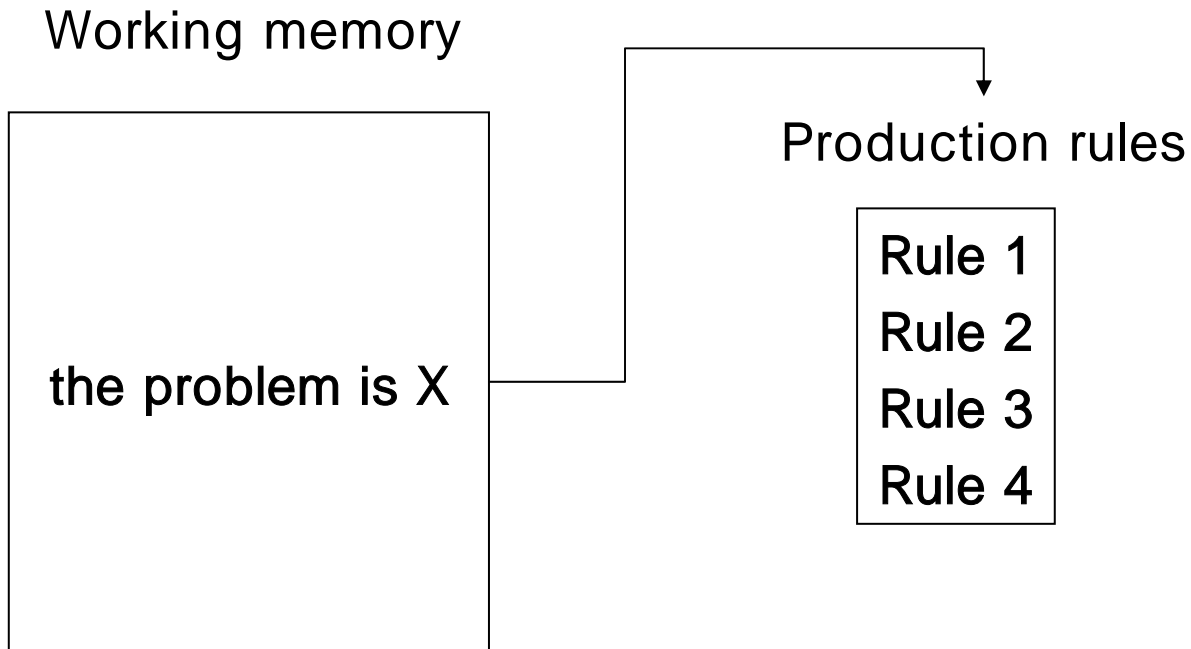


Figure 8.6 The production system at the start of a consultation in the car diagnostic example.

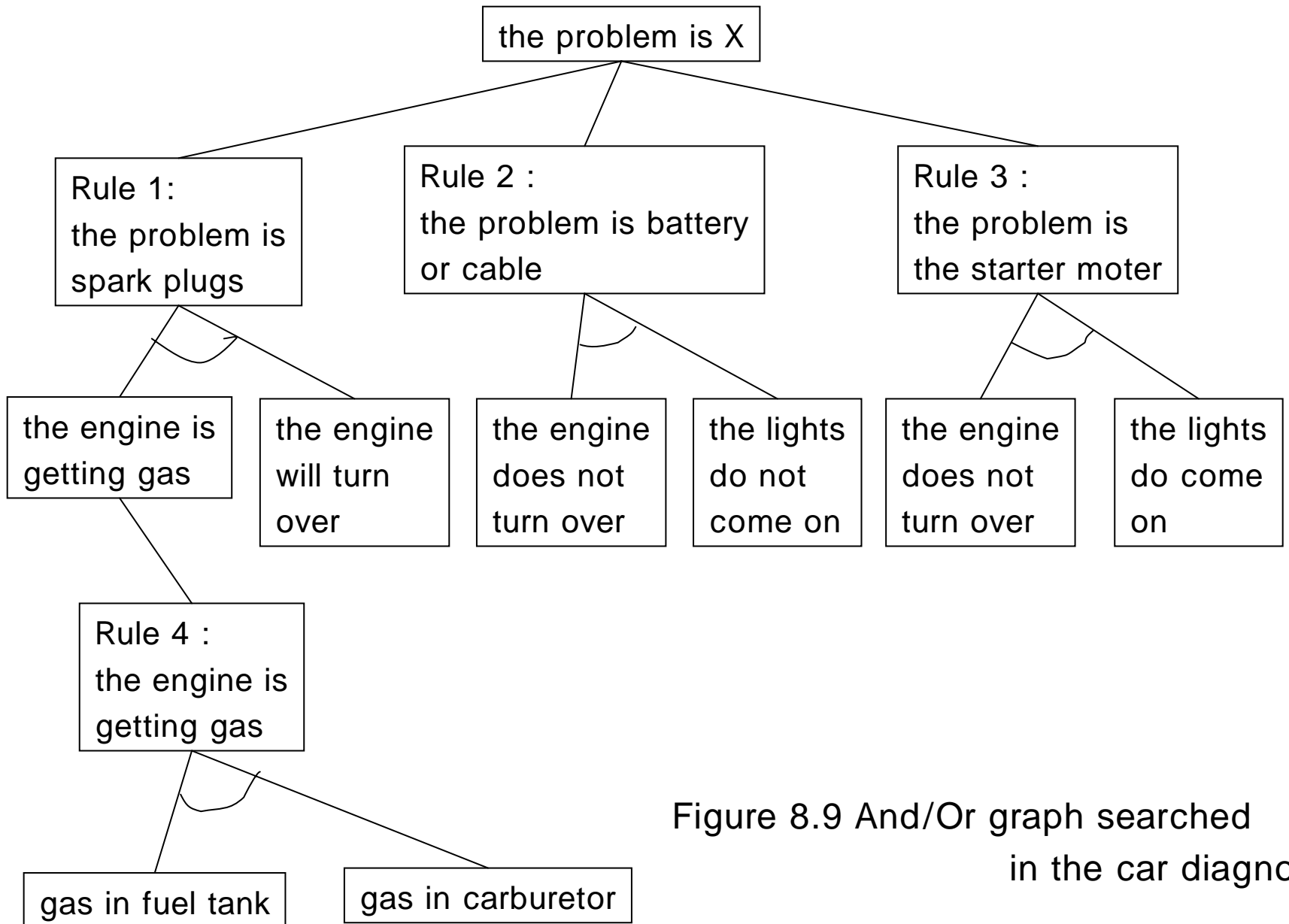


Figure 8.9 And/Or graph searched
in the car diagnostic

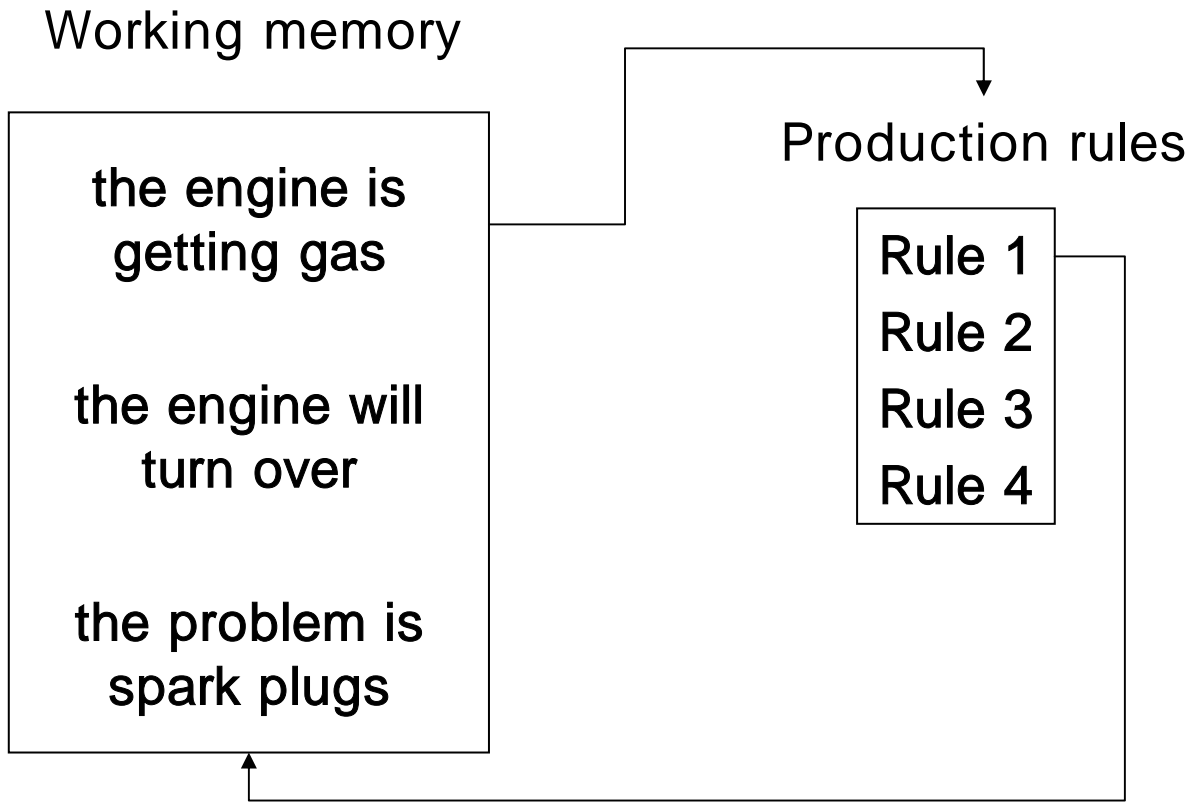


Figure 8.7 The production system after Rule 1 has fired.

Working memory

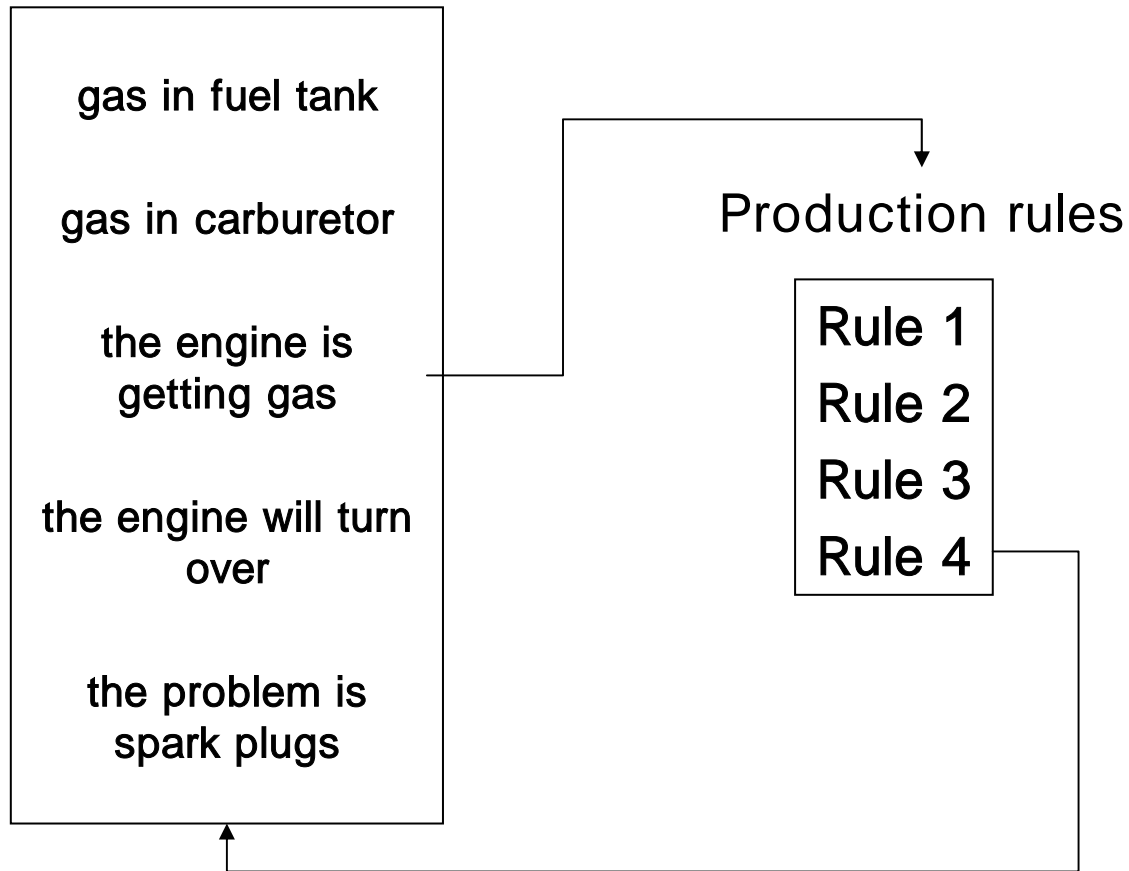


Figure 8.8 The production system after Rule 4 has fired.

(ex) “IDENTIFIER” 7)

(Cheetah, tiger, giraffe, zebra, ostrich, penguin, albatross)

R1 : If animal has hair
 then it is a mammal

R2 : If animal gives milk
 then it is a mammal

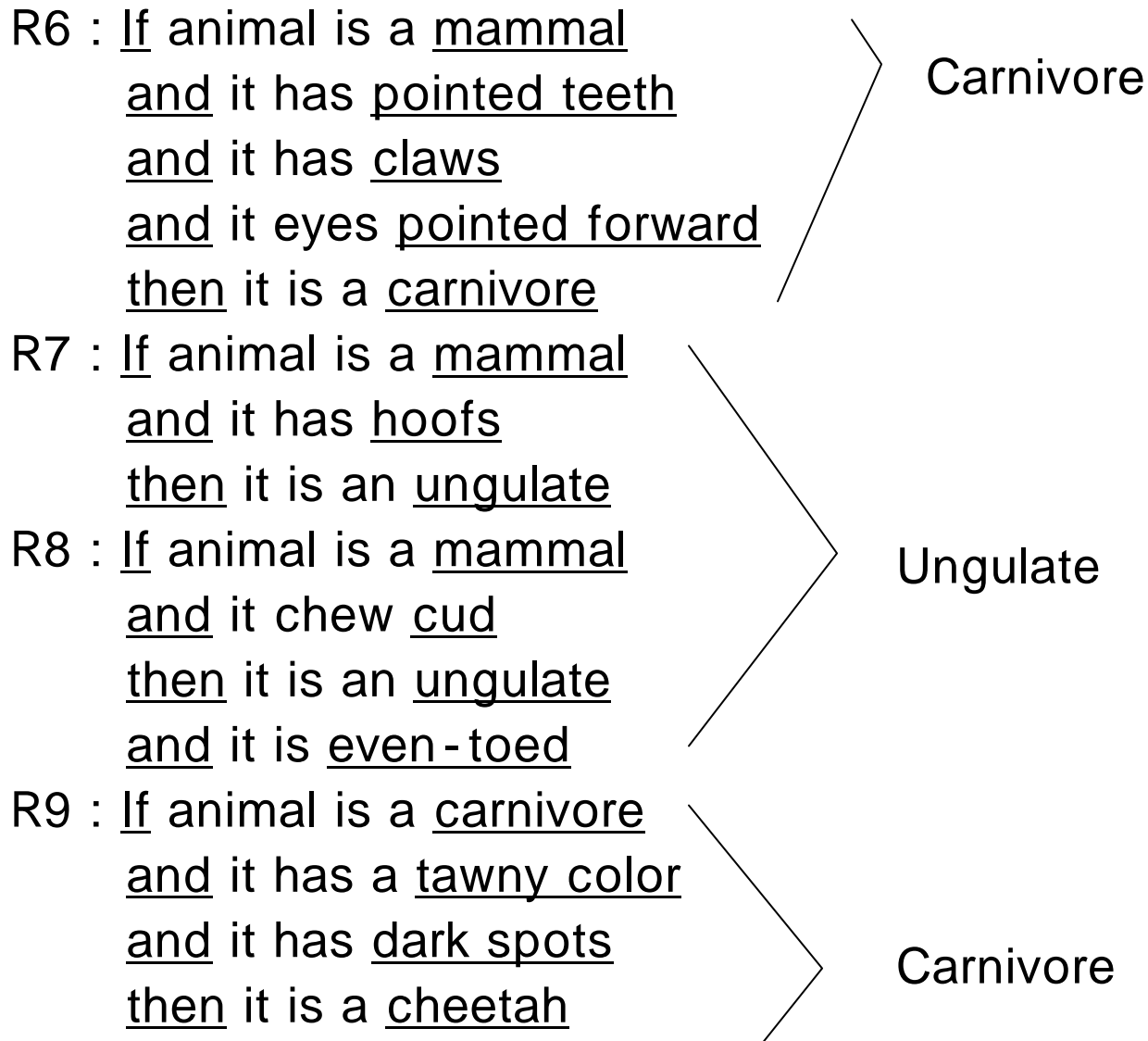
R3 : If animal has feathers
 then it is a bird

R4 : If animal flies
 and it lays egg
 then it is a bird

R5 : If animal is a mammal
 and it eats meat
 then it is a carnivore



Mammal or bird



R10 : If animal is a carnivore
and it has a tawny color
and it has black stripes
then it is a tiger

R11 : If animal is an ungulate
and it has long legs
and it has a long neck
and it has a tawny color
and it has dark spots
then it is a giraffe

R12 : If animal is an ungulate
and it has a white color
and it has black stripes
then it is a zebra



Ungulate

R13 : If animal is a bird
and it does not fly
and it has long legs
and it has a long neck
and it is black and white
then it is an ostrich

R14 : If animal is a bird
and it does not fly
and it swims
and it is black and white
then it is a penguin

R15 : If animal is a bird
and it is a good flyer
then it is an albatross



bird

- Now the following has been observed.

1. Tawny color & dark spots. $\rightarrow ?R_{10} ?R_{11}$

2. Gives a milk & chews its cud. R_2, R_8

3. long legs & long neck. $\rightarrow R_{11}$

The animal is giraffe.

-

1.

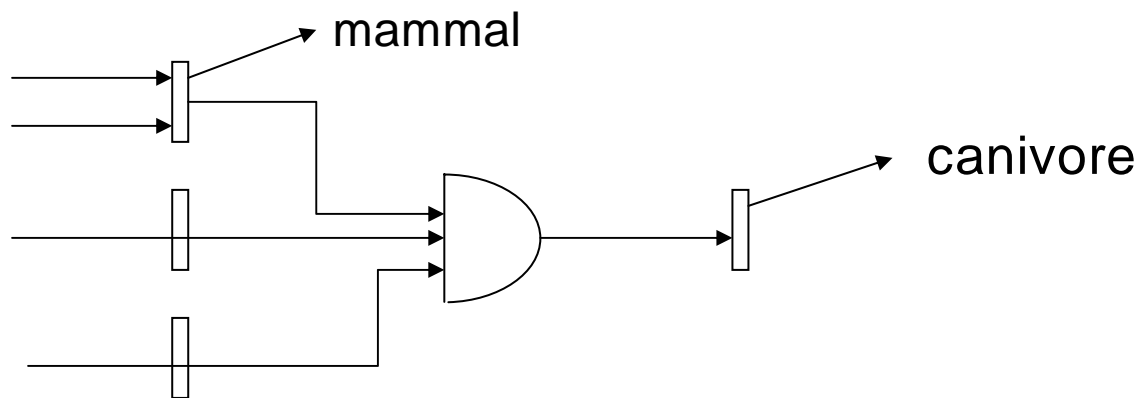
가? (why)

2.

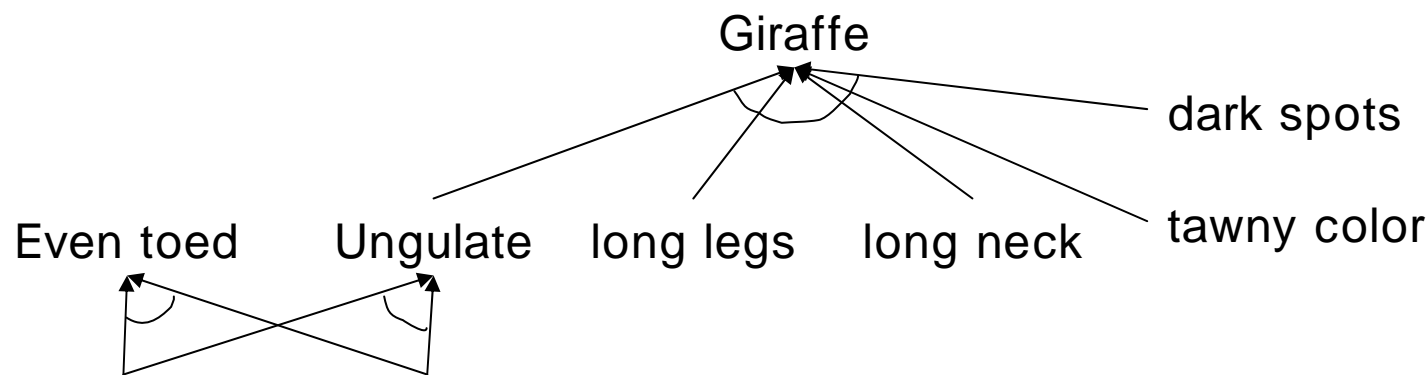
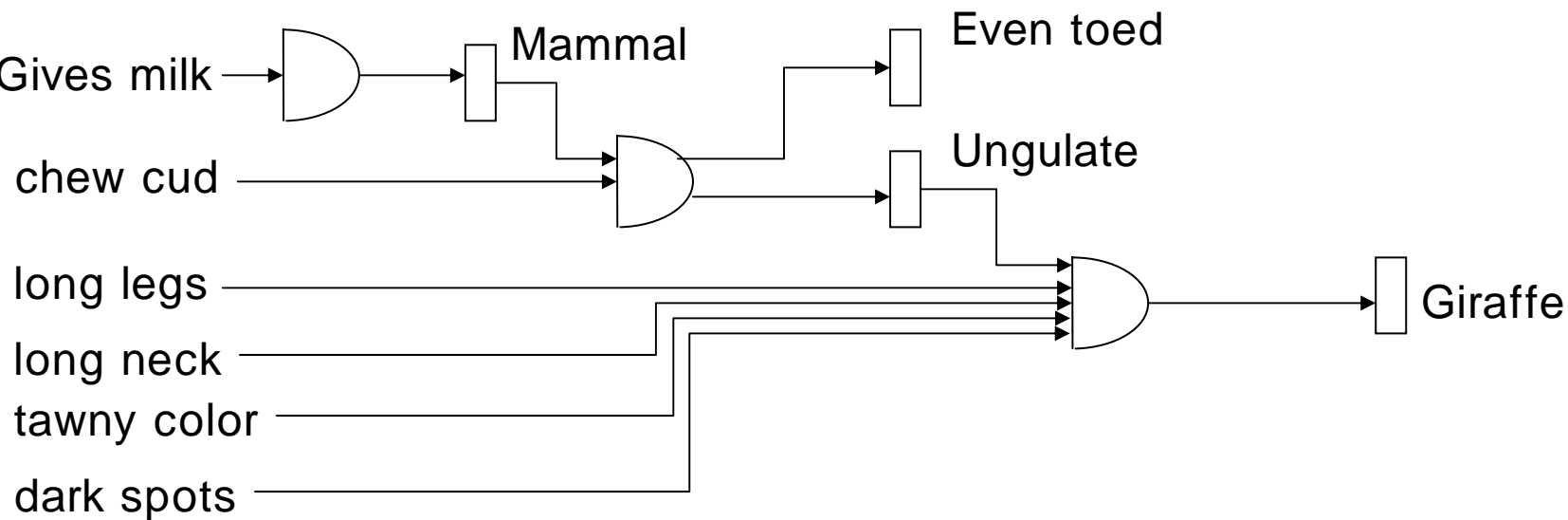
가? (how)

- Answer

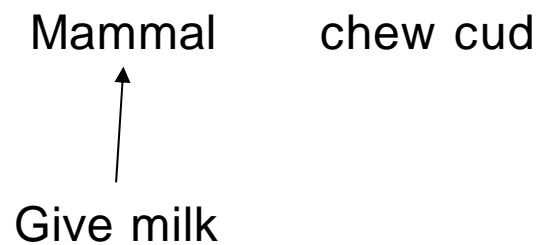
- for why : forward
- for how : backward



< Inference Net >



< AND/OR Graph >



$$P(H_i | E) = \frac{P(E | H_i)P(H_i)}{\sum_{k=1}^n (P(E | H_k)P(H_k))}$$

- $P(H_i|E)$: Prob. that H_i true given evidence E
- $P(H_i)$: Prob. that H true overall
- $P(E|H_i)$: Prob. of observing evidence E when H_i true
- n : number of possible hypotheses

$$P(E) = \sum_{k=1}^n P(E | H_k)P(H_k)$$

< Bayes theorem >

$$P(A | B)P(B) = P(A, B)$$

$$P(B | A)P(A) = P(B, A)$$

$$\therefore P(A | B)P(B) = P(B | A)P(A)$$

$$P(A | B) = \frac{P(B | A)P(A)}{P(B)}$$

- Bayes' rule : known fact \rightarrow unknown fact probability

$$P(X | Y)P(Y) = P(Y | X)P(X)$$

$$P(X | Y) = \frac{P(Y | X)}{P(Y)}$$

$P(X)$: 가

$P(Y)$: 가

$P(X|Y)$:

$P(Y|X)$:

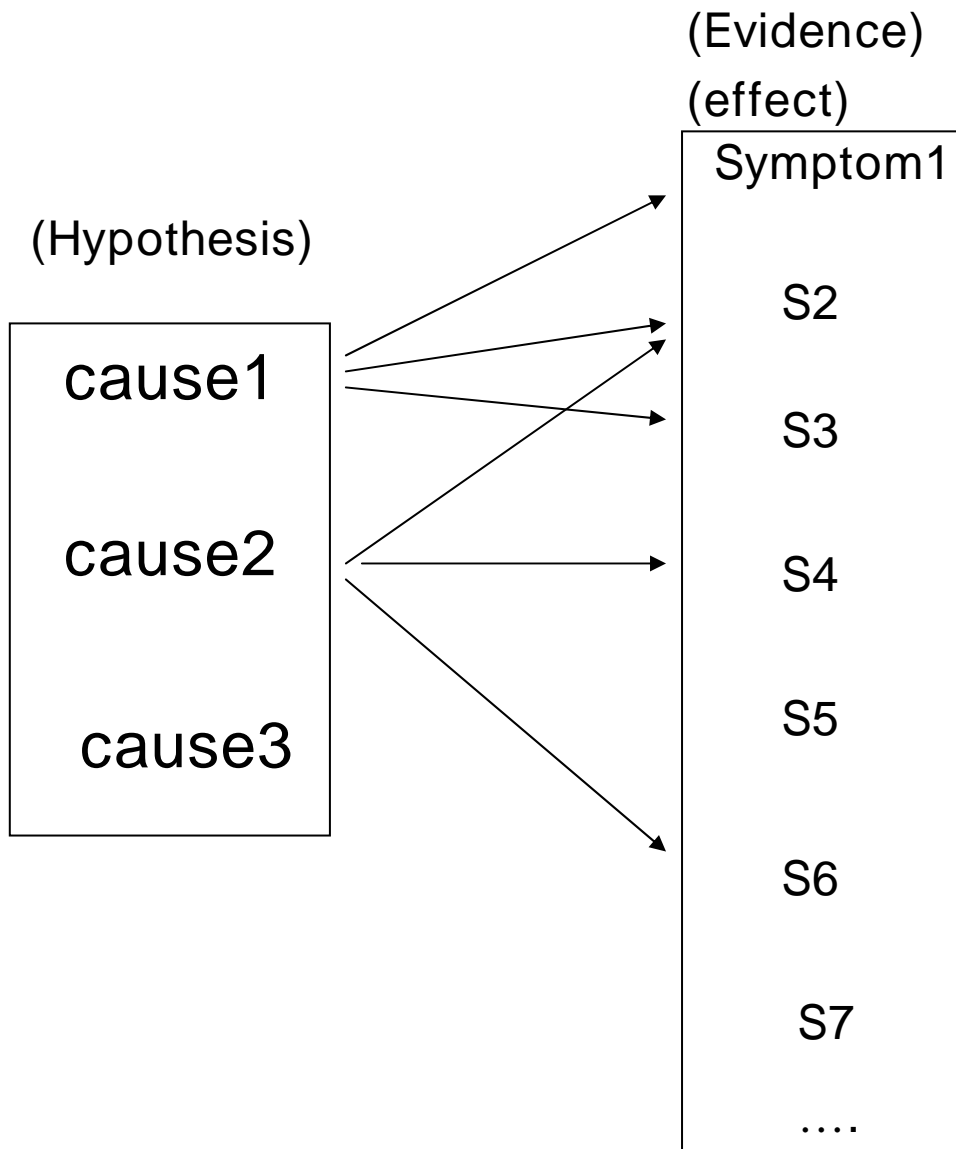
e.g.

$$P(X) = 0.1$$

$$P(Y) = 0.3 \quad P(Y | X) = 0.9$$

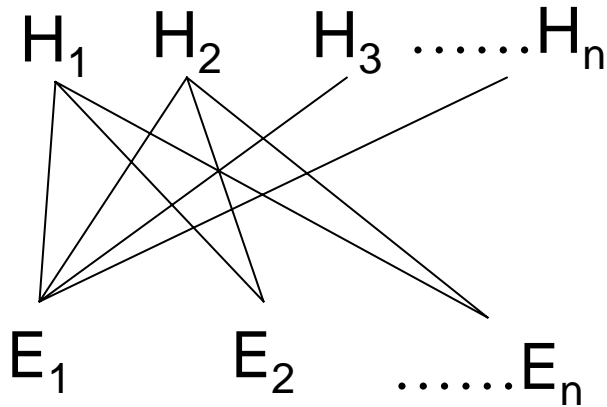
$$P(X | Y) = \frac{0.1 \times 0.9}{0.3} = \frac{0.09}{0.30} = \frac{3}{10} = 0.3$$

< Cause - Symptom(effect) relation >



Two major problems of Bayesian

- All relations of evidences are independent.
- Adding new hypotheses and evidences, it is necessary to rebuild probability tables.



$$P(H_i | E) = \frac{P(E | H_i)P(H_i)}{\sum_{k=1}^n (P(E | H_k)P(H_k))}$$

< Bayes theorem
Multiple evidence case >

$$P(Y | X, E) = \frac{P(X | Y, E)P(Y | E)}{P(X | E)}$$

$$\begin{aligned} P(Z | X, Y) &= \frac{P(X, Y | Z)P(Z)}{P(X, Y)} = \frac{P(X | Z)P(Y | Z)P(Z)}{P(X, Y)} \\ &= \alpha(P(Z)P(X | Z)P(Y | Z)) \end{aligned}$$

If $P(X|Z)$ & $P(Y|Z)$ are independent

α : normalization constant

? Symptom2
cause2

cause1

cause \rightarrow symptom

known

symptom \rightarrow cause

unknown

rule r_1

Bayes rule

$$\boxed{P(C_1 | S_2)} = \frac{P(S_2 | C_1)P(C_1)}{P(S_2 | C_1)P(C_1) + P(S_2 | C_2)P(C_2)}$$

$C_1 :$, $C_2 :$

$S_1 :$

$S_2 :$

$$P(C_1 | S_1) = \frac{P(S_1 | C_1)P(C_1)}{P(S_1 | C_1)P(C_1) + P(S_1 | C_3)P(C_3)}$$

rule r_3

Certainty Factor

- Bayes' theorem
- in traditional prob. theory the sum of confidence for a relationship and confidence against the same relationship must add to 1.
- But, often the expert may have confidence 0.7(e.g.) that some relationship is true & have no feeling about its being not true.

- Certainty Factor

– 가?

| | | |
|---------------------|---|---------------------|
| 0 | ~ | 1 |
| completely wrong | | completely right |

| | | | | |
|---------------------|---|---------|---|---------------------|
| -1 | ~ | 0 | ~ | 1 |
| completely wrong | | unknown | | completely right |

– CF

– CF

– CF

$$c \quad r = c/(1 - c)$$

$$r \quad c = r/(r+1)$$

$$c = 0.5 \rightarrow r = 1$$

(_____ true or false)

$$\text{if } c \geq 0.5 \rightarrow r \geq 1$$

$$c < 0.5 \rightarrow r < 1$$

$$\begin{array}{l} \text{if } c_1 \geq 0.5 \\ c_2 < 0.5 \end{array} \begin{array}{l} \searrow \\ \nearrow \end{array} \Rightarrow \begin{array}{l} r \geq 1 \\ r < 1 \end{array} \begin{array}{l} \searrow \\ \nearrow \end{array} \Rightarrow r_1 \cdot r_2 \leq r_1$$

$$\rightarrow c = (r_1 r_2)/(r_1 r_2 + 1)$$

$$\begin{array}{l} \text{e.g. } c_1 = 0.9 \\ c_2 = 0.25 \end{array} \begin{array}{l} \searrow \\ \nearrow \end{array} \Rightarrow \begin{array}{l} r_1 = 9 \\ r_2 = 0.25/0.75 = 1/3 \end{array}$$

$$r_1 r_2 = 3 \rightarrow c = 3/(1+3) = 0.75 < 0.9$$

:

(0.57† 0.25)

- $C \uparrow$ dependent

Input CF

$$\Rightarrow C_1 \times C_2 \times \dots \times C_n$$

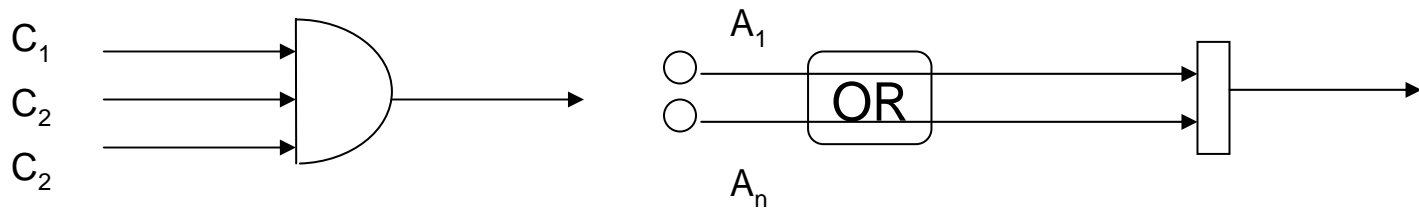
- $A \uparrow$ dependent

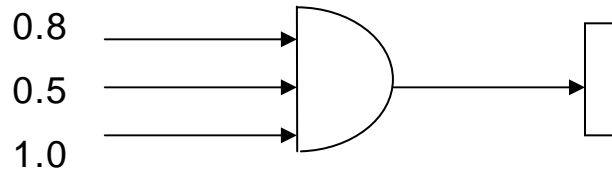
Multiple argued certainty

$$\Rightarrow R_i = A_i / (1 - A_i) ; \text{ Certainty Ratio}$$

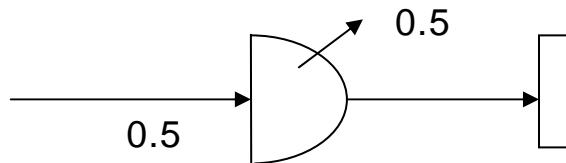
$$\rightarrow R_1 \times R_2 \times \dots \times R_n = R$$

$$\rightarrow A = R / (1 + R)$$

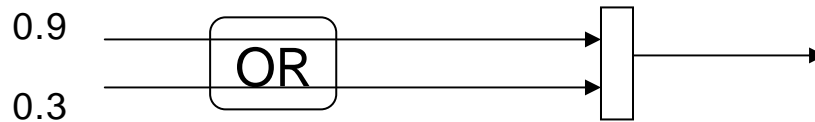




input certainty?



output certainty?

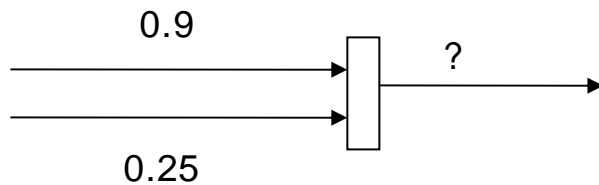
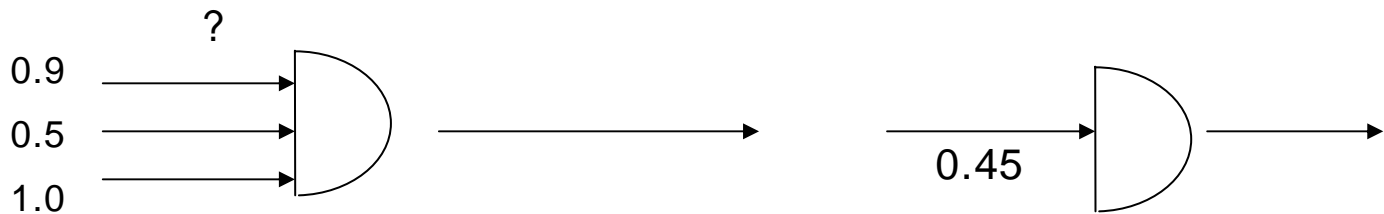


multiple argued
certainty?

$C_i, A_i :$ independent
() ()

Multiple argued

⇒ some conclusion with different rules

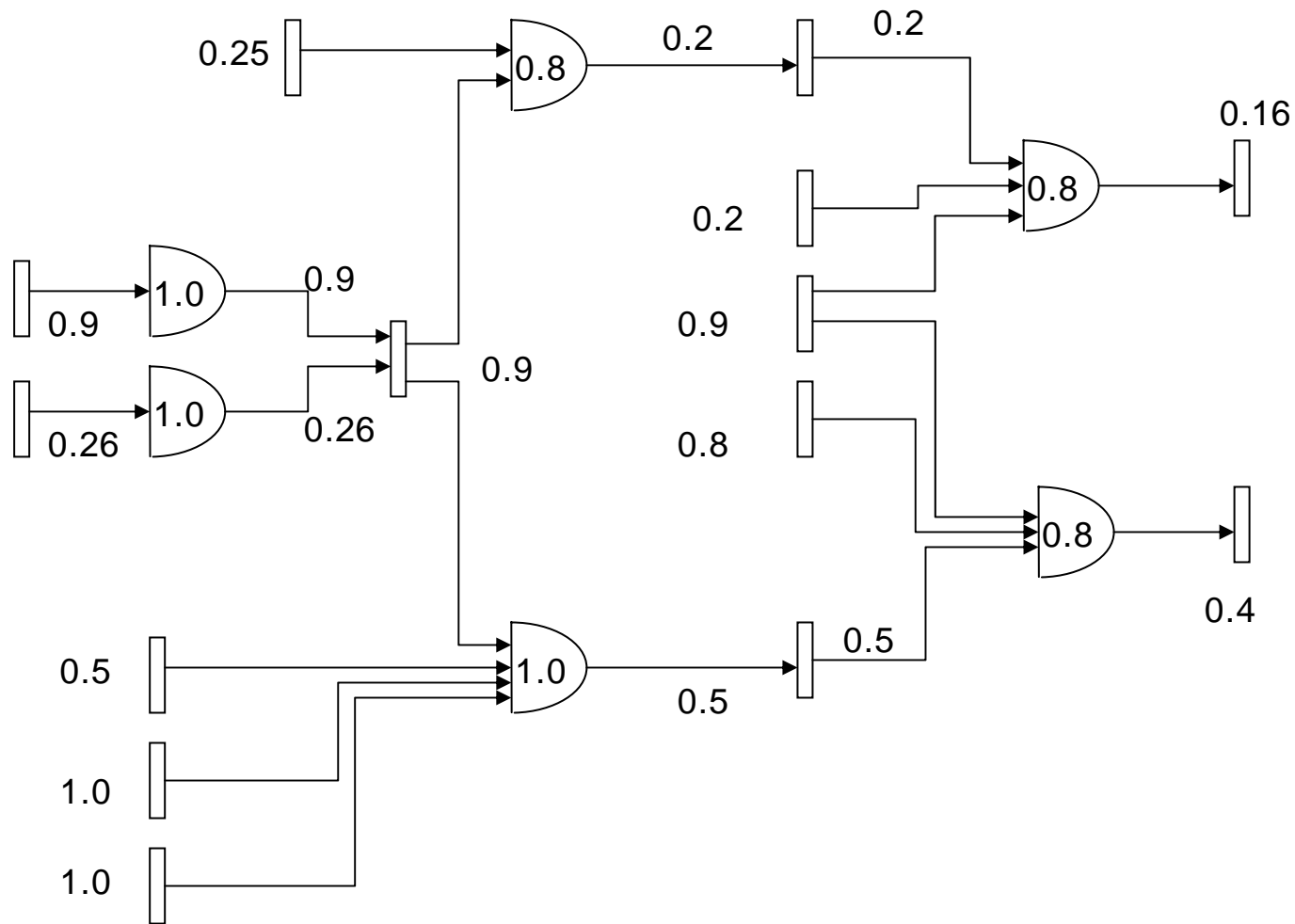


$$CF \quad 0.9 \qquad 0.25$$

$$CR \quad R_1 = \frac{0.9}{1-0.9} = 9 \quad \frac{0.25}{1-0.25} = \frac{1}{3} = R_2$$

$$R_1 \times R_2 = 9 \times \frac{1}{3} = 3 = R$$

$$C = \frac{R}{1+R} = 0.75$$



Conditional problem

| | fever | \neg fever |
|-------------|-------|--------------|
| cold | 0.04 | 0.06 |
| \neg cold | 0.01 | 0.89 |

$$P(\text{cold}) = 0.04 + 0.06 = 0.1$$

$$P(\text{cold} \vee \text{fever}) = 0.04 + 0.01 + 0.06 = 0.11$$

$$= P(\text{cold}) + P(\text{fever}) - P(\text{cold} \wedge \text{fever})$$

$$P(\text{cold} \mid \text{fever}) = \frac{P(\text{fever} \wedge \text{cold})}{P(\text{fever})}$$

$$= \frac{0.04}{0.04 + 0.01} = 0.8$$

Dempster - Shaper theory

- deal with the distinction between uncertainty & ignorance
- computes the probability that the evidence supports the proposition rather than computing the probability of a proposition

→ Belief function

- e.g. Coin flipping

– Head가 belief

(1) since we don't know coin is fair or not
 $\text{Bel}(\text{Head}) = 0$ $\text{Bel}(\neg\text{Head}) = 0$

(2) 가가 fair certainty가 90%
 $\text{Bel}(\text{Head}) = 0.9 \times 0.5 = 0.45$
 $\text{Bel}(\neg\text{Head}) = 0.9 \times 0.5 = 0.45$

Dempster - Shaper theory(cont'd)

(3) ~ 100%

$$\text{Bel}(\text{Head}) = 1 \times 0.5 = 0.5$$

$$\text{Bel}(\neg\text{Head}) = 1 \times 0.5 = 0.5$$

$$\text{Plausibility} = 1 - \text{Bel}()$$

(1) : $\text{Bel}(\text{Head}) = 0$, $\text{Plausibility} = 1 - \text{Bel}(\text{Head}) = 1$

< Probability interval >

(1) [0, 1]

(2) [0.45, 0.55]

(3) [0.5, 0.5]

Dempster - Shaper Theorem

- Belief
 - $b(\phi) = 0$
 - $b(\theta) = 1$
 - for all $A \subset \theta$, $0 \leq b(A) \leq 1$

- Support

$$spt(D_j) = \sum_{D_r \subset D_j} b(D_r)$$

- Plausibility
 - $pl(D_j) = 1 - spt(\neg D_j)$

Example 0

| | | Coin A | |
|--------|------------|------------|-----------|
| | | front(0.8) | back(0.2) |
| Coin B | front(0.6) | 0.48 | 0.12 |
| | back(0.4) | 0.32 | 0.08 |

Belief(Coin A(front) and Coin B(front)) = 0.48

Belief(Coin A(front) and Coin B(back)) = 0.32

Belief(Coin A(back) and Coin B(front)) = 0.12

Belief(Coin A(back) and Coin B(back)) = 0.08

Support(same) = Belief(Coin A(front) and Coin B(front)) +
Belief(Coin A(back) and Coin B(back)) = 0.56

Plausibility(same) = 1 - Belief(Coin A(front) and Coin B(back))
- Belief(Coin A(back) and Coin B(front)) = 0.56

Example 1

| | | Melissa | |
|------|-----------------|---------------|-------------------|
| | | broken(0.9) | don't know(0.1) |
| Bill | not broken(0.8) | 0 | (not broken) 0.08 |
| | don't know(0.2) | (broken) 0.18 | (don't know) 0.02 |

Belief(broken and not broken) = 0

Belief(broken) = $0.18 / 0.28 = 0.643$

Belief(not broken) = $0.08 / 0.28 = 0.286$

Belief(don't know) = $0.02 / 0.28 = 0.071$

Support(broken) = Belief (broken) = 0.643

Plausibility(broken) = $1 - \text{Belief(not broken)} = 0.714$

Support(not broken) = Belief(not broken) = 0.286

Plausibility(not broken) = $1 - \text{Belief(broken)} = 0.357$

Resolution Theorem Proving (Refutation)

- Steps
 1. Put the premises into clause form
 2. Add the negation of what is to be proved in clause form to the set of premises.
 3. Resolve these clauses together, producing new clauses that logically follows.
 4. Produce a contradiction by generating the empty clause.

- Resolution example

1. Whoever can read is literate.
2. Dolphins are not literate.
3. Some dolphins are intelligent cannot read?
→ Some who are intelligent cannot read?

1. $(\forall X)[R(X) \Rightarrow L(X)]$

2. $(\forall X)[D(X) \Rightarrow \neg L(X)]$

3. $(\exists X)[D(X) \wedge I(X)]$

4. $(\exists X)[I(X) \wedge \neg R(X)]$

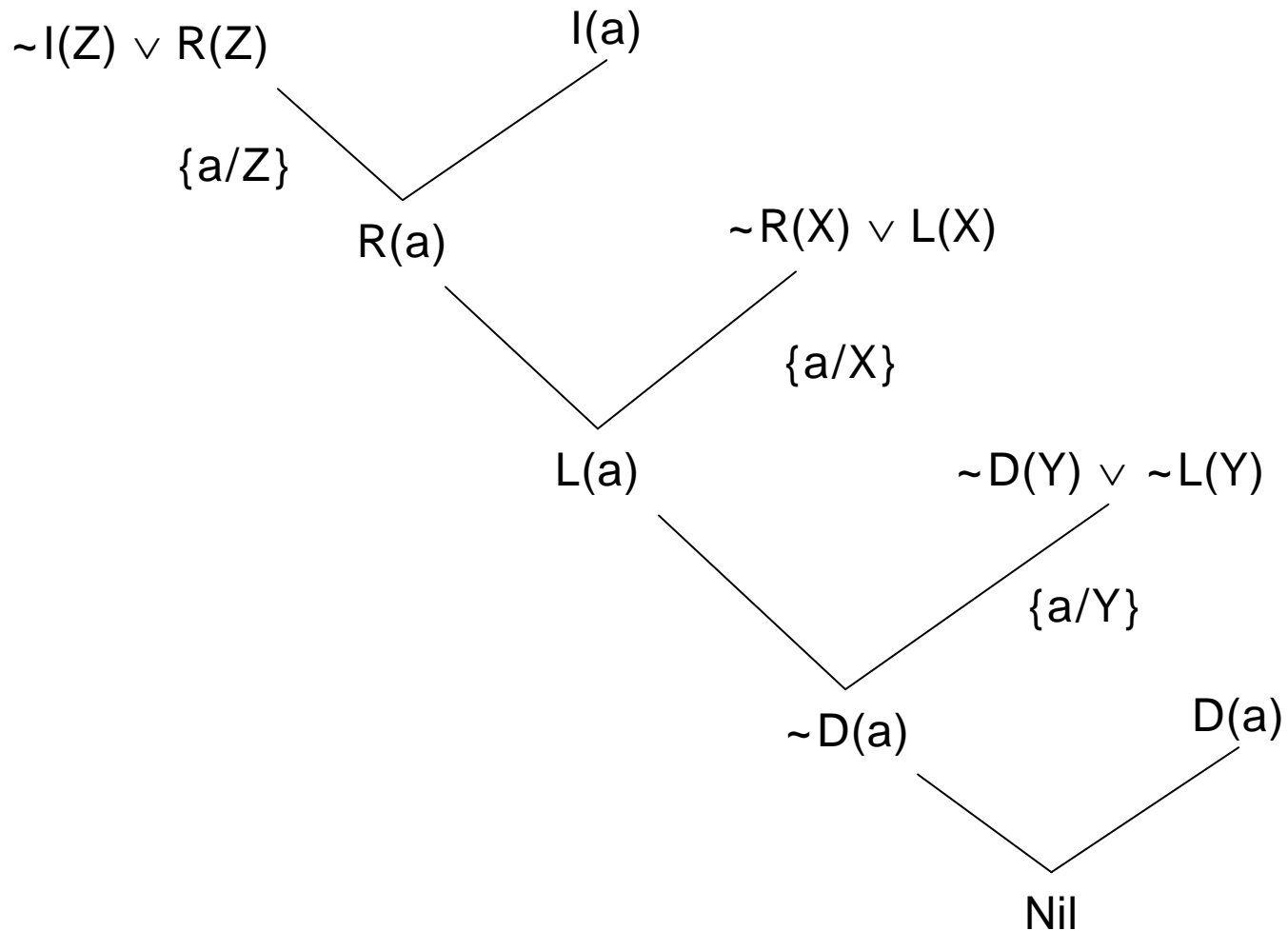
1; $\neg R(X) \vee L(X)$

2; $\neg D(Y) \vee \neg L(Y)$

3; $D(a) \wedge I(a)$

\neg 4; $\neg I(Z) \vee R(Z)$

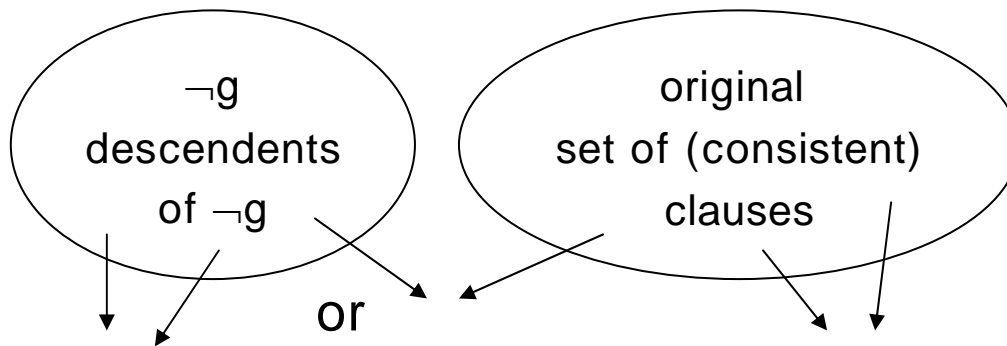
$$\neg \exists X[\sim] = \forall X \neg[\sim]$$



5 : resolution example

therefore $I(a) \wedge \neg R(a)$

- Breadth First Strategy : complete
- Set of support strategy
 - at least one of the resolvents is either the negated goal clause or a clause produced by resolutions on the negated goal \rightarrow complete



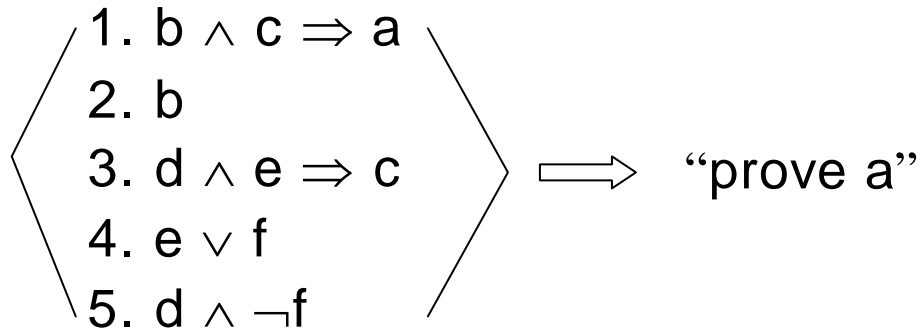
- Unit Preference strategy : complete
 - Unit resolution : not complete
 - \rightarrow one of the resolvents always be a unit clause.
- Linear Input Form Strategy : not complete

- Once the theorem prover shows that the negated goal is inconsistent with the given set of axioms, it follows that the original goal must be consistent → proof of the theorem
- Resolution is a sound inference rule
 - $(P \vee Q) \wedge (\neg G \vee R) \rightarrow P \vee R$
 $P \vee R$ logically follows from $(P \vee Q) \wedge (\neg G \vee R)$
 \therefore sound
 - But it is not complete.
 i.e. given set of axiom logically follow fact
 .
 - refutation complete.
 i.e. the empty or null clause can always be generated wherever a contradiction in the set of clauses exists.
- Refutation
 - inference procedure using resolution, proof by contradiction

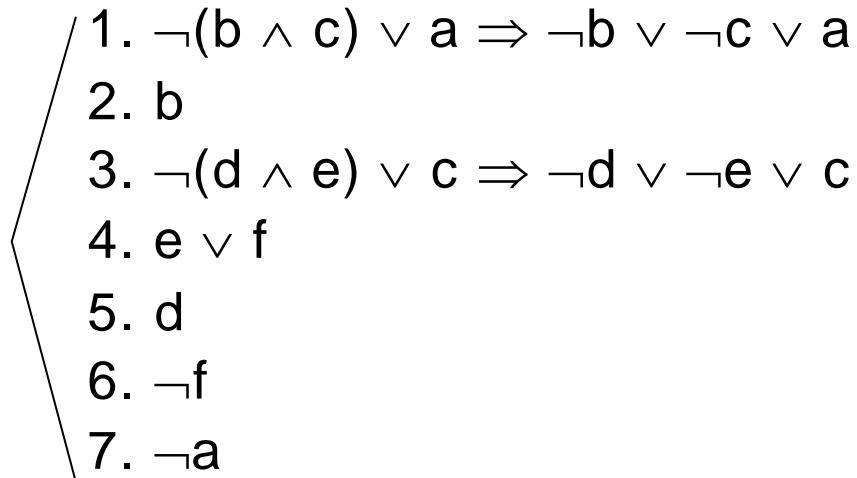
- Refutation
 - proof by contradiction
 - To prove P, assume P is false & prove a contradiction
 - $(S \wedge \neg P \Rightarrow \text{false}) \Leftrightarrow (S \rightarrow P)$

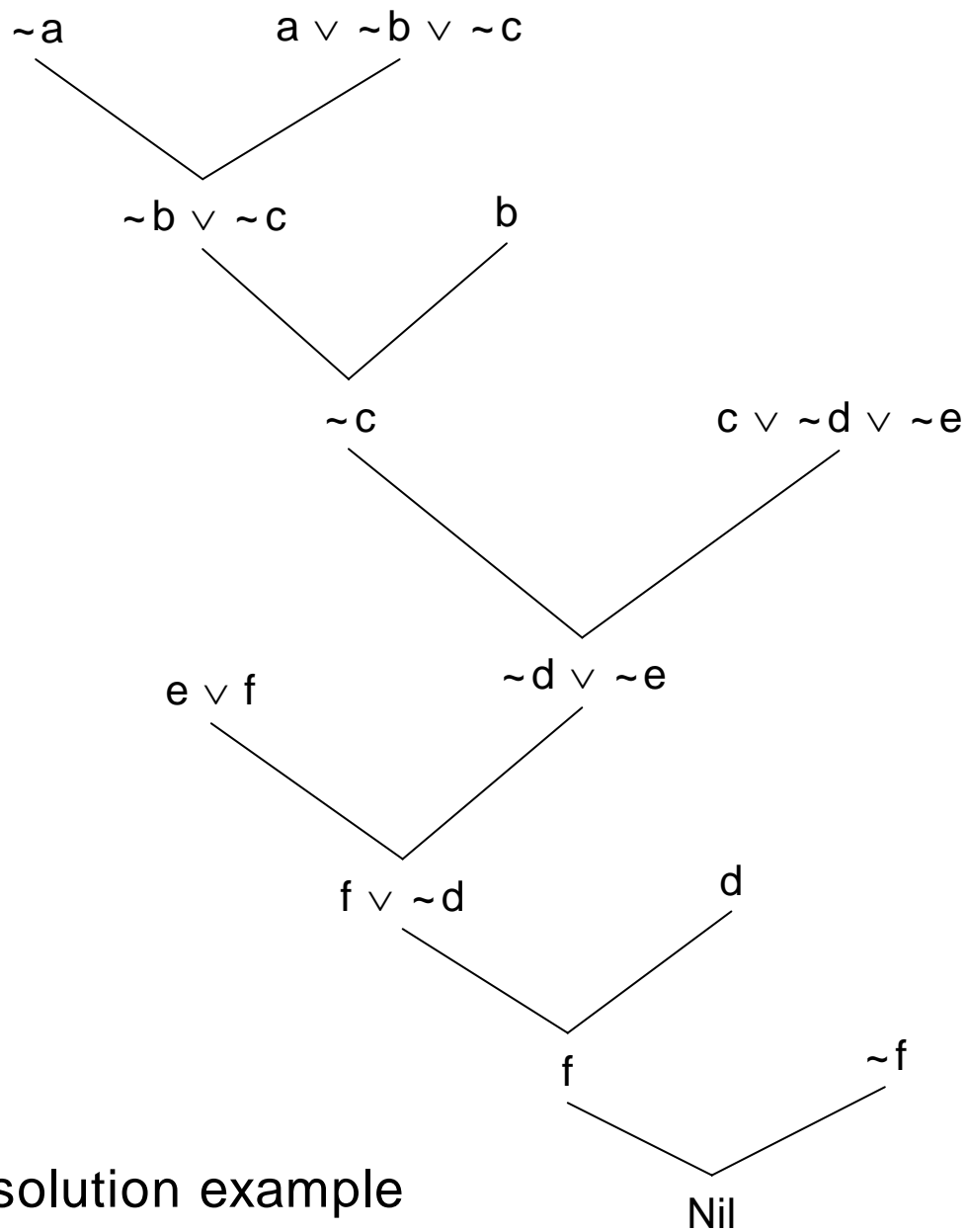
- resolution example

given,



clause form





3 : resolution example

• goal 1 clause

– < Axioms >

$$\neg a(X) \vee f(X) \vee g(f(X))$$

$$\neg f(X) \vee b(X)$$

$$\neg f(X) \vee c(X)$$

$$\neg g(X) \vee b(X)$$

$$\neg g(X) \vee d(X)$$

$$a(g(X)) \vee f(h(X))$$

– < goal >

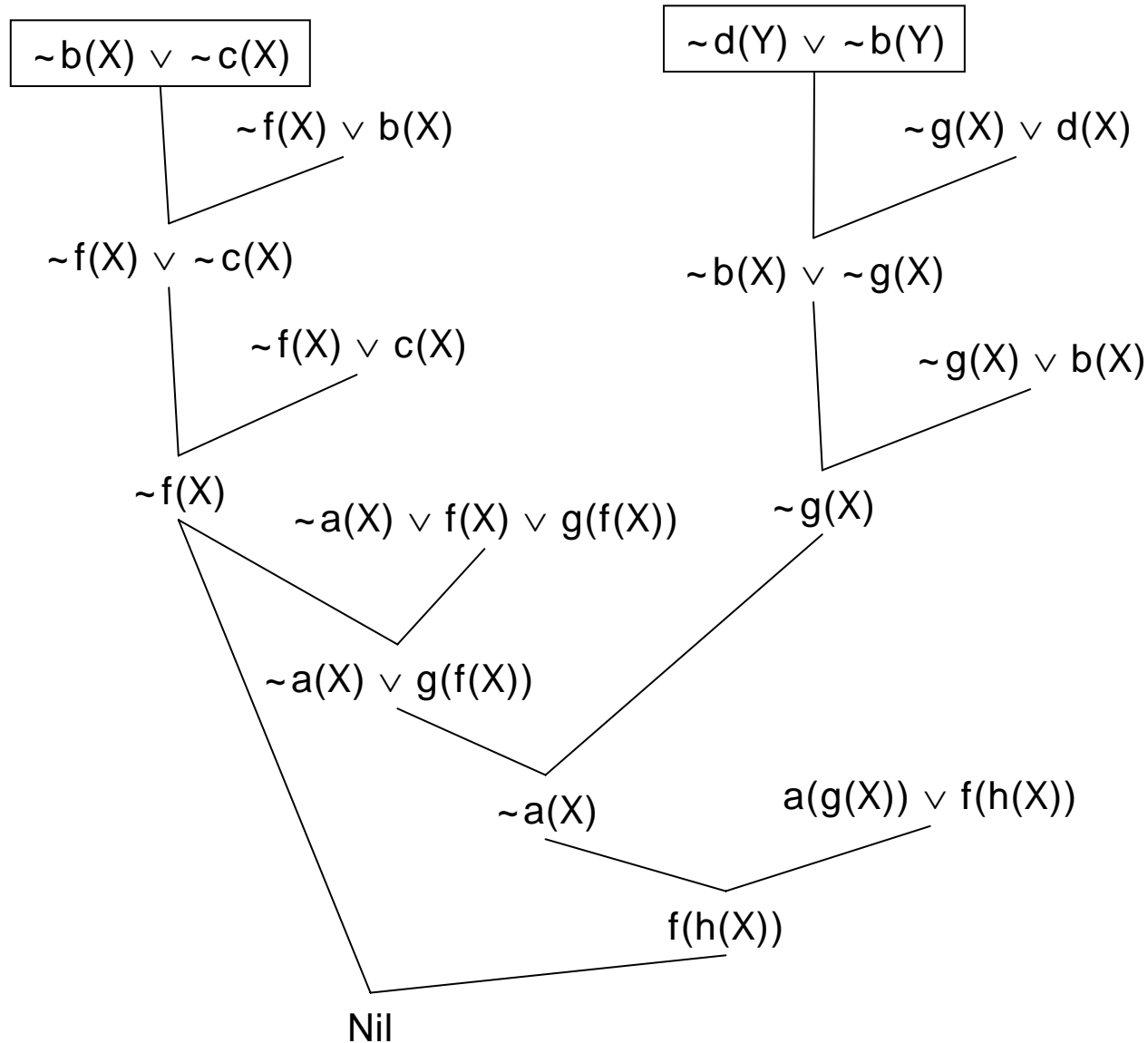
$$(\exists X)(\exists Y)\{[b(X) \wedge c(X)] \vee [d(Y) \wedge b(Y)]\}$$

– < \neg goal >

$$\neg b(X) \vee \neg c(X)$$

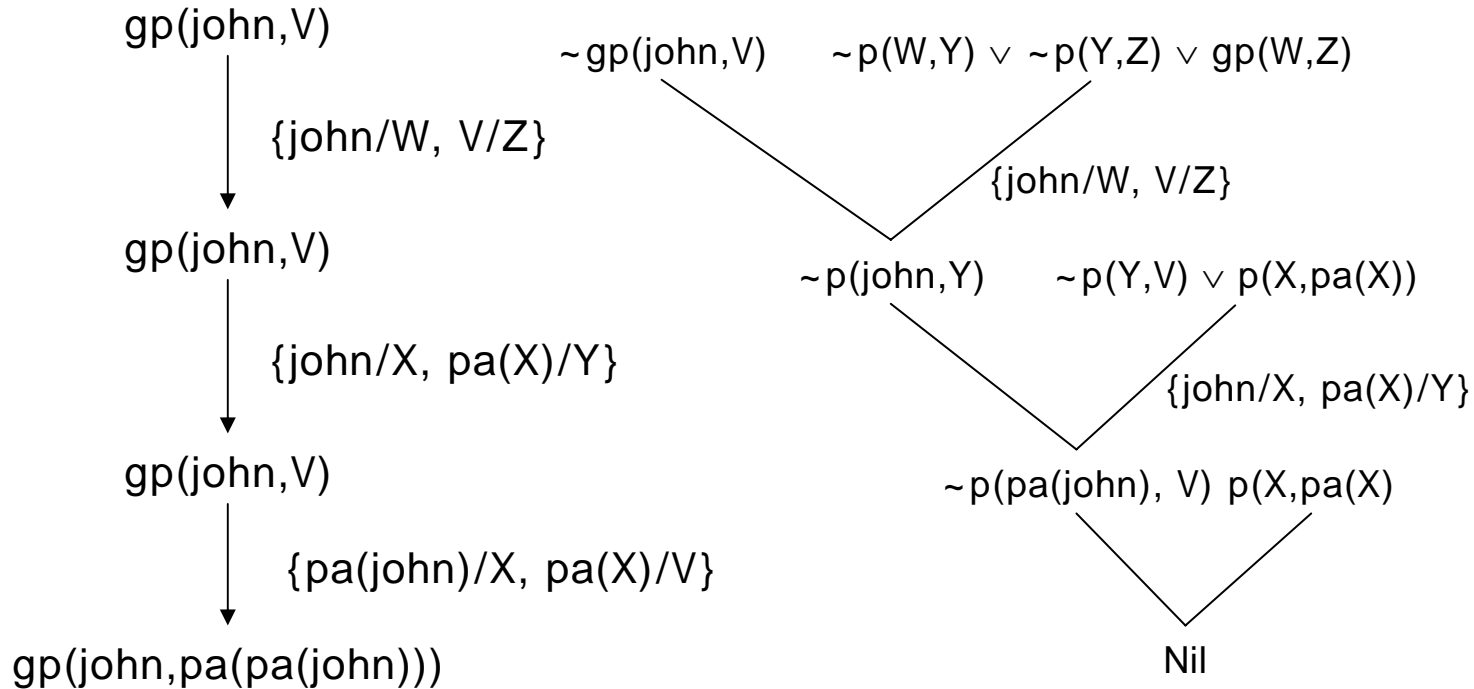
$$\neg b(Y) \vee \neg d(Y)$$

Resolution refutation



- example problem
 - Everyone has a parent.
 1. $(\forall X)(\exists Y)p(X,Y)$
 - A parent of a parent is a grandparent.
 2. $(\forall X)(\forall Y)(\forall Z)p(X,Y) \wedge p(Y,Z) \Rightarrow gp(X,Z)$
 - goal ; John has a grandparent?

$$(\exists W)(gp(john,W))$$
 - negation of goal
 3. $\neg gp(john,W)$
- Clause form
 1. $p(X,pa(X))$
 2. $\neg p(W,Y) \vee \neg p(Y,Z) \vee gp(W,Z)$
 3. $\neg gp(john,V)$



$V \rightarrow pa(X) \rightarrow pa(pa(john))$

* $gp(john, pa(pa(john)))$ is proved

- tautology proof & answer extraction

– goal statement negation tautology

e.g.) $\neg gp(john, V) \rightarrow \neg gp(john, V) \vee gp(john, V)$

$gp(john, V) \vee \sim gp(john, V) \quad \sim p(W, Y) \vee \sim p(Y, Z) \vee gp(W, Z)$

$\{john/W, V/Z\}$

$gp(john, V) \vee \sim p(john, Y) \quad \sim p(Y, V) \vee p(X, pa(X))$

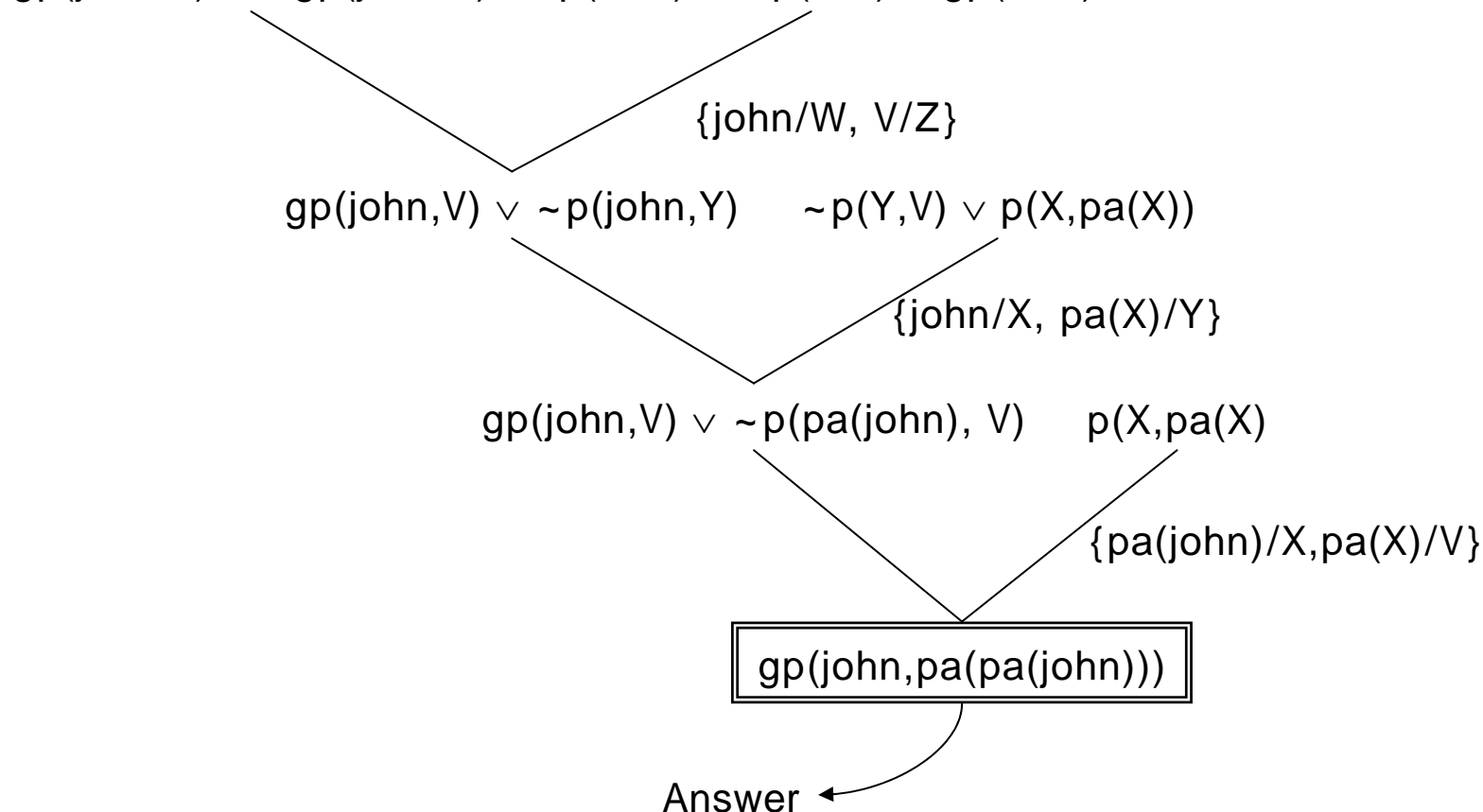
$\{john/X, pa(X)/Y\}$

$gp(john, V) \vee \sim p(pa(john), V) \quad p(X, pa(X))$

$\{pa(john)/X, pa(X)/V\}$

$gp(john, pa(pa(john)))$

Answer



Resolution

- ?tolology
 - $p(X) \vee \neg p(X)$
- subsume clause
 - $p(\text{john})$ subsumes $\forall X p(X)$
 - $p(X)$ subsumes $p(X) \vee q(X)$
- procedure attachment

| | | | | |
|---|---|----------|---------|------------|
| T | F | evaluate | literal | evaluate |
| – | | | literal | T evaluate |
| | 가 | . | | |
| – | | | literal | F evaluate |
| | 가 | . | | literal |

“If Fido goes where John goes & if John is at school, where is Fido?”

1. $(\forall X)[\text{at}(\text{john}, X) \Rightarrow \text{at}(\text{fido}, X)]$

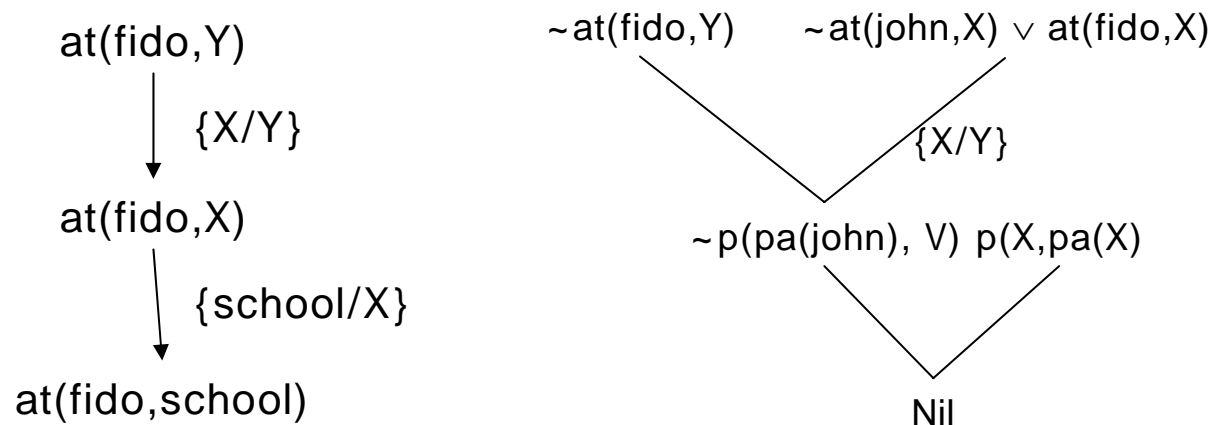
2. $\text{at}(\text{john}, \text{school})$

3. $(\exists X)\text{at}(\text{fido}, X) ? \rightarrow \text{goal}$

1; $\neg \text{at}(\text{john}, X) \vee \text{at}(\text{fido}, X)$

2; $\text{at}(\text{john}, \text{school})$

\neg 3; $\neg \text{at}(\text{fido}, X)$



4 : resolution example

• retain original goal & apply all the substitutions of the refutation to this clause.
we find the answer.

therefore $\text{at}(\text{fido}, \text{school})$
fido school .

- Anyone passing his history exams and winning the lottery is happy

$$\forall X(\text{pass}(X, \text{history}) \wedge \text{win}(X, \text{lottery}) \rightarrow \text{happy}(X))$$

- Anyone who studies or is lucky can pass all his exams

$$\forall X \forall Y (\text{studies}(X) \vee \text{lucky}(X) \rightarrow \text{pass}(X, Y))$$

- John did not study but he is lucky

$$\neg \text{study}(\text{john}) \wedge \text{lucky}(\text{john})$$

- Anyone who is lucky wins the lottery

$$\forall X (\text{lucky}(X) \rightarrow \text{win}(X, \text{lottery}))$$

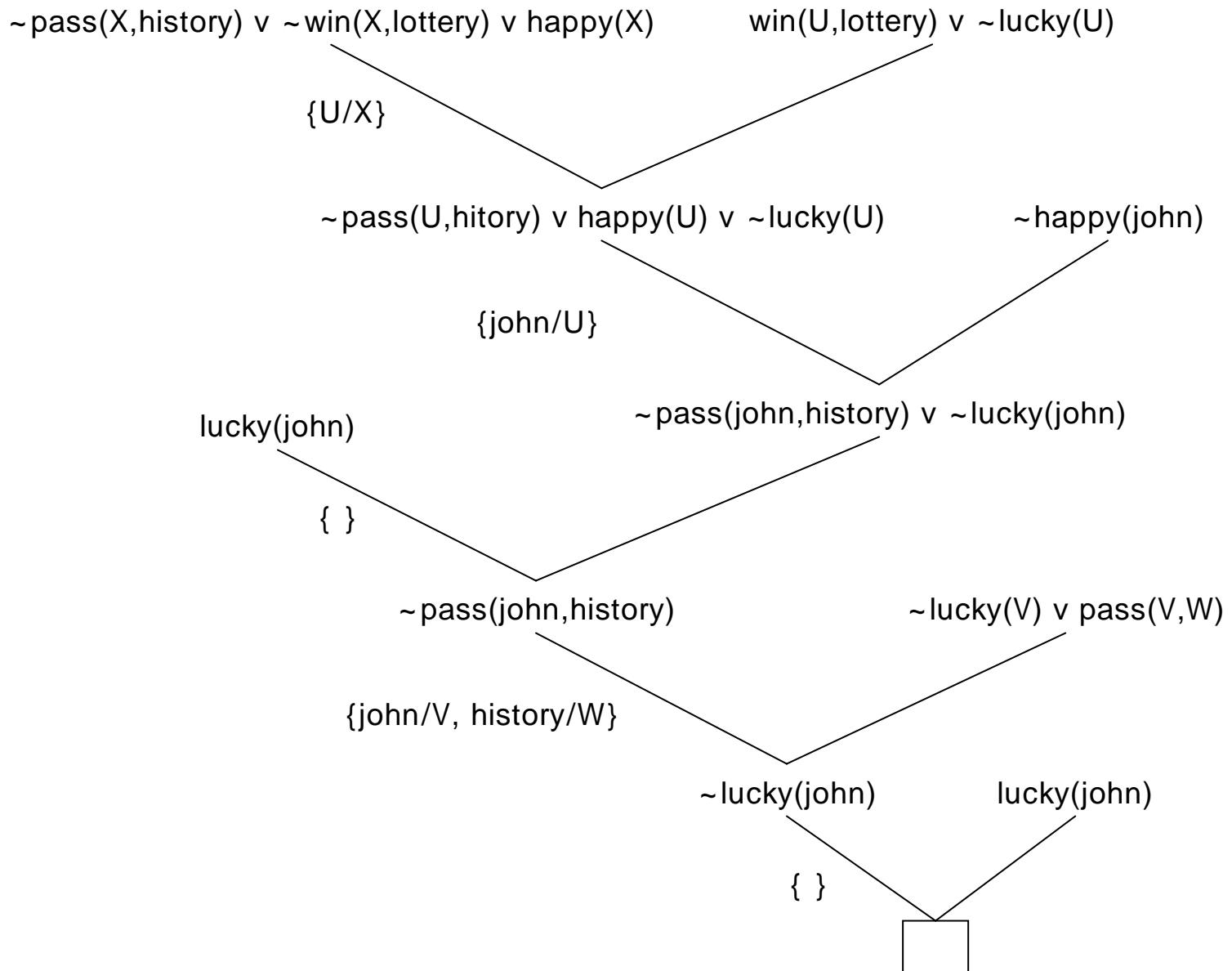
then

- Is John happy?

$\text{happy}(\text{john})$

< Clause form >

1. $\neg \text{pass}(X, \text{history}) \vee \neg \text{win}(X, \text{lottery}) \vee \text{happy}(X)$
2. $\neg \text{study}(Y) \vee \text{pass}(Y, Z)$
 $\neg \text{lucky}(W) \vee \text{pass}(W, V)$
3. $\neg \text{study}(\text{john})$
 $\text{lucky}(\text{john})$
4. $\neg \text{lucky}(U) \vee \text{win}(U, \text{lottery})$
5. $\neg \text{happy}(\text{john})$?



- Resolution strategy

- 1. Breadth - First strategy

- N clause in the original clause set.

- 1st level : N^2 ways of combination

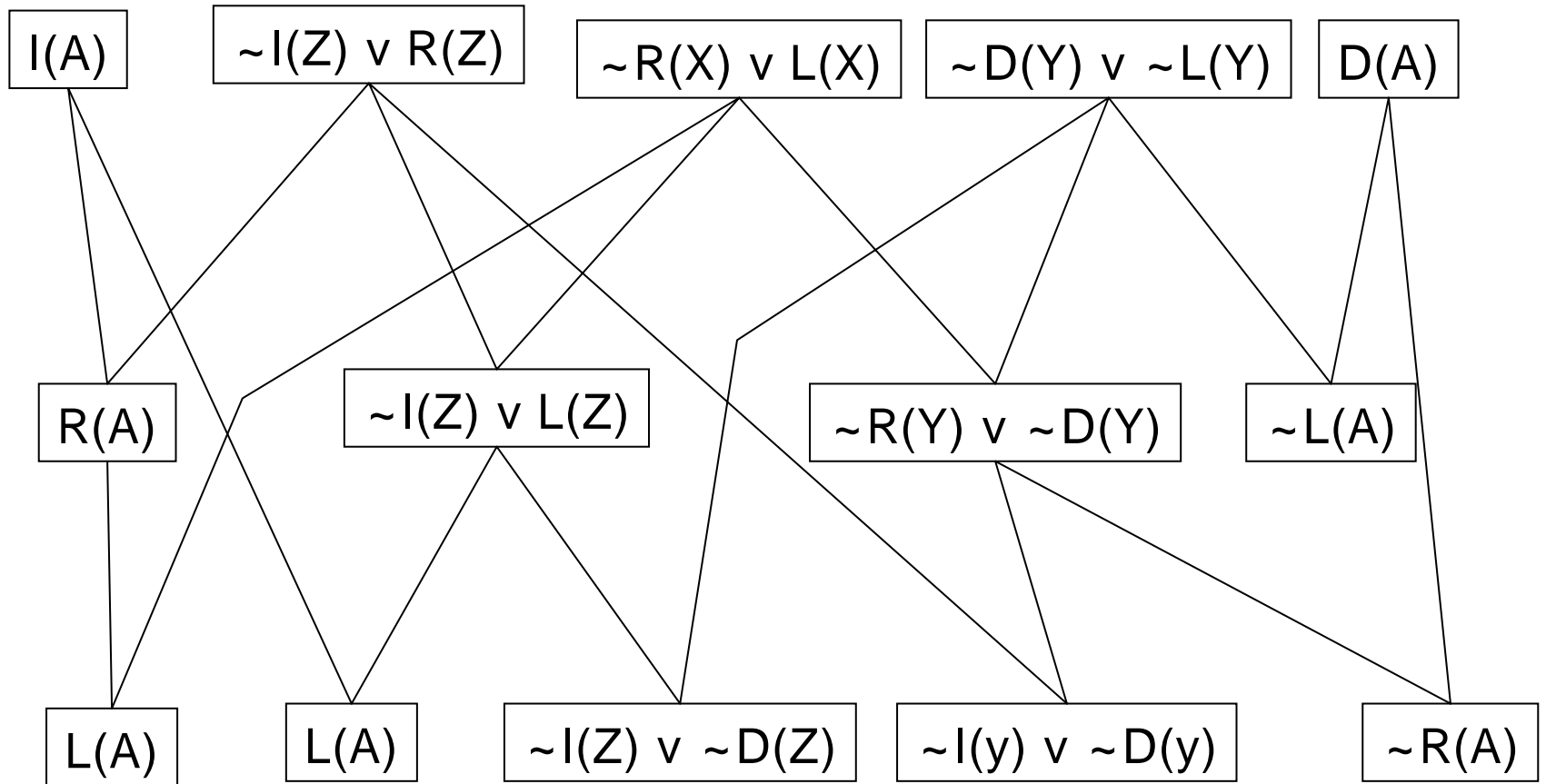
- 2nd level : Resolve the clauses produced at the 1st level with all the original clauses.

- nth level : Resolve all clauses at the level $n - 1$ against the elements of the original clause set & all causes previously produced.

- * large search space

- * find the shortest path solution

- * if refutation exist, it always finds. → complete.



< Breadth - First strategy >

1. The set of support strategy

Suppose a set of input clauses : S

subset of S that includes the negation of the goal : T

S is contradictory \leftrightarrow T is contradictory

* The negation of what we want to prove true is responsible for causing the clause space to be contradictory

```
* Resolution      parent      goal
    descendent      .
```


$$\sim I(Z) \vee R(Z) \quad I(A) \quad \sim R(X) \vee L(X) \quad \sim D(Y) \vee \sim L(Y)$$

$$R(A)$$

$$\sim I(Z) \vee L(Z)$$

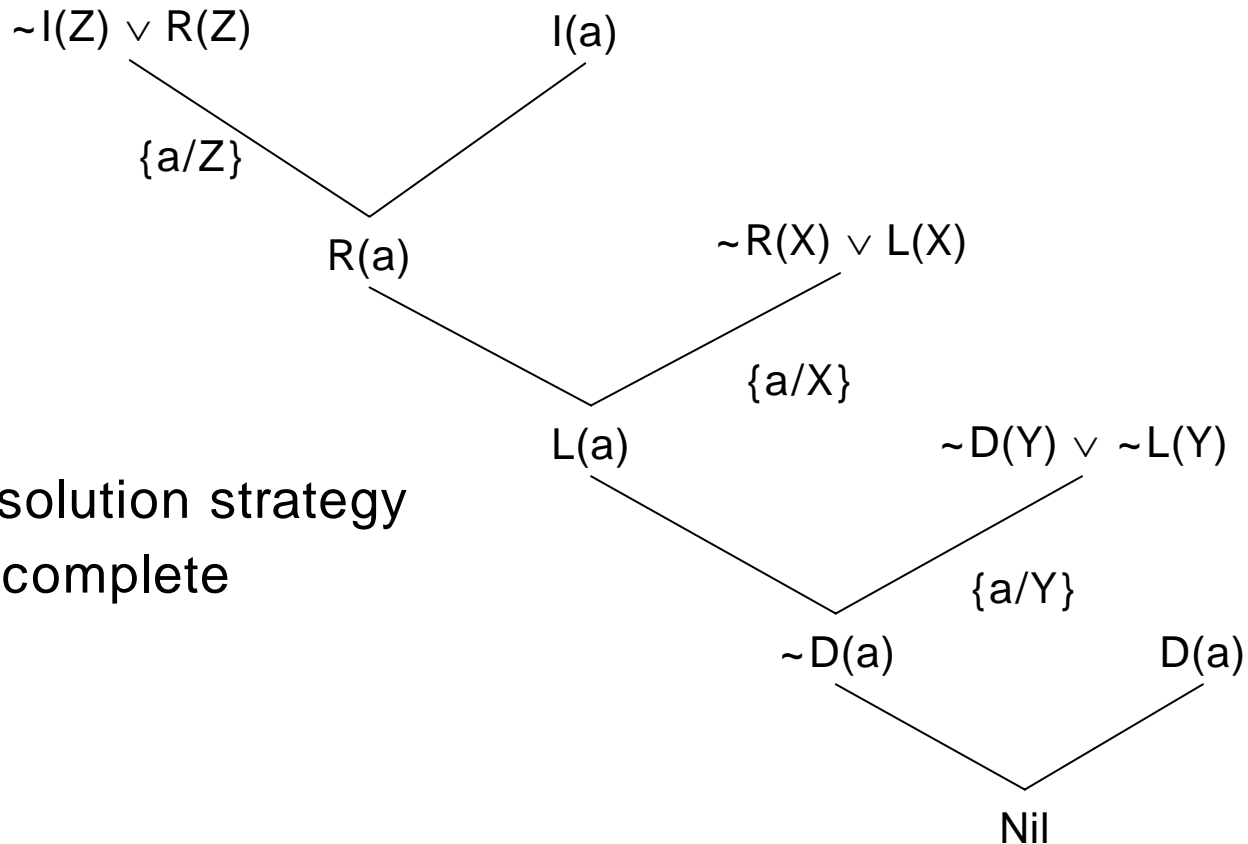
$$L(A)$$

$$\sim D(A)$$

$$\sim I(A)$$

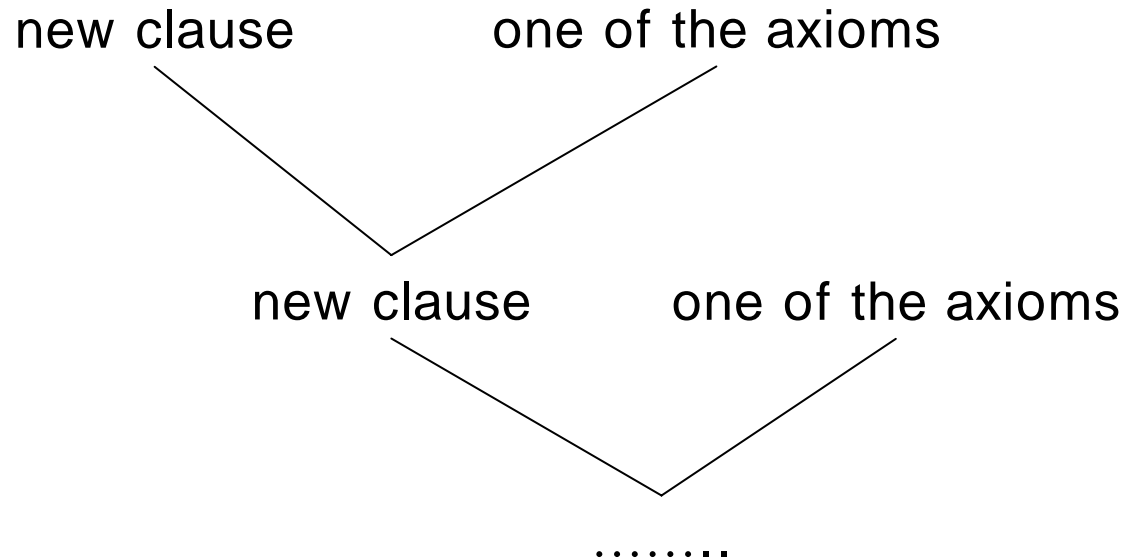
$$\sim D(A)$$

- Unit Preference strategy()
 - literal 가 clause resolution Nil
 - 가 .
 - Resolution literal .
 - complete



Unit resolution strategy
 → not complete

- Linear Input Form strategy
 - negated goal & the original axioms
 - take the negated goal & resolve it with one of the axioms.



- No previously derived clause is used.
- No two axioms are used.
- not complete