Knowledgebase Representation

Uncertainty Management in Rule-based Expert Systems

What is Uncertainty?

- Information available to human experts can be incomplete, inconsistent, uncertain, or all three.
- Uncertainty can be defined as the lack of the exact knowledge that would enable us to reach a perfectly reliable conclusion.
- Classical logic permits only exact reasoning. It assumes that perfect knowledge always exists where:

If A is TRUE then A is not FALSE

If A is FALSE then A is not TRUE

- Most real-world problems, where expert systems could be used, do not provide us with clear-cut knowledge.
- However, expert systems have to be able to handle uncertainty and draw valid conclusions.
- There are variant sources of uncertain knowledge in expert systems.

Sources of Uncertain Knowledge

- **Weak implications:**
  - Difficult task of establishing concrete correlations between IF (condition) and THEN (action) parts of the rules.
  - Expert systems need to have the ability to handle vague associations.

- **Imprecise language:**
  - Our natural language is inherently ambiguous and imprecise.
    - We describe facts with such terms as often and sometimes, frequently and hardly ever.
  - As a result, it can be difficult to express knowledge in the precise IF-THEN form of production rules.
  - However, if the meaning of the facts is quantified, it can be used in expert systems.

Quantification of imprecise terms on a time-frequency scale

Ray Simpson (1944)		Milton Hakel (1968)	
Term	**Mean value**	**Term**	**Mean value**
Always	99	Always	100
Very often	88	Very often	87
Usually	85	Usually	79
Often	78	Often	74
Generally	78	Rather often	74
Frequently	73	Frequently	72
Rather often	65	Generally	72
About as often as not	50	About as often as not	50
Now and then	20	Now and then	34
Sometimes	20	Sometimes	29
Occasionally	20	Occasionally	28
Once in a while	15	Once in a while	22
Not often	13	Not often	16
Usually not	10	Usually not	16
Seldom	10	Seldom	9
Hardly ever	7	Hardly ever	8
Very seldom	6	Very seldom	7
Rarely	5	Rarely	5
Almost never	3	Almost never	2
Never	0	Never	0

Sources of Uncertain Knowledge (cont.)

- **Unknown data:**
  - When the data is incomplete or missing, the only solution is to accept the value ‘unknown’ and proceed to an approximate reasoning with this value.

- **Combining views of different experts:**
  - Large expert systems usually combine the knowledge and expertise of a number of experts.
  - Nine experts participated in the development of PROSPECTOR.
  - Usually, experts have contradictory opinions and produce conflicting rules.
  - To resolve the conflict, the knowledge engineer has to attach a weight to each expert and then calculate the composite conclusion.

Certainty Factors Theory

- Introduced in MYCIN expert system as alternative to Bayesian reasoning.
- A certainty factor (*cf*) measure the expert's belief.
  - The maximum value of the certainty factor was +1.0 (definitely true) and the minimum -1.0 (definitely false).
  - A positive value represented a degree of belief and a negative a degree of disbelief.

- In expert systems with certainty factors, the knowledge base consists of a set of rules in the following syntax:

IF <evidence>

THEN <hypothesis> {*cf*}

where *cf* represents belief in hypothesis *H* given that evidence *E* has occurred.

Uncertain Terms and Their Interpretation

Term	Certainty factor
Definitely not	−1.0
Almost certainly not	−0.8
Probably not	−0.6
Maybe not	−0.4
Unknown	−0.2 to +0.2
Maybe	+0.4
Probably	+0.6
Almost certainly	+0.8
Definitely	+1.0

Certainty Factors Theory (cont.)

- The certainty factors theory is based on two functions:
  - Measure of belief $MB(H, E)$, and
  - Measure of disbelief $MD(H, E)$.
- The measure of belief indicates the degree to which belief in hypothesis H would be increased if evidence E were observed.
- The measure of disbelief indicates the degree to which disbelief in hypothesis H would be increased by observing the same evidence E.
- These functions can be defined in terms of prior and conditional probabilities as follows:

Certainty Factors Theory (cont.)

$$MB(H,E) = \begin{cases} 1 & \text{if } p(H) = 1 \\ \dfrac{max\,[p(H|E), p(H)] - p(H)}{max\,[1,0] - p(H)} & \text{otherwise} \end{cases}$$

$$MD(H,E) = \begin{cases} 1 & \text{if } p(H) = 0 \\ \dfrac{min\,[p(H|E), p(H)] - p(H)}{min\,[1,0] - p(H)} & \text{otherwise} \end{cases}$$

- where:
  - $p(H)$ is the prior probability of hypothesis H being true.
  - $p(H|E)$ is the probability that hypothesis H is true given evidence E.
- Values of $MB(H, E)$ and $MD(H, E)$ range between 0 and 1.

Certainty Factors Theory (cont.)

- The strength of belief or disbelief in hypothesis H depends on the kind of evidence E observed.
  - Some facts may increase the strength of belief, but some increase the strength of disbelief.
- The two measures are combined into one number, the certainty factor, using the following equation:

$$cf = \frac{MB(H,E) - MD(H,E)}{1 - min\,[MB(H,E), MD(H,E)]}$$

- Thus cf, which can range from -1 to +1, indicates the total belief in hypothesis H.

Certainty Factors Theory (cont.)

- The certainty factor assigned by a rule can be propagated through the reasoning chain.
- Propagation of the certainty factor involves establishing the net certainty of the rule conclusion when the evidence in the rule premise is uncertain.
- The net certainty for a single premise rule, cf(H, E), can be easily calculated by multiplying the certainty factor of the premise, cf(E), with the rule certainty factor, cf:

$$cf(H, E) = cf(E) \times cf$$

Certainty Factors Theory (cont.)

- <u>Example:</u>
  - Assume the following rule:

IF	the sky is clear
THEN	the forecast is sunny {*cf* 0.8}

  - And the current certainty factor of *sky is clear* is 0.5, then

$$cf(H,E) = 0.5 \times 0.8 = 0.4$$

  - This result, according to the previous table, would read as 'It may be sunny'.

Certainty Factors Theory (cont.)

- For conjunctive rules such as:

IF <evidence E1 >
AND <evidence E2 >
.
.
.
AND <evidence En >
THEN <hypothesis H > {cf}

- The net certainty of the conclusion, or in other words the certainty of hypothesis H, is established as follows:

$$cf(H, E_1 \cap E_2 \cap \ldots \cap E_n) = min\,[cf(E_1), cf(E_2), \ldots, cf(E_n)] \times cf$$

Certainty Factors Theory (cont.)

- <u>Example:</u>
  - Assume the following rule:

IF	sky is clear
AND	the forecast is sunny
THEN	the action is 'wear sunglasses' {*cf* 0.8}

  - And the certainty of *sky is clear* is 0.9 and the certainty of *the forecast is sunny* is 0.7, then

$$cf(H, E_1 \cap E_2) = min\,[0.9, 0.7] \times 0.8 = 0.7 \times 0.8 = 0.56$$

  - According to the previous table, this conclusion might be interpreted as 'Probably it would be a good idea to wear sunglasses today'.

Certainty Factors Theory (cont.)

- For disjunctive rules such as:

IF <evidence E_1 >
OR <evidence E_2 >
.
.
.
OR <evidence E_n >
THEN <hypothesis H > $\{cf\}$

- The certainty of hypothesis H, is determined as follows:

$$cf(H, E_1 \cup E_2 \cup \ldots \cup E_n) = max\left[cf(E_1), cf(E_2), \ldots, cf(E_n)\right] \times cf$$

Certainty Factors Theory (cont.)

- **Example:**

  - Assume the following rule:

IF	sky is overcast
OR	the forecast is rain
THEN	the action is 'take an umbrella' {cf 0.9}

  - And the certainty of sky is overcast is 0.6 and the certainty of the forecast is rain is 0.8, then

$$cf(H, E_1 \cup E_2) = max[0.6, 0.8] \times 0.9 = 0.8 \times 0.9 = 0.72$$

  - Which can be interpreted as 'Almost certainly an umbrella should be taken today'.

Certainty Factors Theory (cont.)

- When the same conclusion is obtained as a result of the execution of two or more rules:
  - Individual certainty factors of these rules must be merged to give a combined certainty factor for a hypothesis; using the following equation:

$$cf(cf_1, cf_2) = \begin{cases} cf_1 + cf_2 \times (1 - cf_1) & \text{if } cf_1 > 0 \text{ and } cf_2 > 0 \\ \dfrac{cf_1 + cf_2}{1 - min\,[|cf_1|, |cf_2|]} & \text{if } cf_1 < 0 \text{ or } cf_2 < 0 \\ cf_1 + cf_2 \times (1 + cf_1) & \text{if } cf_1 < 0 \text{ and } cf_2 < 0 \end{cases}$$

  - where:

    cf1 is the confidence in hypothesis H established by Rule 1;

    cf2 is the confidence in hypothesis H established by Rule 2;

    |cf1 | and |cf2| are absolute magnitudes of cf1 and cf2, respectively

Certainty Factors Theory (cont.)

- **Example:**

  - Suppose the knowledge base consists of the following rules:

Rule 1:	IF	A is X
	THEN	C is Z {cf 0.8}
Rule 2:	IF	B is Y
	THEN	C is Z {cf 0.6}

  - What certainty should be assigned to object C having value Z if both Rule 1 and Rule 2 are fired?
  - Our common sense suggests that:
    - If we have two pieces of evidence (A is X and B is Y) from different sources (Rule 1 and Rule 2) supporting the same hypothesis (C is Z)
    - Then the confidence in this hypothesis should increase and become stronger than obtaining only one piece of evidence.

Certainty Factors Theory (cont.)

- <u>Case (1):</u>
  - If we assume that: $cf(E_1) = cf(E_2) = 1.0$
  - Then the certainty factor for each rule:

$$cf_1(H, E_1) = cf(E_1) \times cf_1 = 1.0 \times 0.8 = 0.8$$
$$cf_2(H, E_2) = cf(E_2) \times cf_2 = 1.0 \times 0.6 = 0.6$$

  - And the combined certainty factor for the hypothesis:

$$\begin{aligned} cf(cf_1, cf_2) &= cf_1(H, E_1) + cf_2(H, E_2) \times [1 - cf_1(H, E_1)] \\ &= 0.8 + 0.6 \times (1 - 0.8) = 0.92 \end{aligned}$$

  - This example shows an incremental increase of belief in a hypothesis and also confirms our expectations.

Certainty Factors Theory (cont.)

- Case (2): when rule certainty factors have the opposite signs
  - Suppose that:

$$cf(E_1) = 1 \text{ and } cf(E_2) = -1.0$$

  - Then the certainty factor for each rule:

$$cf_1(H, E_1) = 1.0 \times 0.8 = 0.8$$
$$cf_2(H, E_2) = -1.0 \times 0.6 = -0.6$$

  - And the combined certainty factor for the hypothesis:

$$cf(cf_1, cf_2) = \frac{cf_1(H, E_1) + cf_2(H, E_2)}{1 - min\left[|cf_1(H, E_1)|, |cf_2(H, E_2)|\right]} = \frac{0.8 - 0.6}{1 - min\,[0.8, 0.6]} = 0.5$$

  - This example shows how a combined certainty factor is obtained when one rule, Rule 1, confirms a hypothesis but another, Rule 2, discounts it.

Certainty Factors Theory (cont.)

- Case (3): when rule certainty factors have negative signs
  - Suppose that: $cf(E_1) = cf(E_2) = -1.0$
  - Then the certainty factor for each rule:

$$cf_1(H, E_1) = -1.0 \times 0.8 = -0.8$$
$$cf_2(H, E_2) = -1.0 \times 0.6 = -0.6$$

  - And the combined certainty factor for the hypothesis:

$$cf(cf_1, cf_2) = cf_1(H, E_1) + cf_2(H, E_2) \times [1 + cf_1(H, E_1)]$$
$$= -0.8 - 0.6 \times (1 - 0.8) = -0.92$$

  - This example represents an incremental increase of disbelief in a hypothesis.

FORECAST:
An Application of Certainty Factors

- The expert system is required to establish certainty factors for the multi-valued object tomorrow.
- **Knowledge base:**

❖ Rule: 1
if today is rain
then tomorrow is rain {cf 0.5}

❖ Rule: 2
if today is dry
then tomorrow is dry {cf 0.5}

❖ Rule: 3
if today is rain
and rainfall is low
then tomorrow is dry {cf 0.6}

❖ Rule: 4
if today is rain
and rainfall is low
and temperature is cold
then tomorrow is dry {cf 0.7}

❖ Rule: 5
if today is dry
and temperature is warm
then tomorrow is rain {cf 0.65}

❖ Rule: 6
if today is dry
and temperature is warm
and sky is overcast
then tomorrow is rain {cf 0.55}

seek tomorrow

FORECAST (cont.)

The user inputs the facts:

- The weather today is rain
- The rainfall is low with cf=0.8
- The temperature today is cold with cf=0.9

Inference:

Rule: 1
if today is rain
then tomorrow is rain {cf 0.5}

$$cf\,(\text{tomorrow is rain, today is rain}) = cf\,(\text{today is rain}) \times cf = 1.0 \times 0.5 = 0.5$$

tomorrow is rain {0.50}

FORECAST (cont.)

Rule: 3
if today is rain
and rainfall is low
then tomorrow is dry {cf 0.6}

$$cf\,(\text{tomorrow is dry, today is rain} \cap \text{rainfall is low})$$
$$= min\,[cf\,(\text{today is rain}), cf\,(\text{rainfall is low})] \times cf = min\,[1, 0.8] \times 0.6 = 0.48$$

tomorrow is	rain	{0.50}
	dry	{0.48}

FORECAST (cont.)

Rule: 4

if today is rain
and rainfall is low
and temperature is cold
then tomorrow is dry {cf 0.7}

$$cf\,(\text{tomorrow is dry, today is rain} \cap \text{rainfall is low} \cap \text{temperature is cold})$$
$$= min\,[cf\,(\text{today is rain}), cf\,(\text{rainfall is low}), cf\,(\text{temperature is cold})] \times cf$$
$$= min\,[1, 0.8, 0.9] \times 0.7 = 0.56$$

FORECAST (cont.)

tomorrow is dry {0.56}
rain {0.50}

$$cf(cf_{\text{Rule:3}}, cf_{\text{Rule:4}}) = cf_{\text{Rule:3}} + cf_{\text{Rule:4}} \times (1 - cf_{\text{Rule:3}})$$
$$= 0.48 + 0.56 \times (1 - 0.48) = 0.77$$

tomorrow is dry {0.77}
rain {0.50}

- Now we would conclude that the probability of having a dry day tomorrow is almost certain; however we also may expect some rain!

Example 2

Let us consider the following expert system for diagnosing a cold. The database consists of the following facts: (patient's fever is 37.4 , is coughing since less than 24 hours, is sneezing, is having headache with a CF = 0.4 and is having nasal congestion with CF = 0.5).

The rule base consists of the following rules:

Rule 1
IF fever < 37.5
THEN Cold symptoms = true {CF = 0.5}

Rule 2
IF fever > 37.5
THEN Cold symptoms = true {CF = 0.9}

Rule 3
IF cough for more than 24 hours
THEN soar troth = true {CF = 0.5}

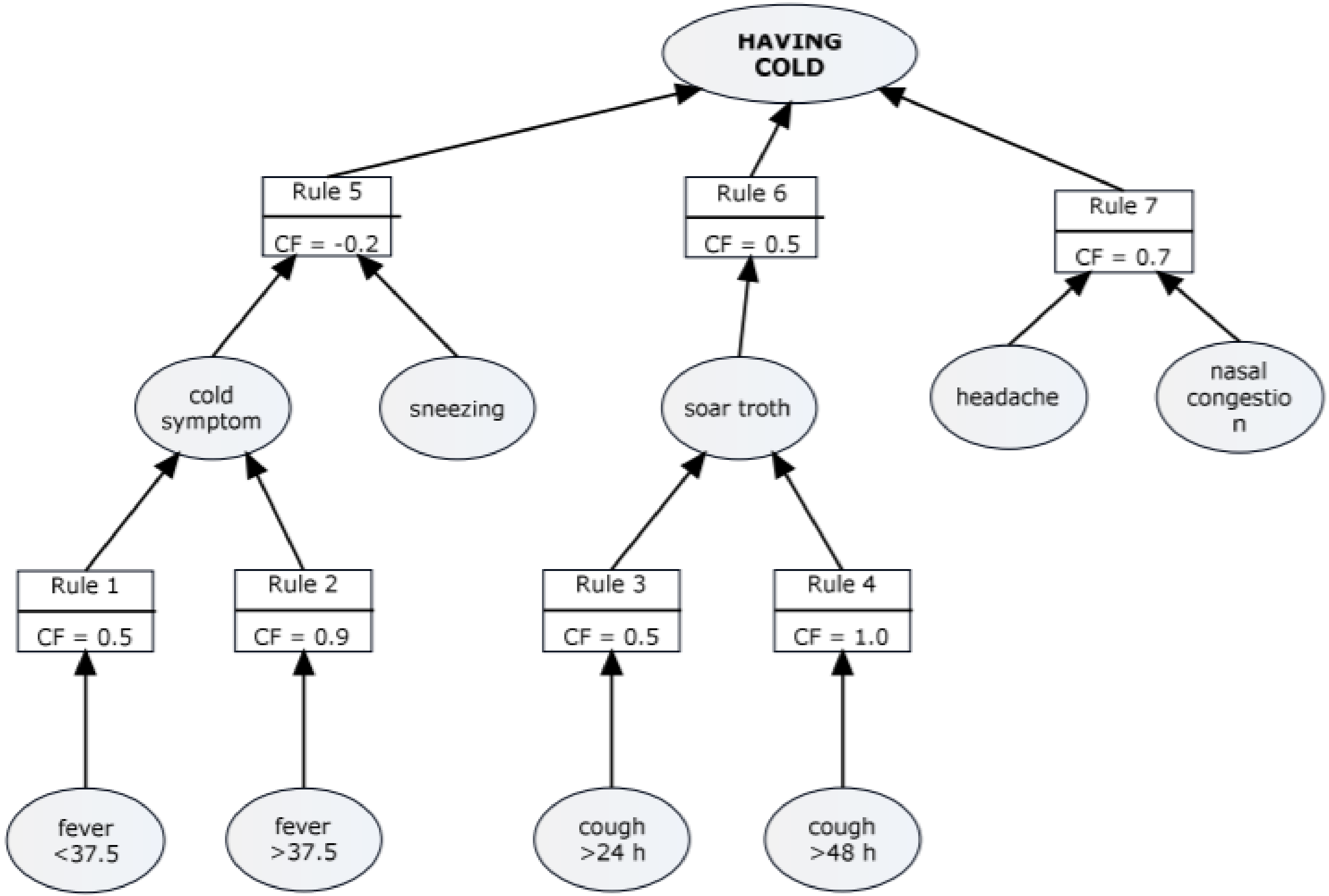
Rule 4
IF cough for more than 48 hours
THEN soar troth = true {CF = 1}

Rule 5
IF Cold symptoms
AND sneezing
THEN having cold {CF = -0.2}

Rule 6
IF soar troth
THEN having cold {CF = 0.5}

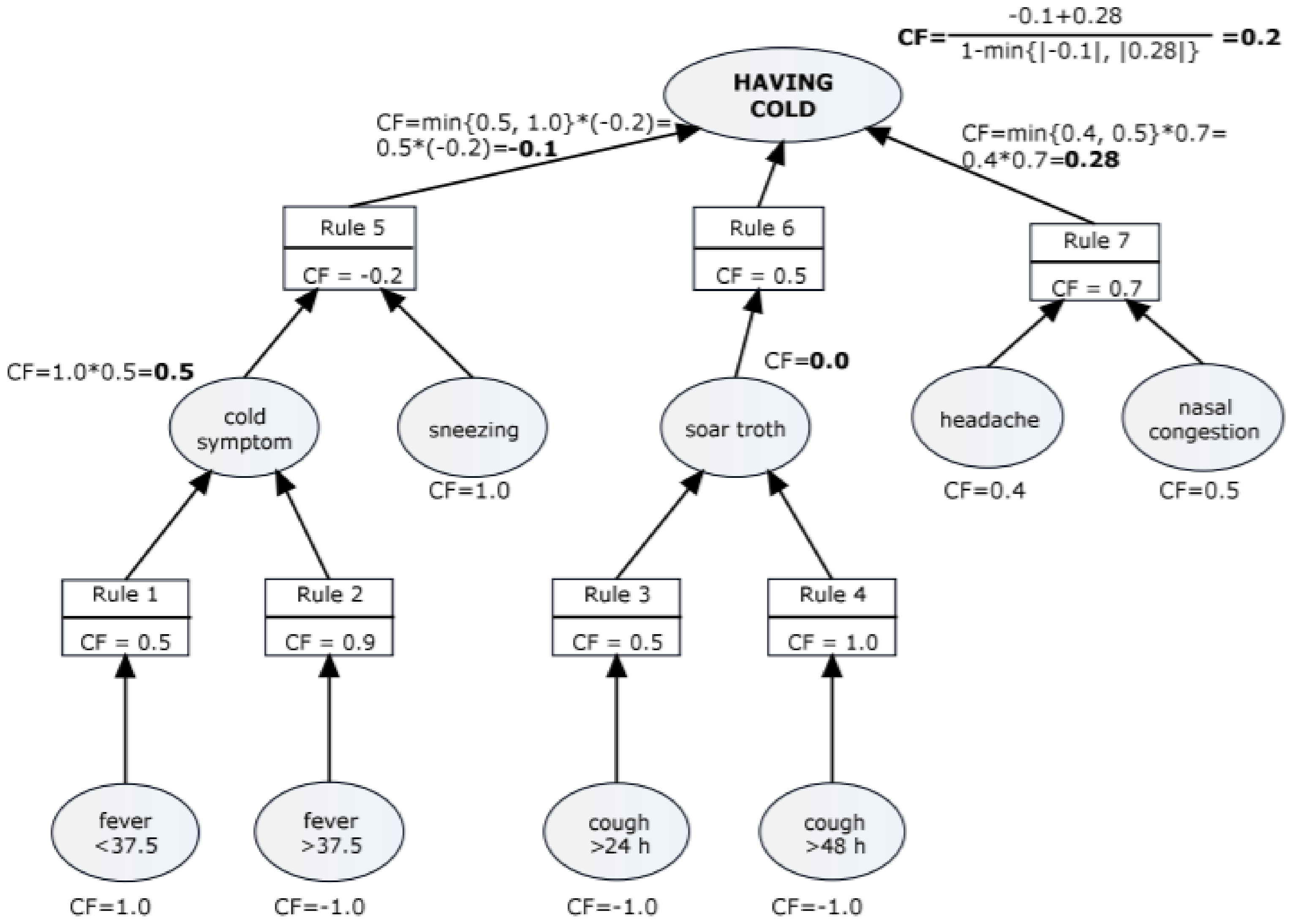
Rule 7
IF headache
AND nasal congestion
THEN having cold {CF = 0.7}

Example 2



HAVING COLD
Rule 5
CF = -0.2
Rule 6
CF = 0.5
Rule 7
CF = 0.7
cold symptom
sneezing
soar troth
headache
nasal congestio n
Rule 1
CF = 0.5
Rule 2
CF = 0.9
Rule 3
CF = 0.5
Rule 4
CF = 1.0
fever <37.5
fever >37.5
cough >24 h
cough >48 h

Example 2



CF= -0.1+0.28 / 1-min{|-0.1|, |0.28|} =0.2
HAVING COLD
CF=min{0.5, 1.0}*(-0.2)=
0.5*(-0.2)=-0.1
CF=min{0.4, 0.5}*0.7=
0.4*0.7=0.28
Rule 5
CF = -0.2
Rule 6
CF = 0.5
Rule 7
CF = 0.7
CF=1.0*0.5=0.5
CF=0.0
cold symptom
sneezing
CF=1.0
soar troth
headache
CF=0.4
nasal congestion
CF=0.5
Rule 1
CF = 0.5
Rule 2
CF = 0.9
Rule 3
CF = 0.5
Rule 4
CF = 1.0
fever <37.5
CF=1.0
fever >37.5
CF=-1.0
cough >24 h
CF=-1.0
cough >48 h
CF=-1.0