Certainty Factor model

Jianqi Huang

SME,CUFE

2022-11-20

The Certainty-Factor Model

- The certainty-factor(CF) model is a method for managing uncertainty in rule-based systems.
- avoid the unreasonable assumptions in Bayes-Model.

The Mechanics of the Model

The Holmers' brief

Mr. Holmes receives a telephone call from his neighbor Dr. Watson stating that he hears a burglar alarm sound from the direction of Mr. Holmes' house. Preparing to rush home, Mr. Holmes recalls that Dr. Watson is known to be a tasteless practical joker, and he decides to first call his other neighbor, Mrs.Gibbons, who, despite occasional drinking problems, is far more reliable.

We can find the rules:

- R_1 :if WATSON's CALL then ALARM, $CF_1 = 0.5$
- R_2 :if GIBBON's CALL then ALARM, $CF_1 = 0.9$
- R_3 :if ALARM's CALL then BURGLARY, $CF_1 = 0.99$

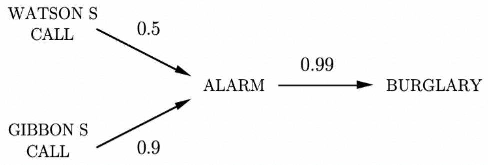


Figure 1: An inference network for Mr. Holmes' situation.

A CF represents a person's *change* in belief.

- The brief increases: a CF between 0 and 1
- The brief decreases: a CF between -1 and 0
- Combination functions: to CF that lie between the evidence and the hypothesis in question.

Step

1. combine CF_1 and CF_2
2. the CFs for R_1 and R_2 to give the CF for the rule R_4

- $R4$: If WATSON'S CALL and GIBBON'S CALL then ALARM, CF_4

Combine CF_1 and CF_2 using the function:

$$CF_4 = \begin{cases} CF_1 + CF_2 - CF_1 CF_2 & CF_1, CF_2 \geq 0 \\ CF_1 + CF_2 + CF_1 CF_2 & CF_1, CF_2 < 0 \\ \frac{CF_1 + CF_1}{1 - \min\{|CF_1|, |CF_2|\}} & otherwise \end{cases} \quad (1)$$

In the Mr.Holmes' case, we have

$$CF_4 = 0.5 + 0.9 - (0.5) \times (0.9) = 0.95$$

the equation (1) is called the parallel-combination function.

Generating another rules R_5: if WATSON'S CALL and GIBBON'S CALL then BURGLARY.

The combination function is

$$CF_5 = \begin{cases} CF_3 CF_4 & CF_3 > 0 \\ 0 & CF_3 \leq 0 \end{cases} \tag{2}$$

In Mr.Holmes' case, we have

$$CF_5 = (0.99) \times (0.95) = 0.94$$

This equation (2) called the serial-combination function.

The conjecture and Disconjecture

- R_6: if CHEST PAIN and SHORTNESS OF BREATH then HEART ATTACK, CF6 = 0.9 Further, suppose that we have indirect evidence for chest pain and shortness of breath:
- R_7: if PATIENT GRIMACES then CHEST PAIN, $CF_7 = 0.7$
- R_8: if PATIENT CLUTCHES THROAT then SHORTNESS OF BREATH, CF8 = 0.9

We can combine CF_6, CF_7, and CF_8 to yield the CF for the rule R9: if PATIENT GRIMACES and PATIENT CLUTCHES THROAT then HEART ATTACK, CF9

The combination function is

$$CF_9 = CF_6 \min(CF_7, CF_8) = (0.9) \min(0.7, 0.9) = 0.63$$

In general, the CF model prescribes that we use the minimum of CFs for evidence in a conjunction, and the maximum of CFs for evidence in a disjunction.

How about the Bayes Inference

Using the example of Mr.Holmes

WATSON'S CALL and GIBBON'S CALL are not conditionally independent given BURGLARY.

Theoretical Problems with CF Model

- Rules the represent logical relationship satisfy the principle of modularity.
- **principle of detachment:** no matter how we established that e is true
- **principle of locality**:no matter what else we know to be true.

Multiple Causes of the Same Effect

- Uncertain reasoning often violates the principles of detachment and locality.
- The inference network does not capture an important interaction among the propositions.

An Earthquake Example

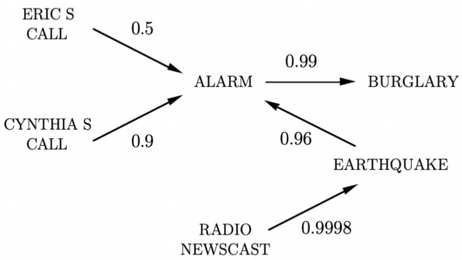


Figure 2: Another inference network for Mr. Holmes' situation.

- R_{11}: if RADIO NEWSCAST then EARTHQUAKE, $CF_{11} = 0.9998$
- R_{12}: if EARTHQUAKE then ALARM, $CF_{12} = 0.95$

He will decreases his brief that there has been a burglary, because the occurrence of an earthquake would account for the alarm sound.

(Pearl 2009) divides uncertain inference into diagnostic and predictive.

- diagnostic: given an effect changing the belief in a cause.
- predictive: given a cause changing the belief in an effect.
- the diagnostic inference can pass through to anther diagnostic inference.
- the predictive inference should not pass through to anther diagnostic inference.

Repair the inference by Adding the Rule

- R_{13}: if EARTHQUAKE then BURGLARY, $CF_13 = -0.7$

Ideally, we would like a representation that encodes only direct relationships among propositions.

Another problem

A Probabilistic Interpretation for Certainty Factors

We can using the likelihood ratio λ

$$\lambda = \frac{o(e|h,\xi)}{p(e|NOTh,\xi)}$$

The $p(e|h,\xi)$ denotes the probability that e is truely given that h is true.

And ξ denotes the background knowledge of the person to whom the belief belongs.

Using Bayes' Theorem

$$\lambda = \frac{O(h|e,\xi)}{O(h|\xi)} = \frac{\frac{p(h|e,\xi)}{1-p(h|e,\xi)}}{\frac{p(h|\xi)}{1-p(h|\xi)}}$$

CF model

$$CF = \begin{cases} \frac{\lambda - 1}{\lambda}, \lambda \geq 1 \\ \lambda - 1, \lambda < 1 \end{cases}$$

- The parallel-combination function follow exactly from Bayes' theorem. The serial-combination are close approximations to it.
- The parallel-combination function to combine CFs for the rules "if e_1 then h" and "if e_2 then h" based on CI(conditionally independent),given h and NOT h.
- Similarly, serial-combination function to CFs for the rules "is a then b" and "if b then c". CI, given b and NOT b.

The assumption of independence are not satisfied by real-world domains. In the parallel-combination function are stronger than Bayes inference.

$$H_0 = \theta_1, H_1 = \theta_2, \theta_1, \theta_2 \in \Theta$$

$$\theta_1 \cup \theta_2 = \Theta, \theta_1 \cap \theta_2 = \phi$$

$$H_0' = \theta_1, H_1' = \theta_i$$

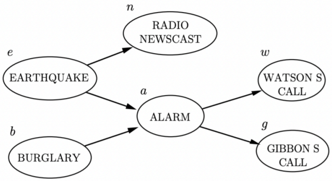
- The CF model could sole the one-to-one problem.
- Hard to solve one-to-n problem

A practical Problem with CF model

- a inference network must trace a trail of rules from observable evidence to hypothesis.
  - (Tversky and Kahneman 1982) found that people are usually most comfortable when they assess predictive rules.
  - the expert physicians prefer to asses the likelihood of a symptom, given a disease.
  - CF model effects are usually observable pieces of evidence.

Belief Network

$$p(e, b, a, n, w, g|\xi) = p(e|\xi)\cdot p(b|\xi)\cdot p(a|e, b, \xi)\cdot p(n|e, \xi)\cdot p(w|a, \xi)\cdot p(g|a, \xi)$$



$p(b+ \,|\, \xi) = 0.0001$

$p(e+ \,|\, \xi) = 0.0003$

$p(w+ \,|\, a-, \xi) = 0.4$
$p(w+ \,|\, a+, \xi) = 0.8$

$p(g+ \,|\, a-, \xi) = 0.04$
$p(g+ \,|\, a+, \xi) = 0.4$

$p(n+ \,|\, e-, \xi) = 0.0002$
$p(n+ \,|\, e+, \xi) = 0.9$

$p(a+ \,|\, b-, e-, \xi) = 0.01$
$p(a+ \,|\, b+, e-, \xi) = 0.95$
$p(a+ \,|\, b-, e+, \xi) = 0.2$
$p(a+ \,|\, b+, e+, \xi) = 0.96$

Figure 4: A belief network for Mr. Holmes' situation.

Belief Network and CF Model

- A knowledge provider can choose the order in which he prefers to assess probability distributions.
- Using a Belief network, the knowledge provider can control the assertions of conditional independence that are encoded in the representation. CF model forces a person to adopt assertions of conditional independence that may be incorrect.
- Using a belief network, a knowledge provider does not have to assess indirect independence.

Conclusion

- CF model as a useful approach to uncertainty management.
- CF model was created to avoid the unreasonable in Bayes model.
- But CF model does not perform well in real-world.
- The parallel-combination has a stronger assumption than conditional Independence in Bayes.
- Using different models to solve the different problems based on the characteristic of problems.

Thanks For your attention.

Reference

Bishop, Christopher M. 2006. *Pattern Recognition and Machine Learning*. Information Science and Statistics. New York: Springer.

Heckerman, David. n.d. "The Certainty-Factor Model," 21.

Pearl, Judea. 2009. *Probabilistic reasoning in intelligent systems: networks of plausible inference*. Rev. 2. ed., transferred to digital printing. The Morgan Kaufmann series in representation and reasoning. San Francisco, Calif: Morgan Kaufmann.

Tversky, Amos, and Daniel Kahneman. 1982. "Causal Schemas in Judgments Under Uncertainty." In, edited by Daniel Kahneman, Paul Slovic, and Amos Tversky, 1st ed., 117–28. Cambridge University Press. https://doi.org/10.1017/CBO9780511809477.009.