When Is a Single “And”-Condition Enough?

Olga Kosheleva[1] **and Vladik Kreinovich**[2]

[1]Department of Teacher Education
University of Texas at El Paso
500 W. University, El Paso, Texas 79968, USA
olgak@utep.edu

[2]Department of Computer Science
University of Texas at El Paso
500 W. University, El Paso, Texas 79968, USA
vladik@utep.edu

Abstract: In many practical situations, there are several possible decisions. Any general recommendation means specifying, for each possible decision, conditions under which this decision is recommended. In some cases, a single “and”-condition is sufficient: e.g., a condition under which a patient is recommended to take aspirin is that “the patient has a fever and the patient does not have stomach trouble”. In other cases, conditions are more complicated. A natural question is: when is a single “and”-condition enough? In this paper, we provide an answer to this question.

Keywords: fuzzy logic; decision making; “and”-condition

1 Formulation of the Problem

1.1 Need to describe conditions

In many practical situations, we need to make a decision. For example, a medical doctor needs to decide what treatment to recommend for a patient. Usually:

- there are several possible decisions;
- so, to describe a general recommendation, we need to list conditions under which each possible decision should be chosen.

In the simplest cases, these conditions are straightforward: e.g.,

“if a patient has a fever, recommend some fever reducer”.

In more complex cases, several such straightforward conditions must be satisfied. For example, the condition to recommend aspirin can be described as

“the patient has a fever *and* the patient does not have stomach trouble”.

We can call such conditions *"and"-conditions*.

In general, there can be several such "and" conditions that lead to the same action. For example, the condition to recommend aspirin can take the following form:

"(the patient has a fever *and* the patient does not have stomach trouble) *or* the patient has a strong headache".

1.2 Conditions are often fuzzy

Word like "fever" are not precise. It is *not* true that:

- 37.9 is not a fever,
- but 38.0 is already a fever.

This Covid-time rule was clearly a simplification.

From the medical doctor's viewpoint, fever is a matter of degree:

- one can have slight fever,
- one can have high fever, etc.

To describe such imprecise ("fuzzy") words in precise terms, Lotfi Zadeh came up with an idea of *fuzzy logic*, where with each such word, we associate a function $m(x)$ that assigns:

- to each possible value x of the corresponding quantity,
- the degree – on the scale from 0 to 1 – to which the given value has this property (e.g., to which a patient has a fever).

This function is known as a *membership function* or, alternatively, a *fuzzy set*; see, e.g., [1, 3, 4, 6, 7, 8].

Comment. The choice of 0-to-1 scale is just a matter of convenience. Instead, we can choose, e.g.:

- the 0-to-10 scale – as in many polls, or
- the -1-to-1 scale – as in some expert systems (see, e.g., [2]).

1.3 A natural question

Since the component conditions are fuzzy, their "and"- and "or"-combinations are also fuzzy. In general, a fuzzy condition means that we assign to each tuple $(x_1,\ldots,x_n)$ a degree $f(x_1,\ldots,x_n)$ to which the given tuple satisfies this condition.

As we have mentioned:

- sometimes, a single "and"-condition is enough,
- sometimes more complex conditions are needed.

So, a natural question is:

When is a single "and"-condition enough?

1.4 What we do in this paper

In this paper, we provide an answer to this question.

2 Let Us Formulate This Problem in Precise Terms

2.1 Towards a precise formulation

- We are given a condition $f(x_1,\ldots,x_n)$, and
- we want to check when this condition can be represent by a single "and"-condition.

By definition, an "and"-condition means

"x_1 has the property m_1 *and* ... *and* x_n has the property m_n"

for some appropriate properties $m_1(x_1)$, ..., $m_n(x_n)$.

In fuzzy logic, our degree of confidence in an "and"-combination is computed by applying the appropriate "and"-operation $f_\&(a,b)$ (also known as *t-norm*) to the degrees of confidence of component statements.

So, our main question takes the following form:

When can a function $f(x_1,\ldots,x_n)$ be represented as

$$f(x_1,\ldots,x_n) = f_\&(m_1(x_1),\ldots,m_n(x_n)) \tag{1}$$

for some "and"-operation $f_\&(a,b)$ and for some membership functions $m_i(x_i)$?

2.2 It makes sense to only consider smooth (differentiable) functions

Usually, small changes in x_i lead to small changes in the degree. So it makes sense to assume that the function $f(x_1,\ldots,x_n)$ is smooth (differentiable). From the practical viewpoint, this assumption make sense; indeed:

- degrees are only known with some accuracy – e.g., hardly any expert can distinguish between his/her degrees of confidence 0.80 and 0.81 – and
- any continuous function on a bounded domain can be approximated, with any given accuracy, by a polynomial – i.e., by a smooth (actually, infinitely many times differentiable) function.

Similarly, it makes sense to require that the membership functions $m_1(x_1)$, ..., $m_n(x_n)$ are smooth.

As for the "and"-operation, there is a theorem (see, e.g., [5]) that:

- any "and"-operation can be approximated, with any given accuracy,

- by a function of the type $g(g^{-1}(a)+g^{-1}(b))$ for some strictly monotonic function $g(x)$.

(Here $g^{-1}(x)$ denotes the inverse function, i.e., a function for which $g^{-1}(a)=b$ if and only if $g(b)=a$.) Thus, it makes sense to restrict ourselves to such "and"-operations.

Similarly to the cases of the original condition and membership function, we can safely assume that the function $g(x)$ is also differentiable.

2.3 Towards a precise formulation (cont-d)

For a smooth "and"-operation of the above type, the equality (1) take the following form:

$$f(x_1,\ldots,x_n)=g(g^{-1}(m_1(x_1)),\ldots,g^{-1}(m_n(x_n))). \tag{2}$$

This expression can be re-written as follows:

$$f(x_1,\ldots,x_n)=g(g_1(x_1)+\ldots+g_n(x_n), \tag{3}$$

where we denoted $g_i(x_i)\stackrel{\text{def}}{=}g^{-1}(m_i(x_i))$.

Vice versa, if we have a description of the form (3), we can represent it in the form (1) if we take $m_i(x_i)\stackrel{\text{def}}{=}g(g_i(x_i))$.

Thus, we arrive at the following definition.

Definition. *We say that a smooth function $f(x_1,\ldots,x_n)$* can be represented by a single "and"-condition *if it can be described by the formula (3) for some smooth functions $g(x)$ and $g_i(x_i)$.*

Comment. Usually, in fuzzy logic, we only consider truth values from the interval $[0,1]$ – and thus, functions whose values are in this interval. However, as we have mentioned earlier, we could as well choose any other interval instead. Because of this, in our result, we do not restrict ourselves to the interval $[0,1]$.

3 Main Result

Proposition. *For any smooth function $f(x_1,\ldots,x_n)$, the following two conditions are equivalent to each other:*

- *this function can be represented by a single "and"-condition;*
- *for every two different indices i and j, we have*

$$\frac{\partial^2}{\partial x_i \partial x_j}\left(\ln\left(\frac{\partial f}{\partial x_i}\right)-\ln\left(\frac{\partial f}{\partial x_j}\right)\right)=0, \tag{4}$$

and for every three different indices i, j, and k, we have

$$\frac{\partial}{\partial x_k}\left(\ln\left(\frac{\partial f}{\partial x_i}\right)-\ln\left(\frac{\partial f}{\partial x_j}\right)\right)=0. \tag{5}$$

Comment. When $n = 2$:

- there is only one pair of different indices, and
- we do not have three different indices.

So, in this case, the following single condition is sufficient:

$$\frac{\partial^2}{\partial x_1 \partial x_2}\left(\ln\left(\frac{\partial f}{\partial x_1}\right) - \ln\left(\frac{\partial f}{\partial x_2}\right)\right) = 0. \tag{4a}$$

Mathematical comment. For our proof, it is sufficient to require that:

- the function $f(x_1,\ldots,x_n)$ is three times differentiable, i.e., that it has derivatives up to the third order:

$$\frac{\partial f}{\partial x_i},\ \frac{\partial^2 f}{\partial x_i \partial x_j},\ \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k},$$

- the function $g(x)$ is one time differentiable, and
- the functions $g_i(x_i)$ are two times differentiable.

Proof.

1°. Let us first prove that:

- if the function $f(x_1,\ldots,x_n)$ can be represented by a single "and"-condition – i.e., can be represented in the form (3),
- then the function $f(x_1,\ldots,x_n)$ satisfies the conditions (4) and (5).

1.1°. Indeed, suppose that the function $f(x_1,\ldots,x_n)$ has the form (3). Then, differentiating the expression in the right-hand side of the formula (3) by x_i, we conclude that

$$\frac{\partial f}{\partial x_i} = A \cdot g_i'(x_i), \tag{6}$$

where:

- we denoted $A \stackrel{\text{def}}{=} g'(g_1(x_1) + \ldots + g_n(x_n))$ and,
- as usual, for functions of one variable, $f'(x)$ means the derivative.

1.2°. By applying logarithm to both sides of the formula (6), we get:

$$\ln\left(\frac{\partial f}{\partial x_i}\right) = \ln(A) + \ln(g_i'(x_i)). \tag{7}$$

So, the difference between such logarithms that appears in the left-hand sides of the desired equalities (4) and (5) has the following form:

$$\ln\left(\frac{\partial f}{\partial x_i}\right) - \ln\left(\frac{\partial f}{\partial x_j}\right) = \ln(g_i'(x_i)) - \ln(g_j'(x_j)). \tag{8}$$

1.3°. Now, we are ready to prove the equality (5).

Indeed, the right-hand side of the expression (8) does not depend on any other variable x_k. So, the derivative of this right-hand side with respect to x_k is equal to 0.

Thus, the condition (5) is satisfied.

1.4°. To prove that the condition (4) is also satisfied, let us first differentiate the right-hand side of the expression (8) with respect to x_j.

The first term in the right-hand side of (8) does not depend on x_j, so its derivative is 0. Thus, we have

$$\frac{\partial}{\partial x_j}\left(\ln\left(\frac{\partial f}{\partial x_i}\right)-\ln\left(\frac{\partial f}{\partial x_j}\right)\right)=-\frac{\partial}{\partial x_j}(\ln(g_j'(x_j))). \tag{9}$$

1.5°. The right-hand side of the expression (9) does not depend on x_i, so its derivative with respect to x_i is equal to 0 – which is exactly what the condition (4) is about.

So indeed, if a function $f(x_1,\ldots,x_n)$ can be represented in the form (3), then this function satisfies the conditions (4) and (5). The first statement is proven.

2°. Vice versa, let us prove that:

- if a smooth function $f(x_1,\ldots,x_n)$ satisfies conditions (4) and (5),
- then this function $f(x_1,\ldots,x_n)$ can be represented in the form (3) for some smooth functions $g(x)$ and $g_i(x_i)$.

2.1°. Indeed, due to the condition (5), the difference

$$D_{i,j}(x_1,\ldots,x_n)\stackrel{\text{def}}{=}\ln\left(\frac{\partial f}{\partial x_i}\right)-\ln\left(\frac{\partial f}{\partial x_j}\right) \tag{10}$$

does not depend on any variables different from x_i and x_j. Thus, the expression $D_{i,j}$ depends only on x_i and x_j:

$$D_{i,j}(x_1,\ldots,x_n)=D_{i,j}(x_i,x_j).$$

2.2°. The condition (4) says that

$$\frac{\partial}{\partial x_i}\left(\frac{\partial D_{i,j}(x_i,x_j)}{\partial x_j}\right)=0. \tag{11}$$

The fact that the derivative with respect to x_i is equal to 0 means that the differentiated function does not depend on x_i, i.e., that it depends only on x_j:

$$\frac{\partial D_{i,j}(x_i,x_j)}{\partial x_j} = F_{i,j}(x_j). \tag{12}$$

for some function $F_{i,j}(x_j)$.

2.3°. For each x_i, we can now integrate both sides of the equality (12) with respect to x_j. Thus, we conclude that

$$D_{i,j}(x_i,x_j) = I_{i,j}(x_j) + C_{i,j}(x_i), \tag{13}$$

where:

- by $I_{i,j}$ we denoted the integral of $F_{i,j}$, and
- by $C_{i,j}(x_i)$, we denoted the integration constant – which, generally speaking, depends on x_i.

2.4°. In particular, for $i = 1$, the expression (13) leads to:

$$D_{1,j}(x_1,x_j) = C_{1,j}(x_1) + I_{1,j}(x_j). \tag{14}$$

In particular, for $j = 2$, we have:

$$D_{1,2}(x_1,x_2) = C_{1,2}(x_1) + I_{1,2}(x_2). \tag{15}$$

2.5°. From the definition (10), we can check that for every j which is different from 1 and 2, we have

$$D_{j,2}(x_j,x_2) = D_{1,2}(x_1,x_2) - D_{1,j}(x_1,x_j). \tag{16}$$

2.6°. Substituting the expressions (14) and (15) into the formula (16), we conclude that

$$D_{j,k}(x_j,x_2) = (C_{1,2}(x_1) - C_{1,j}(x_1)) + I_{1,2}(x_2) - I_{1,j}(x_j). \tag{17}$$

The left-hand side of this equality does not depend on x_1. So the right-hand side should not depend on x_1 either.

Hence, the difference $C_{1,2}(x_1) - C_{1,j}(x_1)$ cannot depend on x_1 and is, thus, a constant. Let us denote this constant by c_i:

$$C_{1,2}(x_1) - C_{1,j}(x_1) = c_i. \tag{18}$$

Then,

$$C_{1,j}(x_1) = C_{1,2}(x_1) - c_i. \tag{19}$$

2.7°. Substituting the expression (19) for $C_{1,j}(x_1)$ into the formula (14), we have

$$D_{1,j}(x_1,x_j) = C_{1,2}(x_1) - c_i + I_{1,j}(x_j), \tag{20}$$

i.e.,

$$D_{1,j}(x_1,x_j) = F_1(x_1) - F_j(x_j), \tag{21}$$

where:

- we denoted $F_1(x_1) \stackrel{\text{def}}{=} C_{1,2}(x_1)$ and
- we denoted $F_j(x_j) \stackrel{\text{def}}{=} c_i - I_{1,j}(x_j)$ for all $j \neq 1$.

2.8°. From (10), we conclude that

$$D_{i,j}(x_i,x_j) = D_{1,j}(x_1,x_j) - D_{1,i}(x_1,x_i). \tag{22}$$

Substituting the expression (21) into this formula, we conclude that for every i and j, we have

$$D_{i,j}(x_i,x_j) = F_i(x_i) - F_j(x_j). \tag{23}$$

2.9°. By definition of $D_{i,j}$, the formula (23) means that

$$\ln\left(\frac{\partial f}{\partial x_i}\right) - \ln\left(\frac{\partial f}{\partial x_j}\right) = F_i(x_i) - F_j(x_j). \tag{24}$$

Moving two terms to the opposite side of this equality, we conclude that

$$\ln\left(\frac{\partial f}{\partial x_i}\right) - F_i(x_i) = \ln\left(\frac{\partial f}{\partial x_j}\right) - F_j(x_j). \tag{25}$$

2.10°. Applying the function $\exp(x)$ to both sides of the equality (25), we conclude that

$$\frac{1}{a_i(x_i)} \cdot \frac{\partial f}{\partial x_i} = \frac{1}{a_j(x_j)} \cdot \frac{\partial f}{\partial x_j} \tag{26}$$

where we denoted

$$a_i(x_i) \stackrel{\text{def}}{=} \exp(F_i(x_i)). \tag{27}$$

2.11°. It is known that $df(x) = f'(x) \cdot dx$ and thus, $a_i(x_i) \cdot x_i = dX_i$, where:

- we denoted $X_i = g_i(x_i)$ and
- we denoted $g_i(x_i) \stackrel{\text{def}}{=} \int a_i(x_i)$.

Thus, the equality (26) takes the form

$$\frac{\partial f}{\partial X_i} = \frac{\partial f}{\partial X_j}. \quad (28)$$

If we introduce new variables

$$u = X_i + X_j \text{ and } v = X_i - X_j,$$

then (28) turns into

$$\frac{\partial f}{\partial v} = 0.$$

Hence f depends only on the sum $X_i + X_j$.

Similarly, we can conclude that f only depends on the sum $X_i + X_j + X_k$, etc., i.e., that

$$f(x_1, \ldots, x_n) = g(X_1 + \ldots + X_n), \quad (29)$$

for some function $g(x)$.

By definition, $X_i = g_i(x_i)$. Substituting $X_i = g_i(x_i)$ into the formula (29), we get exactly the desired equality (3).

The statement is proven, and so is the proposition.

Acknowledgements
This work was supported in part:

- by the National Science Foundation grants:
  - 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science),
  - HRD-1834620 and HRD-2034030 (CAHSI Includes),
  - EAR-2225395 (Center for Collective Impact in Earthquake Science C-CIES),
- and by the AT&T Fellowship in Information Technology.

It was also supported:

- by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478, and
- by a grant from the Hungarian National Research, Development and Innovation Office (NRDI).

The authors are greatly thankful to Imre Rudas and Aniko Szakal for their support.

References

[1] R. Belohlavek, J. W. Dauben, and G. J. Klir, *Fuzzy Logic and Mathematics: A Historical Perspective*, Oxford University Press, New York, 2017.

[2] B. G. Buchanan and E. H. Shortliffe, *Rule Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, Addison-Wesley, Reading, Massachusetts, 1984.

[3] G. Klir and B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.

[4] J. M. Mendel, *Uncertain Rule-Based Fuzzy Systems: Introduction and New Directions*, Springer, Cham, Switzerland, 2017.

[5] H. T. Nguyen, V. Kreinovich, and P. Wojciechowski, “Strict Archimedean t-norms and t-conorms as universal approximators”, *International Journal of Approximate Reasoning*, 1998, Vol. 18, Nos. 3–4, pp. 239–249.

[6] H. T. Nguyen, C. L. Walker, and E. A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2019.

[7] V. Novák, I. Perfilieva, and J. Močkoř, *Mathematical Principles of Fuzzy Logic*, Kluwer, Boston, Dordrecht, 1999.

[8] L. A. Zadeh, “Fuzzy sets”, *Information and Control*, 1965, Vol. 8, pp. 338–353.