Fuzzy Evidential Logic: A Model of Causality for Commonsense Reasoning

Ron Sun
Honeywell SSDC
Minneapolis, MN 55418

Abstract

This paper proposes a fuzzy evidential model for commonsense causal reasoning. After an analysis of the advantages and limitations of existing accounts of causality, a generalized rule-based model FEL (*Fuzzy Evidential Logic*) is proposed that takes into account the inexactness and the cumulative evidentiality of commonsense reasoning. It corresponds naturally to a neural (connectionist) network. Detailed analyses are performed regarding how the model handles commonsense causal reasoning.

Shoham's Causal Theory

The issue of causality has recently received a lot of attentions from various perspectives (cf. Shoham 1987, Iwasaki & Simon 1986, de Kleer & Brown 1986, Pearl 1988, etc.). The issue has wide ranging impact on areas such as learning, control, and recognition. However, most of these logic based models are aimed for modeling truth functional aspects of causal knowledge, and they tend to ignore some important characteristics of commonsense causal reasoning, for example, gradeness of concepts, inexact causal connections, evidentiality of causal rules, etc., while probabilistically motivated models are mainly concerned with the probabilistic aspect of causal events, and they are more computationally complex and oftentime have only marginal cognitive plausibility in terms of mechanisms involved. Connectionism provides a new and different kind of models that might be of help in accounting for causality in commonsense reasoning; these models entertain a number of interesting properties that other models lack (for example, massive parallelism, generalization, fault/noise tolerence, and adaptability; see Waltz & Feldman 1986, Sun & Waltz 1991) and present a new perspective of reasoning as a complex process in a dynamic system; it will be worthwhile to look into the question of how such models can deal with the issue of causality.

Let us look into Shoham's account of causality (Shoham 1987), which is undoubtedly one of the most notable accounts of causality with rule-based formalisms. His temporal modal logic formalism has a close resemblance to Horn clause logic, and therefore is very suitable for use in rule-based systems. According to Shoham's Causal Theory (CT), causes are primary conditions which, together with other conditions, will bring about the effect. These "other" conditions are somewhat secondary. In reasoning, as long as we know that the primary conditions (*causes* or *necessary conditions*) are true and that there is no information that the secondary conditions (*enabling conditions* or *possible conditions*) are false, then we can deduce that effects will follow. The theory is described in terms of modal logic, with one basic modal operator ($\Box$ or *necessity*) for specifying necessary conditions. and one auxiliary modal operator ($\Diamond$ or *possibility*) for specifying possible conditions. The formal definition is as follows:

Definition 1 *A* **Causal Theory** *is a set of formulas of the following form*

$$\wedge_i \Box n_i a_i(t_{i1}, t_{i2}) \wedge_j \Diamond n_j b_j(t_{j1}, t_{j2}) \longrightarrow \Box c(t_1, t_2)$$

where n_i's are either $\neg$ or nothing, $t_2 > t_{i2}$ for all i's, $t_2 > t_{j2}$ for all j's, $n_i a_i$'s are necessary conditions (causes), and $n_j b_i$'s are possible conditions (enabling conditions). C is concluded iff all $n_i a_i$'s are true and none of $n_j b_j$'s are known to be false. [1]

From the standpoint of modeling commonsense knowledge, this model has some advantages, such as that it provides a simple and elegant formalism with efficient inference algorithms, that it is easily representable (and implementable), and that it has

[1] This process is formally described by a *minimization* principle in Shoham (1987).

compatibility with philosophical accounts of causality (Shoham 1987). On the other hand, the model ignores or discounts many aspects of commonsense causal reasoning; for example,

1. All propositions in this theory are binary: either true or false, and there is no sense of gradedness. Commonsense knowledge is certainly not limited to true/false only (Sun 1991, Hink & Woods 1987).
2. Beside the inexactness of individual concepts, reasoning processes in reality are also inexact and evidential. Specifically, the evidential combination process is cumulative (as observed in protocol data; Sun 1991, 1991a); that is, it "adds up" various pieces of evidence to reach a conclusion, with a confidence that is determined from the "sum" of the confidences of the different pieces of evidence. Moreover, different pieces of evidence are weighted, that is, each of them may have more or less impact, depending on its importance or saliency, on the reasoning process and the conclusion reached. We have to find a way of combining evidence from different sources cumulatively and with weights, without incurring too much computational overhead (such as in probabilistic reasoning or Dempster-Shafer calculus; cf. Pearl 1988).
3. Because of the lack of gradedness, the model will make projections too far along a chain of reasoning (or too far into the future; Sun 1991). An example from Shoham (1987):

$$\Box alive(t_0, t_0)$$

$$\Box shoot(t,t) \longrightarrow \Box\neg alive(t+1,t+1)$$

$$\Box alive(t,t)\Diamond\neg shoot(t,t)\Diamond\neg otherwise\text{-}killed(t,t) \longrightarrow \Box alive(t+1,t+1)$$

which means that if one is alive at time t_0, one will continue to be alive as long as not being shot or otherwise killed. So if there is nothing known about "shoot" and "otherwise-killed", then according to the minimal model approach, we will predict that

$$\Box alive(t,t) \quad where\ t \longrightarrow \infty$$

This is certainly not true. The problem is that, along a chain of inference (as well as in temporal projections), the confidence for the conclusions reached should weaken. We can weaken confidence along the way only when gradedness is reinstated into causal theories.
4. The clear-cut *necessity* and *possibility* is a problem, because in reality there is little, if any, qualitative difference between *causes* and *enabling conditions*. The difference is more quantitative (as will be illustrated later), and sometimes the two are interchangeable; for example, "He is shot dead" is expressed in CT as

$$\Box shoot(t,t)\wedge\Diamond\neg wearing\text{-}bullet\text{-}proof\text{-}vest(t,t)\ldots\ldots \longrightarrow \Box dead(t+1,t+1)$$

and "His failure to wear the bullet-proof vest caused his tragic death" is expressed as

$$\Box\neg wearing\text{-}bullet\text{-}proof\text{-}vest(t,t)\wedge\Diamond shoot(t,t)\ldots\ldots \longrightarrow \Box dead(t+1,t+1)$$

So one fact can be both a cause and an enabling condition.
5. Although the model does distinguish two different types of conditions, it does not explain why some conditions are necessary, and some conditions need only to be possible.
6. According to the model, it is necessary to list all causes and all enabling conditions, in order to guarantee correct results. This could be hard to do, because the number of enabling conditions could be infinite.
7. The causal connection between events in the left hand side of an implication and events in the right-hand side of the same implication may not be deterministic. It could be probabilistic, or otherwise uncertain (Suppes 1970).

For reviews of other accounts of causality, see Sun (1991).

Defining FEL

FEL (*Fuzzy Evidential Logic*) is aimed at resolving the problems inherent in existing logical accounts of causality. Like Shoham's formalism, FEL is defined around rules; however, FEL encodes rules with the weighted-sum computation. This formalism is meant to capture, among other things, the gradedness and evidentiality of commonsense reasoning, in a cognitively motivated way. Formal definitions follow (cf. Zadeh 1988):

Definition 2 *A* **Fact** *is an atom or its negation, represented by a letter (with or without a negation symbol) and having a value between l and u. The value of an atom is related to the value of its negation by a specific method, so that knowing the value of an atom results in immediately knowing the value of its negation, or vice versa.* [2]

Now we can define rules and their related weighting schemes:

[2] We will adopt a generic confidence measure as the value of a fact.

Definition 3 *A* **Rule** *is a structure composed of two parts: a left-hand side (LHS), which consists of one or more facts, and a right-hand side (RHS), which consists of one fact. When facts in LHS get assigned values, the fact in RHS can be assigned a value according to a weighting scheme* [3].

Definition 4 *A* **Weighting Scheme** *is a way of assigning a weight to each fact in LHS of a rule, with the total weights (i.e., the sum of the absolute values of all the weights) less than or equal to 1, and of determining the value of the fact in RHS of a rule by thresholded (if thresholds are used) weighted-sum of the values of the facts in LHS (or inner-products of weight vectors and vectors of values of LHS facts). When the range of values is continuous, then the weighted-sum is passed on if its absolute value is greater than the threshold, or 0 if otherwise. When the range of values is binary (or bipolar), then the result will be one or the other depending on whether the weighted-sum (or the absolute value of it) is greater than the threshold or not (usually the result will be 1 if the weighted-sum is greater than the threshold, 0 or -1 if otherwise).* [4]

Definition 5 *A* **Conclusion** *in FEL is a value associated with a fact, calculated from rules and facts by doing the following:*

(1) for each rule having that conclusion in its RHS, obtain conclusions of all facts in its LHS (if any fact is unobtainable, assume it to be zero); and then calculate the value of the conclusion in question using the weighting scheme;

(2) take the MAX of all these values associated with that conclusion calculated from different rules or given in initial input.

Definition 6 *A rule set is said to be* **Hierarchical**, *if the graph depicting the rule set is acyclic; the graph is constructed by drawing a unidirectional link from each fact (atom) in LHS of a rule to the fact (atom) in RHS of a rule.*

Making a rule set hierarchical avoids circular reasoning.

Now FEL can be defined as follows:

Definition 7 *A* **Fuzzy Evidential Logic** *(FEL) is a 6-tuple: $< A, R, W, T, I, C >$, where A is a set of facts (the values of which are assumed to be zero initially), R is a set of rules, W is a weighting scheme for R, T is a set of thresholds each of which is for one rule, I is a set of elements of the form (f, v) (where f is a fact, and v is a value associated with f), and C is a procedure for deriving conclusions (i.e. computing values of facts in RHS of a rule in R, based on the initial condition I).*

We want differentiate FEL into two versions: FEL_1 and FEL_2, which differ in their respective ranges for values associated with facts.

Definition 8 *FEL_1 is FEL when the range of values is restricted to between 0 and 1 (i.e. l=0 and u=1), and the way the value of a fact is related to the value of its negation is:*

$$a = 1 - \neg a$$

for any fact a.

Definition 9 *FEL_2 is FEL when the range of values is restricted to between -1 and 1 (i.e. l=-1 and u=1), and the way the value of a fact related to the value of its negation is:*

$$a = -\neg a$$

for any fact a.

As an illustration of its capability and correctness, we want to show that FEL can implement Horn clause logic as a special case (we will only deal with the propositional version here, and extensions to first order cases is dealt with in Sun 1991). Let us define Horn clause logic first (cf. Chang & Lee 1973):

Definition 10 **Horn clause logic** *is a logic in which all formulas are in the forms of*

$$p$$

or

$$p_1 p_2 \cdots\cdots p_n \longrightarrow q$$

where p's and q are propositions.

Definition 11 *A* **Binary FEL** *is a reduced version of FEL (either FEL_1 or FEL_2), in which values associated with facts are binary (or bipolar), total weights of each rule sum to 1, and all thresholds are set to 1.*

Here is the theorem for the equivalence (see Sun 1991 for proofs):

Theorem 1 *The binary FEL is sound and complete with respect to Horn clause logic.*

We want to show that FEL can simulate Shoham's Causal Theory, to further explore the logical capability of FEL. (We will only consider a non-temporal version of CT, that is, we strip away all

[3]When the value of a fact in LHS is unknown, assign a zero as its value.

[4]This weighting scheme can be generalized, as will be discussed later on.

temporal notations.) We have to find a mapping between truth values of formulas in Causal Theory and values of facts in FEL. Since in CT and in FEL, there is no logical OR and there is only a (implicit) logical AND in the LHS of a rule, which can be taken care of by a weighting scheme as will be discussed later, we do not have to worry about these two connectives in the mapping now. Therefore, we can use a mapping as follows, which can be easily verified to be consistent with regard to logical equivalence (without AND and OR; for example, □ a = ¬◇¬ a, etc.):

(1) M(a= true) = 'a=1'
(2) M(¬ a= true) = 'a=-1'
(3) M(□ a= true) = 'a=1'
(4) M(□¬ a= true) = 'a=-1'
(5) M(◇ a= true) = 'a=0'
(6) M(◇¬ a= true) = 'a=0'
(7) M(¬□ a= true) = 'a=0'
(8) M(¬□¬ a= true) = 'a=0'
(9) M(¬◇ a= true) = 'a=-1'
(10) M(¬◇¬ a= true) = 'a=1'

With the mapping in hand, we can proceed to find a weighting scheme to enable FEL to simulate Causal Theory. The problem is that in FEL we have nodes only for atoms such as a, b, m, n, etc. but not for □a or ◇b, etc. We have two ways of dealing with this:

1. Extending and making more complex the weighting scheme,
2. Adding nodes that can be used to represent atoms with modal operators.

We will adopt the first approach here (the second approach will also work — the difference is insignificant). For a formula in Causal Theory

$$\wedge_i \Box n_i a_i \wedge_j \Diamond n_j b_j \longrightarrow \Box nc$$

we can assign arbitrary weights to atoms: a_i's and b_i's (if there is a negation, the corresponding weight is negative; otherwise, weights are positive), as long as their absolute values sum to 1. However, for b_i's, we will also apply the following function to the link between b_i and c:

$$f_j(b_j) = \begin{cases} 1 & \text{if } b_j = 1 \\ 1 & \text{if } b_j = 0 \text{ and } n_j \neq \neg \\ -1 & \text{if } b_j = 0 \text{ and } n_j = \neg \\ -1 & \text{if } b_j = -1 \end{cases}$$

We will call this function the *elevation* function because it turns all 0's into 1's or -1's. We have thresholds equal to 1 for all rules. We restrict the possible values of facts to -1 or 1.

Now it is easy to verify that a rule in FEL with this specific weighting scheme and thresholds is equivalent to a corresponding formula in Causal Theory: e.g., suppose we have the following formula in CT:

$$\Box a \Box b \Diamond a \Diamond b \longrightarrow \Box e$$

It can be translated into FEL as follows:

$$abc'd' \longrightarrow e \quad (w_1 w_2 w_3 w_4)$$

where $c' = f_c(c)$ and $d' = f_d(d)$ and $\sum_i w_i = 1$, and the threshold equal to 1 for the rule. The equivalence can be verified case by case.

To find a full correspondence between FEL and Causal Theory, we also need a proof procedure that enables the derivation of all correct results (theorems). Here is a proof procedure for CT:

Given a Causal Theory CT, and a set of initial conditions (true events) I:

— For all a ∈ I, infer a, □a, and ◇a.
— **Repeat:**
for $\wedge_i \Box n_i a_i \wedge_j \Diamond n_j b_j \longrightarrow \Box c$ where $n_i a_i$'s are inferred, and $\neg n_j b_j$'s are non-inferable, [5]
infer c, □c, and ◇c.

It is easy to see the correctness of this procedure (see Sun 1991 for all the proofs):

Theorem 2 *The above proof procedure is sound and complete for Causal Theory as defined above.*

We can have a similar proof procedure for FEL:

Given a FEL theory, and a set of initial conditions (true facts) I:

— For all $(a, v_a) \in$ I, infer a with v_a.
— **Repeat:**
for $\wedge_i n_i a_i \longrightarrow c$ where each $n_i a_i$ is inferred with a certain value, or is non-inferable (and therefore a value zero is assumed), [6]
infer c with v_c, where v_c is calculated according to the weighting scheme used.

It is easy to see the correctness of this procedure for FEL, and the correspondence between the two proof procedures:

Theorem 3 *The above proof procedure is sound and complete for hierarchical FEL*

[5]They are not in the RHS of any rule and not in I, or in order to infer it, we have to use a rule which has a fact as a necessary condition in its LHS that is not inferable. Since CT is hierarchical, this is easy to detect. We can preconstruct a "dependency graph" which depicts inferability relations.

[6]According to the weighting scheme used to simulate CT, if a fact is inferred, it must be inferred with a value 1 or -1; if a fact is non-inferable, then its value is 0. When other weighting schemes are used, the results will be different.

Theorem 4 *The proof procedure for FEL carries out exactly the proof procedure for Causal Theory when Causal Theory is implemented in FEL in the aforementioned way.*

Therefore,

Theorem 5 *For every hierarchical, non-temporal Causal Theory, there is a FEL such that CT: w ⊨ a iff FEL: c ⊨ 'a=1', where w is a set of initial conditions for Causal Theory CT, and c is the set of initial conditions for FEL mapped over from w in CT.*

Accounting for Commonsense Causality

Now we are ready to show that FEL extends justifiably Shoham's Causal Theory and solves the problems identified earlier. To extend the FEL version of Shoham's Causal Theory, we first notice that the causes need not be known with absolute certainty, i.e. we should allow a confidence measure associated with each *necessary* fact (i.e. the one with $\Box$), because of the gradedness, uncertainty and fuzziness of our knowledge. By the same token, the conclusions need not be binary either, so that uncertain causes can generate uncertain effects. Moreover, even facts (causes) of absolutely certainty may not guarantee the expected effects (i.e. the idea of uncertain causality; Suppes 1970). Therefore, we will associate a confidence measure with each of the causes (i.e. the facts in LHS of a rule) between -1 and 1, and a confidence measure also with the effect (i.e. the fact in RHS of a rule). We can use weights to create a mapping between confidence measures of causes (i.e. values of the corresponding facts) and confidence measures of effects (i.e. values of the corresponding facts), so from a set of causes and their confidence measures (i.e. a set of facts and their values) we can deduce a confidence measure for an effect (or a value for a fact in RHS). Moreover, the set of weights associated with facts in LHS of a rule should reflect their relative importance: more important causes should have a larger weight associated with them, and since the total weights sum to 1, the value of a weight for a particular fact (condition) reflects its relative importance against a background of all other conditions.

Another issue to consider is how to handle *possible condition* facts (i.e. those with $\Diamond$). As explained before, there is a special function associated with them, which elevates 0 to 1 or -1, according to whether positive or negative forms appear in the causal rule. Since we now extend the binary (or bipolar) space for truth values into a graded, continuous space, there is no more need for that elevation function. It follows from the fact that when a *possible condition* fact is unknown (i.e. its value is 0), the conclusion can still be reached, albeit with a smaller value (in confidence level). Now that we no longer require a binary (or bipolar) outcome, it is fine to have a smaller value for a conclusion when some enabling conditions are unknown. When one of these enabling conditions become known, the value will become higher; that is, we will have more confidence in the conclusion. Normally the weights associated with those enabling conditions will be relatively small anyway, because they are non-essential and close to "*don't care*" conditions. So it is advantageous to remove the elevation functions in the FEL version of Shoham's Causal Theory and assign weights instead.

An alternative perspective of viewing the extension is that of "fuzzifying" the necessity function and the possibility function. Once fuzzified, these new functions wind up to be identity functions. Therefore, combining the above two perspectives, *causes* are those conditions that have *high* weights, and *enabling conditions* are those conditions that have *low* weights.

We can now easily map the FEL terminology into the causal terminology as follows:

Events *are facts in FEL.*
Causal Statements *are rules in FEL.*
Causes *are those conditions of a rule that have high weights associated with them according to some particular weighting scheme.*
Enabling Conditions *are those conditions of a rule that have low weights associated with them according to some particular weighting scheme.*
Effects *are facts in the RHS of a rule.*

Let us go back to the issues we raised before:

- The gradedness is readily taken care of in FEL by the confidence values associated with each fact.
- Because of the introduction of the gradedness and uncertain rules (i.e. total weights sum to less than 1), the confidence we have in the conclusions will weaken along the way in a chaining. For example, here is a FEL rule stating that if one is alive at time t, one will be alive at time $t+1$:
$$alive(t) \longrightarrow alive(t+1)$$
Suppose the weight is equal to 0.99, then if given $alive(0)=1$, we will have $alive(1)=0.99$, $alive(2)=0.98$, $alive(3)=0.97$, and so on.
- There is no more need to tell exactly which condition is necessary and which condition is possi-

ble: they are graded and the difference is only quantitative.

- There is no more need to list all conditions (the total number of which might be infinite), as long as we leave room in the weight distribution (by keeping total weights less than 1). We can list only those conditions that we care about, and by doing so, the sum of weights will then be less than 1, accommodating possible roles of other unlisted conditions in determining the causal outcome.
- The indeterminate or probabilistic nature of causality is readily captured in the weighting scheme: the weights do not have to sum to 1, and not all conditions have to be known for certain in order to deduce a plausible conclusion.

Let us look back to the shooting example. Instead of having two separate causal statements in CT as before,

$$\Box shoot(t,t) \wedge \Diamond\neg wearing\text{-}bullet\text{-}proof\text{-}vest(t,t) \ldots\ldots$$

$$\longrightarrow \Box dead(t+1, t+1)$$

and

$$\Box\neg wearing\text{-}bullet\text{-}proof\text{-}vest(t,t) \wedge \Diamond shoot(t,t) \ldots\ldots$$

$$\longrightarrow \Box dead(t+1, t+1)$$

we will have in FEL one single causal statement for all the situations:

$$\neg wearing\text{-}bullet\text{-}proof\text{-}vest \wedge shoot \ldots\ldots$$

$$\longrightarrow dead \quad (w_1, w_2, \ldots\ldots)$$

and weights are assigned to each fact in LHS (Sun 1991). We assume the values of the unknown facts are zero and calculate the value of the conclusion by inner-products of the weights and the values of the facts in the LHS of the rule.

The Neural Net Connection

An implementation of FEL is a network of elements connected via links, where each element represents an atom and its negation and links represent rules, going from elements representing facts in LHS of a rule to elements representing facts in RHS of a rule; an element is a structure that has multiple sites each of which receives a group of links that represents one single rule, and the weighted-sum computation is carried out for computing and propagating activations. This implementation of FEL is clearly a connectionist network (Sun 1989, 1992).

Concluding Remarks

In this paper, Shoham's modal logic formalism for causal reasoning is critically analyzed; knowing its weakness in expressing graded concepts and other problems resulting from this, we proceed to define a different formalism, FEL, that utilizes weighted-sum computation and corresponds directly to neural net models. We prove that FEL can implement Shoham's logic as a special case as well as Horn clause logic, and furthermore that FEL is a justified extension of Shoham's logic. This work serves to justify a particular connectionist architecture proposed by the author, *CONSYDERR*, in its capabilities for coding rules and for performing commonsense causal reasoning (Sun 1991a).

REFERENCES

Chang, C. and Lee, R.C. 1973. *Symbolic Logic and Mechanical Theorem Proving.* Academic Press.

de Kleer, J. and Brown, J. 1986. Theories of Causal Ordering. *Artificial Intelligence*, 29:33-61.

Hink, R. and Woods, D. 1987. How Human Process Uncertain Knowledge. *AI Magazine*, 8:41-53.

Iwasaki, Y. and Simon, H. 1986. Causality in device behavior. *Artificial Intelligence*, 29:3-32.

Pearl, J. 1988. *Probabilistic Reasoning in Intelligent Systems.* Morgan Kaufman.

Shoham, Y. 1987. *Reasoning about Change.* Ph.D Dissertation, Yale University.

Sun, R. 1989. A discrete neural network model for conceptual representation and reasoning. *Proc.11th Cognitive Science Society Conference*, 916-923. Erlbaum.

Sun, R. and Waltz, D. 1991. Neurally Inspired Massively Parallel Model of Rule-Based Reasoning. in B. Soucek ed. *Neural and Intelligent System Integration*, John Wiley & Sons.

Sun, R. 1991. *Integrating Rules and Connectionism for Robust Reasoning.* Ph.D Dissertation, Brandeis University.

Sun, R. 1991a. Connectionist Models of Rule-Based Reasoning. *Proc.13th Annual Conference of Cognitive Science Society*, 437-442. Erlbaum.

Sun, R. 1992. Beyond associative memories: Logics and varaibles in connectionist models. *Information Sciences.* forthcoming.

Suppes, P. 1970. *A Probabilistic Theory of Causation.* North Holland.

Waltz, D. and Feldman, J. (eds.) 1986. *Connectionist Models and Their implications.* Ablex.

Zadeh, L. 1988. Fuzzy Logic. *Computer*, 21:83-93.