Int. J. Man-Machine Studies (1986) **24**, 65-75

Fuzzy cognitive maps

BART KOSKO

VERAC, Incorporated, 9605 *Scranton Road, Suite* 500, *San Diego, CA* 92121, *U.S.A.*

(*Received* 15 *April* 1985, *and in revised form* 10 *September* 1985)

Fuzzy cognitive maps (FCMs) are fuzzy-graph structures for representing causal reasoning. Their fuzziness allows hazy degrees of causality between hazy causal objects (concepts). Their graph structure allows systematic causal propagation, in particular forward and backward chaining, and it allows knowledge bases to be grown by connecting different FCMs. FCMs are especially applicable to soft knowledge domains and several example FCMs are given. Causality is represented as a fuzzy relation on causal concepts. A fuzzy causal algebra for governing causal propagation on FCMs is developed. FCM matrix representation and matrix operations are presented in the Appendix.

1. Introduction: the knowledge acquisition/processing tradeoff

Most knowledge is specification of classifications and causes. In general, the classes and causes are uncertain (fuzzy or random), usually fuzzy. This fuzziness passes into knowledge representations and on into knowledge bases, where it leads to a *knowledge acquisition/processing tradeoff*. The fuzzier the knowledge representation, the easier the knowledge acquisition and the greater the knowledge-source concurrence. But the fuzzier the knowledge, the harder the (symbolic) knowledge processing.

Fuzzy cognitive maps (FCMs) circumvent the tradeoff. FCMs are fuzzy-graph structures for representing causal reasoning. Their fuzziness allows hazy degrees of causality between hazy causal objects (concepts). Their graph structure allows systematic causal propagation, in particular forward and backward chaining, and it allows knowledge bases to be grown by connecting different FCMs. FCMs are especially applicable in soft knowledge domains (e.g. political science, military science, history, international relations, organization theory) where both the system concepts/relationships and the meta-system language are fundamentally fuzzy.

2. Cognitive maps

Political scientist Robert Axelrod (1976) introduced cognitive maps in the 1970s for representing social scientific knowledge. Axelrod's cognitive maps are signed digraphs. Nodes are variable concepts (like social instability, not like society) and edges are causal connections. A positive edge from node A to node B means A causally increases B. A negative edge from A to B means A causally decreases B. Cognitive maps facilitate *documentary coding*, constructing symbolic representations of expert documents (Henry Kissinger's documents code well; see Fig. 1).

Axelrod exploited the (adjacency) matrix representation of cognitive maps (see Fig. 2). Causal conceptual centrality in cognitive maps can be defined with adjacency-matrix

0020-7373/86/010065+11$03.00/0

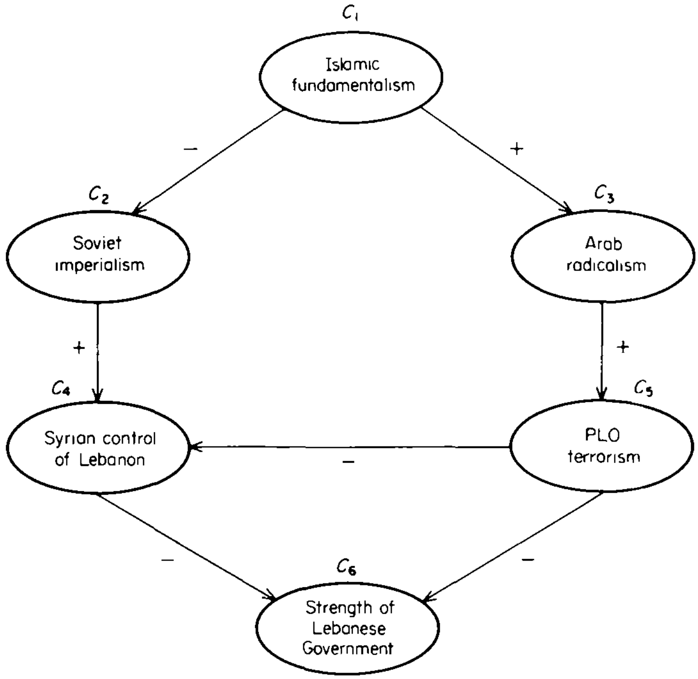


FIG. 1. A cognitive map constructed from Henry A. Kissinger's article "Starting Out in the Direction of Middle East Peace" (printed in the *Los Angeles Times*, Summer 1982). Positive edges represent causal increase. Negative edges represent causal decrease. The "policy variable" is ISLAMIC FUNDAMENTALISM. The "value variable" is STRENGTH OF LEBANESE GOVERNMENT. Other concept nodes are "cognitive variables."

$$
\begin{array}{c c}
 & \begin{array}{cccccc} C_1 & C_2 & C_3 & C_4 & C_5 & C_6 \end{array} \\
\begin{array}{c} C_1 \\ C_2 \\ C_3 \\ C_4 \\ C_5 \\ C_6 \end{array} &
\begin{pmatrix}
0 & -1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & -1 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\end{array}
$$

FIG. 2. The adjacency-matrix representation of Kissinger's cognitive map in Fig. 1. $e_{ij} = e(C_i, C_j)$ is the causal edge function value, the causality causal concept node C_i imparts to C_j. C_i causally increases C_j if $e_{ij} = 1$, causally decreases C_j if $e_{ij} = -1$, and imparts no causality if $e_{ij} = 0$.

components and much causal chaining information can be obtained from reachability matrices. These techniques are reviewed and fuzzified in the Appendix.

In general, cognitive maps are too binding for knowledge-base building. For, in general, causality is fuzzy. Causality admits of degrees, and vague degrees at that. It occurs partially, sometimes, very little, usually, more or less, etc. More generally still, the knowledge-base building promise of cognitive maps is combining knowledge sources' cognitive maps, but the fuzziness of the combined knowledge rises to the level of fuzziness of the fuzziest knowledge source. Fuzzy cognitive maps accomodate this knowledge-base building property. An example is the FCM in Fig. 3, where the causal

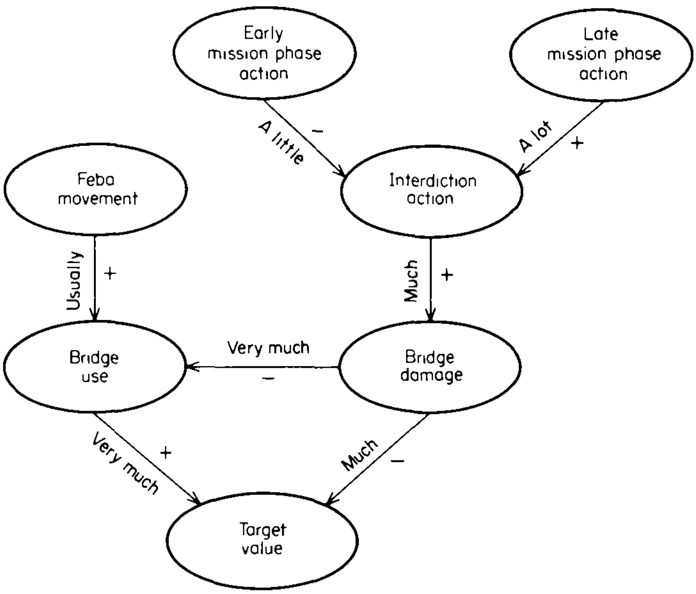


FIG. 3. Bridge Target Value FCM. Strategic objectives, mission tactics, and facts of the battlefield are related by fuzzy causality to produce a net utility effect. (FEBA stands for Forward Edge of Battle Area and hence as it moves, the probability that a random bridge will be used increases.)

relationships affecting a bridge's tactical target value are fuzzy in the manner of military science.

The next three sections formally develop FCMs and a fuzzy causal algebra for propagating causality on a FCM.

3. Representing causal reasoning

David Hume notwithstanding, causality is more complicated than logical implication. Consider causal increase, or positive causality. If "A causes B" is represented as "A implies B," then, by contraposition, "A causes B" is everywhere replaceable with "not-B causes not-A." But though smoking causes lung cancer, not having lung cancer does not cause non-smoking. Rather what can be inferred is that not smoking tends to cause lung non-cancer, so to speak. And this is a quite general relationship among causally increasing quantities (positive correlates). If A causes B, then increasing A increases B, and decreasing A decreases B. The inverse relationship holds between causally decreasing quantities (negative correlates). If A causally decreases B, then increasing A decreases B, and decreasing A increases B.

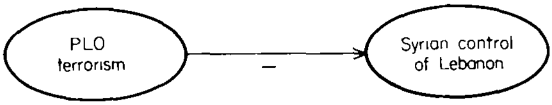
The non-implication nature of causal reasoning can be represented in a fuzzy set (logic) framework. Since causal objects are variable concepts, they can be represented as fuzzy subsets of some concept space, where change in fuzzy-set membership degree represents concept variation. Define a *concept* C_i as the fuzzy union (disjunction) of some fuzzy *quantity* set Q_i and associated *dis-quantity* set $\sim Q_i$:

$$C_i = Q_i \cup \sim Q_i.$$

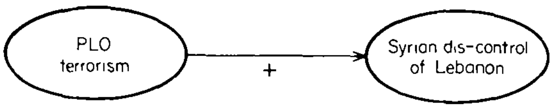
$\sim Q_i$ can be thought of as the abstract negation, or local fuzzy set complement, of Q_i. The "negation" operator $\sim$ simply indicates a set partition. (The negation of the concept is still its set complement.) It need only be a set index for the pair $(Q_i, \sim Q_i)$ that obeys double negation: $\sim\sim Q_i = Q_i$. Otherwise Q_i and $\sim Q_i$ are arbitrary fuzzy sets. Fuzzy causality can then be defined in terms of fuzzy set-theoretical (logical) relationships among fuzzy concepts. Let $C_j = Q_j \cup \sim Q_j$.

Definition. C_i *causes* C_j iff $Q_i \subset Q_j$ and $\sim Q_i \subset \sim Q_j$;
C_i *causally decreases* C_j iff $Q_i \subset \sim Q_j$ and $\sim Q_i \subset Q_j$, where "$\subset$" stands for fuzzy set inclusion (logical implication).

Hence negative causality can be defined with the same fuzzy quantities and relationships as positive causality; i.e. negative causality is eliminable. Hence the negative causal relationship



is equivalent to the positive causal relationship

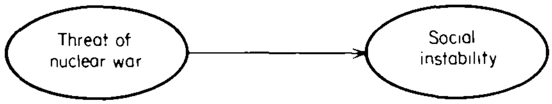


This observation leads to a general rule of replacement in FCM construction.

Rule. Replace every $C_i \underset{-}{\rightarrow} C_j$ *with* $C_i \underset{+}{\rightarrow} \sim C_j$.

Henceforth the negative causal arrow $\underset{-}{\rightarrow}$ will not occur. The unsigned arrow $\rightarrow$ will mean positive causality. The cognitive maps in Figs 1 and 3 are displayed transformed in Figs 4 and 5.

A second, more subtle point when representing causal reasoning is that *modifiers* of causal quantities need not be negated (complemented). Not smoking causes lung dis-cancer, not non-lung cancer. To develop a fuzzy sociological example, observe that the cognitive-map fragment



is equivalent to (or, more generally, functionally related to) the fragment

Here the generic causal quantities are THREAT and STABILITY (between which causal decrease still holds). The modifiers are NUCLEAR_WAR and SOCIAL. No assumption is made that modifier (fuzzy) sets are closed under abstract negation (complementation) in the underlying cognitive space. In particular, there need be no DIS_NUCLEAR_WAR or NOT_SOCIAL fuzzy subsets. Hence, more generally, causal concepts are built out of fuzzy set relations among quantity, dis-quantity, and modifier fuzzy subsets. Figure 6 pictures this situation in the sociological example.

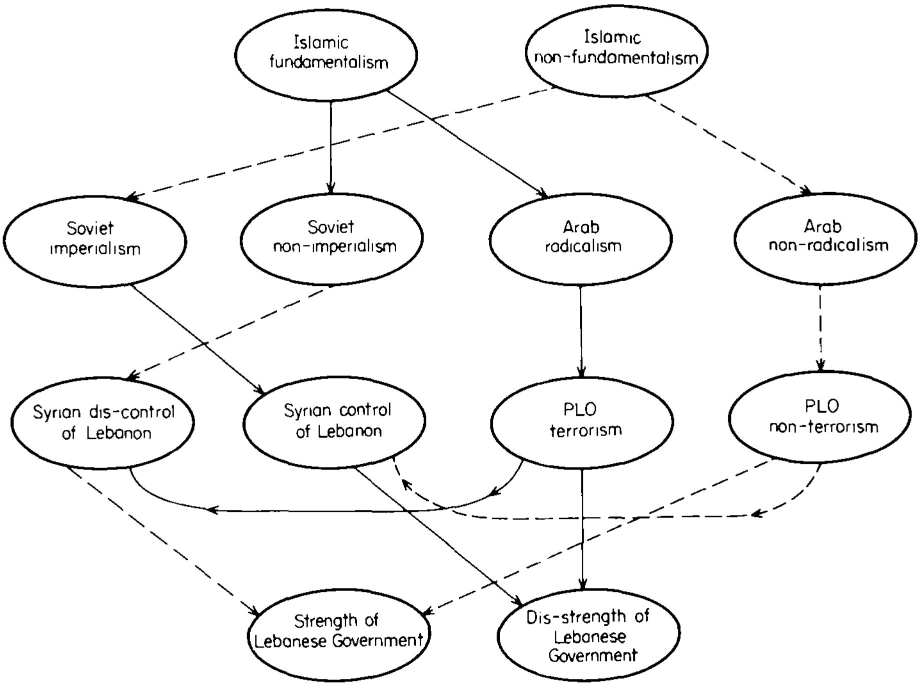


FIG. 4. Positive-causality representation of Kissinger's cognitive map in Fig. 1. All edge arrows indicate positive causality (causal increase). Edge arrows from new dis-concepts are dashed.

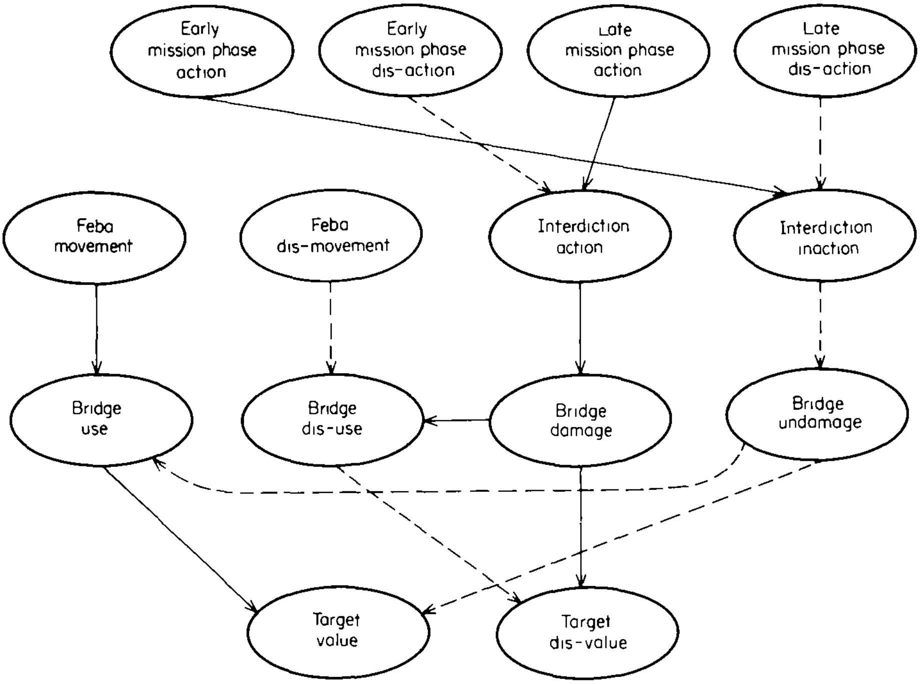


FIG. 5. Positive-causality representation of bridge target value FCM in Fig. 2. Fuzzy causal weights have been omitted for convenience. Edge arrows from new dis-concepts are dashed. (The default assumption is that the fuzzy causal weight between transformed dis-concepts is the same weight as between the untransformed concepts. In general, the weights may differ.)

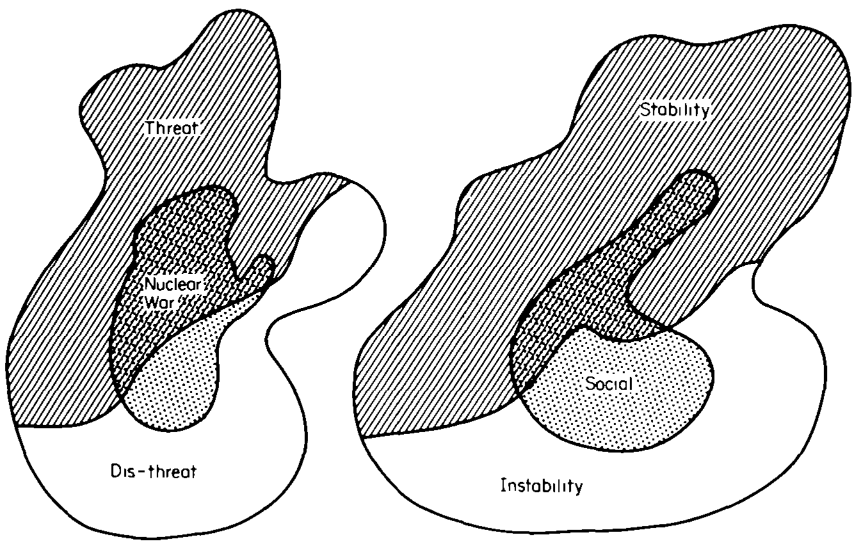


FIG. 6. Fuzzy-set representation of causal concepts. A modifier set fuzzily intersects the fuzzy union of a quantity and dis-quantity set. (In general, causal concept components need not be connected.) Causal increase or decrease is represented by the appropriate inclusion of intersections.

The cognitive-map fragment in the sociological example can be symbolically represented as set inclusions of set intersections:

$$\text{THREAT} \cap \text{NUCLEAR\_WAR} \subset \text{INSTABILITY} \cap \text{SOCIAL},$$

$$\text{DIS\_THREAT} \cap \text{NUCLEAR\_WAR} \subset \text{STABILITY} \cap \text{SOCIAL},$$

where "$\cap$" stands for fuzzy set intersection (conjunction).

More generally, let Q_i, $\sim Q_i$ and M_i be the respective *i*th quantity, dis-quantity, and modifier fuzzy sets. (The default M_i is the set universe.) Then the *i*th *causal concept* C_i can be defined as $C_i = (Q_i \cup \sim Q_i) \cap M_i$. This leads to the final definitions of causal increase and decrease.

Definition. C_i *causes* C_j iff $(Q_i \cap M_i) \subset (Q_j \cap M_j)$ and $(\sim Q_i \cap M_i) \subset (\sim Q_j \cap M_j)$; C_i *causally decreases* C_j iff $(Q_j \cap M_i) \subset (\sim Q_j \cap M_j)$ and $(\sim Q_i \cap M_i) \subset (Q_j \cap M_j)$.

4. The abstract FCM framework

FCMs are fuzzy causal graphs (fuzzy graphs). The apparatus needed to define them requires embedding the concepts and definitions of the previous section in abstract fuzzy spaces. The fuzzy logical analogue is immediate but will not be mentioned.

Let X be some non-empty set. Let $F(2^X)$ denote the fuzzy power set of X—the set of all fuzzy subsets of X. For fuzzy subsets A, $B \in F(2^X)$, define the *degree of subsethood* (Bandler & Kohout, 1980) of A in B by $m_{F(2^B)}(A)$, i.e. the degree to which A belongs to B's fuzzy power set. Degree of subsethood will used to represent fuzzy causality.

Call the fuzzy-subset class $\mathscr{Q} \subset F(2^X)$ a *quantity space* on X if every $A \in \mathscr{Q}$ can be represented as $A = Q \cup \sim Q$ for some Q, $\sim Q \in F(2^X)$. Again "$\sim$" stands for abstract negation or local complementation. Call some fuzzy-subset class $\mathscr{M} \subset F(2^X)$ a *modifier space* on X if $X \in \mathscr{M}$. (X is the default modifier.) Then call the class $\mathscr{C} \subset F(2^X)$ a *concept space* on X if $\mathscr{C} = \mathscr{Q} \cap \mathscr{M}$, i.e., if $\mathscr{C} = \{(Q \cup \sim Q) \cap M\colon Q \cup \sim Q \in \mathscr{Q}, M \in \mathscr{M}\}$. Say that the concept space $\mathscr{C}$ is *causal* if, for all C_i, $C_j \in \mathscr{C}$,

(1) $Q_i \cap M_i \subset Q_j \cap M_j \Rightarrow \sim Q_i \cap M_i \subset \sim Q_j \cap M_j$,

(2) $Q_i \cap M_i \subset \sim Q_j \cap M_j \Rightarrow \sim Q_i \cap M_i \subset Q_j \cap M_j$, for some $Q_i \cup \sim Q_i$, $Q_j \cup \sim Q_j \in \mathscr{Q}$ and some M_i, $M_j \in \mathscr{M}$.

Then call the abstract pair $\mathscr{F} = (X, \mathscr{C})$ a *fuzzy cognitive space* on X if $\mathscr{C}$ is causal (and contains fuzzy sets).

$\mathscr{C}$ contains the fuzzy nodes of the abstract FCM (fuzzy causal graph). The graph is specified when the causally connected subsets of $\mathscr{C} \times \mathscr{C}$ are specified. This amounts to defining a fuzzy edge function. (The edge function is fuzzy if its range set contains more than two objects.) Formally, e: $\mathscr{C} \times \mathscr{C} \to P$ is a *fuzzy causal edge function* on $\mathscr{C}$ if $e_{ij} = (C_i, C_j) = m_{F(2^{C_j})}(C_i)$, the fuzzy-set membership of concept C_i in concept C_j's fuzzy power set, i.e. the degree of subsethood of C_i in C_j. The range set P can be any partially ordered set; classically, $P = [0, 1]$. To generalize unit-interval conventions, assume $e(C_i, C_i) \leq p$ for all $p \in P$. So if P is the unit interval, $e_{ii} = 0$ and degenerate cycles are prohibited. Then, finally, the abstract pair (e, $\mathscr{F}$) is a *fuzzy cognitive map* on X if the fuzzy causal graph (e, $\mathscr{C}$) is cycle-free.

5. Fuzzy causal algebra

A fuzzy causal algebra governs causal propagation and causal combination on a FCM. It thus governs forward and backward chaining on a FCM. The algebra developed below depends only on the partial ordering on P, the range set of the fuzzy causal edge function e, and on general fuzzy-graph properties (e.g. path connections). The algebra extends to any digraph knowledge-representation scheme.

Axelrod speaks of indirect and total causal effects on cognitive maps. Fix some causal path from concept node C_i to concept node C_j, say $C_i \to C_{k_1} \to \cdots \to C_{k_n} \to C_j$, which can be denoted with ordered indices as $(i, k_1, \ldots, k_n, j)$. Then the *indirect effect* from C_i to C_j is the causality C_i imparts to C_j (the degree of subsethood of C_i in C_j) via the path $(i, k_1, \ldots, k_n, j)$. The *total effect* of C_i on C_j is all the indirect-effect causality C_i imparts to C_j. Hence if there is only one causal path from C_i to C_j, the total effect of C_i on C_j reduces to the indirect effect.

The operations of indirect and total effect correspond naively to multiplication and addition of real numbers (field elements). Axelrod (1976) employs a causal calculus of signs (+ and −) that exploits this correspondence. The indirect effect of a path from C_i to C_j is negative if the number of negative causal edges in the path is odd, positive if the number is even. The total effect of C_i on C_j is negative if all indirect effects on C_i on C_j are negative, positive if they are all positive, indeterminate otherwise. Hence indeterminacy tends to dominate in this sign scheme. It can be removed, for a price, with a numeric weighting scheme. If the causal edges are weighted with positive or negative real numbers w_{ij}—i.e. if $w_{ij} = e_{ij} \in \mathfrak{R}$—as might occur if concept A causes B five times as much as C causes B, then the indirect effect of C_i on C_j on path $(i, k_1, \ldots, k_n, j)$ is the product $w_{ik_1} \times w_{k_{12}} \times \cdots w_{k_n j}$ and the total effect is the sum of the path products. This weighting scheme generalizes the sign calculus, removes indeterminacy from the total-effect operation, and allows (requires!) finer causal discrimination between concepts. But the real-valued requirement of P makes knowledge acquisition difficult: forced numbers from insufficient decision information, different numbers from the same knowledge source on different days, etc. The difficulties increase with the number of knowledge sources and with the amount of knowledge-source responses—the knowledge acquisition/processing tradeoff in this context.

A fuzzy causal algebra, and hence an FCM, bypasses the knowledge acquisition/processing tradeoff. It allows fuzzy inputs to be processed as systematically as real-valued inputs. The only price paid is a fuzzy output!

A fuzzy causal algebra is created by abstracting operations from multiplication and addition that are defined on a (fuzzily) partially ordered set P of causal values. Let $\mathscr{C}$ be a causal concept space (on some underlying nonempty set X) and let e: $\mathscr{C} \times \mathscr{C} \to$ P be a fuzzy causal edge function. (A really fuzzy e maps into fuzzy subsets of some class P fuzzily partially ordered by fuzzy set inclusion (degree of subsethood).) Then the simplest abstract operations are got by interpreting the indirect-effect operator I as some minimum (infimum) operator and the total-effect operator T as some maximum (supremum) operator—these operators depending only on P's partial order—and the simplest of these operators are the min (inf) and the max (sup). Formally, let there be m-many causal paths from C_i to C_j: $(i, k_1^l, k_2^l, \ldots, k_{n_l}^l, j)$ for $1 \le l \le m$. Let $I_l(C_i, C_j)$ denote the indirect effect of concept C_i on concept C_j on the lth causal path. Let $T(C_i, C_j)$ denote the total effect of C_i on C_j over all m causal paths. Then

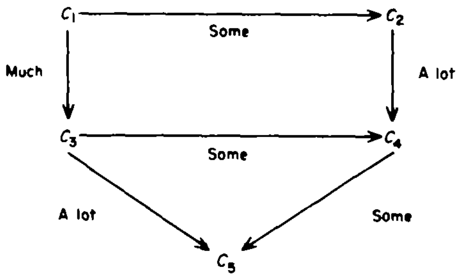
$$I_l(C_i, C_j) = \min\{e(C_p, C_{p+1}) \colon (p, p+1) \in (i, k_1^l, \ldots, k_{n_l}^l, j)\},$$

$$T(C_i, C_j) = \max_{1 \le l \le m} I_l(C_i, C_j),$$

where p and $p+1$ are contiguous left-to-right path indices.

Hence the indirect-effect amounts to specifying the weakest causal link in a path and the total-effect operation amounts to specifying the strongest of the weakest links.

For example, suppose the causal values are given by $P = \{\text{none} \lesssim \text{some} \lesssim \text{much} \lesssim \text{a lot}\}$ and the FCM is given by



The three causal paths from C_1 to C_5 are $(1, 3, 5)$, $(1, 3, 4, 5)$, $(1, 2, 4, 5)$. So the three indirect effects of C_1 on C_5 are

$$\begin{aligned} I_1(C_1, C_5) &= \min\{e_{13}, e_{35}\} = \min\{\text{much, a lot}\} \\ &= \text{much}, \end{aligned}$$

$I_2(C_1, C_5) = \text{some}$, $I_3(C_1, C_5) = \text{some}$. Thus the total effect of C_1 on C_5 is

$$\begin{aligned} T(C_1, C_5) &= \max\{I_1(C_1, C_5), I_2(C_1, C_5), I_3(C_1, C_5)\} \\ &= \max\{\text{much, some}\} = \text{much}. \end{aligned}$$

In words, C_1 imparts much causality to C_5.

Finally, in fullest generality, I and T can be any respective t-norm (triangular-norm) operator t and t-conorm s (see Klement, 1981; Yager, 1981, for definitions and properties). If P is the unit interval, then the two triangular norms are related by

$$t(x, y) = 1 - s(1 - x, 1 - y),$$

and by similar transformations for more general range sets P. Yager (1981) proves the following maximality/minimality property of min and max:

$$t(x, y) \le \min(x, y) \le \max(x, y) \le s(x, y).$$

Thus t-norms other than min are never more causally lenient than min and t-conorms other than max are never more causally stringent than max. This property justifies selecting min and max as default I and T operators when little or nothing is known about how leniently I or how stringently T should causally behave.

References

AXELROD R. (1976). *Structure of Decision: the Cognitive Maps of Political Elites.* Princeton, NJ: Princeton University Press.

BANDLER, W. & KOHOUT, L. (1980). Fuzzy power sets and Fuzzy implication operators. In *Fuzzy Sets and Systems*, vol 4, pp. 13–30.

KLEMENT, E. P. (1981). Operations on fuzzy sets and fuzzy numbers related to triangular norms. *11th International Conference on Multi-valued Logic.* University of Oklahoma.
YAGER, R. R. (1981). Comments on "Textured Sets. An Approach to Aggregation Problems with Multiple Concerns". *IEEE Transactions on Systems, Man, and Cybernetics,* vol 11, pp. 730–731.

Appendix: FCM matrix properties

Let $C_1, C_2, \ldots, C_n$ be causal concepts and let $e_{ij} = e(C_i, C_j)$ be the causal edge function value, the amount of causality C_i imparts to C_j. For a cognitive map, $e_{ij} = 0$, 1, or -1 (0 represents an absence of a causal relationship). Let $E = (e_{ij})_{1 \le i,j \le n}$ represent the matrix of causal edge values for the given FCM.

Suppose the FCM is a cognitive map. Then E is an *adjacency matrix.* It lists all one-edge paths on the cognitive map. $E^2 = [e_{ij}^{(2)}] = E \times E$ lists all two-edge paths on the cognitive map. For

$$e_{ij}^{(2)} = \sum_{k=1}^{n} e_{ik} \times e_{kj}$$

is nonzero only if there is a k′ such that $e_{ik'}$ and $e_{k'j}$ are nonzero. Similarly $E^3, E^4, \ldots, E^{n-1}$ list the effect of summing all three-edge, four-edge, ..., (n − 1)-edge indirect effects. (Since FCMs are acyclic, there are no paths with more than n − 1 edges.) Then the *total-effect matrix* T is the sum of the powered matrices E^i:

$$T = \sum_{i=1}^{n-1} E^i.$$

(A useful criterion for the existence of cycles in E is the following: E is acyclic if and only if T's main diagonal is everywhere zero (otherwise, some concept is affecting itself).) T is more generally known as the *reachability matrix.* If the above process is repeated with E replaced $\bar{E}$, the matrix of absolute values of E, then $\bar{e}_{ij}^{(k)}$ is nonzero if and only if there are $\bar{e}_{ij}^{(k)}$-many *k*-edge paths from C_i to C_j. Such information is useful when searching for forward and backward chains.

The *conceptual centrality* of causal concept node C_i is denoted by $CEN(C_i)$ and defined by

$$CEN(C_i) = IN(C_i) + OUT(C_i),$$

where

$$IN(C_i) = \sum_{k=1}^{n} \bar{e}_{ik},$$

$$OUT(C_i) = \sum_{k=1}^{n} \bar{e}_{ki}.$$

The column sum of absolute values $IN(C_i)$ represents the number of concepts causally impinging on concept C_i. Similarly the row sum $OUT(C_i)$ represents the number of concepts concept C_i causally impinges on. Hence the conceptual centrality $CEN(C_i)$ represents the importance of concept node C_i to the causal flow on the cognitive map.

These concepts can be fuzzified and applied to more general FCMs. Suppose the causal edge function is classically fuzzy, i.e. suppose it takes values in the unit interval.

Then construction of fuzzy adjacency and fuzzy reachability matrices proceeds as above by replacing e_{ij} with a causal indicator function (which is 1 if and only if $e_{ij} > 0$). The fuzzy conceptual centrality of concept C_i is computed directly as in the non-fuzzy case. The advantage now is that causal quality counts. A concept node can be connected to fewer nodes than another concept yet still have greater conceptual centrality if its connections are more heavily weighted.

When e maps into more general partially ordered sets, indicator functions can still be used to compute adjacency and reachability matrices, since only cardinality of connections matters. Causal conceptual centrality can be computed with the above identities by replacing addition with maximum (supremum).