MODELING AETIOLOGICAL KNOWLEDGE AND CAUSAL-EVIDENTIAL REASONING USING Q/C-E NETWORKS

Pietro Baroni*, Giovanni Guida*, Silvano Mussi°*

* Università di Brescia, Dipartimento di Elettronica per l'Automazione, Via Branze 38, 25123 Brescia, Italy
tel: +39 30 3715469 fax: +39 30 380014 - e-mail: {baroni, guida}@bsing.ing.unibs.it

° Consorzio Interuniversitario Lombardo per l'Elaborazione Automatica, Via Sanzio 4, 20090 Segrate (MI), Italy
tel: +39 2 26922434 fax: +39 2 2135520 - e-mail: mussi@icil64.cilea.it

1. Introduction

Causal reasoning is one of the most popular forms of common-sense reasoning, which has been widely studied both from theoretical and application points of view. Many attempts have been made to develop models and formalisms able to capture the essential features of this form of reasoning. These include, among others, different types of causal networks [10] [11], the debate about causality in qualitative physics [5] [6] [7] [8], and theories of causation [14] [9] [13]. However, no one of these attempts seems to have received an incontestable consensus. According to [15] "The topic of causality remains controversial." Among the many open problems in the field of causal reasoning, we focus in this paper on some fundamental issues which concern:

- the epistemological analysis of knowledge underlying the concept of causality;
- the study of the indissoluble relation between causality and uncertainty;
- the characterization of the general processes underlying the intuitive concept of reasoning about causes and effects, and the investigation of the relationships existing between knowledge representation and reasoning, often obscured by the use of ambiguous terms such as "causal knowledge" and "causal reasoning".

2. Aetiological knowledge: a definition

Causality has often been related to some non-causal model of the considered system. For instance in [5] it is assumed that components can be modeled through qualitative confluence equations, whereas in [7] equations are used to express functional relations. De Kleer and Brown claim that "the engineer's notion of causality derives from the relationship between the equations and their underlying components which comprise the modeled system" [6], and in [7] causal ordering is defined as "an asymmetric relation among the variables and equations of a set of simultaneous equations".

Our goal is to capture a more general concept of causality, which is independent and preexistent with respect to equations. We advocate that the notion of causality is not related, in its origin, to that of physical knowledge (of course, causal reasoning can not contradict physical evidence), but proceeds along a different path and exploits an essentially different epistemological type of knowledge. Our standpoint is that there exists a form of knowledge, that we call *aetiological knowledge*, which humans derive from experience through application of a general mental scheme and then use for supporting various reasoning tasks. Intuitively, aetiological knowledge has to do with the attribution of reciprocal roles (the role of *cause* or the role of *effect*) to observed phenomena. Consider, for example, the case of a complex system, such as the human body, where:

- internal physico-chemical phenomena are very far from being completely known;
- experimentation on working systems (i.e., living persons) is subjected to strong limitations;
- the classical modeling scheme using separate interconnected components is inadequate (consider, for example, the role and the nature of blood).

Equations to model the human body (or significant parts of it) are not available and it is fairly dubious that they would in any case be an adequate modeling tool. For instance, there is no equation relating the quantity of swallowed food to indigestion. However, everyone of us is able to relate the ingestion of an abnormal quantity of food at dinner with an indigestion the day after. This ability is related to the use of aetiological knowledge.

In general, when dealing with poorly understood phenomena, humans try to apply to them some form of simple mental scheme, in order to derive useful knowledge from observations. Note that knowledge may be useful for practical tasks, even if it is not complete nor correct. Indeed, aetiological knowledge is especially useful in domains where incompleteness and possibly incorrectness affects most of the available information. Therefore it appears that uncertainty is a fundamental ontological attribute of aetiological knowledge, as it will discussed later in more detail. Intuitively, the *mental scheme* we can assume as underlying the existence of aetiological knowledge is quite simple:

> Given an observed phenomenon (say O-phen), one tries to find out among previously observed phenomena (say PREV-phen), at least one phenomenon, (say P-phen), to which he can ascribe the role of cause of O-phen, which in turn is ascribed the role of effect of P-phen.

For instance, in presence of an indigestion one ascribes the role of cause to the food ingested the day before, discarding, for instance, the fact that the patient smoked many cigarettes or that he was sitting at the head of the table.
The attribution of *cause* and *effect* roles to phenomena generates aetiological knowledge (aetiology is the study of the causes), which, in its very nature, is made up by cause-effect relations between phenomena.
The crucial point here is, of course, how P-phen is chosen in PREV-phen. Aetiological knowledge may be derived from repeated observations, from ad hoc experimentation, from considerations about system structure, etc. The detailed analysis of this issue which concerns the formation of aetiological knowledge is, however, beyond the scope of the present paper. What is important to note here is that aetiological knowledge is the result of the application of a mental scheme to an observed system; therefore, in principle, is not necessarily related with knowledge about the physical properties of the observed system. For instance, when relating food to indigestion, we may totally ignore that an organ named stomach exists, that indigestion is a disease involving stomach, and that food is digested by stomach.
The reason why we ascribe roles to phenomena is basically pragmatical. For instance we will be able to prevent indigestion by avoiding its cause (excessive food ingestion) and we will be able to foresee an indigestion in presence of its cause. Therefore, since aetiological knowledge arises from practical needs, it may prove more or less adequate with respect to the goal of correctly performing tasks relevant to such needs. For instance, one could erroneously recognize a different cause for indigestion, or ignore other possible different causes for the same effect, or neglect factors enabling or preventing the causal relation, and thus fail to prevent indigestion on the occasion of another dinner or prevent it by excessive precautionary measures.
Since aetiological knowledge originates from a purely mental scheme and lacks, in general, a physical support, its adequacy can not be definitively proved. Moreover, as already mentioned, completeness and correctness are usually unachievable requirements in most contexts where aetiological knowledge may prove useful. Therefore, aetiological knowledge is intrinsically affected by uncertainty.

3. Causal and evidential reasoning

Aetiological knowledge may be used for two different, but substantially related, basic reasoning tasks, namely evidential and causal reasoning. *Evidential reasoning* consists in deriving causes from effects: given an observed (or hypothesized) phenomenon, one tries to explain it by hypothesizing the presence of other phenomena which may assume the role of its causes. *Causal reasoning* consists in deriving effects from causes: given an observed (or hypothesized) phenomenon, one is induced by it to expect the presence of other phenomena, which have the role of its effects.
The distinction between causal and evidential reasoning has been firstly advocated by Pearl [12] who claimed that "almost every default rule falls into one of two categories: expectation-evoking or explanation-evoking". In our opinion, this claim should, however, be somewhat adjusted. First of all, it is possible to imagine default rules which consist in purely empirical associations, without any causal or evidential connotation, such as "red tomatoes are tender". The proposed classification applies correctly only to rules expressing aetiological knowledge. Second, the indication that every rule belongs to one of two separate classes is somewhat misleading. In fact, the proposed classification refers to different uses that can be made of a piece of knowledge, rather than to intrinsically different properties of knowledge itself. This is rather clear if one examines the well-known example about rain and sprinkler proposed in [12]. Knowledge about this example is represented through 8 rules: 4 of them are causal (e.g., "rain causes wet grass") and the other 4 are their evidential counterpart (e.g., "wet grass is an evidence for rain"). As a matter of fact, it is more natural to think that behind each pair of rules there is only one chunk of aetiological knowledge. For example, the two rules presented above are summarized by a single relation where rain assumes the role of cause and wet grass that of effect.
Causal and evidential reasoning are two different, although inseparable, ways of exploiting aetiological knowledge. In fact, it is definitely impossible to imagine any piece of aetiological knowledge which could not be exploited in both these ways. Moreover any interesting "causal reasoning" task always includes some combination of both these forms of reasoning. In fact, generation of hypotheses through evidential reasoning would not be really useful if the effects of these hypotheses could not be explored. For instance, in a diagnostic task, starting from a given observation, it is necessary to explain it, in order to verify if it suggests the presence of a fault, but it is also necessary to derive all the consequences of what has been observed and of its hypothesized causes, in order to check if there is a reasonable expectation for any fault or whether an hypothesized fault is consistent with new observations.

4. Quantified causal-evidential networks

Quantified causal-evidential networks (*Q/C-E networks*) are a knowledge representation formalism purposely designed to represent aetiological knowledge and to reason with it.

Nodes of a Q/C-E network are associated to propositions concerning the domain of interest. Uncertainty about a proposition P is represented by a belief state namely a pair bel[(P, true),bel(P, false)] denoted as bels(P). The belief state represents how much one is authorized to believe in the association between a given proposition and its possible truth values, on the basis of the available evidence, The two components of the belief state, called belief degrees, are independent and may assume values in the real interval [0, 1] or in a finite set of linguistic labels [4].

Turning now to *relations* between propositions, they are represented by a double link between a pair of nodes, let say P and Q: one link, qualified as *causal link* (say *C-link*) and directed, for instance, from P to Q, indicates that P has the role of cause of Q; the other link, qualified as *evidential link* (say *E-link*) and directed in the opposite direction, indicates that Q has the role of effect for P. In Q/C-E networks, links indicate that belief states in the linked propositions have a reciprocal influence; they do not express a relation between propositions, but rather between their belief states. In fact, if you have reasons for believing that P is true, then you are, at some extent, authorized to believe that also Q is true, and vice versa.

A belief degree, called *belief level*, is associated to causal and evidential links. The belief level of a relation R between propositions is denoted by bell(R). Considering again the propositions P and Q mentioned above, the value of the belief level of the C-link from P to Q represents how much one is prompted to believe in Q, supposed that P is totally believed to be true; similarly, the value of the belief level of the E-link from Q to P represents how much one is prompted to believe in P, supposed that Q is totally believed to be true. A relation with its own associated belief level is called *quantified relation*.

5. Reasoning with Q/C-E networks

In this section we illustrate propagation and aggregation in Q/C-E networks with reference to the preventive diagnosis task. In particular, our discussion will be based on the experience gathered in the design of ASTRA, a knowledge-based system for preventive diagnosis of power transformers [1].

We will focus on a concrete example: Figure 1 shows a fragment of a Q/C-E network used in ASTRA, which captures a piece of the expert's aetiological knowledge about transformer insulating oil.

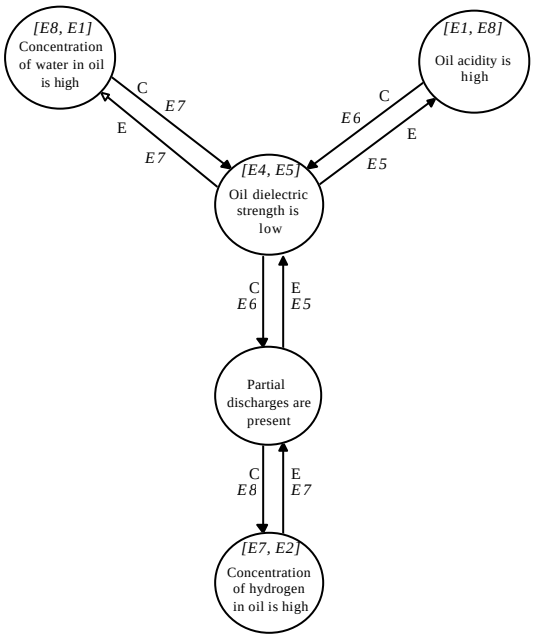

Figure 1 - A fragment of a Q/C-E network.

Belief degrees are represented in ASTRA through linguistic labels, whose semantics is based on fuzzy numbers. A scale of nine linguistic labels has been adopted, derived from [4], namely: [UNINFORMED, EXTREMELY UNLIKELY, MOST UNLIKELY, UNLIKELY, IT MAY, LIKELY, MOST LIKELY, EXTREMELY LIKELY, CERTAIN]. Each label in this set is represented in short form as E_i, with i = 1, 2, ..., 9. E_1 (UNINFORMED, the minimum) represents negligible belief (including null belief) in the truth or falsity of a proposition P, while E_9 (CERTAIN, the maximum) represents full belief in the truth or falsity of P. In the following subsections we will describe the main steps of uncertain reasoning within Q/C-E networks.

5.1. Evidence collection and translation into belief states of the relevant propositions

Uncertain reasoning starts with the collection of evidences from the real world and the translation of evidences into belief states of the relevant propositions. These phases are clearly domain dependent and their description is beyond the scope of this paper. Therefore, let us suppose that, starting from collected evidences, four initial belief states are generated for the corresponding nodes of the Q/C-E network. They are showed in Figure 1.

5.2 Propagation of belief states through the Q/C-E network

Initial belief states have then to be propagated throughout the network. For the sake of clarity, we examine propagation in two stages: local propagation (i.e., from one node to the adjacent nodes) and global propagation (i.e., propagation through the entire Q/C-E network).

5.2.1 Local propagation

Suppose that P and Q are two adjacent nodes in a Q/C-E network, that P is a cause of Q, and that a belief state has been associated to P, namely: bels(P) = [bel(P, true), bel(P, false)]. Propagating uncertainty from P to Q implies being able of establishing which belief state should be associated to Q, given the belief state associated to P and the belief levels of the C- and E-links connecting P and Q. To formally face the problem let us consider the goal of calculating the pair bels(Q) = [bel(Q, true), bel(Q, false)] on the basis of [bel(P, true),bel(P, false)], bell(P $\xrightarrow[C]{}$ Q), and bell(Q $\xrightarrow[E]{}$ P).

As far as bel(Q, true) is concerned, the following intuitive proportion may be considered:

bel(Q, true) : bel(P, true) = bell(P $\xrightarrow[C]{}$ Q) : 1,

where 1 stands for "total belief" (about the truth of P, i.e. it is certain that P is true). Therefore we may state that:

bel(Q, true) = bel(P,true) $*$ bell(P $\xrightarrow[C]{}$ Q),

where the symbol "$*$" denotes for usual multiplication between real numbers (or a T-norm operator [3] between fuzzy numbers representing belief degrees, in case of linguistic labels being used to denote belief degrees).

As far as bel(Q, false) is concerned, consider that, given a certain belief degree for $\neg$ P, the strongest is the dependence of the belief in Q on belief in P, the more the belief in $\neg$ P should give rise to belief in $\neg$ Q. An intuitive indication of how strong is the dependence of the belief in Q on belief in P is given just by bell(Q $\xrightarrow[E]{}$ P): if the presence of an effect strongly suggests the presence of a certain cause, this means that the effect is strictly tied to this cause and, therefore, reasons for disbelieving in the cause should also have a primary influence on the disbelief in the effect. For a better understanding of this point, consider the example, already presented above, of a disease giving occasionally rise to a very particular symptom: the causal link is weak, but the evidential link is strong, accordingly to the fact that the symptom strictly depends on that disease. In this case, it is reasonable to assume that if there are reasons to exclude the presence of the disease, there are also reasons to exclude that the symptom will appear. On the contrary, if bell(Q $\xrightarrow[E]{}$ P) is low, the effect is loosely bound to the cause and reasons for disbelieving the cause should have less influence on disbelief in the effect. Therefore we may assume that:

bel(Q, false) = bel(P, false) $*$ bell(Q $\xrightarrow[E]{}$ P).

The discussion of local propagation presented above has considered the case of deriving the belief state of an effect from that of a cause. Parallel considerations can be made for the problem of deriving the belief state of a cause from that of an effect. In this case the following relations hold:

bel(P, true) = bel(Q, true) $*$ bell(Q $\xrightarrow[E]{}$ P)

bel(P, false) = bel(Q, false) $*$ bell(P $\xrightarrow[C]{}$ Q).

5.2.2 Global propagation

Global propagation, is based on three main points:

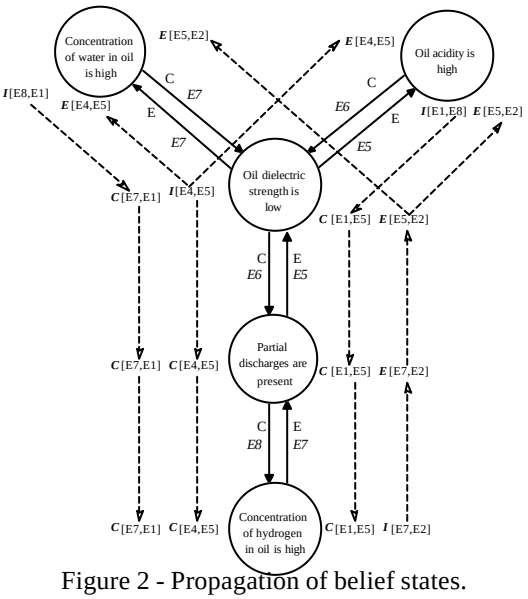

Figure 2 - Propagation of belief states.

- Belief states with different origins are propagated separately. In particular three types of belief states are defined:
  - I belief states, deriving directly from initial evidences (initial belief states);
  - E belief states, deriving from an evidential link;
  - C belief states, deriving from a causal link.
- Propagation of I and E belief states is carried out along both causal and evidential links and, therefore, can generate both C and E belief states. Propagation of C belief states, instead, is carried out only along causal links, and not along evidential links. In other words, from causes one can only infer effects, and from effects further effects. This allows avoiding in a simple and natural way the well-known problem of *illegal inferences* [12].
- Each propagated belief state is labelled with the list of all previously visited nodes, and propagation is blocked when an already visited node is reached again. This allows avoiding in a general and effective way the well-known problem of *cyclic dependencies* [2].

Figure 2 shows the results of propagation through the fragment of Q/C-E network introduced in Figure 1.

5.3. Aggregation

After propagation has been carried out, each node of the Q/C-E network will have more than one belief state associated to it (of I, E, or C type). Assuming therefore that to each node a list of belief states is associated, called *belief list*, the

purpose of aggregation is to produce a final synthetic judgement about the global belief that one is reasonably authorized to associate to a given proposition. Aggregation is done in two steps, namely:

- First of all, *partial aggregation* combines together separately the belief states having different origins, namely I, E, and C belief states. This way, the belief list associated to each node in the Q/C-E network is reduced to a triple (possibly, with some missing element), called the *aggregated IEC triple*.
- Later, *final aggregation* combines together the elements of each aggregated IEC triple into a global, final judgment about the belief state of the relevant proposition.

In order to define appropriate aggregation operators for partial and final aggregation, a basic distinction should be made between monotonic and non-monotonic operators.

Monotonic aggregation ensures that a belief state expressing disbelief must not produce a weakening effect when aggregated with a belief state expressing belief. This type of aggregation is perfectly according to common sense in the case of partial aggregation of belief states of type C, i.e. originating from causal links.. Intuitively, the disbelief in one of the causes of a certain effect, should not weaken the belief in the effect due to the presence of other causes.

According to the intuitive concept introduced above, monotonic aggregation of a set of belief states, can be simply obtained by selecting in the set the belief state which is closest to the B point.

Non-monotonic aggregation allows the results of subsequent aggregation operations to freely go in the direction of a stronger or weaker belief, taking into account both the distribution of the different belief states and their number. This type of aggregation is perfectly according to common sense in the case of aggregation of belief states of type I and E, i.e. originating from initial evidence or from evidential links and, also, in the case of final aggregation. In fact, a belief based on a collected evidence, should be retractable if other contrasting evidences arise. This can be done by defining a non-monotonic aggregation operator based on a principle of elastic equilibrium. More precisely, we first represent all the belief states of P in the correct positions on the belief plane (i.e. the plane representing all possible belief states) and consider them as fixed points. Then, we provide ideal elastic springs connecting each of these points with a mobile point representing the (desired) result of non-monotonic aggregation (i.e., the aggregated belief state of P), and we define appropriate elastic constants (K_i) for such springs, depending both on the origin of the belief state and on its position in the belief plane (we skip here the details due to space limitations). Finally, we compute the position of elastic equilibrium of P in the belief plane and interpret it as the aggregated belief state of P.

More formally, given the set = {[t1 f1], [t2 f2], ..., [tn fn]}, the non-monotonic aggregation operator computes the belief state [tx fx], where tx and fx are the roots of the following two equations:

$$K1*(tx - t1) + K2*(tx - t2) + ... + Kn*(tx - tn) = 0$$

$$K1*(fx - f1) + K2*(fx - f2) + ... + Kn*(fx - fn) = 0.$$

Results of partial aggregation (labelled with I, E, C, respectively) and final aggregation (labelled with F) are shown in figure 3. It can be noted here that there is a significant suspect about the presence of partial discharges, owing to both casual and evidential inferences. Moreover, the initial dubious judgment about dielectric strength is reversed by inferences, even if it still remains somewhat controversial. These results have proved to be perfectly in agreement with expert's evaluations, during the validation of the ASTRA system.

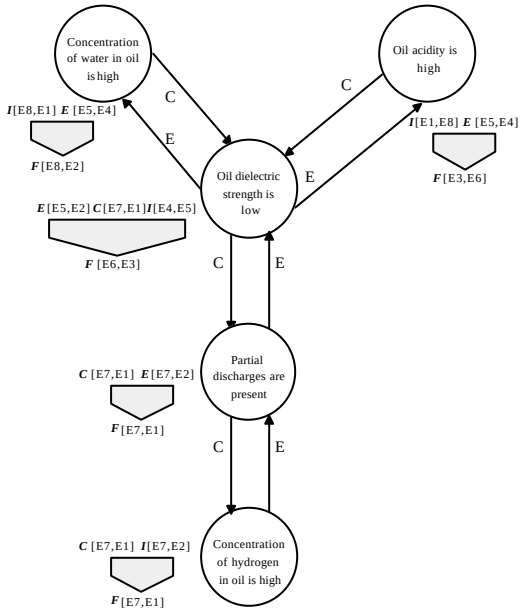

Figure 3 - The results of aggregation.

A detailed comparison of Q/C-E networks with other related approaches to causal reasoning is beyond the limits of the present paper. However we remark that their definition is grounded on a coherent set of cognitively plausible general considerations and they have shown a very satisfactory behavior in the context of a real application.

References

[1] P. Baroni, F. Cremonesi, G. Guida, S. Mussi, and S. Yakov, ASTRA: A knowledge-based system for state assessment, preventive diagnosis, and intervention planning of power transformers, *Proc. ISAP 94 Int. Conf. on Intelligent System Application to Power Systems*, Montpellier, F, 1994, 463-470**.**

[2] A. Bonarini, E. Cappelletti, and A. Corrao, Network-based management of subjective judgments: A proposal accepting cyclic dependencies, *IEEE Trans. on Systems, Man, and Cybernetics* SMC-22(5), 1992, 1008-1023.

[3] P. Bonissone, Summarizing and propagating uncertain information with triangular norms, *International Journal of Approximate Reasoning* 1, 1987, 71-101.
[4] P.P. Bonissone and K.S. Decker, Selecting uncertainty calculi and granularity: an experiment in trading-of precision and complexity, in L.Kanal and J.Lemmer (Eds.), *Uncertainty in Artificial Intelligence*, Elsevier, Amsterdam, NL, 1986, 217-247.
[5] J. de Kleer and J.S. Brown, A qualitative physics based on confluences, *Artificial Intelligence* 24, 1984, 7-83.
[6] J. de Kleer and J.S. Brown, Theories of causal ordering, *Artificial Intelligence* 29, 1986, 33-62.
[7] Y. Iwasaki and H.A. Simon, Causality in device behavior, *Artificial Intelligence* 29, 1986, 3-32.
[8] Y. Iwasaki and H.A. Simon, Theories of causal ordering: Reply to de Kleer and Brown, *Artificial Intelligence* 29, 1986, 63-67.
[9] D. Lewis, Causation, *Journal of Philosophy* 70, 1973, 556-567.
[10] R.S. Patil, P. Szolovits, and W.B. Schwartz, Causal understanding of patient illness in medical diagnosis, *Proceedings 7th IJCAI*, Vancouver, BC, 1981, 893-899.
[11] J. Pearl, Fusion, propagation and structuring in belief networks, *Artificial Intelligence* 29, 1986, 241-288.
[12] J. Pearl, Embracing causality in default reasoning, *Artificial Intelligence* 35, 1988, 259-271.
[13] Y. Shoham, Nonmonotonic reasoning and causation, *Cognitive Science* 14, 1990, 213-252.
[14] P. Suppes, *A Probabilistic Theory of Causation*, North Holland, Amsterdam, NL, 1970.
[15] D.S. Weld and J. de Kleer, Causal explanations of behavior, in D.S. Weld and J. de Kleer (Eds.), *Readings in Qualitative Reasoning about Physical Systems*, Morgan Kaufmann, San Mateo, CA, 1990, 611-616.