Granular Causality Speculations

Lawrence J. Mazlack

Applied Computational Intelligence Laboratory
University of Cincinnati
Cincinnati, Ohio 45221-0030

mazlack@uc.edu

***Abstract* - In many ways, causality deals with granular descriptions.. This is true for commonsense reasoning as well as for mathematical and scientific theory. At a very fine-grained level, the physical world itself may be made up out of granules. Our commonsense perception of causality is often granular. Knowledge of at least some causal effects is imprecise. In our commonsense world, it is unlikely that all possible factors can be known. Our commonsense understanding of the world deals with imprecision, uncertainty and imperfect knowledge. People recognize that a complex collection of elements causes a particular effect, even if the precise elements of the complex are unknown. They may not know what events are in the complex; or, what constraints and laws the complex is subject to. Sometimes, the details underlying an event can be known to a fine level of detail, sometimes not. The level of detail can reasonably be called the event's grain size. In general, commonsense reasoing is more successful in reasoning about a few large grain sized events than many fine grained events. The central question is: To what extent can we usefully vary the causal grain size?**

INTRODUCTION

In many ways, causality is granular. This is true for commonsense reasoning as well as for mathematical and scientific theory. At a very fine-grained level, the physical world itself may be granular. Our commonsense perception of causality is often large grained while the underlying causal structures may be somehow described in a more fine-grained manner.

Causal relationships exist in the commonsense world. "Positive" causal relationships are: *if* α *then* β or ($\alpha \rightarrow \beta$). For example: When an automobile driver fails to stop at a red light and there is an accident, it can be said that the failure to stop was the accident's cause. (Although, failing to stop at a red light is not a *certain* cause of an accident; sometimes no accident occurs.) Counterfactually, usually (but not always), stopping for a red light avoids an accident. (Sometimes, an automobile stopped at a red light is involved in an accident.) So, knowledge of at least some causal effects is imprecise for both positive and counterfactual descriptions. Perhaps, complete knowledge of all possible factors might lead to a crisp description of whether an effect will occur. However, in our commonsense world, it is unlikely that all possible factors can be known. It is also unlikely that it may be possible to fully know, with certainty, all of the elements involved.

We do things in the world by exploiting our commonsense *perceptions* of cause and effect. Manipulating perceptions has been explored [Zadeh, 1999] but is not the focus of this paper. The interest here is how perceptions affect commonsense causal reasoning, granularity, and the need for precision.

When trying to precisely reason about causality, we need complete knowledge of all of the relevant events and circumstances. In commonsense, every day reasoning, people use approaches that do not require complete knowledge. Often, approaches follow what is essentially a *satisficing* [Simon, 1955] paradigm. The use of non-optimal mechanisms does not necessarily result in ad hocism [Goodrich, 2000], for example, "Zadeh [1998] questions the feasibility (and wisdom) of seeking for optimality given limited resources. However, in resisting naive optimizing, Zadeh does not abandon the quest for justifiability, but instead resorts to modifications of conventional logic that are compatible with linguistic and fuzzy understanding of nature and consequences."

In commonsense reasoning, people recognize that a complex collection of elements can be involved causally in a particular effect, even if the precise elements of the complex are unknown. People may not know what events are in the complex; or, what constraints and laws the complex is subject to. Sometimes, the details underlying an event are known to a fine level of detail, sometimes not. The level of detail can reasonably be called the event's grain size.

In general, people are more successful in reasoning about a few large grain sized events than many fine grained elements that might make up a complex completely describing the large grained element. When using commonsense reasoning, people do not always even need to know the extent of the underling complexity. This is also true for situations not involving common sense reasoning; for example, when designing an electric circuit, designers are rarely concerned with the precise properties of the materials used; instead, they are concerned with the devices functional capabilities and take the device as a larger grained object. Complexes may be best handled on a black-box, large grained basis. That is, it can be recognized that a complex exists; but we do not necessarily need to deal with the deals internal to the complex.

A question concerning complexes is: To what extent can we increase the causal grain size and still have useful causal information. Conversely, can we start with a large grained causal event and then derive the finer grained structure? If we start with a large grained structure and resolve it, will our computational complexity burdens be reduced?

COMPLEXES

When events happen, there are usually other related events. The entire collection of events can be called a complex. The events can be called the elements of the complex.

When an automobile's ignition switch is turned on, it is natural to say that this causes the engine to start. But, it would not happen if a large system of other conditions were not in place. The wiring has to connect the switch to the starter and ignition system. The engine has to be in good working order. The switch has to be connected to the battery so electricity can flow through it. The battery has to be operational. There has to be available fuel; and so forth. Turning the ignition switch on is only one action in a complex of conditions required for the engine to start.

A "mechanism" [Simon, 1991] or a "causal complex" [Hobbs 2001, 2003] is a collection of events whose occurrence or non-occurrence results in a consequent event happening. Hobbs' causal complex is the *complete* set of events and conditions necessary for the causal effect (consequent) to occur. Hobbs suggests that using a causal complex does not require precise, complete knowledge of the complex. (Different workers use the terms "mechanism and "causal complex" differently; I am using them as these author's use them.)

An issue is how to distinguish between what is in a complex and what is not. Another issue is how to distinguish between the things that deserve to be called "causes" and those that do not. Hobbs suggests that a consideration of causal complexes can be divided into:

- Distinguishing what events are in a causal complex from those outside of it. [Lewis, 1973] [Oritz, 1999] [Simon, 1952, 1991] [Pearl, 2000]
- Within a causal complex, recognizing the events that should be identified as causes from those that are not. [Macke, 1993] [Shoham, 1990]

The causal complexes explicitly considered by Hobbs and Pearl have a required structure that may be overly restrictive for commonsense causal understanding, namely:

- If *all* of the events in the causal complex happen, then the effect will occur
- There is nothing in the causal complex that is irrelevant to the effect

These requirements are probably too precise and extensive to be realized in a commonsense world. Sometimes, only some of the events need to happen. For example, someone may be able to save more money if their taxes are lowered or if they earn more money. Either even may lead to greater savings. However, neither may result in increased savings if they also have to pay a large divorce settlement. So, if all of the events happen, the effect may happen. If some of the events happen, the effect may happen. In the commonsense world, we rarely whether all of the events are in a complex are necessary. For example, a man may want to attract the attention of a woman. He may do a large number of things (e.g., hair, clothes, etc.). If he does attract the woman, he may never know which things were relevant and which were not.

Sometimes, it is enough to know what happens at a large grained level; at other times it is necessary to know the fined grained result. For example, if *Bill believes that turning the ignition key of his automobile* <u>*causes*</u> *the automobile to start*, it is enough if Bill engages an automobile mechanic when his automobile does not start when he turns the key on. On the other hand, the automobile mechanic needs to know a finer grained view than does Bill.

Instead of being concerned with all of the fined grained detail, a better approach is to incorporate granulation using rough sets and/or fuzzy sets to soften the need for preciseness. And then accept impreciseness in the description. Each complex can be considered to be a granule. Larger complexes can be decomposed into smaller complexes. Thus, going from large grained to small grained.

Recognizing Causality Is Of Interest In Many Domains

Recognizing causality has is of interest in many areas. Of particular interest to this paper are areas where the analysis is non-experimental. The world is taken as it is and not subject to experimentation. Data mining is of concern to the computational sciences. An area not well known to people working in the computational sciences is economics.

Economics

Perhaps, the applied area that has the greatest history of attempting to deal with causality and non-observational data is economics. Econometrics is distinguished from statistics by econometrics interest in establishing causation [Hoover, 2003].

How and if causality can be recognized has been a significant area of discussion. Some of this discussion mirrors discussion that has gone on in the computational sciences. Hoover [2003] provides a good entry to the discussion of causality in economics.

The issue of causal ordering is often of importance to those modeling causality in data discovery. Granger [1969] defined causality depends on one-way, time ordered conception of causality. In contrast, Simon [1952, 1953] provides an analysis of causality that does not rely on time order. Some believe [Hausman, 1998, 1] that causal relations are mostly indicated by asymmetric relationships. An abbreviated version of the relationships that Hausman lists is:

- *Time-order:* Effects do not come before causes
- *Probabilistic Independence*
- *Agency* or *manipulability:* Causes can be used to manipulate their effects, but effects cannot be used to manipulate their causes. Effects of a common cause cannot be used to manipulate one another.
- *Counterfactual dependence:* Effects counterfactually depend on their causes, while causes do not counterfactually depend on their effects.
- *Overdetemination:* Effects over determine their causes, while causes rarely over determine their effects
- *Invariance:* Dependent variables in an equation are effects of the independent variables.
- *Screening-off:* Causes screen off their effects
- *Robustness:* The relationship between cause and effect is invariant with respect to the frequency of the cause.
- *Connection dependence:* If the connection between cause and effect is broken, only the effect would be affected.

Friedman [1949] argues that any cause that we isolate is never the whole cause and that every direct cause itself has its own direct causes, so that networks of causation spread synchronically across the economy and diachronically back into the mists of time. If this is true, granules must necessarily be imprecise as separation trough truncation from a network would be required.

In Data Mining

There are several different data mining products. The most common are *conditional rules* or *association rules*. Conditional rules are most often drawn from induced trees while association rules are most often learned from tabular data.

- *Conditional rule:*

```
IF Age is old
   THEN vote frequency is: often
      with {belief = high}
```

• *Association rule:*

*Customers who buy **Strawberries***
*also tend to buy **Whipped Cream***
with {confidence = 0.8}
in {support = 0.15}

At first glance, these structures seem to imply a causal or cause-effect relationship. That is: *A voter's old age* <u>*causes*</u> *them to vote;* or: *A customer's purchase of strawberries* <u>*causes*</u> *the customer to also buy whipped cream.* In fact, when typically developed, rules do not *necessarily* describe causality. Also, the strength of any causal dependency may be very different from that of a possibly related association value. All that can be said is that associations describe the strength of joint co-occurrences. The confidence measure is simply an estimate of the conditional probability. Support indicates how often the joint occurrence happens. The joint occurrence count is symmetric; that is, it does not matter what we count first. Recognizing causality in mined data is very important to the utility of the results

Sometimes, the association might be causal; for example, if *someone eats salty peanuts* and *then drinks beer,* there is likely a causal relationship. On the other hand, if a *rooster grows* and *then the sun rises*, there is probably not a causal relationship..

Association rules are sometimes used is to aid in making retail marketing decisions. However, the use of simple association rules may lead to errors. Errors might occur; either if causality is recognized where there is no causality; or if the direction of the causal relationship is wrong. For example:

"A study of past customers shows that 94% are sick"

is it the rule:

"Our customers are sick, so they buy from us"

or is it:

"If people use our products, they are likely to become sick"?

It is not enough to know that people both buy are products and are sick; what is needed is knowledge of what causes what.

If causality is not recognized, the naive application of association rules can result in bad decisions [Silverstein, 1998*a*]. This can be seen in the more extensive example from Mazlack [2003]:

Example: At a particular store, a customer buys:

- *hamburger 33%* of the time
- *hot dogs* 33% of the time
- both *hamburger* and *hot dogs* 33% of the time
- *sauerkraut*[*] only if *hot dogs* are also purchased

This would produce the transaction matrix:

	hamburger	hot dog	sauerkraut
t_1	1	1	1
t_2	1	0	0
t_3	0	1	1

This would lead to the associations:

- (hamburger, hot dog) = 0.5
- (hamburger, sauerkraut) = 0.5
- (hot dog, sauerkraut) = 1.0

If the merchant:

- Reduced price of hamburger (as a sale item)
- Raised price of sauerkraut to compensate (as the rule *hamburger fi sauerkraut* has a high confidence.
- The offset pricing compensation would not work, as the sales of sauerkraut would not increase with the sales of hamburger. Most likely, the sales of hot dogs (and consequently, sauerkraut) would likely decrease as buyers would substitute hamburger for hot dogs.

GRANULAR SPACE-TIME

Space-time impacts the idea that causality can be described by a time separation (timeline) between cause and effect. The malleability of Einstein's space-time that has the effect that what is "now" and "later" is local to a particular observer; another observer may have contradictory views.

One of the key principles of space-time is that of *background independence*. This principle says that the geometry of space-time is not fixed. Instead, the geometry is an evolving, dynamical quantity. A closely related principle is *diffeomorphism invariance*. This principle implies that unlike theories prior to general relativity, one is free to choose any set of coordinates to map space-time and express the equations. A point in space-time is defined only by what physically happens at it, not by its location according to some special set of coordinates (no coordinates are special).

Modern physics has developed a theory that entails that space and time are granular [Smolin, 2004]. This is an extension of quantum theory. Quantum mechanics require that certain quantities, such as the energy of an atom, can only come in specific, discrete units. Over the last few years, theory has evolved concerning quantum gravity and quantum space-time. This area of endeavor is sometimes called *loop quantum gravity.* (The term *loop* arises from how some computations in the theory involve small loops marked out in space-time.) The work is concerned with quantum theory of the structure of space-time at the smallest size scales.

What concerns us in this paper is that there are apparently limits on fine grain size. These limits apply to areas, volumes, and time [Smolin, 2004]. There is a non-zero minimum area (about one square Planck length, or 10^{-66} square centimeter) and a discrete series of allowed quantum areas. Similarly, there is a non-zero absolute minimum volume (about one cubic Planck length, or 10^{-99} cubic centimeter) and it restricts the set of larger volumes to a discrete series of numbers. Time is also discrete; it comes in "clicks" of 10^{-43} seconds (approximately the Planck time). Time does not exist between the clicks; there is no "in between," in the same way that there is no water between adjacent molecules of water.

This information should influence how we think about causality. If the universe is fundamentally granular, causal descriptions need to somehow deal with granularity. How to do this is unclear. Rough sets might be the best way of handling the granularity of causal complexes. Similarly, they seem to be a good tool to initially approach the granularity of space-time.

Recognizing many things with absolute certainty is problematic. As this is the case, our causal understanding is based on a foundation of inherent uncertainty and incompleteness. Consequently, causal reasoning models must accommodate inherent ambiguity. Mazlack [2003] lists:

- *Quantum Physics:* In particular, Heisenberg's uncertainty principle

[*] Sauerkraut is a form of pickled cabbage. Some people greatly enjoy using sauerkraut as a garnish with sausages. However, it is rarely consumed as a garnish with hamburger. For more about sauerkraut, see: *http://www.sauerkraut.com/*

- Knowledge of the world might never be complete because we, as observers, are integral parts of what we observe
- *Gödel's Theorem:* Showed that in any logical formulation of arithmetic that there would always be statements whose validity was indeterminate. This strongly suggests that there will always be inherently unpredictable aspects of the future.
- *Turing Halting Problem:* Turning showed that any problem solvable by a step-by-step procedure could be solved using a Turing machine. However, there are many routines where you cannot ascertain if the program will take a finite, or an infinite number of steps. Thus, there is a curtain between what can and cannot be known mathematically.
- *Chaos Theory:* Chaotic systems appear to be deterministic; but are computationally irreducible. If nature is chaotic at its core, it might be fully deterministic, yet wholly unpredictable [Halpern, 2000, 139].
- *Space-Time:* The malleability of Einstein's space time that has the effect that what is "now" and "later" is local to a particular observer; another observer may have contradictory views.
- *Arithmetic Indeterminism:* Arithmetic itself has random aspects that introduce uncertainty as to whether equations may be solvable. Chatin [1987, 1990] discovered that Diophantine equations may or may not have solutions, depending on the parameters chosen to form them. Whether a parameter leads to a solvable equation appears to be random.

It may well be that a precise and complete knowledge of causal events is not possible or at least uncertain. On the other hand, we have a commonsense belief that causal effects exist in the real world. If we can develop models tolerant of imprecision, it would be useful. Also, to some degree, the degree of importance that some of these items has decreases as grain size increases.

Causality Recognition

Different aspects of causality have been examined. The idea of "positive" causation ($\alpha \rightarrow \beta$) is at the core of common sense reasoning. Negation or counter factuals ($\neg\alpha \rightarrow \neg\beta$) also have a place; although it may result in errors in reasoning. For example, the rule *If a person smokes, they will get cancer* cannot be simply negated to *If a person <u>does not</u> smoke, they will <u>not</u> get cancer.* (Effects can be *overdetermined;* that is: more than one item can cause an effect. In this case, people who do not smoke may also get cancer.)

Other ideas that are sometimes involved in causal reasoning are *causal uncorrelatedness* [Shafer, 1999] where if two variables have no common cause they are causally uncorrelated. This occurs if there are no single events that cause them to both change. Similarly, causal independence occurs when speaking about probabilities:

$$P_S(X=x \; \& \; Y=y) = P_S(X=x)P_S(Y=y) \text{ for all } S$$

Similarly, Dawid [1999] focuses on the negative; i.e., when α does not affect β. Dawid speaks in terms of *unresponsiveness* and *insensitivity*. β is unresponsive to α if whatever the value of α might be set to, β will be unchanged. β is insensitive to α if whatever the value α may be set, the uncertainty about β will be unaffected. Along the same vein, Shoham [1990, 1991] distinguishes between *causing, enabling,* and *preventing.* The enabling factor is often considered to be a causal factor by others. However, Shoham distinguished between background (enabling) conditions and foreground conditions. The background (enabling) conditions are inferred by default. For example, "If information is present that the key was turned and nothing is mentioned about the stated about the state of the battery, then it is inferred that the motor will start, because the battery is assumed, by default to be alive. Given this distinction, causing is taken to refer to the foreground conditions where enabling and preventing refer to the background conditions (in this example, turning the key causes the motor to start, the live battery enables it, the dead battery prevents it)"

Statistics is the traditional tool used to discover causality when handling experimental (observational) data. However, the data of greatest interest in the computational sciences is non-observational. In this domain, traditional statistical methods are either not useful an/or are too computationally complex. This having been said, some work has been done in data mining using chi-squared testing to reduce the search space [Silverstein, 1998].

Various Bayesian based methods have been suggested to recognize causality. Probably the best known is the class of methods based on Directed Acyclic Graphs (DAGs). The most fully developed approach is Pearl [2000]. Silverstein [1998] followed a similar approach. Pearl [1991] and Spirtes [1993] make the claim that it is possible to infer causal relationships between two variables from associations found in observational (nonexperimental) data without substantial domain knowledge. Spirtes claim that directed acyclic graphs can be used if (a) the sample size is large and (b) the distribution of random values is faithful to the causal graph. Robins [1999] argues that their argument is incorrect. Lastly, Scheines [1994] only claims that in some situations will it be possible to determine causality.

Without going deeply into this debate, from the commonsense causal reasoning view, the directed graph methods all have similar liabilities, specifically:

- *Discrete or continuous data must be reduced to Boolean values.*

  Objection: This is an early technique that was and is used in data mining when analyzing market basket data. However, it is essentially flawed. Quantities do matter; some data co-occurrences are conditioned on there being a sufficiency of a co-occurring attribute. Also, some relationships may be non-linear based on quantity.

  ***Example**: Situation: Customers frequently buy either wine or beer for themselves in varying amounts. However, when buying for a party, they often purchase both beer and wine and they usually purchase in larger quantities.*

Actual basket:		Binary basket:	
Beer	Wine	Beer	Wine
6	0	1	0
0	1	0	1
12	0	1	0
0	3	0	1
24	4	1	1
24	5	1	1
48	2	1	1

Missed rule: When at least 24 beers purchased, wine also purchased;
otherwise, there is no relationship between beer and wine

Naively constructing an association rule would find a rule that misleadingly represents the situation; i.e.,

misleading rule: When beer is purchased, wine is also purchased
{confidence = 0.6} {support = 0.43}

This rule is misleading. To the naive, it implies that purchase probabilities are uniform, in fact they are not. Under one set of conditions, *beer* and *wine* are <u>never</u> purchased together under one set of conditions; and, under another set of conditions always purchased together.

- *There is no missing data.*

  Objection: There is almost always missing data of some sort. Data collection is rarely fully representative and complete. Incremental data is often acquired that is at variance with previously acquired data. What is needed is a methodology that is not brittle in the face of incompleteness.

- *Causal relationships are not cyclic, either directly or indirectly (through another attribute).*

  Objection: This is at variance with our commonsense understanding of the world.

  For example: I tell Jane that I love her. Then, she tells me that she loves me. Then, I tell Jane that I love her more than before; then, she … and so forth and so forth. Clearly, the cyclic reinforcement would be substantial.

  Another example is shown below:

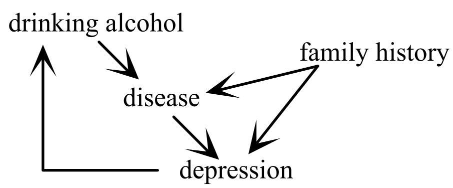

Figure 1. Cyclic relationship.

Simon [1991] and Shoham [1991] identify cases where causality is simultaneous.

MARKOFF MODELS

One possible way of describing the causal relationships is with Markoff (Markov) Models. They treat a system as a series of states with specific, constant rate transitions between them. At all times, the system is in exactly one state. (Transitions are considered to be instantaneous.) The only information available is the current state, the allowed transitions, and the probability of these transitions. Such a system is referred to as memoryless, as is said to posses the *Markoff property*. This means that the system is totally characterized by its current state. None of the past states or transitions have any effect on the transitions out of the current state.

Quantitatively describing the relationships between the nodes can be complex. One possibility is an extension of the random Markoff model, shown in Figure 2. The state value is 1/0 as an event either happens or does not.

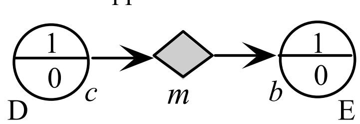

Figure 2. Random Markoff model: c = P(D), m = the possibility/probability that when D is present, the causal mechanism brings about E, b = the probability that some other (unspecified) causal mechanism brings about E.

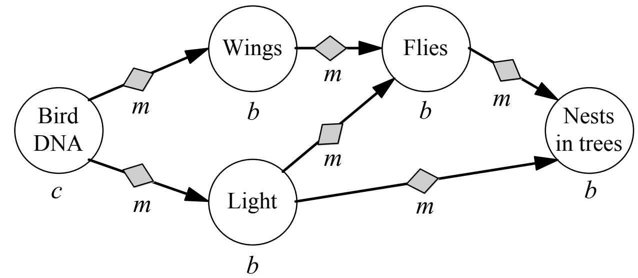

Figure 3. Markoff model applied to a "bird" example [Rehder, 2002]

Perhaps, fuzzy sets and fuzzy Markoff algorithms in particular might be a good way of approaching causality represented by Markoff models. Fuzzy sets are useful for dealing with any situation where the exact value of a variable is unknown. Instead of a guess of the value of the variable (which can easily be wrong), or a distribution of its possible values (which is usually unknown, so this problem reduces to a guess), fuzzy logic deals with the *possibility* of the variable taking on a set of values. In this way, it assumes less, and shows explicitly both what is and is not known.

Increasing Grain Size

Depending on the goal of the work, it may be more useful to work on a larger grain size. For example, if instead of a complex in the form of *Figure 3*, perhaps a diagram similar to *Figure 4*, would be sufficient. This form of representation is widely used in automata design; its extension to causal reasoning and imprecise reasoning would seem to be reasonable.

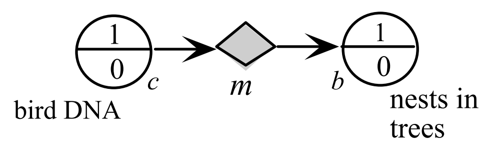

Figure 4. Greater grain size (higher level) abstraction of Figure 3.

The greater grain size would be more useful in commonsense reasoning. It also might be more computationally tractable in the discouvery process. Rough sets might be the best way of handling the granularity of causal complexes. Similarly, they seem to be a good tool to initially approach the granularity of space-time.

EPILOGUE

Causal reasoning occupies a central position in human reasoning. In many ways, causality deals with granular descriptions. This is true for commonsense reasoning as well as for mathematical and scientific theory.

At a very fine-grained level, the physical world itself may be composed of granules.

Causal relationships exist in the commonsense world. Our commonsense perception of causality is often granular. Knowledge of at least some causal effects is imprecise. Perhaps, complete knowledge of all possible factors might lead to a crisp description of whether an effect will occur. However, in our commonsense world, it is unlikely that all possible factors can be known. In commonsense, every day reasoning, we use approaches that do not require complete knowledge.

Our commonsense understanding of the world tells us that we have to deal with imprecision, uncertainty and imperfect knowledge. Need an algorithmic way of handling imprecision if we are to computationally handle causality. A difficulty is striking a good balance between precise formalism and commonsense imprecise reality People recognize that a complex collection of elements causes a particular effect, even if the precise elements of the complex are unknown. They may not know what events are in the complex; or, what constraints and laws the complex is subject to. Sometimes, the details underlying an event are known to a fine level of detail, sometimes not. The level of detail can reasonably be called the event's grain size. In general, people are more successful in reasoning about a few large grain sized events than many fine grained events. The central question is: To what extent can the causal grain size be varied and still have useful causal information os the question.

References

G. Chatin [1987] **Algorithmic Information Theory**, Cambridge University Press, Cambridge, United Kingdom

G. Chatin [1990] "A Random Walk In Arithmetic," New Scientist 125, n 1709 (March, 1990), 44-66

A. Cheok, S. Chevalier, M. Kaynak, K. Sengupta, K. Chung [2002] "Use of a Novel Generalized Fuzzy Hidden Markov Model for Speech Recognition," *Proceedings. IEEE International Conference on Fuzzy Systems,* December 2-5, 2002 Melbourne, Australia

A. Dawid [1999] "Who Needs Counterfactuals" in **Causal Models and Intelligent Data Management** (ed) A. Gammerman) Springer-Verlag, Berlin

M. Friedman [1949] "The Marshallian Demand Curve," *Journal of Political Economy,* v 57, 463-495

M. Goodrich, W. Stirling, E. Boer [2000] "Satisficing Revisited," *Minds and Machines,* v 10, 79-109

C. Granger [1969] "Investigating Causal Relations By Econometric Models And Cross-Spectral Methods," *Econometrica,* v 37, 424-438

P. Halpern [2000] **The Pursuit Of Destiny**, Perseus, Cambridge, Massachusetts

D. Hausman [1998] **Causal Asymmetries**, Cambridge University Press, Cambridge, United Kingdom

J. Hobbs [2001] "Causality," *Proceedings, Common Sense 2001, Fifth Symposium on Logical Formalizations of Commonsense Reasoning*, New York University, New York, May, 145-155

J. Hobbs [2003] "Causality And Modality: The Case Of 'Would'," to appear in *Journal of Semantics*

K. Hoover [2003] "Lost Causes," *HES Conference,* Presidential Address, Durham, North Carolina

D. Lewis [1973] **Counterfactuals**, Harvard University Press, Cambridge University Press

L. Mazlack [2003] "Causality Recognition For Data Mining In An Inherently Ill Defined World," 2003 BISC FLINT-CIBI International Joint Workshop On Soft Computing For Internet And Bioinformatics, December, 2003

C. Oritz [1999] "Explanatory Update Theory: Applications Of Counterfactual Reasoning To Causation," *Artificial Intelligence,* v 108, n 1-2, 125-178

J. Pearl, T. Verma [1991] ""A Theory Of Inferred Causation," *Principles Of Knowledge Representation And Reasoning: Proceedings Of The Second International Conference,* Morgan Kaufmann, 441-452

J. Pearl [2000] **Causality**, Cambridge University Press, New York, NY

B. Rehder [2002] "A Causal-Model Theory of the Conceptual Representation of Categories," FLAIRS 2002 Workshop on Causality, Pensacola, May

R. Robins, L. Wasserman [1999], "On The Impossibility Of Inferring Causation From Association Without Background Knowledge," in (eds) C. Glymour, G. Cooper, **Computation, Causation, and Discovery** AAAI Press/MIT Press, Menlo Park, 305-321

R. Scheines, P. Spirtes, C. Glymour, C. Meek [1994] **Tetrad II: Tools For Causal Modeling**, Lawrence Erlbaum, Hillsdale, NJ

G. Shafer [1999] "Causal Conjecture," in **Causal Models and Intelligent Data Management** (ed) A. Gammerman) Springer-Verlag, Berlin

Y. Shoham [1990] "Nonmonotonic Reasoning And Causation," *Cognitive Science,* v14, 213-252

Y. Shoham [1991] "Remarks On Simon's Comments," *Cognitive Science*, v15, 301-303

C. Silverstein, S. Brin, R. Motwani [1998*a*] "Beyond Market Baskets: Generalizing Association Rules To Dependence Rules," Data Mining And Knowledge Discovery, v 2, 39-68

C. Silverstein, S. Brin, R. Motwani, J. Ullman [1998*b*] "Scalable techniques for mining causal structures," *Proceedings 1998 VLDB Conference,* New York, NY, August 1998, 594-605

H. Simon [1952] "On The Definition Of The Causal Relation," *The Journal Of Philosophy,* v 49, 517-528. Reprinted in Herbert A. Simon, **Models Of Man**, John Wiley, New York, 1957

H. Simon [1953] "Causal ordering And Identifiability," Reprinted in Herbert A. Simon, **Models Of Man**, John Wiley, New York, 1957

H. Simon [1955] "A Behavioral Model Of Rational Choice," *Quarterly Journal Economics,* v 59, 99-118

H. Simon [1991] "Nonmonotonic Reasoning And Causation: Comment," *Cognitive Science,* v 15, 293-300

L. Smolin [2004] "Atoms of Space And Time," *Scientific American,* January, 2004, 66-75

P. Spirtes, C. Glymour, R. Scheines [1993] **Causation, Prediction, and Search,** Springer-Verlag, New York

L. Zadeh [1998] "Maximizing Sets And Fuzzy Markoff Algorithms," *IEEE Transactions on Systems, Man, and Cybernetics - Part C: Applications and Reviews,* v 28, 9-15

L. Zadeh [1998] "From computing With Numbers To Computing With Words - From Manipulation Of Measurements To Manipulation Of Perceptions," *IEEE Transactions on Circuits and Systems,* v 45, n 1, 108-119