Autonomous Synthesis of Fuzzy Cognitive Maps from Observational Data: Preliminaries[1,2]

Daryl Fletcher[a,c]
Duong Nguyen[b,c]
Krzysztof Cios[c] krys.cios@cudenver.edu
[a]Science Applications International Corporation (SAIC), daryl.p.fletcher@saic.com
[b]Raytheon IIS/Space Systems, dnguyen1@raytheon.com
[c]Department of Computer Science, University of Colorado-Denver

Abstract—We review current research in fuzzy cognitive map (FCM) parameter estimation techniques and present results for a method based on constrained optimization as motivation for a discussion of the core issue of *causal discovery* in intelligent systems. While FCM parameter estimation techniques show utility for automatically determining causal strength from state observations, their success relies heavily on background knowledge about the causal structure of the system, acquired from interviews with domain experts. In this paper we identify causal discovery as a fundamental task for automated FCM synthesis and present a survey of related research in automatic causal discovery that provides a basis for future work in the area of autonomous FCM synthesis.

TABLE OF CONTENTS

1. INTRODUCTION

There are a number of ways an intelligent agent can reason about its environment and all have strengths and weaknesses under certain circumstances. If an agent's objective is to be descriptive about the input-output relationships of its world, it can build a neural network model or employ any one of a variety of equation discovery algorithms to build a purely quantitative model of its environment. For reasoning about cause-effect relationships, a Bayesian network (BN) or structural equation model might be appropriate, but these approaches are limited in that the former does not allow for feedback between nodes and the latter is used to confirm a hypothesis about an existing causal structure rather than learn one from observational data. Fuzzy cognitive maps (FCMs) are a relatively young methodology for modeling the cause-effect relationships of complex (nonlinear) systems where the causal structure of the system is represented as a signed, directed *cyclic* graph with feedback. In an FCM, the nodes (vertices) represent concepts and the sign, direction and magnitude of the edges between vertices characterize the causal relationship between nodes. Under certain circumstances, an FCM or a hierarchy of FCMs can provide an effective means for an agent to reason about the causal behavior of its environment, but to be most useful in an intelligent system, an agent would need to be able to synthesize an FCM with little if any preprogrammed knowledge about its environment. To synthesize an FCM, an agent must be capable of identifying concepts, identifying cause-effect relationships (*causal discovery*) and learning *causal strength* from observational data.

Fuzzy cognitive maps are similar to neural networks (NNs) and both can be used to model complex systems, with subtle differences. Primarily, they differ in the way they are constructed. FCMs are usually constructed manually using extensive background knowledge about the system being

[2] IEEEAC paper #1451, Version 1, Updated August 26, 2004

modeled while NNs are created from observational data using supervised or unsupervised learning with little if any background knowledge about the system. A second important distinction is that FCMs are a "grey box" modeling technique where traditionally the model structure embodies cause-effect relationships between variables and is somewhat explanatory in that the sign, magnitude and direction of edge weights convey some knowledge about the causal structure of the system. In an FCM, there is an edge between nodes a and b if and only if there is a direct cause-effect relationship between them and the direction, sign and magnitude of that edge reflect the properties of that relationship. Generally speaking, NNs provide a "black box" model of the system where the relationships between observed variables are not clearly reflected in the network topology nor are the sign and magnitude of the neural connections usually thought of as representing causal structure.

The *Universal Approximation Theorem* for neural networks ensures that under certain conditions, multilayer feedforward networks with a single hidden layer and an appropriately smooth hidden layer activation function are capable of approximating an arbitrary function and its derivatives to an arbitrary degree of accuracy[30]. A multilayer feedforward network with a sufficient number of hidden units can model the *behavior* of a nonlinear system but not necessarily reflect the true nature of the relationships between variables in its topology. While neural networks and FCMs both exhibit *model-data consistency*, applying background knowledge during FCM construction results in a *model-reality consistency* not usually reflected in the topology of a neural network.

There is general consensus in the research community that if FCMs are to prove useful in real-world, large-scale application domains, FCMs need to be automatically constructed from observational data with minimal input from domain experts. This notion of "autonomous synthesis" is especially important for an intelligent agent. Two main tasks in any system modeling effort are parameter estimation and structure identification[7]. While automatic parameter estimation is an important aspect of autonomous FCM synthesis, the FCM literature scarcely addresses using automatic causal discovery to assist in the structure identification of FCMs for large and small-scale systems.

In this document, we refer to the vertices of an FCM, all of which represent causal concepts, as variables and we refer to the values of the edges that connect the vertices as parameters, which represent causal strength. By convention, FCM parameters take on values in either the continuous interval [-1, 1], or from the set {-1, 0, 1}, while FCM variables take on values in the continuous interval [0, 1] or from the set {0, 1}. The causal structure of an FCM is the set of variables and directed edges that comprise the FCM without assigned values for the parameters and variables.

Current parameter estimation techniques for automatic FCM synthesis rely heavily on the existence of a pre-defined causal structure for the system defined by domain experts. A pre-defined causal structure serves to constrain the search space of possible solutions for parameters that will lead the FCM to the desired steady state. The unconstrained search space of possible causal structures is enormous. If we assume a trivalent FCM (edge values are in the set {-1, 0, 1}) with no self-feedback, there are 3^{n^2-n} directed cyclic graphs in the unconstrained search space. For an FCM with only 5 nodes (a 5x5 connection matrix) there are 3,486,784,401 possible causal structures! However, an intelligent agent cannot rely on preprogrammed knowledge about the causal structure to constrain the search space.

A widely discussed tenet of cause-effect analysis is:

$$\text{Causality} \Rightarrow \text{Correlation} \tag{1}$$

and the contra positive

$$\sim\text{Correlation} \Rightarrow \sim\text{Causality} \tag{2}$$

Note that the converse of (1), Correlation ⇒ Causality, is *not true*. Cause-effect relationships are inherently temporal in nature (*a* can be a possible cause of *b* only if *a* occurs before *b*) and given a "long enough" time series of state observations, it is easier to infer a *lack* of causality from a *lack* of correlation than to prove a causal relationship between variables, particularly if we are not allowed to experiment with the system. If an intelligent agent can infer a lack of correlation between concepts with a high degree of certainty, it can use this knowledge to constrain the search space by eliminating causal structures that exhibit non-physical relationships.

The task of inferring causal influences that are not spurious covariances from observational data is difficult, but not impossible. A large body of research addresses causal discovery methods for directed *acyclic* graphs (DAGs)[8 – 20], particularly Bayesian networks, and to a limited extent for directed cyclic graphs[21-24], though not specifically for FCMs. Structural equation modeling (SEM) is not a causal discovery technique but is rather *confirmatory* in that it attempts to show whether or not the causal assumptions embedded in a model match the observed data[1]. While these methods have shown promise for certain domains, the assumption of a *Markov condition* in DAG causal discovery theory for the most part precludes application of the theory, without careful modification, to systems with feedback and a SEM approach generally relies on an existing hypothesis of system structure, much like existing FCM parameter estimation techniques. However, that is not to say that these areas of research cannot yield insight into the problem of autonomous FCM synthesis.

When an FCM represents a fuzzy *causal* map, an intelligent agent can attack the problem of FCM structure synthesis from two angles: eliminating connections between concepts that are uncorrelated and discovering causal relationships from data. In section 2 of this paper we present an example of FCM parameter estimation from observational data based on a traditional constrained optimization approach. We use our results to motivate the discussion of automatic causal discovery for FCM structure identification presented in section 3.

2. FCM Parameter Estimation based on Constrained Optimization

The most basic FCM "learning" technique of adding weighted, augmented connection matrices produced manually by domain experts is described by Kosko[2] and combines both the structure identification and parameter estimation aspects of model construction. In later work Dickerson and Kosko[3] used observational data and Differential Hebbian Learning (DHL) to adapt FCM connection matrices by *correlating changes* in concepts as the system evolves. It is reported[4] that this method is good at generating spurious causal connections between concepts, an undesirable characteristic for a knowledge representation structure that is basically explanatory in nature, such as an FCM. A. V. Heurga[4] presented a Balanced Differential Learning algorithm that addressed the perceived shortcomings of Kosko and Dickerson's DHL approach by taking into account more than one concept to calculate the weights for the cause-effect links between variables. From a theoretical standpoint, the DHL algorithms are flawed in that they implicitly assume an underlying causal structure in the data when in fact one might not exist. One could supply the DHL algorithms with a data set generated by a dynamical system where the changes over time are correlated, but without a single true cause-effect relationship between the variables, and the result would be a connection (causal strength) matrix with non-zero entries.

More recently, Koulouriotis et. al.[5] presented an FCM learning approach based on Evolutionary Strategies (ES) while Parsopoulos et. al.[6] applied Particle Swarm Optimization (PSO) for adjusting the parameters of a control system FCM based on observational data. Both of these approaches showed promising results on a limited number of test cases, however, their success relies on knowing much about the causal structure beforehand, i.e. the number and location of the "zero" entries in the FCM connection matrix. In the ES approach the algorithm starts with a population of individuals that are vectors with *n* elements with *n* equal to the number of cause-effect relationships to be estimated. The PSO algorithm was applied to update only the values of the non-zero entries in the FCM connection matrix.

To illuminate the importance of starting with a "good" causal structure where the zero entries in the connection matrix are known and motivate discussion of the role of structure identification, we present a simple FCM parameter estimation technique based on constrained optimization.

FCM Parameter Estimation using Matlab®

Given a data set generated by a known FCM, our auto-synthesis algorithm should be able to infer the FCM that generated the data. While this is certainly not sufficient to prove correctness over a general class of problems, it serves as a good starting point and illustrates the core issues of creating an FCM strictly from observational data. Consider an FCM with 7 concepts represented by the following connection matrix:

$$W = \begin{pmatrix} 0 & 0 & 0.6 & 0.9 & 0 & 0 & 0 \\ 0.1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.7 & 0 & 0 & 0.9 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 0 & -0.9 & -0.9 \\ -0.3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.8 & -0.9 \end{pmatrix}$$

Figure 1 - FCM connection matrix for 7 concepts

The matrix shown in Figure 1 is the connection matrix for the fuzzyfied Public Health Issues FCM first appearing in Tsadiras and Margaritis[26]. Using an arbitrary initial concept vector $\underline{\mathbf{C}}\mathbf{o}$ = [0.4 0.1 0.39 0.6 0.7 0.66 0.23] and sigmoid activation function, $1/(1 + e^{-kx})$ with k = 1, we evolved the FCM represented by the connection matrix in Figure 1 to produce the following "observational" data set for the dynamical system:

Step	C1	C2	C3	C4	C5	C6	C7
1	0.5528	0.5922	0.6525	0.7231	0.7410	0.5533	0.5349
2	0.6097	0.7406	0.7279	0.7722	0.7905	0.5780	0.6269
3	0.6249	0.7773	0.7491	0.7893	0.8093	0.5910	0.6480
4	0.6284	0.7861	0.7547	0.7944	0.8151	0.5941	0.6525
5	0.6292	0.7883	0.7562	0.7958	0.8167	0.5945	0.6534
6	0.6294	0.7888	0.7565	0.7961	0.8172	0.5944	0.6535
7	0.6295	0.7889	0.7566	0.7962	0.8173	0.5943	0.6535
8	0.6295	0.7889	0.7566	0.7962	0.8173	0.5942	0.6535
9	0.6295	0.7889	0.7566	0.7962	0.8173	0.5942	0.6535

Figure 2 - "Observational" data set generated by a known FCM

where the columns represent concept values as they evolve over time. The FCM converges to a fixed point,

$$\underline{\mathbf{C}}_{\text{final}} = [0.6295\ 0.7889\ 0.7566\ 0.7962\ 0.8173\ 0.5942\ 0.6535]$$

after only 8 iterations. This simple step reveals the first issue in our discussion of automatic FCM synthesis from observational data, i.e. that the data can govern the activation function selected for the synthesized FCM. Two common FCM activation functions are the sigmoid function mentioned above, which maps its input to the interval [0, 1] and a simple step-type activation function that maps its input to the set {0, 1}. Another alternative is to use the hyperbolic tangent function for a mapping to the interval [-1, 1]. If our observational data is binary, a step-type activation function would be appropriate while if the data are in the interval [0, 1] or [-1, 1] (or could be transformed as such), a sigmoid or hyperbolic tangent function should be used.

We use the data set shown in Figure 2, generated by the known dynamical system represented by the connection matrix in Figure 1 as the "observational" data from which an intelligent agent will synthesize an FCM. Note that we have a relatively short "time series" to work with. This is by design, since for an intelligent agent the mechanism by which measurements are obtained may very well be a scarce resource. While a large-scale system may embody thousands of parameters the temporal span of the observational data may be limited due to resource constraints. How much data an intelligent agent needs to synthesize an FCM is a question posed for future research.

An FCM represents a dynamical system where the state transitions are governed by:

$$\mathbf{x}(k+1) = \boldsymbol{f}(\underline{\mathbf{x}}(k) * \mathrm{W} + \underline{\mathbf{x}}(k)) \tag{3}$$

where $f(\beta)$ is an activation function, W is the time-invariant connection matrix representing causal strength between the concepts and the right-most $\underline{\mathbf{x}}(k)$ term in the argument to f is a bias. Showing a couple of iterations

$\underline{\mathbf{x}}(1) = \boldsymbol{f}(\underline{\mathbf{x}}^{\mathrm{T}}(0)*\mathrm{W} + \underline{\mathbf{x}}(0))$

$\underline{\mathbf{x}}(2)=\boldsymbol{f}(\underline{\mathbf{x}}(1)*\mathrm{W}+\underline{\mathbf{x}}(1)) = \boldsymbol{f}(\boldsymbol{f}(\underline{\mathbf{x}}^{\mathrm{T}}(0)*\mathrm{W}+\underline{\mathbf{x}}(0))*\mathrm{W}+\boldsymbol{f}(\underline{\mathbf{x}}^{\mathrm{T}}(0)*\mathrm{W}+ \underline{\mathbf{x}}(0)))$

Figure 3 - First two iterates of FCM evolution

it becomes clear that the trajectory $\underline{\mathbf{X}} = \{\underline{\mathbf{x}}_1, \underline{\mathbf{x}}_2, \ldots, \underline{\mathbf{x}}_K\}$ is entirely determined by the starting point $\underline{\mathbf{x}}(0)$.

Since the values of the variables in the observational data are continuous in [0, 1], it is safe to assume a sigmoid activation function for the purpose of synthesizing the FCM. Some algebra transforms the equation for the first iterate shown in Figure 3 into the form

$$\mathbf{x}(0)^{\mathrm{T}} * \mathrm{W} = \boldsymbol{f}^{-1}(\underline{\mathbf{x}}(1)) - \underline{\mathbf{x}}(0) \tag{4}$$

where

$$\mathrm{f}^{-1}(\mathrm{y}) = -\ln((1/\mathrm{y}) - 1) \tag{5}$$

The left-hand side of equation (4) can be manipulated to yield the familiar matrix equation form $\mathbf{A}*\underline{\mathbf{y}} = \underline{\mathbf{b}}$ where $\underline{\mathbf{y}}$ is an n^2 vector of entries in W and n is the number of concepts. This is an underdetermined system of n equations in n^2 unknowns so we must constrain the system in order to find a unique solution.

Without background knowledge about the structure of the system, we form a set of inequality constraints by assuming that concepts have no self-feedback (we're simply assuming that a variable doesn't influence itself) and that the lower and upper bounds of the edge weights are –1 and +1, respectively. We also introduce an equality constraint based on the first iterate of the dynamical system we are trying to model.

Our objective function to minimize is the sum of the error norms of the trajectories calculated with the parameter estimates and the true "observed" trajectory shown in Figure 2. The optimization problem is of the form:

$$\min F(\mathbf{x}) = \sum_{i=2}^{K} \frac{1}{2} \| \mathrm{A}\mathbf{x} - \mathbf{d}_{\mathrm{i}} \|^2 \tag{6}$$

Subject to:

inequality constraints: $[\mathrm{LB}] \leq \underline{\mathbf{X}} \leq [\mathrm{UB}]$ (7)

equality constraint: $\mathrm{A}*\underline{\mathbf{x}} = \mathrm{f}^{-1}(\underline{\mathbf{C}}(1)) - \underline{\mathbf{C}}(0)$ (8)

The sum of the error norms in (6) starts with the 2nd observation (we use the first observation in the equality constraint) and ends with the index of the fixed point in the observation data set shown in Figure 2. The [LB] and [UB] notation in (7) represents vectors of the lower and upper bound inequality constraints mentioned above. The equality constraint in (8), derived in equations (4) and (5), forces the parameter estimates to approximate the first step in the trajectory as closely as possible since the trajectory is entirely determined by the starting point $\underline{\mathbf{x}}(0)$. The vectors $\underline{\mathbf{C}}$ (1) and $\underline{\mathbf{C}}(0)$ are the observed values at $t = 0$ and $t = 1$.

The Matlab Optimization Toolbox function *fmincon*() is supplied with the objective function, constraints, and an initial starting point of $\underline{\mathbf{x}}(0) = \underline{\mathbf{0}}$. The constraints are "loose" since the only zeros specified are those corresponding to self-feedback, i.e., the diagonal entries in the edge matrix.

After 15 iterations the constrained optimization problem given by (6), (7) and (8) yields the following estimates for the FCM parameters:

$$\mathrm{W}_{est} = \begin{pmatrix} 0 & 0.1 & 0.1 & 0.1 & 0.2 & -0.1 & 0 \\ 0.1 & 0 & 0.1 & 0.1 & 0.2 & 0.5 & 0.1 \\ 0 & 0.2 & 0 & 0.2 & 0.1 & 0 & 0 \\ -0.1 & 0.1 & 0.1 & 0 & 0.2 & -0.3 & 0 \\ -0.1 & 0 & 0.1 & 0.2 & 0 & -0.4 & 0 \\ -0.1 & -0.1 & 0.1 & 0.1 & 0 & 0 & 0 \\ 0 & 0.2 & 0.1 & 0.1 & 0.1 & 0.1 & 0 \end{pmatrix}$$

Figure 4 - Estimated edge matrix using "loose" constraints

Note that the edge weights shown in Figure 4 estimated by the constrained optimization method are significantly different from those in Figure 1 belonging to the FCM that generated the "observed" data. The strengths and directions of the relationships between variables are generally inconsistent with what we know to be correct. Simulation of the system using the estimated edge weights yields the following trajectory:

Step	C1	C2	C3	C4	C5	C6	C7
1	0.5528	0.5922	0.6525	0.7231	0.7410	0.5533	0.5349
2	0.6089	0.7402	0.7273	0.7713	0.7891	0.5747	0.6237
3	0.6248	0.7772	0.7485	0.7886	0.8089	0.5892	0.6468
4	0.6286	0.7861	0.7546	0.7942	0.8151	0.5938	0.6525
5	0.6294	0.7883	0.7562	0.7959	0.8169	0.5949	0.6538
6	0.6295	0.7888	0.7567	0.7963	0.8174	0.5951	0.6541
7	0.6296	0.7890	0.7568	0.7965	0.8175	0.5952	0.6542
8	0.6296	0.7890	0.7568	0.7965	0.8176	0.5952	0.6542
9	0.6296	0.7890	0.7569	0.7965	0.8176	0.5952	0.6542

Figure 5 - FCM trajectory using estimated edge weights

The concept values of the trajectory resulting from simulating the system using the estimated edge weights are nearly identical to the observed trajectory shown in Figure 2, and the approximated system converges to a fixed point in the same number of iterations. We have closely reproduced the *behavior* of the system, but the system structure represented by our estimated edge matrix *does not* closely approximate the true structure of the system. Furthermore, we cannot characterize the edge weights as representing *causal* strength nor the form of the edge matrix as a *causal* structure since there has been no notion of causation during the synthesis process. We could analyze the observational data we used to create the edge matrix for correlations, but we *cannot* infer causality from correlation.

In the next experiment, we tighten the constraints to include zeros for the edge weights between concepts that do not have a direct causal relationship. The rule that ~correlation ⇒ ~causation does not help with this particular set of observational data because a correlation analysis of the trajectory shows strong correlation for all variables. In order to constrain the search space, an agent would have to apply background knowledge about the system. The upper and lower bound vectors are the same as before with the exception of the additional zeros. Using the same objective function and starting point but with tighter constraints, after 10 iterations our optimization method yields an edge matrix identical to the one used to generate the observed trajectory.

$$W_{est} = \begin{pmatrix} 0 & 0 & 0.6 & 0.9 & 0 & 0 & 0 \\ 0.1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.7 & 0 & 0 & 0.9 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 0 & -0.9 & -0.9 \\ -0.3 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.8 & 0.9 \end{pmatrix}$$

Figure 6 - Estimated edge matrix using "tight" constraints

The notion of causality has been introduced into the synthesis process by asserting knowledge about the *causal* nature of the relationships between some (but not all) of the variables, so this estimated edge matrix accurately reflects the structure of the entire system and the causal structure of the parts of the system represented *only by those particular variables*.

Summary of Experimental Results

We generated an observational data set for a dynamical system using a known FCM and let the simulation converge to a fixed point. The data set was intentionally kept small (the "time series" was only about as long as the number of parameters) to mimic a resource-constrained environment for an intelligent agent. With this limited amount of data, we used a constrained optimization algorithm in an attempt to infer the parameters and structure of the FCM that generated the data, assuming that the parameters were time-invariant. The first step in our FCM synthesis required selecting an appropriate activation function. With the exception of assuming no concept self-feedback, we formed a loose set of constraints using no background knowledge about the system. Optimization using these constraints resulted in an estimated FCM edge matrix that accurately modeled the behavior of the system but did not accurately reflect the true structure of the system, suggesting that we found a local, but not necessarily a global solution. Since we had no background knowledge about the nature of the relationships between variables in the observed data from which we constructed our edge matrix, we could not assert that our model had a *causal* structure. We then tightened the constraints using background knowledge about the lack of direct causality between concepts (the number and location of the zeros in an FCM edge matrix). Using the fact that ~correlation ⇒ ~causation did not aid in constraining the search space because all of the variables in the observational data were highly correlated. Using the tight constraints, our optimization method yielded parameters and structure that accurately reflected the parameter values and structure of the dynamical system that generated the observed data. We introduced the notion of causation into the synthesis process by making an explicit statement regarding the lack of a causal relation between a subset of variables enabling us to make statements about the causal structure of the parts of the system represented by those variables. We cannot say with any degree of certainty that the non-zero entries in the estimated connection matrix represent causal behavior. This is the subject of the next section.

3. Causal discovery methods related to FCM

Synthesis

If an intelligent agent is tasked with finding a satisfactory *explanation* for a given set of observations, that explanation is intimately related to the notion of *causation*[9]. Fuzzy cognitive maps are directed cyclic graphs (DCGs) that enable an intelligent agent to represent knowledge about cause and effect relationships in its environment in the presence of feedback. Bayesian networks are directed *acyclic* graphs (DAGs) and provide a similar method for representing causal influences but their acyclic nature limits their use to domains without feedback.

Our experiments in section 2 illustrated some of the issues an intelligent agent would encounter during autonomous FCM synthesis. The search space for causal structures is exponential in the number of concepts and perhaps the simplest way to limit the search space is to identify variables that aren't correlated. However, this approach doesn't always work, and when it does, we can't make positive statements about the causal relationships of the remaining variables, only the *lack* of causality between the uncorrelated ones. Once an agent has knowledge about the causal structure of its environment, estimating the strength of the cause-effect relationships is a relatively straightforward parameter estimation problem as presented in this paper and elsewhere[5,6].

In order to autonomously synthesis an FCM from observational data, an intelligent agent must be capable of *causal discovery*. In the next section we introduce related research in the area of causal discovery that can be used as a basis for future research in autonomous FCM synthesis.

A Survey of Related Research in Causal Discovery

Defining and identifying causality[25] from sample data is notoriously difficult, but for our work in autonomous FCM synthesis we can leverage the wealth of knowledge about causal discovery and the nature of causation resulting from research in fields such as Bayesian networks, structural equation modeling (SEM) and data mining. For example, Pearl[8] devotes an entire chapter to learning causal structure from observational data in the context of Bayesian networks. Pearl and Verma[9] present a theory of inferred causation that addresses the issue of distinguishing spurious covariance from genuine causation using statistical analysis. Tian and Pearl[10] propose a method for discovering causal relationships in data based on the detection and interpretation of local spontaneous changes in the environment. Heckerman, Meek and Cooper[12] compare a Bayesian to a constraint-based approach for the discovery of acyclic causal models from data. Spirtes, Glymour and Scheines[15] present the PC and FCI (Fast Causal Inference) algorithms, which form the basis of the TETRAD [30] tools that assist a user in searching the space of causal models represented as Bayesian networks or linear SEMs.

SEM focuses on covariance analysis with the important distinction of the assumption of causality in structural equation models[1]. The emphasis of SEM research is on hypothesis testing of manually specified causal models, rather than on the automated search over the space of models[11] but much of the underlying theory developed in the context of relating causal structure and observations is pertinent to how we think about the task of causal discovery in FCM synthesis. A recursive SEM is a type of DAG model while a non-recursive SEM refers to a type of DCG model. Since there's an assumption of a causal structure in both SEMs and FCMs and FCMs are represented as DCGs, fuzzy cognitive maps are related to non-recursive structural equation models and we might be able to apply certain aspects of the causal theory for SEMs to FCMs.

Cooper[18] developed a Local Causal Discovery (LCD) algorithm that is a specialization of the PC and FCI constraint-based causal discovery algorithms presented in Spirtes, Glymour and Scheines[15] and later applied it to automated causal discovery using text data mining in[11]. Silverstein et.al.[19] develop a causal discovery algorithm based on Cooper's LCD algorithm that uses a chi-squared statistic to substitute for the LCD dependence and independence tests, and use their algorithm for determining causal relationships from market basket data.

Some work has been done in the area of causal discovery for DCGs. Richardson[21,22,23] presents a cyclic causal discovery algorithm and an algorithm for deciding Markov equivalence of DCG models as an extension to earlier work done by Pearl and Spirtes with DAGs. Pearl and Dechter [24] extend the d-separation criterion used as a test for conditional independence relationships in Bayesian networks to feedback systems involving discrete variables and claim that the results "should have direct application in programs that learn the structure of feedback systems". Harwood and Scheines[17] develop a genetic algorithm search over causal models in the context of linear DAGs (SEMs without feedback) that should extend to cyclic (non-recursive) SEMs.

An important cumulative result of this research is that it *is* possible to characterize the statistical signature of causal structure, and though the majority of the results pertain to DAGs, the body of knowledge resulting from this work has the potential to provide the underpinnings of an approach for causal discovery for autonomous synthesis of FCMs. One must be careful, however, of blindly applying methods developed in the context of Bayesian networks to causal discovery for FCMs. For example, the PC and FCI algorithms and their derivatives rely on the assumption of a *Causal Markov Condition*[15] for establishing conditional independence relationships which holds by definition in causal Bayesian networks but not necessarily for causal DCGs.

4. CONCLUSIONS AND FUTURE WORK

We showed that without adequate constraints, estimating FCM edge weights from observational data can result in a model that accurately represents the behavior of the system but not necessarily the true relationships between the variables that generated the data, highlighting the importance of constraining the search space through proper structure identification. FCM structure traditionally embodies the notion of causality, the nature of which has a rich history of philosophical debate that we choose to not elaborate on but point out the special role it plays in FCM synthesis from observational data. Automatic causal discovery is an active area of research in disciplines related to FCMs such as Bayesian networks and structural equation models. We present a survey of literature that can serve as a foundation for future work in FCM structure identification, though caution must be exercised when applying results from closely related disciplines since certain assumptions may fail to hold.

A domain expert acquires his or her knowledge over time and there is always some degree of uncertainty associated with expert opinion. An intelligent agent should go through a similar learning process when autonomously synthesizing an FCM and should have a way of expressing a degree of uncertainty associated with its learned view of the world. An expert's domain knowledge can be represented by a rule-based expert system, so a reasonable approach to autonomous FCM synthesis incorporates a fuzzy expert system to represent knowledge about relationships learned from observations over time. Confidence factors associated with the learned production rules can represent the agent's degree of uncertainty regarding the cause-effect relationships with the confidence factor increasing or decreasing as new facts about the system are learned.

Our future work on autonomous FCM synthesis for intelligent agents will build on ideas introduced in this paper, exploring the use of hybrid techniques for learning and knowledge representation in a complex environment in the broader context of exploring novel methods for intelligent identification and monitoring of large-scale, systems-of-systems. Our preliminary approach will attempt to determine the feasibility of the DCG causal discovery approaches suggested by Richardson[21,22,23], Pearl and Dechter[24] and Harwood and Scheines[17] in the context of an intelligent agent architecture. We will further study the integration of rule-based fuzzy expert systems and text data mining techniques into the FCM learning process and introduce the notion of indeterminacy into autonomous FCM synthesis using neutrosophic cognitive and relational maps [31] in an effort to create a framework for autonomous FCM synthesis that closely resembles the cognitive process of human experts.

APPENDIX A – A BRIEF INTRODUCTION TO FUZZY COGNITIVE MAPS

A fuzzy cognitive map[2] is a signed, directed graph with feedback where the nodes represent concepts and a directed edge e_{ij} measures how much concept C_i causes C_j. A time varying concept $C_i(t)$ measures the degree of occurrence of some fuzzy event, such as the degree to which a component has failed or the "strength" of subsystem health, and can take on values in the fuzzy interval [0, 1]. The edges e_{ij} take on values in the fuzzy interval [-1, 1] where $e_{ij} = 0$ indicates no causality from C_i to C_j, $e_{ij} > 0$ indicates causal correlation in the same direction (C_j increases as C_i increases or C_j decreases as C_i decreases) and $e_{ij} < 0$ indicates negative causal correlation (C_j decreases as C_i increases or C_j increases as C_i decreases). A concept cannot cause itself (no self-feedback), so the entries on the diagonal of the edge matrix are always zero.

An FCM is usually constructed by a knowledge engineer who acquires domain knowledge from systems experts and uses that knowledge to define the concepts, causal directions and fuzzy values of the edges of the graph. Edge matrices resulting from interviews with multiple domain experts can be combined to yield an edge matrix that collectively encodes the background knowledge of the cause and effect relationships of a system. A weighting function can be used to give more weight to an FCM constructed from an interview with a more experienced systems engineer and a lesser weight to one constructed on advice from a less experienced systems engineer[32]. The resulting FCM is a linear combination of the separate FCMs.

Fletcher et.al.[32] construct a simple FCM to model high-level systems health for the Command and Data Handling (C&DH), Electrical Power System (EPS) and Thermal Control System (TCS) for the International Space Station.

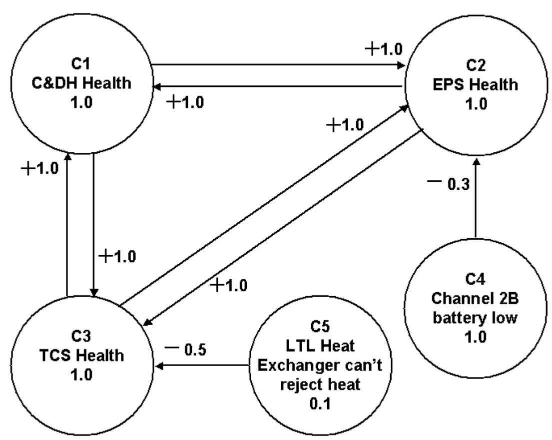

Figure 7 - A High-level ISS Systems Health Monitor implemented as a Fuzzy Cognitive Map

The FCM shown in Figure 7 is represented by a concept vector $\underline{\mathbf{C}}_t(0)$ = [1.0, 1.0, 1.0, 1.0, 0.1] and the connection matrix:

$$\begin{pmatrix} 0.0 & 1.0 & 1.0 & 0.0 & 0.0 \\ 1.0 & 0.0 & 1.0 & 0.0 & 0.0 \\ 1.0 & 1.0 & 0.0 & 0.0 & 0.0 \\ 0.0 & -.3 & 0.0 & 0.0 & 0.0 \\ 0.0 & 0.0 & -.5 & 0.0 & 0.0 \end{pmatrix}$$

Figure 8 – Connection (Edge) matrix for the FCM shown in Figure 7

An FCM is a dynamical system that can simulate the behavior of the system being modeled. Successive matrix-vector multiplications are performed using the concept vector and edge matrix with the output of one operation being used as the input to the next. The FCM simulation will either diverge or converge to a fixed point (a single vector) or limit cycle (repeating pattern of vectors). During the simulation, the edge values remain fixed while the concept values vary.

References

[1] K. Bollen, *Structural Equations with Latent Variables*, John Wiley & Sons, Inc., 1989.

[2] B. Kosko, *Neural Networks and Fuzzy Systems*, Prentice Hall, Inc., 1992.

[3] J.A. Dickerson, B. Kosko, “Virtual Worlds as Fuzzy Cognitive Maps”, *Presence,* Volume 3, Number 2, pp. 173-189, 1994.

[4] A.V. Huerga, “A Balanced Differential Learning Algorithm in Fuzzy Cognitive Maps”, *Technical Report, Departament de Llenguatges i Sistemes Informatics, Universitat Politecnica de Catalunya (UPC)*, C\Jordi Girona 1-3, E0834, Barcelona, Spain http://monet.aber.ac.uk:8080/monet/docs/pdf_files/qr_02/qr2002alberto-vazquez.pdf

[5] D. E. Koulouriotis, I.E. Diakoulakis, D.M. Emiris, “Learning Fuzzy Cognitive Maps using Evolution Strategies: a Novel Schema for Modeling and Simulating High-Level Behavior”, *IEEE Congress on Evolutionary Computation (CEC2001*), pp. 364-371, Seoul, Korea, 2001.

[6] K.E. Parsopoulos, E.I. Papageorgiou, P.P. Groumpos, M.N. Vrahatis, “A First Study of Fuzzy Cognitive Maps Learning Using Particle Swarm Optimization”, *Proceedings of the IEEE 2003 Congress on Evolutionary Computation,* Canberra, Australia, pp. 1440-1447, ISBN: 0-7803-7804-0, IEEE Catalog Number: 03TH8674, 2003.

[7] R. Stolle, E. Bradley, “Multimodal Reasoning for Automatic Model Construction”, *Proceedings Fifteenth National Conference on Artificial Intelligence 1998 (AAAI-98)*, Madison, Wisconsin, July 1998.

[8] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann Publishers, Inc., San Francisco, CA., 1988.

[9] J. Pearl, T.S. Verma, “A Theory of Inferred Causation”, *Second International Conference on the Principles of Knowledge Representation and Reasoning*, Cambridge, MA., April, 1991.

[10] J. Tian, J. Pearl, “Causal Discovery from Changes: A Bayesian Approach”, Citation unknown. This is Part-II of a two-part paper submitted to UAI-01. Part-I is entitled Causal Discovery from Changes.

[11] S. Mani, G.F. Cooper, “Causal Discovery from Medical Textual Data”, *Proceedings of the AMIA Annual Fall Symposium*, pp. 542-546, Hanley and Belfus Publishers, Philadelphia, PA., 2000.

[12] D. Heckerman, C. Meek, G. Cooper, “A Bayesian Approach to Causal Discovery”, In C. Glymour and G. Cooper, editors, *Computation, Causation and Discovery*, MIT Press, Cambridge, MA., 1999.

[13] H. Dai, K. Korb, C. Wallace, X. Wu, “A Study of Causal Discovery with Small Samples and Weak Links”, *Proceedings of the 15th International Joint Conference on Artificial Intelligence IJCAI’97*, Morgan Kaufmann Publishers, Inc., pp. 1304-1309, 1997.

[14] C. Wallace, K. Korb, H. Dai, “Causal Discovery via MML”, *Proceedings of the 6th Pacific-Asia Conference on Knowledge Discovery and Data Mining (PAKDD-2002)*, Taiwan, 304-315, 2002.

[15] P. Spirtes, C. Glymour, R. Scheines, *Causation, Prediction and Search*, The MIT Press, 2000.

[16] P. Spirtes, C. Glymour, “An Algorithm for Fast Recovery of Sparse Causal Graphs”, *Social Science Computer Review*, 9:1, Spring 1991.

[17] S. Harwood, R. Scheines, “Genetic Algorithm Search over Causal Models”, *Technical Report No. CMU-PHIL-131*. Dept. of Philosophy, Carnegie Mellon Univ., Pittsburgh, PA. 15213, 2002.

[18] G. F. Cooper, “A Simple Constraint-Based Algorithm for Efficiently Mining Observational Databases for Causal Relationships”, *Data Mining and Knowledge Discovery*, 1, pp. 203-224, 1997.

[19] C. Silverstein, S. Brin, R. Motwani, J. Ullman, “Scalable Techniques for Mining Causal Structures”, *Proceedings of the 24th VLDB Conference*, pp. 594-605, August 1998.

[20] T. Richardson, P. Spirtes, “Automated Discovery of Linear Feedback Models”, *Technical Report CMU-75-Phil*., 1996.

[21] T. Richardson, “A Discovery Algorithm for Directed Cyclic Graphs”. In *Proceedings of the 12th Conference on Uncertainty in Artificial Intelligence*, Portland, Oregon. E. Horvitz and F. Jensen (eds.), Morgan Kaufmann, San Francisco, CA. 1996.

[22] T. Richardson, “A Polynomial-Time Algorithm for Deciding Markov Equivalence of Directed Cyclic Graphical Models”, *Technical Report CMU-PHIL-63*, Philosophy

Department, Carnegie Mellon University.

[23] T. Richardson, “Equivalence in Non-Recursive Structural Equation Models”, In Proceedings of the 11th Symposium on Computational Statistics, COMPSTAT, 20-26, Vienna, Austria, August 1994.

[24] J. Pearl, R. Dechter, “Identifying Independencies in Causal Graphs with Feedback”, *UAI 1996*, Portland, Oregon, pp. 420-426, 1996.

[25] L.A. Zadeh, “Causality is Undefinable – Toward a Theory of Hierarchical Definability”, *2001 IEEE International Fuzzy Systems Conference*.

[26] A. K. Tsadiras, K. G. Margaritis: “Cognitive Mapping and Certainty Neuron Fuzzy Cognitive Maps”, *Information. Sciences,* **101**, pp. 109-130, 1997

[27] C. D. Stylios and P.P. Groumpos, "Fuzzy Cognitive Maps in Modeling Supervisory Control Systems“, *Journal of Intelligent & Fuzzy Systems*, 8, No2, pp. 83-98, 2000.

[28] Z.Q. Liu, “Fuzzy Cognitive Maps: Analysis and Extensions”, *Soft Computing and Human-Centered Machines*, Z.-Q. Liu and S. Miyamoto (Eds.), Springer-Verlag Tokyo, 2000.

[29] R. Scheines, P. Spirtes, C. Glymour, C. Meek, T. Richardson, “TETRAD 3 Tools for Causal Modeling User’s Manual”, web site: http://www.phil.cmu.edu/projects/tetrad/tet3/master.htm

[30] K. Hornik, M. Stinchcombe, H. White, “Universal Approximation of an Unknown Mapping and its Derivatives using Multilayer Feedforward Networks”, *Neural Networks*, 3:551-560, 1990.

[31] W.B. Vasantha Kandasamy, F. Smarandache, “*Fuzzy Cognitive Maps and Neutrosophic Cognitive Maps*”, Xiquan, 2003.

[32] D.P. Fletcher, F. Akkawi, R. Alena, D. Duncavage, “From Research to Operations, Integrating Components with an Aspect-Oriented Framework and Ontology”, *2004 IEEE Aerospace Conference Proceedings*, March 6–13, 2004.

BIOGRAPHY

Daryl Fletcher is a Computer Scientist with Science Applications International Corporation (SAIC) where he works as a Technical Lead on Advanced Diagnostic Systems for the International Space Station at NASA-Ames Research Center. Mr. Fletcher has a B.S. degree in Applied Mathematics and an M.S. degree in Civil Engineering, both from the University of Colorado-Boulder and is currently a Ph.D. student in Computer Science at the University of Colorado-Denver. His research interests include automatic model generation and refinement and soft computing techniques applied to large-scale systems modeling.

Dr. Duoug Nguyen is a Senior Principal Systems Engineer with Raytheon IIS/Space Systems in Aurora, Colorado. Dr. Nguyen is an adjunct faculty member in the Department of Computer Science and Engineering at the University of Colorado-Denver, a former engineering Professor at Colorado State University and has held positions as Chief Scientist at Logicon/Geodynamics, Senior Technical Consultant at Northrup Grumman/TRW and as a consultant to the Centre National d'Etudes Spatiales (CNES/France), NASA and NIH. He is a member of IEEE and AIAA.

Krzystof Cios (Ph.D., D.Sc.) is a Professor in the Department of Computer Science and Engineering at the University of Colorado-Denver, and Affiliated Professor with the Department of Preventive Medicine and Biometrics at the University of Colorado Health Sciences Center and Associate Director of the University of Colorado Center for Computational Biology, University of Colorado Bioenergetics Institute. His research interests are in the areas of data mining and knowledge discovery, bioinformatics, machine learning, and in neural networks of spiking neurons.