Hierarchical Latent Class Models and Statistical Foundation for Traditional Chinese Medicine

Nevin L. Zhang[1], Shihong Yuan[2], Tao Chen[1], and Yi Wang[1]

[1] Hong Kong University of Science and Technology, Hong Kong, China
{lzhang, csct, wangyi}@cs.ust.hk,
[2] Beijing University of Traditional Chinese Medicine, Beijing, China
yuanshih@yahoo.com.cn

Abstract. Traditional Chinese medicine (TCM) is an important avenue for disease prevention and treatment for the Chinese people and is gaining popularity among others. However, many remain skeptical and even critical of TCM because a number of its shortcomings. One key shortcoming is the lack of a scientific foundation and hence objective diagnosis standards. When viewed as a black box, TCM diagnosis is simply a classifier that classifies patients into different classes based on their symptoms. A fundamental question is: Do those classes exist in reality? To seek an answer from the machine learning perspective, one would naturally use cluster analysis. Previous clustering methods are unable to handle the complexity of TCM. We have therefore developed a new clustering method in the form of hierarchical latent class (HLC) models. In this paper, we provide a brief review of HLC models and present a case study to demonstrate the possibility of establishing a statistical foundation for TCM using HLC models.

Area: machine learning, data mining, kdd; **Paradigm**: probabilistic or numeric; **Technique**: bayesian networks; **Application**: traditional Chinese medicine.

Introduction

In TCM Diagnosis, patient information is collected through an overall observation of symptoms and signs rather than micro-level laboratory tests. The conclusion of TCM diagnosis is called *syndrome* and the process of reaching a diagnostic conclusion from symptoms is called *syndrome differentiation*. There are several syndrome differentiation systems, each focusing on a different perspective of the human body and with its own theory. The theories describe relationships between syndrome factors and symptoms, as illustrated by this excerpt:

> KIDNEY YANG [3] (Yang *et al.* 1998) is the basis of all YANG in the body. When KIDNEY YANG is in deficiency, it cannot warm the body and the

[3] Words in small capital letters are reserved for TCM terms.

> patient feels cold, resulting in intolerance to cold, cold limbs, and cold lumbus and back. Deficiency of KIDNEY YANG also leads to SPLEEN disorders, resulting in loose stools and indigested grain in the stool. [4]

Here syndrome factors such as KIDNEY YANG FAILING TO WARM THE BODY and SPLEEN DISORDERS DUE TO KIDNEY YANG DEFICIENCY are not directly observed. They are similar in nature to concepts such as 'intelligence' and are indirectly measured through their manifestations. Hence we call them *latent variables*. In contrast, symptom variables such as 'cold limbs' and 'loose stools' are directly observed and we call them *manifest variables*. TCM theories involve a large number of latent and manifest variables. Abstractly speaking, they describe relationships among latent variables, and between latent variables and manifest variables. Hence they can be viewed as *latent structure models* specified in natural language.

TCM is an important avenue for disease prevention and treatment for ethnic Chinese and is gaining popularity among others. However, it suffers a serious credibility problem especially in the west. One reason is the lack of rigorous randomized trials in support for the efficacy of TCM herb treatments (Normile 2003). Another equally important reason, on which this paper focuses, is the lack of scientific validations for TCM theories. Researchers in China have been searching for such validations in the form of laboratory tests for more than half a century, but there has been little success. We propose and investigate a statistical approach. In the next three paragraphs, we explain the premise and the main idea of the approach.

We human beings often invoke latent variables to explain regularities that we observe. Here is an experience that many might share. I (the first author) was looking at some apartment buildings nearby one night. I noticed that, for a period of time, the lighting from several apartments was changing in brightness and color at the same times and in perfect synchrony. This caught my attention and my brain immediately concluded that there must be a common cause that was responsible for changes. My brain did so without knowing what the common cause was. So, a latent variable was introduced to explain the regularity that I observed. What I tried to do next was to find the identity of the latent variable.

We conjecture that, in a similar vein, latent syndrome variables in TCM were introduced to explain observed regularities about the occurrence of symptoms. Take the concept KIDNEY YANG FAILING TO WARM THE BODY as an example. We believe that in ancient times it was first observed that symptoms such as intolerance to cold, cold limbs, and cold lumbus and back often occur together in patients, and then, to explain the phenomenon, the latent variable KIDNEY YANG FAILING TO WARM THE BODY was created.

When explaining the phenomenon of synchronous change in lighting, I resorted to my knowledge about the world and concluded that the common cause must be that residents in those apartments were watching the same TV channel. Similarly, when explaining patterns about the occurrence of symptoms, ancient

[4] Through out this paper, citations about TCM theories are mostly from (Yang *et al.* 1998), a bilingual textbook.

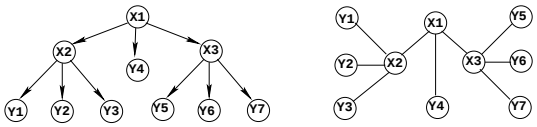


Fig. 1. An example HLC model and the corresponding unrooted HLC model. The X_i's are latent variables and the Y_j's are manifest variables.

Chinese resorted to their understanding of the world and the human body. This explains why concepts from ancient Chinese philosophy such as YIN and YANG are prevalent in TCM theories. Words such as KIDNEY and SPLEEN also appear in TCM theories because there was primitive anatomy in ancient times. However, the functions that TCM associates with KIDNEY and SPLEEN are understandably different from the functions of kidney and spleen in modern western medicine.

Thus, the premise of our work is that TCM theories originated from regularities ancient Chinese doctors observed in their experiences with patients. The main idea of our approach, called *the latent structure approach*, is to collect patient symptom data systematically, analyze the data based on statistical principles, and thereby obtain mathematical latent structure models. If the mathematical latent structure models match the relevant aspects of TCM theories, then we would have validated those aspects of TCM theories statistically. This would also suggest the possibility of laying a statistical foundation for TCM through data analysis and thereby turning it into a modern science.

In the past few years, we have developed a class of latent structure models called *hierarchical latent class (HLC) models*, and have used HLC models in a case study to test the idea of the latent structure approach. In this paper, we briefly review HLC models, describe the case study and report the findings.

1 HIERARCHICAL LATENT CLASS MODELS

Hierarchical latent class (HLC) models (2004) are tree-structured Bayesian networks where variables at leaf nodes are observed and are hence called *manifest variables*, while variables at internal nodes are hidden and hence are called *latent variables*. All variables are assumed discrete. HLC models generalize latent class (LC) models (Lazarsfeld and Henry 1968) and were first identified as a potentially useful class of Bayesian networks by Pearl (1988).

Fig. 1 shows an example HLC model (left diagram). A *latent class (LC) model* is an HLC model where there is only one latent node. We usually write an HLC model as a pair $M = (m, \theta)$, where θ is the collection of parameters. The first component m consists of the model structure and cardinalities of the variables. We will sometimes refer to m also as an HLC model. When it is necessary to distinguish between m and the pair (m, θ), we call m an *uninstantiated HLC model* and the pair (m, θ) an *instantiated HLC model*.

Two instantiated HLC models $M{=}(m, \theta)$ and $M'{=}(m', \theta')$ are *marginally equivalent* if they share the same manifest variables Y_1, Y_2, ..., Y_n and

$$P(Y_1, \ldots, Y_n|m, \theta) = P(Y_1, \ldots, Y_n|m', \theta'). \quad (1)$$

An uninstantiated HLC models m *includes* another m' if for any parameterization θ' of m', there exists parameterization θ of m such that (m, θ) and (m', θ') are marginally equivalent, i.e. if m can represent any distributions over the manifest variables that m' can. If m includes m' and vice versa, we say that m and m' are *marginally equivalent*. Marginally equivalent (instantiated or uninstantiated) models are *equivalent* if they have the same number of independent parameters. One cannot distinguish between equivalent models using penalized likelihood scores.

Let X_1 be the root of an HLC model m. Suppose X_2 is a child of X_1 and it is a latent node. Define another HLC model m' by reversing the arrow $X_1{\rightarrow}X_2$. In m', X_2 is the root. The operation is hence called *root walking*; the root has walked from X_1 to X_2. Root walking leads to equivalent models (Zhang 2004). This implies that it is impossible to determine edge orientation from data. We can learn only *unrooted HLC models*, which are HLC models with all directions on the edges dropped. Fig. 1 also shows an example unrooted HLC model. An unrooted HLC model represents a class of HLC models. Members of the class are obtained by rooting the model at various nodes. Semantically it is a Markov random field on an undirected tree. The leaf nodes are observed while the interior nodes are latent. Marginal equivalence and equivalence can be defined for unrooted HLC models in the same way as for rooted models. From now on when we speak of HLC models we always mean unrooted HLC models unless it is explicitly stated otherwise.

Let $|X|$ stand for the cardinality of a variable X. For a latent variable Z in an HLC model, enumerate its neighbors as $X_1, X_2, \ldots, X_k$. An HLC model is *regular* if for any latent variable Z, $|Z| \leq \prod_{i=1}^{k} |X_i| / \max_{i=1}^{k} |X_i|$, and when Z has only two neighbors, strict inequality holds and one of the neighbors must be a latent node. Note that this definition applies to both instantiated and uninstantiated models.

Given an irregular instantiated model m, there exists a regular model that is marginally equivalent to m and has fewer independent parameters (Zhang 2004). The process of obtaining the regular model is called *regularization*. It is evident that if penalized likelihood is used for model selection, the regularized model is always preferred over m itself.

Assume that there is a collection $\mathbf{D}$ of i.i.d samples on a given set of manifest variables that were generated by an unknown regular HLC model. The learning task is to reconstruct the regular unrooted HLC models that corresponds to the generative model.

Although not using the terminology of HLC models, Connolly (1993) proposed the first, somewhat *ad hoc*, algorithm for learning HLC models and tested it on one toy example with 4 manifest variables. A more principled hill-climbing algorithm was developed by Zhang (2004). The algorithm consists of two search routines, one optimizes model structure while the other optimizes cardinalities of

latent variables in a given model structure. It is hence called *double hill-climbing (DHC)*. It can deal with data sets with about one dozen manifest variables. Zhang and Kočka (2004) proposed another algorithm called *heuristic single hill-climbing (HSHC)*. HSHC combines the two search routines of DHC into one and incorporates the idea of structural EM (Friedman 1997) to reduce the time spent in parameter optimization. HSHC can deal with data sets with dozens of manifest variables.

Results presented in this paper were obtained using the HSHC algorithm. The algorithm hill-climbs in the space of all unrooted regular HLC models for the given manifest variables. We assume that the BIC score is used to guide the search. The BIC score of a model m is:

$$BIC(m|\mathbf{D}) = logP(\mathbf{D}|m, \theta^*) - \frac{d(m)}{2} logN$$

where θ^* is the ML estimate of model parameters, $d(m)$ is the *standard dimension* of m, i.e. the number of independent parameters, and N is the sample size.

Geiger *et al.* (1996) argued that, when latent variables are present, it is inappropriate to use standard model dimension in the BIC score. One should instead of use what is called *effective model dimension*. Unfortunately, it is rather difficult to compute effective model dimensions. In this paper, we use standard model dimensions and leave it to future work to investigate the impact of this choice. Empirical results of Zhang (2004) show that the DHC algorithm is able to obtain high quality models with standard model dimensions. As will be seen, results presented in this paper indicate that the HSHC algorithm is also able to obtain high quality models with standard model dimensions.

2 Data and Data Analysis

The data set used in our case study involves 35 symptom variables, which are considered important when deciding whether a patient suffers from the so-called KIDNEY DEFICIENCY syndrome, and if so, which subtype. Each variable has four possible values: none, light, medium, and severe. The data were collected from senior citizen communities, where the KIDNEY DEFICIENCY syndrome frequently occurs. There are totally 2,600 records. Each record consists of values for the 35 symptom variables, but there is no information about syndrome types.

We refer the relevant TCM theory that explains the occurrence of the 35 symptoms as the *TCM* KIDNEY *theory*. As mentioned earlier, this is a latent structure model specified in natural language. The objective of the case study is to induce a mathematical latent structure model from the data based on statistical principles and compare it with the TCM KIDNEY theory to see whether and how well they match.

The KIDNEY data were analyzed using the HSHC algorithm. The best model that we obtained is denoted by M. Its BIC score is -73,860 and its structure is shown in Fig. 2. In the model, Y_0 to Y_{34} are the manifest variables that appear in the data, while X_0 to X_{13} are the latent variables introduced in the process data analysis.

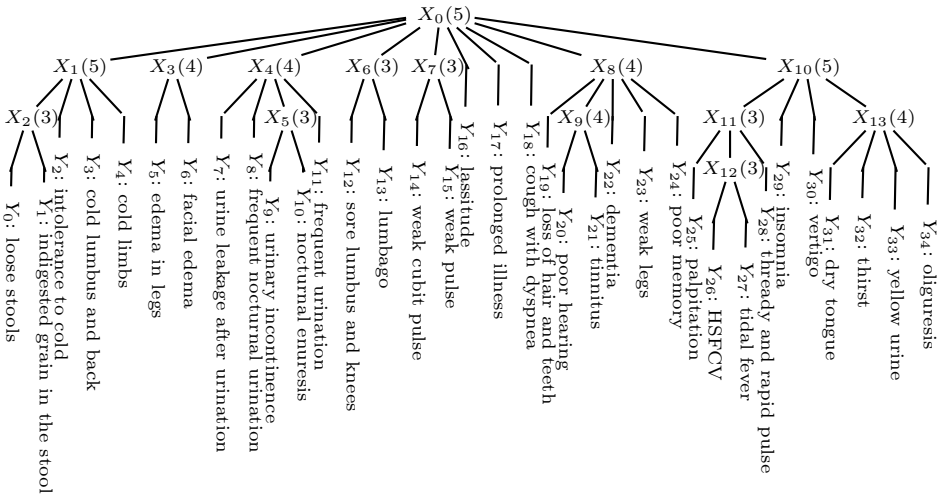


Fig. 2. The structure of the best model M found for KIDNEY data. The abbreviation HSFCV stands for Hot Sensation in Five Centers with Vexation, where the five centers refer to the centers of two palms, the centers of two feet, and the heart. The integer next to a latent variable is the number of possible states of the variable.

3 Latent Variables

We now set out to compare the structure of model M with the TCM KIDNEY theory. According to the semantics of HLC models, the left most part of model M states that there is a latent variable X_1 that is (1) directly related to the symptoms intolerance to cold (Y_2), cold lumbus and back (Y_3), and cold limbs (Y_4); and (2) through another latent variable X_2 indirectly related to loose stools (Y_0) and indigested grain in the stool (Y_1). On the other hand, the TCM KIDNEY theory asserts that when KIDNEY YANG is in deficiency, it cannot warm the body and the patient feels cold, resulting in manifestations such as cold lumbus and back, intolerance to cold, and cold limbs. Deficiency of KIDNEY YANG also leads to SPLEEN disorders, resulting in symptoms such as loose stools and indigested grain in the stool.

Here model M and the TCM KIDNEY theory both mention the same five symptom variables; they both describe how two latent variables are related to those five symptoms; and the relationships described in the two cases share the same structure. The only difference is that the latter has named the two latent variables and given an explanation to the relationships, while model M simply states the relationships. Therefore, we have a good match between model M and the TCM KIDNEY theory here. The latent variable X_1 can be interpreted as KIDNEY YANG FAILING TO WARM THE BODY, while X_2 can be interpreted as SPLEEN DISORDERS DUE TO KIDNEY YANG DEFICIENCY (KYD).

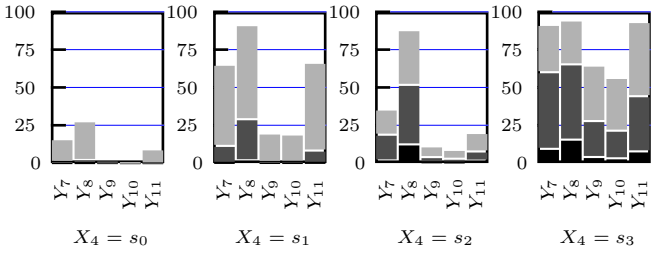
To the right of X_1, model M states that there is a latent variable X_3 that is directly related to the symptoms of edema in legs (Y_5) and facial edema (Y_6).

On the other hand, the TCM KIDNEY theory asserts that when KIDNEY YANG is in deficiency, it cannot control WATER, which overflows to the surface of the face and the legs, resulting in facial edema and edema in legs. Here we see another good match. The latent variable X_3 can be interpreted EDEMA DUE TO KYD.

To the right of X_3, model M states that there is a latent variable X_4 that is (1) directly related to the symptoms of urine leakage after urination (Y_7), frequent nocturnal urination (Y_8) and frequent urination (day) (Y_{11}); and (2) through another latent variable X_5 indirectly related to urinary incontinence (day) (Y_9) and nocturnal enuresis (Y_{10}). On the other hand, the TCM KIDNEY theory asserts that when KIDNEY fails to control the urinary bladder, one would observe clinical manifestations such as frequent urination, urine leakage after urination, frequent nocturnal urination, and in severe cases urinary incontinence and nocturnal enuresis. Once again, there is a good match between this part of M and the relevant aspect of the TCM KIDNEY theory. The latent variable X_4 can be interpreted as KIDNEY FAILING TO CONTROL UB, where UB stands for the urinary bladder.

According to the TCM KIDNEY theory, clinical manifestations of the KIDNEY ESSENCE INSUFFICIENCY syndrome includes premature baldness, tinnitus, deafness, poor memory, trance, declination of intelligence, fatigue, weakness, and so on. Those match the symptom variables in model M that are located under X_8 fairly well and hence X_8 can be interpreted as KIDNEY ESSENCE INSUFFICIENCY. The clinical manifestations of the KIDNEY YIN DEFICIENCY syndrome includes dry throat, tidal fever or hectic fever, fidgeting, hot sensation in the five centers, insomnia, yellow urine, rapid and thready pulse, and so on. Those match the symptom variables under X_{10} fairly well and hence X_{10} can be interpreted as KIDNEY YIN DEFICIENCY. Finally, the TCM KIDNEY theory asserts that KIDNEY DEFICIENCY can be caused by prolonged illness and manifests as one or more sub-KIDNEY DEFICIENCY syndromes such as KIDNEY YANG DEFICIENCY, KIDNEY FAILING TO CONTROL UB, KIDNEY ESSENCE DEFICIENCY, and KIDNEY YIN DEFICIENCY. Moreover, patients suffering from KIDNEY DEFICIENCY usually share common symptoms such as lumbago, sore and weak lumbus and knees, and mental and physical fatigue. Those and the structure of M suggest that X_0 should be interpreted as KIDNEY DEFICIENCY.

All symptom variables in the case study are those that a TCM doctor would consider when making a diagnostic decision about KIDNEY DEFICIENCY. Hence there is no surprise that one of the latent variables can be interpreted as KIDNEY DEFICIENCY. However, it is very interesting that some of the latent variables in model M correspond to syndrome factors such as KIDNEY YANG FAILING TO WARM THE BODY, SPLEEN DISORDERS DUE TO KYD, EDEMA DUE TO KYD, KIDNEY FAILING TO CONTROL UB, KIDNEY ESSENCE DEFICIENCY, and KIDNEY YIN DEFICIENCY, as each of them is associated with only a subset of the symptom variables in the TCM KIDNEY theory. As the latent variables were introduced by data analysis based on a statistical principle, the case study had provided statistical validation for the introduction of those syndrome factors to the TCM



Cluster	Most Typical Members				
	Y_7	Y_8	Y_9	Y_{10}	Y_{11}
X_4=s_0	0	0	0	0	0
X_4=s_1	1	1	0	0	1
X_4=s_2	0	3	0	0	0
X_4=s_3	3	3	3	3	3

Fig. 3. Characteristics of the 4 clusters identified by the states of the latent variable X_4: The diagrams show the probability distributions of the symptom variables Y_7, Y_8, Y_9, Y_{10}, and Y_{11} in each of the clusters. The table shows the most typical members of the clusters.

KIDNEY theory and for what are asserted about their relationships with symptom variables.

There are also aspects of M that do not match TCM theory well. One mismatch is in the scope: model M involves fewer symptom variables than all those TCM associates with KIDNEY. Another kind of mismatch is more technical. Consider the symptoms insomnia and vertigo. According to the TCM KIDNEY theory, they can be caused both by KIDNEY YIN DEFICIENCY and KIDNEY ESSENCE INSUFFICIENCY. However, in M they are directly connected only to X_{10}, and not to X_8. Such mismatches are due to the restriction that HLC models must be tree-structured. Those mismatches suggest future research directions. However, they do not defeat the conclusion that the case study has provided statistical validation for the TCM KIDNEY theory.

4 Latent Classes

By analyzing the KIDNEY data using HLC models, we have not only obtained a latent structure, but also clustered the data in multiple ways. As indicated in Fig. 2, the latent variable X_1 has 5 states. This means that the data has in one way been grouped into 5 clusters, with one cluster corresponding to each state of X_1. Similarly, the fact that X_4 has 4 states means that the data has in another way been grouped into 4 clusters, with one cluster corresponding to each state of X_4. In the following, we examine the meaning of those *latent classes* and point out that they, like the latent variables, provide statistical validation for aspects of the TCM KIDNEY theory. We use the 4 clusters given by X_4 as an example.

Fig. 3 shows the probability distributions of the 5 symptom variables Y_7 to Y_{11} in each of the 4 clusters given by X_4. There is one diagram for each cluster. In each diagram, there are 5 bars, each corresponding to one of the 5 symptom variables. The bars consist of up to 4 segments, each corresponding to one state of the symptom variables. Note that the white segments are located at the top and they reach up to the top boundary of the diagrams. White and black segments represent the probabilities of the symptoms being absent and severe respectively, while the two shades of grey represent those for two intermediate states. The

bar diagrams provide an overall characterization of the clusters. In contrast, information about the most typical members of the clusters gives us a more concrete idea about their members. The *most typical member* of the cluster $X_4 = s_0$, for instance, is the configuration of the states of the variables Y_7 to Y_{11} that maximizes the probability $P(X_4 = s_0|Y_7, Y_8, Y_9, Y_{10}, Y_{11})$. This configuration might not be unique. Nonetheless, we still speak of 'the most typical member' for simplicity.

We can now digest the meanings of the 4 clusters. First of all, the meaning of X_4 is KIDNEY FAILING TO CONTROL UB (KFCUB). Hence the clusters can be viewed as different states of KFCUB. Five symptoms are involved here, namely urine leakage after urination (ULU) (Y_7), frequent nocturnal urination (FNU) (Y_8), urinary incontinence (UI) (Y_9), nocturnal enuresis (NE) (Y_{10}), and frequent urination (FU) (Y_{11}). In the cluster $X_4 = s_0$, the five symptoms almost never occur and the most typical member has none of those symptoms. Hence, the cluster can be interpreted as *no* KFCUB. In the cluster $X_4 = s_3$, on the other hand, the five symptoms have high probabilities of occurring and for the most typical member, all the symptoms occur at the severe level. Hence, this cluster can be interpreted as *severe* KFCUB.

Next consider the clusters $X_4 = s_1$ and $X_4 = s_2$. The overall probability of the symptoms occurring is higher in $X_4 = s_1$, while the probability of the symptoms occurring at the medium or severe levels are higher in $X_4 = s_2$. For the most typical member of $X_4 = s_1$, three of the symptoms, namely FU, ULU, and FNU, occur only at the light level, while for the most typical member of $X_4 = s_2$, only one of the symptom, namely FNU, occurs at the severe level. Therefore, both clusters can be interpreted as light KFCUB. We interpret $X_4 = s_1$ as *light* KFCUB *(A)* and $X_4 = s_2$ as *light* KFCUB *(B)*.

The latent classes contain information that validates aspects of the TCM KIDNEY theory. As mentioned in the previous section, the TCM KIDNEY theory associates the symptoms FU, ULU and FNU with KFCUB and it asserts that UI and NE occurs only in the case of severe KFCUB. This is consistent with the bar diagrams in Fig. 3. We see that, in the first three clusters where KFCUB is not severe, UI and NE either do not occur at all or occur only at the light level with low probabilities. In the last cluster where KFCUB is severe, the probabilities of UI and NE occurring are high, but still comparatively lower than those of the other three symptoms. The TCM KIDNEY theory also asserts that FNU is a typical symptom of KIDNEY DEFICIENCY, meaning that it occurs frequently among patient suffering from the syndrome. This again is consistent with the bar diagrams. We see that the probability of FNU occurring is high in the three clusters where KFCUB is present, namely the last three clusters. Moreover, it occurs for the most typical members of all the three clusters.

In the TCM KIDNEY theory, there are other qualitative assertions about relative symptom occurrence frequencies in addition to those mentioned in the previous paragraph. All such assertions that are relevant to model M are consistent with model M. Therefore, our case study has provided statistical validation

for those assertions. Moreover, it has refined the assertions by providing probabilistic quantifications.

5 Conclusion and Future Work

The TCM KIDNEY theory was formed in ancient times, while model M was obtained through modern day data analysis. It is very interesting that they match each other well. This shows that, contrary to popular perception, there are scientific truths in TCM theories. It also suggests the possibility of laying a statistical foundation for TCM through data analysis. The statistical model and latent clusters resulting from data analysis can be used in future research to improve diagnosis and treatment as follows: (1) Study the characteristics of each cluster and decide a treatment for it, (2) collect symptom information about a patient and use the model to compute the posterior probabilities of the patient belonging to various clusters (this is model-based diagnosis), and (3) combine the treatments prescribed for the individual clusters based on the probability values to formulate a treatment for the patient.

Acknowledgement

Research on this work was supported by Hong Kong Grants Council Grant #622105 and The National Basic Research Program (aka the 973 Program) under project No.2003CB517101.

References

1. Friedman, N. (1997). Learning belief networks in the presence of missing values and hidden variables. In *Proc. of 14th Int. Conf. on Machine Learning (ICML-97)*, 125-133.
2. Geiger D., Heckerman D. and Meek C. (1996). Asymptotic model selection for directed networks with hidden variables. . In *Proc. of the 12th Annual Conference on Uncertainty in Artificial Intelligence (UAI-96)*, 283-290.
3. Lazarsfeld, P. F., and Henry, N.W. (1968). *Latent structure analysis*. Boston: Houghton Mifflin.
4. Normile, D. (2003). The new face of traditional Chinese Medicine. *Science* 299: 188-190.
5. Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference* Morgan Kaufmann Publishers, Palo Alto.
6. Yang, W., F. Meng and Y. Jiang (1998). *Diagnostics of Traditional Chinese Medicine*. Academy Press, Beijing.
7. Schwarz, G. Estimating the dimension of a model, Annals of Statistics, 6(2): 461-464, 1978.
8. Zhang, N.L. (2004) Hierarchical latent class models for cluster analysis, *Journal of Machine Learning Research*, 5(6): 697–723.
9. Zhang, N.L. and T. Kocka (2004). Efficient Learning of Hierarchical Latent Class Models. *Proc. of the 16th IEEE International Conference on Tools with Artificial Intelligence*, Boca Raton, Florida.