MENTOR: A Bayesian Model for Prediction and Intervention in Mental Retardation

Subramani Mani and Marco Valtorta*
Department of Computer Science
University of South Carolina
Columbia, SC 29208

Suzanne McDermott
Department of Family and Preventive Medicine
University of South Carolina School of Medicine
Columbia, SC 29208

1 Problem Statement and Background

Mental Retardation(MR) or mental deficiency is a complex medical and social problem. The prevalence is estimated to be about 2.5 per cent of the population [Bats93], [StSu92]. Various studies have reported somewhat different figures(between 2% and 5%) depending on the definition of MR adopted and the age group surveyed [StSu92]. It is a developmental disability with a complex etiology. The causative factors and mechanisms are not well understood. According to the American Association on Mental Retardation (AAMR), "Mental Retardation is characterized by significantly subaverage intellectual functioning" [AAMR92, page 5]. The AAMR has recommended that people scoring below two Standard Deviations (SD) in a standardized IQ test be classified as retarded [AAMR92, page 5]. These tests are normalized to a mean of 100 with a SD of 15. Those with scores below 50 are considered severly retarded. Scores in the category of 50-69 fall in the classification of Mild Mental Retardation (MMR). Though AAMR suggests inclusion of limitation of adaptive skills also [AAMR92, page 6], many studies have used cognitive tests (IQ scores) for classification [StSu92], [McDe93].

A category called Borderline Mental Retardation (BMR)–scores falling between one and two standard deviations, was in vogue previously. But due to the social stigma attached to MR and concerns about test errors, it was de-emphasized subsequently.

We shall go by IQ scores and keep the category of BMR for understanding causal mechanisms. For severe MR a cause can be found in the majority of cases. In MMR, which forms 85% of MR, a cause cannot be put down in half the cases [Bats93].

So here we have a complex web of unknown causal mechanisms, disagreement among experts, controversies (the large literature of nature versus nurture) and serious gaps in the experts' understanding of the etiological factors.

1.1 Why a Bayesian Model?

A Bayesian modeling approach will shed some light on the causal mechanisms, give us a tool for prediction of MR and open up avenues for early intervention—medical and social.

Bayesian Networks (also referred to as Causal Probabilistic Networks) came into prominence in the eighties [Pear88], [Neap90], [LaSp88], [Coop84]. They continue to generate

*E-mail address for correspondence: mgv@cs.scarolina.edu

a lot of interest in the nineties. There are many Bayesian Expert Systems in the medical arena— Munin [AWFA87], ACORN [Wyat89], Expert Systems(ES) for Hematologic diagnosis [NDPS92], Diagnostica [BCWB93] and PATHFINDER [Heck92]. ES shells have also been designed to facilitate easy construction of Bayesian Network applications. HUGIN [AOJJ89], IDEAL [SrBr90] and BAIES [Cowe92] are three such shells.

A Bayesian Network(BN) is a Directed Acyclic Graph(DAG) with nodes representing variables and the directed edges(arrows) signifying a cause and effect relationship [Pear88], [Neap90]. This qualitative structure is quantified by assigning the conditional probabilities [Pear88], [Neap90]. There is some controversy regarding the use of the term causal network and a caution has been expressed in this regard by authors who prefer to emphasize dependency and avoid the more suggestive term causality [Sobe93]. We expect a Bayesian model to shed light on some of the causal mechanisms and will take the Bayesian network to be a causal model, unless we have reason to believe otherwise by examining it.

Causal Models are potential intervention models. It is possible to identify intervention links and optimize the benefits of intervention using such models.

> We believe the special value of DAG causal models is in predicting the results of interventions that change the distribution of variable values in a population. ...They are often the very point of causal models in studies that aim to influence policy. Such predictions can be made if one knows the causal structure of the systems in the population and understands the direct effects of intervention. Unlike prediction within a fixed distribution, predictions of the outcomes of interventions absolutely require the use of the causal relations represented in the directed graph. Regression or other methods which take no account of causal structure will not suffice. [GlSp93, page 255]

2 MR models from literature

The most recent model from the literature is the one developed by McDermott and Altekruse [McAl94]. It is a dynamic model explaining prevalence of MR and changes in prevalence based on demographic factors and child health policy decisions. The model stresses socio-economic variables—poverty and deprivation as contributing to MMR. But this model cannot be applied to individual cases.

MR-Expert [HCGD93] is a rule based expert system. It has been implemented as a decision support system to impart guidance in handling violent behaviors displayed by certain individuals with MR. It has the capacity to provide on-line literature review also.

Claire and Greenspan & Gransfield discuss conceptual models of MR [Clai89], [GrGr92]. These two models contribute to the definition of mental retardation but do not fall into the class of models we are interested in.

3 Our approach to model building

Basically there are two methods of building a Bayesian Expert System (BES):

1. Asking the domain expert to construct the network(DAG) and assign the prior probabilities.
2. Building the network from data. There are a few algorithms available to accomplish this— BIFROST [LaTS93], K2 [CoHe92] and CB [SiVa93]. The prior probabilities can also be computed from data. The models are validated by comparing with the performance of an expert [SDLC93].

We use a combination of the two strategies—to capture the network from data using the CB algorithm and prune the DAG with the help of the expert and published literature. Prior probabilities are obtained from data and fine tuned by the expert.

3.1 Model building tools

We now present some salient features of our dataset, the algorithm selected for our model building and the Bayesian Expert System shell to be used.

3.1.1 Dataset: Child Health & Development Studies

The Child Health & Development Studies(CHDS) is a prospective longitudinal study of pregnant mothers and the outcome of the pregnancies. The offsprings were also followed up and studied at various stages of growth and development. The study was initiated by the School of Public Health, University of California at Berkeley, in collaboration with the Kaiser Foundation. The area covered was the Bay area and the number of pregnancies studied exceeds 20,000. The very poor and the immensely affluent fell outside the purview of the study. Mothers were interviewed at the time of their first ante-natal visit and enrollment started. The data is based on personal interviews and medical charts. The study started in 1959 and continued into the eighties. The quality of the study is borne out by the more than hundred publications from this dataset until 1987. Various kinds of data pertaining to the mother, father and child are spread over twenty six files [CHDS87].

Variable selection About fifty relevant variables thought to play a role in MR have been identified from the following files.

BASIC: Family Background & Pregnancy Outcome
MATCOND: Maternal Conditions and Drugs Administered
PNVISIT: Prenatal Visits
DELIVERY: Labor and Delivery
ANOMALY: Congenital Anomalies in the Newborn
FIVEYR2: Developmental Examination in Five Year Olds, which includes IQ scores of children
NINETO11: Nine to Eleven Year-Old Examination, which includes IQ scores of children and their mothers.

The above information is taken from [CHDS87]. The total number of children—five year-olds and nine year-olds with IQ scores is approximately six thousand. There are about 3000 mothers with IQ scores.

3.1.2 Algorithm: CB Algorithm

The CB algorithm uses Conditional Independence tests (χ^2 tests) for ordering the nodes and uses a slightly modified K2 algorithm [CoHe92] to capture the DAG [SiVa93].

K2 is reported to be able to handle missing data [CoHe92] but no experimental evaluation on a large data set with missing values is given. The CB algorithm has been evaluated on some datasets with missing values and the performance is reported to be satisfactory.

3.1.3 Bayesian Expert System Shell: HUGIN

HUGIN provides a graphical interface for representing the nodes(domain variables) and the directed edges(causal relationships between the variables). A user friendly mechanism for naming the variables, entering the states of the variables and also to assign the conditional probabilities is provided. HUGIN has implemented the Lauritzen & Spiegelhalter method of probability propagation in DAGs [LaSp88]. The HUGIN shell was developed by Andersen, Olesen, F.V.Jensen & F.Jensen in Denmark [AOJJ89].

3.2 Modeling Strategy and Validation

The dataset will be randomly partitioned into two—one containing 80% and the other 20% of records. The bigger partition will be used for constructing the network and the smaller set for validation. (We note that the CB algorithm uses cross-validation to set certain parameters.)

A important issue that we are addressing is that of missing values [DeLR77]. We have 6000 children with IQ scores but only 3000 mothers with IQ scores. We have run the CB algorithm on the major partition of the whole dataset and built Net_One. Another dataset will be created comprising the records with mothers' IQ scores. This will also be randomly partitioned as discussed above. The algorithm will be run on the major partition building Net_Two. Net_One and Net_Two will be superimposed, compared and refined with the help of the expert. We have also implemented an algorithm (which we called IMP) to impute missing values in the original dataset. Initial results with IMP are very promising.

The Network will be validated by the records not used in building the Net. Further fine tuning will be done under the guidance of the domain expert. A second validation will be done using data on MR from South Carolina. The following three rules will be used to evaluate the network: (1) the rule of chronology; (2) the rule of commonsense; (3) the domain rule, which in this case is the rule of biological plausibility.

As we write this extended abstract, we have identified the data to be used and the medical expert (Dr. McDermott), we have built a Bayesian network from the data using the CB algorithm, and we have presented the network to the expert for critiquing. We have also implemented an algorithm to impute missing data, which we have used to preprocess our data set. We expect to further refine and test the Bayesian network, using Hugin, before the January workshop.

References

[AAMR92] American Association on Mental Retardation. *Mental Retardation* American Association on Mental Retardation, 1719 Kalorama Road, NW Washington DC, 1992.

[AOJJ89] Andersen, Stig K., Kristian G. Olesen, Finn V. Jensen, and Frank Jensen. "HUGIN–a Shell for Building Bayesian Belief Universes for Expert Systems." *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence,* Detroit, 1080-1085, 1989.

[AWFA87] Andreassen, S., M. Woldbye, B. Falck, and S.K. Andersen. "MUNIN–A Causal Probabilistic Network for Interpretation of Electromyographic Findings." *Proceedings of the Tenth International Joint Conference on Artificial intelligence (IJCAI-87),* Milan, 366-372, 1987.

[Bats93] Batshaw, Mark L. "Mental Retardation." In: *Pediatric Clinics of North America–The Child With Developmental Disabilities, 40, 3, 507-522.* Batshaw, Mark L, ed. Philadelphia, PA: W.B Saunders Company 1993.

[BCWB93] Blinowska, Aleksandra., Gilles Chattellier, Adam Wojtasik, and Jacques Bernier. "Diagnostica—A Bayesian Decision-Aid System—Applied to Hypertension Diagnosis." *IEEE Transactions on Bio-Medical Engineering, 40,* 3, 230-236, March 1993.

[CHDS87] *Data Archive and Users Manual of the Child Health and Development Studies, 1* School of Public Health, University of California at Berkeley, CA, March 1, 1987.

[Clai89] Claire, Lindsay St. "A Multidimensional Model of Mental Retardation: Impairment, Subnormal Behavior, Role Failures, and Socially Constructed Retardation." *American Journal on Mental Retardation, 94,* 1, 88-96, 1989.

[CoHe92] Cooper, Gregory F. and Edward Herskovits. "A Bayesian Method for the Induction of Probabilistic Networks from Data." *Machine Learning, 9,* 309-347, 1992.

[Coop84] Cooper, Gregory F. "NESTOR: A Computer Based Medical Diagnostic Aid that integrates Causal and Probabilistic Knowledge." Technical Report HHP-84-48, Medical Computer Science Group, Stanford University, Stanford, CA, 1984.

[Cowe92] Cowell, R.G. "BAIES: A Probabilistic Expert System Shell with Qualitative and Quantitative Learning." In: *Bayesian Statistics 4,* J.M. Bernardo, J.O. Berger, A.P. David and A.F.M. Smith, eds. 595-600. Clarendon Press, Oxford.

[DeLR77] Dempster, A.P., N.M. Laird, and D.B. Rubin. "Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society Section B,* 39, 1, 1–28.

[GlSp93] Glymour, Clark and Peter Spirtes. "Comment: Conditional Independence and Causal Inference." *Statistical Science,* 8(1993), 3, 250-255.

[GrGr92] Greenspan, Stephen and James M. Granfield. "Reconsidering the Construct of Mental Retardation: Implications of a Model of Social Competence." *American Journal on Mental Retardation, 96,* 4, 442-453, 1992.

[HCGD93] Hile, Mathew G., Donna M. Campbell, Bagher B. Ghobary and Marcie N. Desrochers. "Development of knowledge bases and the reliability of a decision support system for behavioral treatment consultation for persons with mental retardation: The Mental Retardation-Expert." *Behavior Research Methods, Instruments, & Computers,25(1993),* 2, 195-198.

[Heck92] Heckerman D.E., E.J. Horvitz and B.N. Nathwani. "Toward Normative Expert Systems: Part I The Pathfinder Project." *Methods of Information in Medicine, 31(1992),* 90-105.

[LaSp88] Lauritzen S.L. and D.J. Spiegelhalter. "Local Computation with probabilities on Graphical Structures and their Applications to Expert Systems." *J.R.Statist. Soc.,* B 50, 157-224, 1988.

[LaTS93] Lauritzen S.L., B. Thiesson, and D. Spiegelhalter. "Diagnostic Systems Created by Model Selection Methods–A Case Study." *Preliminary Papers of the Fourth International Workshop on Artificial Intelligence and Statistics,* Ft. Lauderdale, FL, January 3-6, 1993, 93-105.

[McAl94] McDermott, Suzanne W. and Joan M. Altekruse. "Dynamic Model for Preventing Mental Retardation in the Population: The Importance of Poverty and Deprivation." *Research in Developmental Disabilities, 15,* 1, 49-65, 1994.

[McDe93] McDermott, Suzanne W., Ann L. Coker and Robert E. McKeown. "Low birthweight and risk of mild mental retardation by ages 5 and 9 to 11." *Pediatric and Perinatal Epidemiology, 1993, 7,* 195-204.

[NDPS92] Nguyen D.T., L.W. Diamond, G. Priolet and C. Sultan. "Expert System Design in Hematology Diagnosis." *Methods of Information in Medicine 1992; 31:* 82-89.

[Neap90] Neapolitan, Richard E. *Probabilistic Reasoning in Expert Systems: Theory and Algorithms.* New York: John Wiley and Sons, 1990.

[Pear88] Pearl, Judea. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* San Mateo, CA: Morgan Kaufmann, 1988.

[SDLC93] Spiegelhalter, David J., A. Philip Dawid, Steffen L. Lauritzen, and Robert G. Cowell. "Bayesian Analysis in Expert Systems." *Statistical Science,* 8(1993),3, 219-283.

[SiVa93] Singh, Moninder and M.Valtorta. "A New Algorithm for the Construction of Bayesian Network Structures from Data." *Proceedings of the 9th Conference on Uncertainty in Artificial Intelligence (UAI-93),* Washington, DC, July 1993, 259-264.

[Sobe93] Sobel, Michael E. "Comment: Bayesian Analysis in Expert Systems." *Statistical Science,* 8(1993),3, 269-273.

[SrBr90] Srinivas S. and Breese J. "IDEAL: a software package for the analysis of influence diagrams." In: *Uncertainty in Artificial Intelligence.* L.N. Kanal, J. Lemmer and T.S. Levitt, eds. Elsevier North-Holland, 6, 1990, 212-219.

[StSu92] Stein, Zena A. and Mervyn W. Susser. "Mental Retardation." In: *Public Health & Preventive Medicine 13th ed.* Last, John M. and Robert B. Wallace, eds. Appleton & Lange, San Mateo, CA, 1992.

[Wyat89] Wyatt, Jeremy and David Spiegelhalter. "The Evaluation of Medical Expert Systems." In: *Advanced Models of Cognition for Medical Training and Practice. NATO ASI Series Vol.F97,* David A. Evans and Vimla L. Patel eds. Springer-Verlag NY 1992.