Graphical Modelling in Mental Health Risk Assessments

Olufunmilayo Obembe, Christopher D. Buckingham
School of Engineering and Applied Science
Aston University
Aston Triangle, Birmingham B4 7ET, UK
obembeo@aston.ac.uk, c.d.buckingham@aston.ac.uk

***Abstract* - Probabilistic models can be a combination of graph and probability theory that provide numerous advantages when it comes to the representation of domains involving uncertainty. In this paper, we present the development of a chain graph for assessing the risks associated with mental health problems, which is a domain that has high amounts of inherent uncertainty. The Galatean mental health Risk and Social care Tool, GRiST, has been developed to support mental-health risk assessments by using a psychological model to represent the expertise of mental-health practitioners. It is a hierarchical knowledge structure based on fuzzy sets for reasoning with uncertainty. This paper describes how a chain graph can be developed from the psychological model to provide a probabilistic evaluation of risk that complements the one generated by GRiST's clinical expertise.**

Keywords- mental health risk assessment; probability graphs; chain graphs.

I. Introduction

Risk assessment is a fundamental part of life, whether it be a mundane decision about the chance of rain or a much more vital one about the risk of a nuclear power station malfunctioning. In the mental-health domain, predicting whether someone is going to commit suicide or engage in an act of violence is extremely difficult, partly because the likelihoods are so low but also because of the lack of statistical data. The Galatean mental-health Risk and Social care assessment Tool (GRiST, [1]) was developed to address these problems by modelling how expert mental-health practitioners make risk assessments. However, its accumulating database of risk data has become a resource for more probabilistic approaches such as probability graphs, which are well-suited for capturing and reasoning with uncertainty where there is prior knowledge structuring [2].

In the past, mental health risk assessment was predominantly carried out using unstructured clinical approaches but it has since been realised that the best results can be obtained by using a combination of both structured clinical judgements and actuarial tools, such as one based on a probability graphical model [3]. This paper explores how a probability chain graph can be developed from the GRiST model of expertise to perform mental health risk assessments. In Section II, a brief overview of related work is given. In Section III, the background to mental-health risk assessment and GRiST is briefly introduced, followed by a discussion of the types of building blocks used in probability graphs in Section IV. In Section V, the development of the composite GRiST probability chain graph from these building blocks is presented. The paper then concludes with an outline of future work for the research.

II. Related Work

The challenge of providing effective risk assessment in the mental health domain is not a new one. The two broad approaches used are the clinical and actuarial approaches. The clinical approach can be split into structured and unstructured methods. Both are based on a clinician's experience but the unstructured method has no other input and is thus highly subjective. The structured clinical approach is more formal because it links the clinician's judgement with data-gathering tools that help guide the collection of information. In contrast, the actuarial approach uses statistical methods to provide risk assessments and is the least subjective of the approaches. Proponents of each method have argued about the various advantages and disadvantages over many years [4-6] but current policy is to integrate them where possible.

One of the actuarial methods (the category to which this paper belongs) uses an important technique from computer science: the multiple iterative classification tree (ICT) model, which is part of the MacArthur Violence Risk Assessment method [7]. This tool was developed to predict the risk of violence behaviour among recently discharged patients [7]. The results obtained from the ICT model had a high level of accuracy for the specific population group but the tool is resource and time intensive [8]. Another model uses tree mining [9] but a problem with all these approaches is the difficulty of obtaining clinical data that covers multiple risks and probabilities that can be generalised across populations.

Although GRiST is a model based on structured clinical judgement, it collects comprehensive and precisely defined data for all risks that are automatically stored in a database and thus available for probabilistic analyses. The idea is to link its clinical judgements to actuarial analysis and create a risk tool that explicitly connects the two risk-assessment approaches. Furthermore, the hierarchical structuring of GRiST's knowledge base provides the potential for

 ISBN: 978-1-61208-108-3

informing the structure of corresponding probabilistic graphs, which have limited current use in the mental health domain. GRiST allows both causal and non-causal relationships to be modelled between the various risk factors (cues), which helps provide a more accurate domain representation. This paper explores how to exploit GRiST for creating a probabilistic graph model of risk assessment and thus link probabilistic analyses to its fuzzy-set modelling of clinical reasoning. See [10] and [11] for a review of some identified advantages for using graphical models in risk assessments.

III. THE GALATEAN RISK SCREENING TOOL, GRIST

GRiST is a decision support system for mental health risk assessments [12] and is based on a psychological model of classification known as the Galatean model [13]. The variables modelled in the GRiST structure and the relationships between them are represented by a hierarchical tree structure (see Figure 1) and uncertainty is processed using fuzzy-set membership grades (MGs). In the GRiST knowledge structure (GKS), there are two main types of nodes (variables), namely concept nodes and datum nodes. Datum nodes are leaf nodes in the tree and thus do not have children. They represent measureable input data such as a person's *insight into behaviour* (Figure 1). Concept nodes are made up of two or more component nodes, which could be datum nodes or other concept nodes (e.g. *substance misuse* in Figure 1)

.

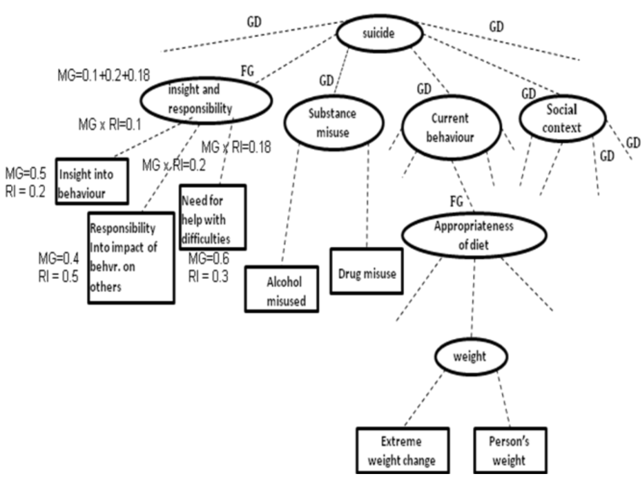

Figure 1. Small subsection of the GRiST Knowledge structure with datum and concepts nodes represented as rectangles and ovals respectively and generic nodes by g and generic distinct nodes by gd.

Both datum and concept nodes can be subcategorised into types that describe their locations and contextual behaviour in the GKS. Nodes that occur only once are called non-generic whereas those that have multiple occurrences are named generic. Some generic nodes are additionally distinguished as "generic distinct" because these have different uncertainty parameters for different locations in the tree whereas the plain generic nodes have exactly the same internal uncertainty representation wherever they occur.

The full GKS was originally elicited from 46 domain experts and comprised of 3,026 nodes with 338 unique concept nodes and 692 unique datum nodes [12]. Subsequent knowledge validation reduced it considerably so that only about 220 datum nodes are identified for data collection.

A. *GRiST Data Structures and Uncertainty Processing*

In GRiST, uncertainty is encapsulated by MGs and relative influences (RIs). An MG represents the degree of membership of an object in a node of the tree, with each nodes's MG ultimately contributing to the top-level risk membership (e.g., *suicide* and *self harm*). The actual contribution depends on the node's RI, which weighs the node's influence compared to its siblings. For example, from Figure 1, *insight into behaviour* has three children with RIs of 0.2, 0.5, and 0.3, which filter their MG contributions to the parent node as shown.

The MG distributions of the datum nodes were specified by clinical experts and enable patient cues to map to MGs that feed through the entire GRiST tree, from leaf nodes to the top-level risks. Equation 1 shows the formula used but see [13] for further details on the MG propagation process:

$$MG(X) = \sum_{p=1}^{P} (MG(datum_p) \times \prod_{l=1}^{L} RI_{lp}) \qquad (1)$$

Using Equation (1), the MG of each concept node is calculated by multiplying the MG of the datum node along each path p with all the RIs along that path and up each level l leading to the concept, and then summing the value obtained with all the corresponding values obtained along each path to the concept. An example of this is depicted on the left branch of the tree in Figure 1, to show how the MG is generated for the *insight and responsibility* concept.

B. *GRiST Knowledge Structure (GKS) Constraints and Independence Properties*

Having described the GKS and its data types, we now briefly outline the constraints incorporated in this knowledge structure and their correlation to various independence properties. A brief description is given below but see [14] for a more in-depth discussion. In this paper, we extend the work presented in the earlier paper [14] and expand on the constraint mapping and structure combination rules. Semantically there are three different types of relationships that can exist between any two nodes in the GKS, as follows:

- *IS-A* relationships refer to a 'kind-of' relationship, where the children nodes are a type of the parent node and are thus associated through their common parent. An example of this type of relationship is the parent node *substance misuse* and its children nodes *alcohol misuse* and *drug misuse*.
- *Contribute-to* relationships refer to those where the children nodes 'contribute to' and influence the parent node. For example the relationship type between the parent node *constraints on suicidal behaviour* and its internal nodes *insight and responsibility* and *religious values/beliefs affecting suicide risk* (see Figure 1) are of type *contribute-to*

because the children nodes directly contribute to the value of the parent node. Of the three types of relationship, this is the only causal one.

- *Wrapper* relationships occur when the parent node serves as a form of container for the children nodes (the parent 'wraps' the variables together) rather than being a cohesive variable in its own right. In this type of relationship there is a correlation between the children and parent nodes but not one that can be assumed to be causal. For example the relationship types between the parent concept *general current behaviour* and its children nodes *appropriateness of diet*, *challenging behaviour*, *daily activity*, *reckless risk taking*, *sleep disturbance*, *unintentional risk making* and *uncharacteristic recent change in behaviour* is of the type *wrapper* (it is partially depicted in Figure 1).

For the purpose of mapping to the probability graphs, it is important to note that the *is-a* and *wrapper* relationship types map to non-causal links, whereas the *contribute-to* relationship type is causal. As a result of these semantic relations between the various node types in the GKS, it is possible to give a concise list of the various constraints that exist within the model. These constraints in turn aid in the formal definition of the set of component structures that the GKS can be decomposed into. These component structures are then mapped into probability graphs that will eventually form the building blocks for the resultant probability chain graph that will provide the model for inferring the final risk assessments.

Detailed discussion of the GRiST component structures can be found in [14] but for clarity some of the constraints are outlined next. In the definitions, a root concept refers to the highest ancestor node in the structure under discussion. So, for example, if the entire diagram depicted in Figure 1 is taken to be one structure, then its root concept node is the *suicide* node. The constraints are for certain aspects of the structure that must hold wherever it is located in the GKS.

Constraints Related to Generic Root node structures:

- The RI value of the root node varies.
- The MG value of the root node is fixed.
- The RI values of the internal nodes are fixed.
- Given that the MGs and RIs of all internal nodes are fixed then the point of reference (i.e. the context) for the internal nodes is their root node.
- Generic root structures must be kept as cohesive wholes. It is possible to have a cohesive whole within another cohesive node i.e., root concept of type generic with internal nodes of type generic.
- As every node within a root concept of type generic has a fixed RI and MG everywhere the root concept occurs, if one of the internal structures is of type generic distinct (defined later), it will also need to have fixed RIs and MGs values within the context of the root concept everywhere it occurs. This is not the default or usual behaviour of these nodes, as explained next, but is in fact a special case.

Constraints Related to Generic Distinct Roots

- The RI value of the root node varies.
- The MG value of the root node varies.
- The RI values of the internal nodes vary.
- If a generic distinct node has at least one node of generic type as an ancestor then the context (or point of reference) for the generic distinct node is the nearest ancestor of type generic. Otherwise, the context for the generic distinct node is its neighbouring nodes.
- In the case where all the internal nodes of a generic distinct root concept are of type generic, if all the MGs and RIs of these internal nodes are always fixed it is obvious that the root concept MG value cannot vary and will itself always be fixed too, which is incorrect behaviour for a generic distinct root concept. This therefore leads to the constraint that a root concept of type generic distinct cannot have all its internal nodes to be of type generic. To make it possible for the variation in the root concept's MG value in various locations, there must be at least one internal node of type generic distinct. This is seen to be true in the GKS and is a good test of the validity for generic distinct node definitions.

From these constraints we obtain two structures that the GKS can be broken down into and for which probabilistic equivalents can be acquired.

1. For the generic root node that always has the same uncertainty values regardless of its location in the GKS, the context (i.e., point of reference) for all contained nodes is its own root concept. Within (and only within) the root concept the uncertainty values of the internal nodes are fixed and always remain the same regardless of location.

The name given to this type of structure is the **fixed generic component structure** (FG). An example from Figure 1 is the generic root node *insight and responsibility* and its internal nodes.

2. The generic distinct structures have varying internal RIs and varying root concept MG. The context for these nodes are the neighbouring nodes where the neighbouring nodes refer to the root concept, all its internal nodes (descendents), the root concept siblings and the root concept parents (as the root concept MG varies). It is also dependent on the top risk in which it occurs (e.g., suicide, self harm and so on). This type of structure has been named the **generic distinct component structure** (GD). This structure has no generic ancestor or, more to the point, if it did, it behaves as an FG node and can be ignored as a GD concept.

Both the FG and GD structures are composite wholes that can contain other composite variables within them. An example of this is seen in Figure 1 where the GD structure

 ISBN: 978-1-61208-108-3

with root concept node *current behaviour* contains within it the FG structure with root node *appropriateness of diet*.

The next section explores the relationship between these component structures and the independence properties they represent and maps them to different probability graphs, which will serve as building blocks for the final probability chain graph.

IV. THE BUILDING BLOCK PROBABILITY GRAPHS

In this section we develop the building blocks by examining the independence properties of each of the component structures discussed in Section III.

A. Mapping the Fixed Generic Structure

As mentioned earlier, because in a FG structure the relevant context for the determination of the uncertainty values is the root concept node of the FG structure itself, the independence property of the FG structure is as follows;

The nodes in a FG structure (i.e., its root node and all its internal structures) are conditionally independent of all other nodes in the GKS.

The above can be modelled by a Markov Random Field (MRF) [15]. The local Markov property for a MRF structure (the property outlining its independence status) states that a variable is conditionally independent of all other variables given its neighbours [15]. This directly correlates to the independence property of a FG structure, if we replace the variable in the MRF with the FG structure and in a similar manner also define the neighbours to be the FG root concept and all its descendents. An example FG structure from Figure 1 is the *insight and responsibility* root concept and its internal nodes *insight into behaviour*, *need for help with difficulties* and *responsibility into impact of behaviour on others*.

We, however, also need to consider the different relationship types (i.e., *is-a*, *wrapper* and *contribute-to*) between the nodes. Normally the *contribute-to* type should be represented by a directed edge, because the relationship is causal, and both the *is-a* and *wrapper* types are represented with non-directed edges as there is no implied causality.

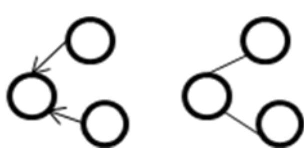

Figure 2. Directed and Undirected Graphs

However, the two diagrams in Figure 2 represent two different independence relations (see [16] for more on independence relations between the different graph types) but for FG structures these are not relevant within the MRF equivalent. Hence the relationships are all represented as undirected edges even when the relationship type between nodes is of type *contribute-to* (i.e., causal).

B. Mapping the Generic Distinct Structure

The GD structure is highly context sensitive and as such its independence properties are dependent on the identified relationship types between the various nodes in the structure. There are four possible relationship structures that can be obtained from the GD, namely: the non-causal to non-causal; the non-causal to causal; the casual to non-causal; and finally the causal to causal hierarchical relationship links. Recall that *is-a* and *wrapper* relationships are non-causal and the *contribute-to* relationship type is causal.

The independence properties of both the causal to non-causal and the non-causal to causal map to causally linked MRFs, which are in effect chain graphs (these will be discussed in more detail in the next section). The non-causal to non-causal on the other hand map to an undirected graph (MRF). For cases where this mapping is hierarchical, tree structured MRFs (TS-MRFs) [17] allow us to model such cases. And finally, the causal to causal links map to directed graphs (Bayesian Belief Networks) [18].

C. Summary of Building Block Probability Approach

To summarise the building block probability approach, a two step method is used. Initially each identified component structure (i.e., FG or GD) is represented in the overall graphical model as a composite variable. This results in an embedded model where each node itself represents a graphical model. The second layer, is then reached when we model and consider an individual sub-tree. For the FG structures we map the nodes to MRFs or their variants (as discussed earlier). The uncertainty contribution from each node (i.e., embedded graphical model) is then plugged into the overall graph. See [19] and [20] for other work involving different aspects of embedded graphs and mixture trees.

V. COMPOSITE PROBABILITY CHAIN GRAPH

Chain graphs are graphical models, which allow both directed and undirected graphs with the constraint that they do not have semi-directed cycles [21]. Linking two variables in a chain graph with a directed edge implies that the relationship between them is causal, and the direction of the edge is from cause to effect. On the other hand variables that are linked with an undirected edge do not have a causal relationship but however have an associative relationship (in a similar manner to MRFs). As a result of the inherent causal and associative relationships contained within the GKS, which are also clearly seen in the mapping to the building block probability graphs (discussed in Section IV), it makes logical sense to model this knowledge structure using a probability chain graph. More in-depth discussions on the chain graph can be seen in [22]

A. Development of the GRiST Chain Graph

The GRiST chain graph was developed from the building block probability graphs using a step by step combination of the graphs. The combining of multiple probabilistic graphical models requires care and one of the important considerations is to ensure that the independence properties represented in the different models being combined is maintained in the composite graphs (see [23] for a discussion on a framework for probabilistic graphical model combination).

B. Maintaining the Integrity of the Independence Properties

To ensure that the composite structure we have developed correctly models both a chain graph and a chain graph with the correct independence properties, some conditions have to be fulfilled.

- During the combination process, cycles and semi cycles must be avoided (to circumvent violating the chain graph constraint). This condition can be fulfilled because as a direct result of the GKS constraints and the GRiST hierarchical form, the graphical building blocks that are to be combined do not contain any cycles or semi-cycles. As such the only way that cycles can be introduced into the combined structure is if the components are combined in an opposite direction to the uncertainty propagation, which would not make logical sense.
- Secondly a chain graph needs to map into subsections (or blocks), with variables within a subset being linked via undirected edges and variables between blocks being linked via directed edges. Again this is possible because of the hierarchical form of the GRiST structure (and hence the building blocks' probability structures). The different levels in these knowledge structures map directly to the notion of subsets in chain graphs.

In addition to the above, to ensure that the integrity of the independence properties are maintained throughout the development process, the possible combination options between the various graphical building block graphs and their combination rules must be clearly defined, as described next.

Combining two FG Structures

Any two structures that need to be combined must be directly linked in the GKS and one of the two FG root nodes will be an ancestor to the other. Semantically this means that given the ancestor root concept and all the other nodes comprising the two FG structures (including the other FG root concept, which in effect is seen as an internal node to the ancestor root concept) the combined MRF structure (i.e., from both FG structures) is independent of all other nodes in the model. Therefore regardless of the relationship types between the nodes in these two structures, they should be linked using an undirected edge to obtain an MRF or TS-MRF (depending on the type of FGs combined). For the FG and FG structure combinations, if the combination rule does not override the relationship types between nodes in some instances, a directed graph might be used to combine two FGs and this will completely change the independencies represented in the GKS. Figure 3 depicts an example of the change in conditional independencies caused by the use of the wrong link type.

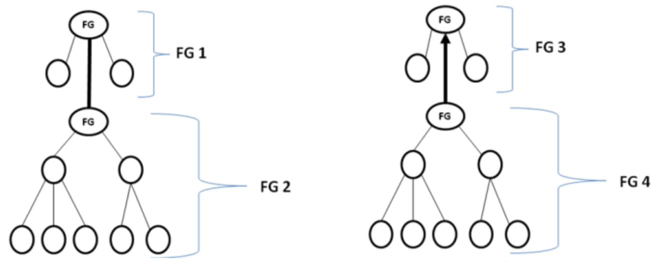

Figure 3. Change in conditional independence as a result of wrong link

In Figure 3, the combination of structures FG 1 and FG 2 results in a composite FG structure that is independent of all other nodes in the model. However, in the second structure (right hand diagram) as a result of the arrow used to link FG structures FG 3 and FG 4, a set of nodes are obtained within the composite combined FG that are conditionally independent of each other given their prior and concurrent nodes.

Combining a FG Structure and GD Structure

Semantically, when a FG structure is linked to a GD structure, from the definition of the FG structure, we see that it remains independent of the GD structure (recall that the composite FG structure is independent of all other nodes in the model). The GD structure on the other hand is dependent on its neighbours and thus will not be independent of the FG structure. The challenge here is defining the link in such a way that the GD dependence relations with the FG structure remain consistent. In this case the order of the combination is important. Where the FG structure is an ancestor to the GD structure, if an undirected link type is used to combine these two structures, the composite structure will be a FG structure (see Figure 4).

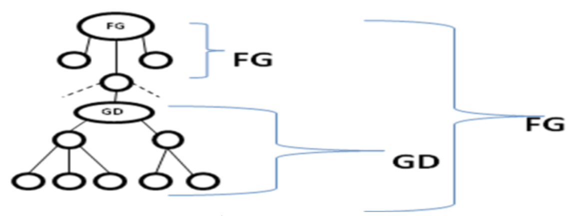

Figure 4. Combination of FG and GD structures resulting in a composite FG structure.

However where the GD structure is the ancestor structure, the relationship type between the two structures is needed to determine the link type. It is a directed link for the *contribute-to* relationship and undirected links for all other relationship types.

C. Summary of GRiST Chain Graph Development Steps

The following summarises the translation process from the GKS to the final composite GRiST chain graph.

Step 1: Partition the GKS into graphical building blocks (i.e., FG or GD).

Step 2: For each type of structure, identify the relationship types that exist between the various nodes (i.e., *is-a*, *wrapper* or *contribute-to*).

Step 3: Next identify the graphical building blocks that each component structure maps to, based on its independence properties.
Step 4: Repeat the above for the entire structure, until you are left with just graphical building blocks needing combination.
Step 5: Identify the type of combination to be carried out.
Step 6: Apply the relevant combination rules.
Step 7: The resultant chain graph should contain a mixture of explanatory, intermediate and response variables, and also maintain the independence properties of the original GKS.

VI. CONCLUSION AND FUTURE WORK

The development of a probability chain graph for mental health risk assessment has been presented in this paper. We have shown how the knowledge encapsulated in the psychological fuzzy-set based GRiST model can be mapped into initial component structures based on the inherent constraints in the model and how these in turn can be mapped to building block probability structures that can be eventually moulded together to construct a chain graph for mental health risk assessments.

The present solution is ongoing and future work will focus on using data collected by the tool in clinical use to validate the chain graph structure and learn its parameter settings. Examination of the dependencies between variables and their ontological definitions within GRiST will help determine whether the structural relationships are correct. Parameter estimation will then take place in two stages: estimating the potential functions for related cues in the MRF structures followed by estimating the conditional probability distributions for the nodes in the directed segments of the chain graph. Finally the identification of the most effective and efficient inference algorithms for the developed structure will be carried out. The methods discussed in this paper could be applicable to other systems based on hierarchical expertise, especially ones that contain both causal and non-causal relations.

REFERENCES

[1] See www.galassify.org/grist

[2] P. J. Lucas, “Bayesian Networks in biomedicine and health-care”, Artificial Intelligence in Medicine, vol 30, 2004, pp. 201-214.

[3] Department of Health, Best Practice in Managing Risk, Department of Health, London, 2007.

[4] B. Littlechild and C. Hawley, “Risk Assessments for Mental Health Service Users : Ethical, Valid and Reliable?”, Journal of Social Work, vol 10(2), pp 211-229, April 2010.

[5] T. R. Litwack, “Actuarial versus Clinical Assessments of Dangerousness”, Psychol, Pub Pol, Law, vol 7, pp 409-443, 2001.

[6] M. E. Rice, G. T. Harris, and V. L. Quinsey,”The Appraisal of Violence Risk”, Current Opinion in Psychiatry, vol 15(6), pp 589-593, 2002.

[7] J. Monahan, H. J. Steadman, P. C. Robbins, P. Appelbaum, S. Banks, T. Grisso, K. Heilbrun, E. P. Mulvey, L. Roth, and E. Silver, “An Actuarial Model of Violence Risk Assessment for Persons With Mental Disorders”, Psychiatric Services, vol 56(7), pp 810 – 815, 2005.

[8] J. Monahan, H. J. Steadman, P. S. Appelbaum, P. C. Robbins, E. P. Mulvey , E. Silver , L. Roth , and T. Grisso, “Developing a Clinically Useful Actuarial Tool for Assessing Violence Risk”, British Journal of Psychiatry, vol 176, pp 312-319., 2000.

[9] M. Hadzic, F. Hadzic, and T. Dillon, “Tree Mining in Mental Health Domain”, Proceedings of the Proceedings of the 41st Annual Hawaii International Conference on System Sciences, pp 230, 2008.

[10] R. Anderson, F. Camacho, F. Camacho, and R. Balkrishnan, “A Comparison of Risk Classification Methods in Medicare HMO Enrollee Health Risk Assessment”, Academy Health Update Health Services Research, vol 19(5), 2002.

[11] S. Ferson, “Bayesian Methods in Risk Assessment”, Applied Biomathematics Technical Report, 2005, Available: http://www.ramas.com/bayes.pdf. Last accessed 21.08.2010.

[12] C. D. Buckingham, A. Ahmed, and A. E. Adams, “Using XML and XSLT for Flexible Elicitation of Mental-health Risk Knowledge”, Medical Informatics and the Internet in Medicine, vol 32(1), pp 65-81, 2007.

[13] C. D. Buckingham, “Psychological Cue Use and Implications for a Clinical Decision Support System”, Medical Informatics and the Internet in Medicine, vol 27(4), pp. 237-251, 2002.

[14] O. Obembe and C. D. Buckingham, “Developing a Probabilistic Graphical Structure from a Model of Mental-Health Clinical Risk Expertise”, in Knowledge-Based and Intelligent Information and Engineering Systems, 14th International Conference, KES2010, R. Setchi, I. Jordanov, R. J. Howlett and L. C. Jain, Eds. in press.

[15] P. Perez, “Markov Random Fields and Images”, CWI Quarterly, vol 11(4), pp. 413-437, 1998.

[16] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.

[17] C. D'Elia, G. Poggi, and G. Scarpa, “A Tree-Structured Markov Random Field Model for Bayesian Image Segmentation”, IEEE Transactions on Image Processing, vol 12(10), October 2003, pp. 1259-1273.

[18] F. V. Jensen and T. D. Nielsen, Bayesian Networks and Decision Graphs, 2nd ed., Springer Publishing Company, Incorporated , 2007.

[19] K. Murphy and A. Nefian, “Embedded Graphical Models”, Intel Research Technical Report, June 2001.

[20] M. Meila and M. I. Jordan, “Learning with Mixture of Trees”, The Journal of Machine Learning, vol 1, pp. 1 – 48, 2001.

[21] M. Drton, “Discrete Chain Graph Models”, Bernoulli, vol 15(3), pp. 736-753, 2009.

[22] S. L. Lauritzen and N. Wermuth, “Graphical Models for Associations between Variables, Some of which are Qualitative and Some Quantitative”, Ann. Statist., vol 17, pp. 31-57, 1989.

[23] C. A. Jiang, T. Y. Leong, and K. L. Poh, “PGMC: A Framework for Probabilistic Graphic Model Combination”, in C. P. Friedman, J. Ash and P. Tarczy-Hornoch, Eds. Proceedings of the American Medical Informatics Association Annual Symposium (AMIA) 2005, pp 370-374.