PROBABILISTIC GRAPHICAL MODEL SUPPORTING EARLY DIAGNOSIS OF AUTISM SPECTRUM DISORDER

Justyna Szczygieł[1], Agnieszka Oniśko[1], Jolanta Świderska[2], Elżbieta Krysiewicz[2], Jerzy Sienkiewicz[3]

[1] Faculty of Computer Science, Bialystok University of Technology, Białystok, Poland

[2] Fundacja Oswoić Świat, Białystok

[3] Towarzystwo Przyjaciół Dzieci, Oddział Białystok

Abstract: Bayesian networks are recognized as a suitable tool for modelling diagnostic problems. The power of this modelling is that it can combine knowledge coming from different sources. For example, in case of medical domain, the expert knowledge can be merged along with the medical data. This paper presents a Bayesian network model for early diagnosis of autism. The model was built based on the medical literature and then was revised by two domain experts. Our tool is dedicated to parents that can perform an early diagnosis of their child before visiting a specialist.

Keywords: Bayesian networks, medical diagnosis, autism spectrum disorder

1. Introduction

The support of medical diagnosis by computer-based tools has a long history with the first approaches proposed in the 1960s and 1970s (e.g., [6,9]). The medical diagnostic support systems built in the last few decades were based on various approaches that can be divided into two categories: (1) statistical modeling and (2) artificial intelligence modeling that includes fuzzy sets, neural networks, decision trees, or probabilistic graphical models. Probabilistic graphical models such as Bayesian networks have proven to be powerful tools for modeling complex diagnostic problems involving uncertain knowledge. They have been employed in solving a variety of medical diagnostic problems reaching the size of hundreds or thousands of variables (e.g., [1,2,3,4,10]).

Advances in Computer Science Research, vol. 11, pp. 151-164, 2014.

Autism spectrum disorder (ASD) is a neurodevelopmental disorder with genetic origins that leads to an impaired social interaction. In the last decades, a dramatic increase of the ASD prevalence has been observed [5]. For example, the U.S. 2010 statistics show that the prevalence of ASD in children aged 8 is 1 in 68 [8]. ASD is not easy to diagnose especially in children before the age of 24 months. We know that the disease is more prevalent in males than in females with the ratio 4.5 : 1 [8]. However, there is no definitive cause of ASD identified, i.e., usually there is a combination of different risk factors and symptoms that have to be present to establish a final diagnosis. Early diagnosis is important since different types of therapy can improve child's development. For example, the therapeutic interventions can help the child to talk, walk, and communicate with others and then can increase child's chances for living independently in a society when they are adult.

There is no standard procedure or screening examination for early ASD diagnosis in Poland. Parents are often not aware of this disease and may overlook its first symptoms. We have proposed a tool — a web-based application to support early diagnosis of ASD. This tool is dedicated to parents that observe an odd behaviour in their child. The core of our application is the AutismNET model, a Bayesian network that was built based on the medical literature and then revised by two domain experts. The model allows for estimating the probability of developing ASD based on the observed signs and symptoms entered into the model. This probability can be further interpreted by parents suspecting ASD in their child. The AutismNET model was developed and presented in [13].[4]

The reminder of this paper is structured as follows. Section 1. provides a brief introduction to the problem of the ASD diagnosis. Section 2. presents an overview of Bayesian networks. Section 3. describes the AutismNET model and its application. Section 4. concludes the paper.

2. Bayesian networks

Bayesian networks [14] are acyclic directed graphs modeling probabilistic influences among variables. The graphical part of a Bayesian network reflects the structure of a modeled problem, while conditional probability distributions quantify local interactions among neighboring variables.

Figure 1 captures a simple Bayesian network model. This example model includes one risk factor and two symptoms of autism. A left hand side of the figure

[4] Justyna Pawłowska is a maiden name of Justyna Szczygieł.

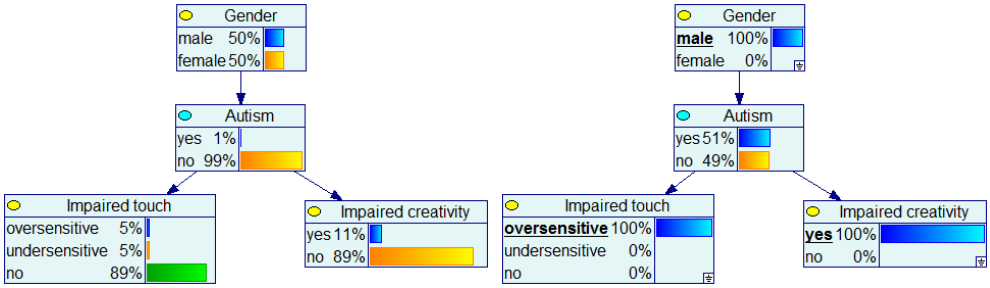

Fig. 1. A simple example of a Bayesian network

shows the model along with marginal probabilities for each node while a right hand side of the figure shows the same model but with observed three nodes and a posteriori probability distribution for the node *Autism*. Each arc of this graph represents a probabilistic relationship. For example, the arc between the variables *Gender* and *Autism* indicates that autism in males is more prevalent, i.e., males are around four times more probable to be diagnosed with autism than females. Furthermore, this simple example captures two possible symptoms of autism: (1) impaired touch and (2) impaired creativity. The numerical parameters of a Bayesian network model include a conditional probability distribution for the nodes that have parents (e.g., *Autism*, *Impaired touch*, and *Impaired creativity*) and a prior distribution for the nodes without parents (e.g., *Gender*). These probability distributions can be learned from the data or can be assessed by the domain experts.

After creating a Bayesian network model, we can perform a reasoning that involves calculating a posteriori probability distribution for the node *Autism* given the observations that were entered into the model. This calculation consists of repetitive application of a Bayes theorem that spreads over the network and leads to a derivation of conditional posterior probabilities in every node of the network. A right hand side of Figure 1 shows the result of such probabilistic reasoning and answers the question: *What is a probability of developing autism for a boy that has impaired creativity and that has oversensitive touch?* The probability of developing autism in this example model is equal to 51%.

3. The AutismNET model

The following section describes the process of building the AutismNET model for early diagnosis of ASD. The first part of the section shows a graphical structure of the

model while the second presents a quantitative part that includes conditional probability distributions. While referring to the nodes of a Bayesian network model we will use three terms: *target*, *observation*, and *auxiliary* indicating three different types of nodes. The category *target* stands for the nodes representing diagnoses, *observation* represents all these nodes that we would usually observe. For example, the nodes *Impaired touch*, *Impaired creativity*, and *Gender* in Figure 1 have a status *observation*. A type *auxiliary* indicates the nodes that we would never observe.

3.1 Graphical structure of the model

We have started building the model from browsing and studying the medical literature related to the ASD diagnosis.

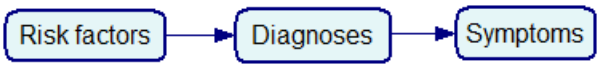

Fig. 2. The three layers of AutismNET

First version of the model The first version of the model was built based on knowledge encountered from the medical literature. A knowledge engineer identified 85 variables that were modeled in the framework of a Bayesian network model. The variables have belonged to three categories: (1) risk factors, (2) diagnoses, and (3) signs and symptoms. These three categories were mapped into three layers of the AutismNET model (see Figure 2). Additionally, we decided to group the variables within these three layers into submodels. This procedure helped us to organize the models' variables and to facilitate the process of navigation. The concept of a submodel is implemented in GeNIe [11] and it is simply a logical concept that does not introduce any additional relationships in the model.

Figure 3 presents the first version of the AutismNET model. The model consists of 85 nodes grouped in 14 submodels. A top layer of the model includes 9 submodels representing 43 risk factors while a bottom layer consists of 5 submodels representing 31 different symptoms. A middle layer represents two diseases: *Autism spectrum disorder* and *ADHD* (Attention Deficit Hyperactivity Disorder). We included *ADHD* as a part of differential diagnosis for *Autism spectrum disorder*.

One of the problems that we encountered during building the model was the number of parents per node. For example, the node *Autism spectrum disorder* had

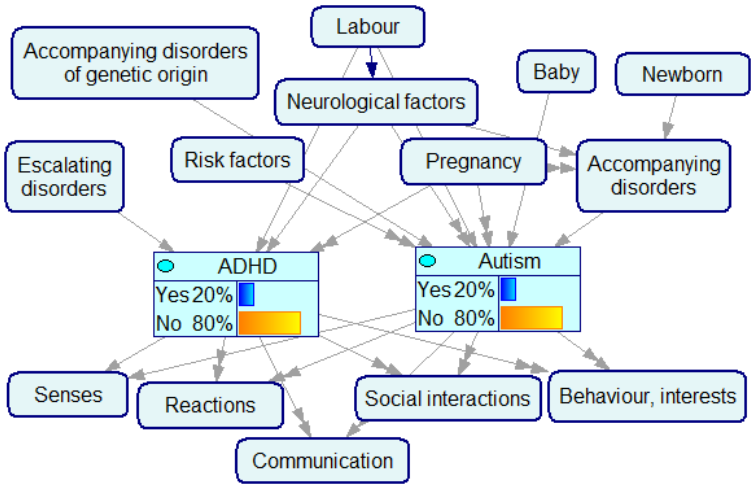

Fig. 3. The first version of the AutismNET model

initially 20 parents. Since it would be almost impossible to estimate the numerical parameters for the node with 20 parents,[5] we have applied a technique called in Bayesian networks as "parent divorcing" [12]. This technique helps to decrease a complexity of the network by reducing the number of parents per node. We have created 6 auxiliary nodes that divorced the parents of the node *Autism spectrum disorder*. The result of this procedure was a decrease of the number of parents for the node *Autism spectrum disorder* from 20 to 10. All the auxiliary nodes that we have created were modeled as the NoisyMAX gates. The advantage of applying the NoisyMAX gates is that we can estimate conditional probability distribution of a node with a smaller number of numerical parameters [7].

Figure 4 shows an example of "parent divorcing" that we have performed in the AutismNET model. The left hand side of the figure captures the four out of 20 parents of *Autism spectrum disorder*, while the right hand side of the figure shows the result of divorcing these parents. The auxiliary node *Labour complications* that divorced the parents was further modeled as a Noisy MAX gate.

[5] Assuming that all nodes are binary, the node with 20 parents needs 2^{20} independent probabilities to elicit.

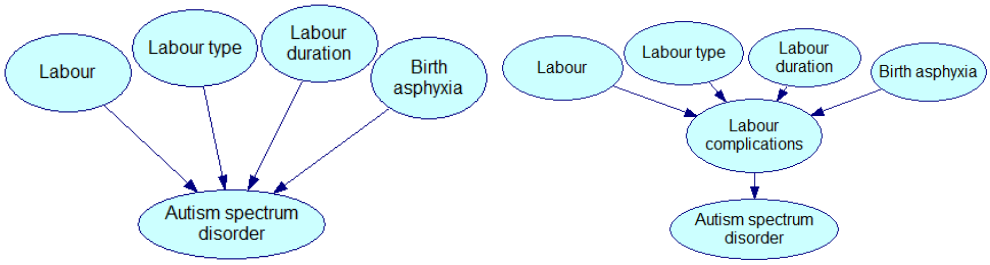

Fig. 4. An example of parent divorcing in AutismNET

Second version of the model After building the first version of the AutismNET model, we have scheduled five meetings with the experts[6] to verify it. Each meeting lasted around two hours. The first two meetings were devoted to verification of the model variables, while during the next two meetings we elicited the numerical parameters from the experts. The last meeting was devoted to model evaluation.

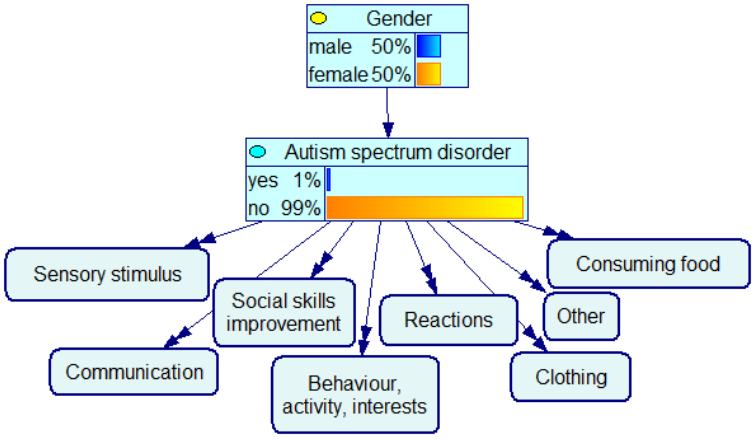

Fig. 5. The second version of the AutismNET model

For the first meeting we have prepared a list of the model's variables printed for each of the two experts — the variables were grouped by submodels. During this session with experts we went through this list and performed a clarity test for each

[6] The third and fourth author of this paper.

variable. The experts have excluded all the variables representing the risk factors of ASD, but the gender i.e., 43 nodes were removed from the model along with 9 auxiliary nodes. The experts claimed that risk factors that we had included in the model do not have much diagnostic value and that the ASD diagnosis should be performed mainly based on signs and symptoms. At the same time, the experts proposed to include in the model new variables, for example: *No response to name*, *No response to reading books by parents*, *Gestural communication*, *Impaired speech*. We included in the model 22 additional variables that along with previous variables were grouped into 8 submodels. The resulting model had three new submodels: *Clothing*, *Consuming food*, and *Other*. Three out of five submodels in *Version 2.0* changed their name, for example, the submodel *Senses* was changed to *Sensory stimulus*. The experts excluded also the variable *ADHD* claiming that this disease is not crucial in differential diagnosis of *Autism spectrum disorder*. During the second meeting with the experts, we again verified the model's variables – although this time it involved verification of variables states. For example, the variable *Impaired sleep* was initially modeled as a binary node with two states: *Short sleep* and *Normal sleep*. The experts claimed that it should be modeled as the variable with three states: (1) *Short sleep*, (2) *Interrupting sleep*, and (3) *Normal sleep*.

Figure 5 presents the second version of the model after two meetings with the experts. In fact, after removing from the model all the variables representing risk factors, the second version of AutismNET became a naive Bayesian network. The model consists of 50 nodes: one target node and 49 observation nodes.

Third version of the model Although our experts believed initially that only signs and symptoms play a significant role in a diagnosis of ASD, we agreed after a short discussion with them that it would be interesting to include the risk factors in the AutismNET model. Therefore, we have created the third version of the model that includes again three layers of the variables, i.e., risk factors, diagnoses, and signs and symptoms. Figure 6 presents the third version of the AutismNET model. The model consists of 100 nodes grouped in 16 submodels. Similarly to Figure 3, a top layer of the model includes the nodes representing risk factors while a bottom layer captures the signs and symptoms. The third version of the AutismNET model is essentially a hybrid of two previous versions: with a top layer of risk factors from *Version 1.0* and two bottom layers from *Version 2.0*.

Tables 1 and 2 present the properties for the three versions of the AutismNET model. Table 1 contains the information about the nodes and submodels of the three versions of AutismNET. For example, *Version 1.0* of the model has in total 76 nodes

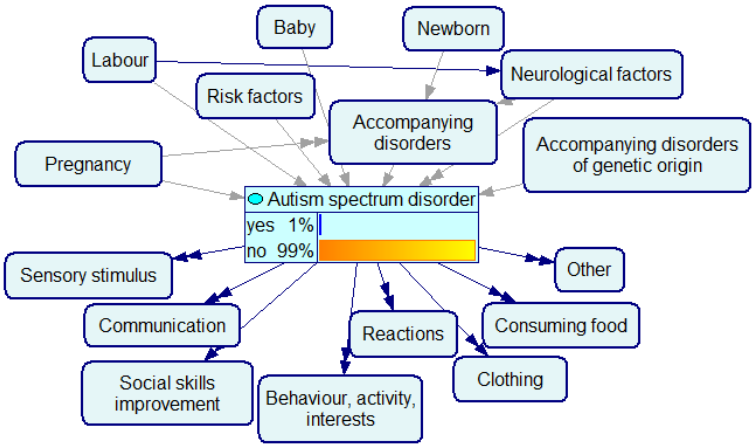

Fig. 6. The third version of the AutismNET model

modeled as CPTs[7] and 9 nodes modeled as NoisyMAX distributions. There were 2 target nodes, 74 observation nodes, and 9 auxiliary nodes; this version of the model was grouped into 14 different submodels. Table 2 captures additional structural statistics of AutismNET and shows a complexity of the models. For example, *Version 1.0* of the model has 92 arcs, on average 1.08 parents per node (*Avg indegree*), and a maximal number of parents equal to 10 (*Max indegree*). The model has on average 2.12 outcomes per node and a maximal number of outcomes per node is equal to 4.

Table 1. Characteristics of the AutismNET models; #nodes indicates the number of all nodes, #CPT stands for the number of nodes with the CPT distributions, #NoisyMAX stands for the number of nodes with the NoisyMAX distributions, etc.

Version	#nodes	#CPT	#NoisyMAX	#target	#observation	#auxiliary	#submodels
Version 1.0	85	76	9	2	74	9	14
Version 2.0	50	50	0	1	49	0	8
Version 3.0	100	90	10	1	89	10	16

[7] CPTs stands for Conditional Probability Tables

Table 2. Structural statistics of the AutismNET models; #arcs stands for the number of arcs

Version	#arcs	Avg indegree	Max indegree	Avg outcomes	Max outcomes
Version 1.0	92	1.08	10	2.12	3
Version 2.0	49	0.98	1	2.28	4
Version 3.0	102	1.02	10	2.19	4

3.2 Elicitation of numerical parameters

Quantitative part of Bayesian network model includes conditional probability distributions. These probability distributions can be learned from the data or assessed by a domain expert. There were no data available to us, therefore, we had to rely on the expert opinion while quantifying the model. In the first version of the model we simply assigned the distribution $(0.2, 0.8)$ for binary nodes, or a uniform distribution for non-binary nodes. The quantification of the model was conducted for the second version of the model. Two meetings with the experts were devoted to elicitation of the numerical parameters. During the first meeting we were posing the following type of questions: *What is the probability that a symptom is present if a child has ASD?* For example, we asked the following question:

What is the probability that oversensitive touch is present if a child has ASD?

During the second meeting we were posing the following type of questions: *What is the probability that a symptom is present if a child does not have ASD?* For example, we asked the following question:

What is the probability that oversensitive touch is present if a child does not have ASD?

During these two meetings, that lasted four hours together, we elicited 122 independent numerical parameters. We have noticed that it was easier for the experts to assess the parameters for the first scenario, e.g., when we were asking for the probability of a symptom being present if a child had ASD.

During elicitation of the numerical parameters we have identified the variable that was not significant in diagnosis of ASD. While assessing the probabilities for the variable *Mood swinging* we have noticed that the experts specified the same probability distribution for autistic and non autistic population. This led us to removing this variable from the model.

Figure 7 presents a fragment of AutismNET along with its probability distributions. The node *Gender* with two states *male* and *female* is described by a prior

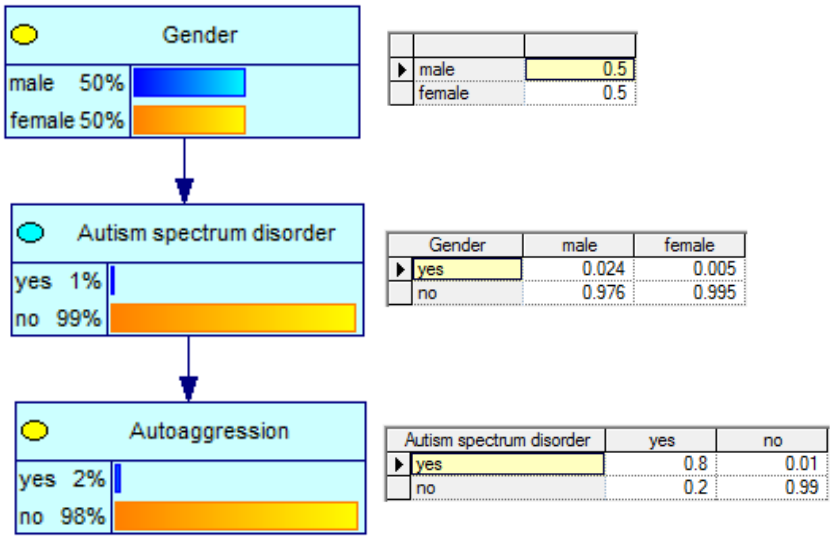

Fig. 7. A fragment of the AutismNET model along with conditional probability tables

probability distribution and it represents a general population distribution. The node *Autism spectrum disorder* is described by a conditional probability distribution and it reflects the relationship between gender and ASD. These numerical parameters were specified based on the published statistics of the ASD prevalence in a general population depending on a gender. The last node *Autoaggression*, again, is described by a conditional probability distribution that was elicited from our experts. In fact, to quantify this distribution, the experts had to specify only two independent numerical probabilities: 0.8 and 0.01.

Table 3 shows a summary of quantitative part of the AutismNET models along with the number of probabilities for target, observation, and auxiliary nodes. The table shows also the number of dependent probabilities that are part of the CPT or NoisyMAX distributions. For example, *Version 1.0* of the model has 2,551 dependent probabilities modeled by CPT and 96 dependent probabilities modeled by NoisyMAX.

Since the third version of the model is a hybrid of the two previous versions, only the variables from a bottom layer has the probabilities elicited by the experts.

Table 3. Numerical parameters of the AutismNET models

Version	CPT	NoisyMAX	Target	Observation	Auxiliary
Version 1.0	2,551	96	2,112	417	118
Version 2.0	228	0	2	226	0
Version 3.0	1,527	100	1,024	503	104

3.3 Model evaluation

We did not have any access to objective data to evaluate the AutismNET model, therefore, we have performed only a subjective expert evaluation. We were entering the data representing a typical autistic child and then the experts were observing how the a posteriori probability of developing ASD is changing. During this evaluation we also looked at the variables with the highest diagnostic value (calculated in GeNIe based on cross-entropy measure) and asked the experts whether indeed these variables are important in a diagnostic process. The experts had confirmed that the indicated variables have a high diagnostic value.

While playing with AutismNET, we have noticed that the model is too sensitive with respect to observed symptoms, i.e., after observing a few symptoms as present, the probability of ASD was approaching the value of 1.0. For example, after we had observed the following symptoms in the model: an oversensitive touch, unusual preoccupation with toys, and repeating unusual movements or actions, the calculated model probability of developing ASD was 99.9%. This value suggested that the model's probabilities need additional revision and refinement.

3.4 Application of AutismNET

We have built a web-based interface for the AutismNET model. This interface allows to access the model through Internet and to perform a diagnosis by answering the questions of a survey. These questions correspond to the variables modeled in AutismNET and they are grouped by submodels. This interface is dedicated to parents that would like to perform an initial diagnosis of their child. The model could possibly indicate a need for a more detailed diagnosis by a specialist. Figure 8 presents a screen shot of the application. A list of 8 elements on a left hand side of the window corresponds to 8 submodels from *Version 2.0* of AutismNET.[8] A right hand side

[8] Currently, only *Version 2.0* is fully quantified. Therefore, it is used in a web-based interface of AutismNET.

of the window shows a list of questions of the survey that correspond to the variables of the submodel *Behaviour, activity, interests*. The application is available at `http://www.autismnet.pl`.

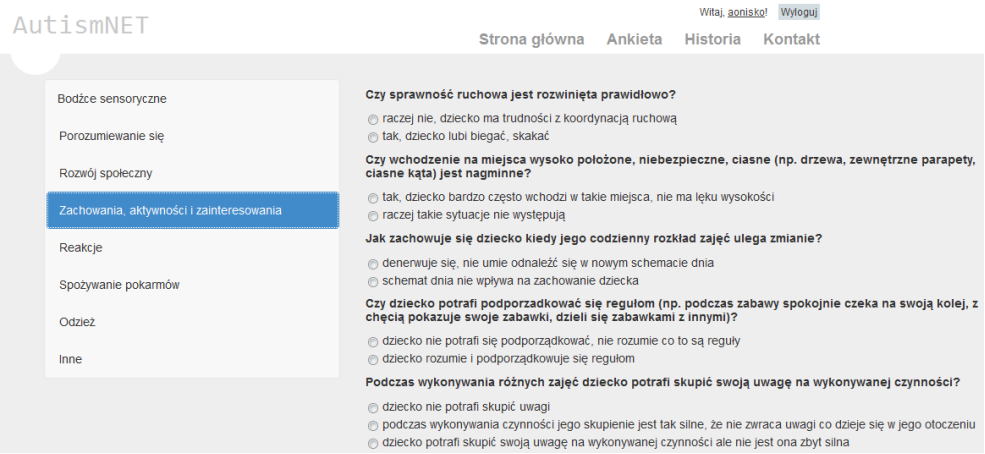

Fig. 8. Screen shot of AutismNET web-based application (in Polish)

4. Conclusions

We have built the AutismNET model for early diagnosis of ASD. The model was entirely built based on the medical literature and experts' knowledge. The model calculates the posteriori probability of developing ASD given entered observations. AutismNET has also a web-based interface that facilitates the interaction with the model. This tool is dedicated to parents that observe an odd behaviour in their child and suspect ASD.

Our project has several shortcomings that we plan to address in a future version of the model. The model requires a revision of the numerical parameters since they are too sensitive towards observed symptoms. A sensitivity analysis has to be performed to identify these numerical parameters of the model that should be further tuned. We also need to refine the variables representing risk factors modeled in AutismNET and then elicit the numerical parameters for the corresponding nodes.

It is an interesting research question whether a simple diagnostic model would perform better than a complete extended model. We plan to answer this question by

comparing the two models: the AutismNET model in *Version 2.0* consisting of 50 nodes and the AutismNET model in *Version 3.0* including 100 nodes.

We also plan to implement AutismNET user interface for a mobile device – this will even increase the availability of the model to its potential users.

Acknowledgments

The model was created and tested using SMILE, an inference engine, and GeNIe, a development environment for reasoning in graphical probabilistic models, both developed at the Decision Systems Laboratory, University of Pittsburgh and available at `https://dslpitt.org/genie/`.

References

[1] S. Andreassen, M. Woldbye, B. Falck, and S. K. Andersen. MUNIN – A causal probabilistic network for interpretation of electromyographic findings. In J. McDermott, editor, *Proceedings of the Tenth International Joint Conference on Artificial Intelligence*, pages 366–372, Los Altos, CA, 1987. Morgan Kaufmann Publishers.

[2] I.A. Beinlich, H.J. Suermondt, R.M. Chavez, and G.F. Cooper. The ALARM monitoring system: A case study with two probabilistic inference techniques for belief networks. In *Proceedings of the Second European Conference on Artificial Intelligence in Medical Care*, pages 247–256, London, 1989.

[3] Gregory F. Cooper. *NESTOR: A Computer-based Medical Diagnostic Aid that Integrates Causal and Probabilistic Knowledge*. PhD thesis, Stanford University, Computer Science Department, 1984.

[4] F. J. Díez, J. Mira, E. Iturralde, and S. Zubillaga. DIAVAL, a Bayesian expert system for echocardiography. *Artificial Intelligence in Medicine*, 10:59–73, 1997.

[5] E. Fombonne. Epidemiological trends in rates of autism. *Molecular Psychiatry*, 7:S4–S6, 2002.

[6] G. Anthony Gorry. Computer-assisted clinical decision-making. *Methods of Information in Medicine*, 12:45–51, 1973.

[7] Max Henrion. Some practical issues in constructing belief networks. In L.N. Kanal, T.S. Levitt, and J.F. Lemmer, editors, *Uncertainty in Artificial Intelligence 3*, pages 161–173. Elsevier Science Publishers B.V., North Holland, 1989.

[8] Developmental Disabilities Monitoring Network Surveillance Year 2010 Principal Investigators, Centers for Disease Control, and Prevention (CDC). Prevalence of autism spectrum disorder among children aged 8 years - autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *MMWR Surveill Summ*, 63(2):1–21, 2014.

[9] R. S. Ledley and L. B. Lusted. Reasoning foundations of medical diagnosis. *Science*, 130(3366):9–21, July 1959.

[10] Peter J. F. Lucas, Linda van der Gaag, and Ameen Abu-Hanna. Bayesian networks in biomedicine and health-care. *Artificial Intelligence in Medicine*, 30:201–214, 2004.

[11] Decision Systems Laboratory University of Pittsburgh. GeNIe and SMILE software available at `https://dslpitt.org/genie/`.

[12] K. G. Olesen, U. Kjaerulff, F. Jensen, F. V. Jensen, B. Falck, S. Andreassen, and S. K. Andersen. A MUNIN network for the median nerve — a case study on loops. *Applied Artificial Intelligence*, 3:384–403, 1989.

[13] Justyna Pawlowska. Probabilistic graphical model for diagnosis of autism (in Polish). Master's thesis, Faculty of Computer Science, Bialystok University of Technology, Bialystok, Poland, July 2014.

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

PROBABILISTYCZNY MODEL WSPIERAJĄCY WCZESNE DIAGNOZOWANIE AUTYZMU

Streszczenie: Sieci bayesowskie są często używanym narzędziem w rozwiązywaniu problemów diagnostycznych. Jedną z zalet tego narzędzia jest możliwość łączenia wiedzy pochodzącej z różnych źródeł. Na przykład, wiedza ekspertów może być połączona z danymi. W naszym artykule prezentujemy model sieci bayesowskiej wspomagający wczesne diagnozowanie autyzmu. Model został zbudowany w oparciu o literaturę medyczną, a następnie zweryfikowany przez ekspertów. Narzędzie, które stworzyliśmy jest dedykowane rodzicom, którzy mogą dokonać wstępnej diagnozy zanim skontaktują sie ze specjalistą.

Słowa kluczowe: sieci bayesowskie, diagnozowanie medyczne, autyzm

Artykuł zrealizowano w ramach pracy badawczej S/WI/2/2013.