БЪЛГАРСКА АКАДЕМИЯ НА НАУКИТЕ . BULGARIAN ACADEMY OF SCIENCES

КИБЕРНЕТИКА И ИНФОРМАЦИОННИ ТЕХНОЛОГИИ • Том **2**, № 2
CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **2**, No 2

София . 2002 . Sofia

A Generalized Net, Description for Laryngeal Pathology Detection Excluding the Refusal from Classification Option*

Stefan Hadjitodorov[1], *Petar Mitev*[1], *Krassimir Atanassov*[1], *Georgi Gluhchev*[2], *Vassil Vassilev*[2], *Anthony Shannon*[3]

[1]*Centre for Biomedical Engineering, 1113 Sofia, E-mails: sthadj@argo.bas.bg, pmmi@clbme.bas.bg, krat@argo.bas.bg*
[2]*Institute of Information Technologies, 1113 Sofia, E-mails: gluhchev@iinf.bas.bg, vasko@iinf.bas.bg*
[3]*KvB Institute of Technology, North Sydney, 2060, University of Technology, Sydney, 2007, AUSTRALIA, E-mail: tony@kub.edu.au*

Abstract: *A Generalized Net (GN — an extension of the Petri net) model of the process of laryngeal pathology detection is described. The model is the second one from a series of investigations on the application of the GNs to the speech analysis problems.*

Keywords: *Generalized net, Laryngeal pathology detection, Model.*

1. Introduction

Most of the laryngeal pathologies produce a change in the voice of the patient. An effective and non-invasive method for early diagnostics of the voice-producing system diseases is the acoustical analysis of the voice. That is why several systems and approaches [1–7] to laryngeal pathology detection based on voice acoustical analysis have been developed in the last few years. The researchers have used various parameters describing the pathological voices – Fundamental frequency F_0 and Pitch period T_0; various statistics of T_0 and F_0; Amplitude perturbation – *shimmer*; Pitch perturbation – *jitter*; Ratio of the harmonics energy to noise energy in the time and spectral domains – HNR_Y and HNR_Q; Degree of hoarseness (DH); Normalized noise energy (NNE); Turbulent noise index (TNI); Ratio of the first harmonic energy to the energy of the rest of harmonics – NFHE; Duration ratio of the non-vocalized to the vocalized part of the signal – DUV [8]; etc., and different recognition methods for classification of the patients.

* This work was supported partly by the National BulgarianFoundation for Scientific Research in the frameworks of project I-804/98.

One of the main drawbacks of these systems and approaches is the presence of a classification error and in particular the most dangerous error – classification of a patient with laryngeal disease as a normal speaker, the so called "false negative". In order to increase the accuracy of laryngeal pathology detection, some researchers use two level classification schemes. At the first level a number of classifiers make their classification decisions, and at the second level their results are combined in a proper way to obtain the final classification. In simpler schemes the classification decision may include the option (or class) "refusal to classify". In more sophisticated schemes a final definite decision (excluding the class of "refusals") is sought. Here, the process of acoustical analysis and a simple two level classification scheme excluding the refusal option have been modelled by generalized nets.

2. The model

Below we shall construct a reduced Generalized Net (GN; see [9] and short remarks – in [10]) model without temporal components, without transitions, places and tokens priorities and without places and arcs capacities (Fig. 1). In this GN the tokens keep all their history.

We shall describe the transition condition predicates and the tokens characteristics not fully formally for the sake of easier understanding of the formalism in use.

Initially, tokens β and γ are placed in places l_6 and l_9 with initial characteristics, respectively:

"DB of patients with known classification",
"list of classifiers"

while tokens noted by α enter sequentially place l_1 with initial characteristic

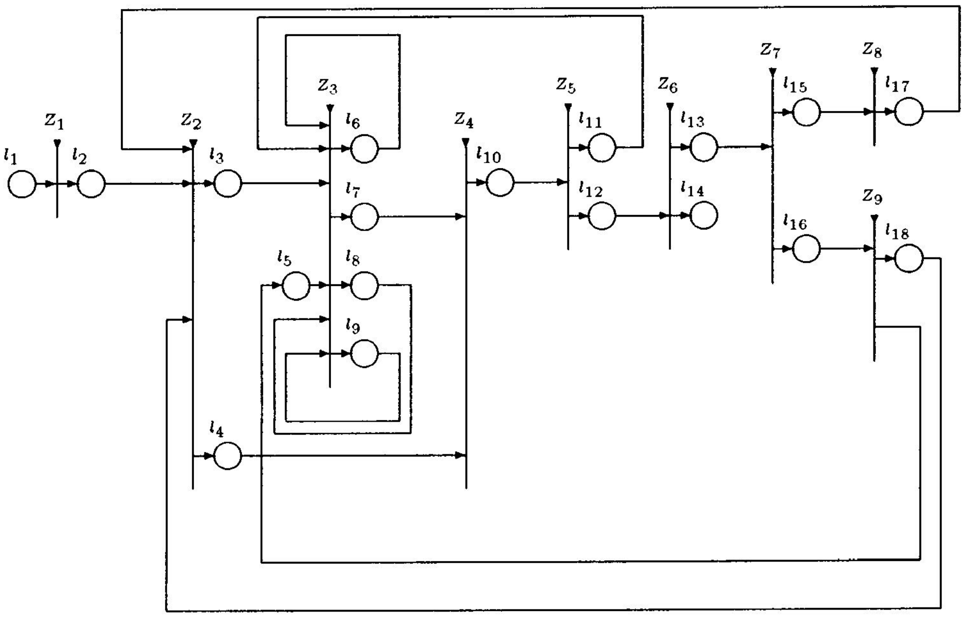

Fig. 1. GN-model

"list of patient's speech parameters; estimation of patient's status, (normal pathology, unknown)".

As well, for brevity we shall use the notation α_i for other tokens.

GN-transitions have the following forms.

$$Z_1 = \langle \{l_1\}, \{l_1\}, \begin{array}{c|c} & l_2 \\ \hline l_1 & true \end{array}, \vee(l_1) \rangle.$$

Each token α obtains characteristic

"vector with the results of a digital acoustical signal analysis of the patient's speech (feature vector, describing the current patient's speech)"

in place l_2.

$$Z_2 = \langle \{l_2, l_{17}, l_{18}\}, \{l_3, l_4\}, \begin{array}{c|cc} & l_3 & l_4 \\ \hline l_2 & W_{2,3} & W_{2,4} \\ l_{17} & false & true \\ l_{18} & false & true \end{array}, \vee(l_2, l_{17}, l_{18}) \rangle,$$

where $W_{2,3}$ = "there is information about the patient's status (healthy or ill)";
$W_{2,4} = \neg W_{2,3}$,
where $\neg\, p$ is the negation of predicate p.

Token α does not obtain characteristic on entering any of the places l_3 and l_4.

$$Z_3 = \langle \{l_3, l_5, l_6, l_8, l_9, l_{11}\}, \{l_6, l_7, l_8, l_9\}, \begin{array}{c|cccc} & l_6 & l_7 & l_8 & l_9 \\ \hline l_3 & true & false & false & false \\ l_5 & false & false & true & false \\ l_6 & true & false & false & false \\ l_8 & false & false & false & true \\ l_9 & false & W_{9,7} & W_{9,8} & true \\ l_{11} & true & false & false & false \end{array},$$

$$\wedge(l_6, \vee(l_8, l_9), \vee(l_3, l_5, l_{11})) \rangle,$$

where $W_{9,7}$ = "the patient (a token in place l_4) must be classified";
$W_{9,8}$ = "the DB of patients with known classification is sufficient for classifiers' training".

Token α arriving from place l_3 or l_{11} enters place l_6 and then unites with token β that stays there and obtains the characteritic

$$\text{“ } x_{cu}^{\beta} = x_{cu-1}^{\beta} \cup x_{cu}^{\alpha} \text{ ”},$$

where x_i^{ω} denotes the i-th characteristic of token ω and cu shows that the characteristic is the current (the last) one.

Token γ' arriving from place l_5 enters place l_8 and obtains characteristic

"the classifier from place l_5 trained at the current time-step".

If $W_{9,7}$ = *true*, token γ splits to two tokens – the same token γ that stays in place l_9 and token δ with characteristic

"combination of classifiers' results concerning token α in place l_4"

in place l_7. If $W_{9,8}$ = *true*, token γ from place l_9 enters place l_8 with characteristic

"list of the classifiers trained at the current time-step".

Token γ arriving from place l_8 enters place l_9 without any characteritic.

We must note that this transition realizes three different functions simultaneously. First, it represents the process of data accumulating in the DB of patients with known classification; second – the process of a classifier's training; and third – classification by means of each trained classifier and combining of classifiers' results for classification of new patient.

Each of these three functions can be described in details by GNs, too.

$$Z_4 = \langle \{l_4, l_7\}, \{l_{10}\}, \begin{array}{c|c} & l_{10} \\ \hline l_4 & true \\ l_7 & true \end{array}, \wedge(l_4, l_7) \rangle.$$

Tokens α and δ unite in place l_{10}, generating a token α that obtains characteristic

"result of the data classification".

$$Z_5 = \langle \{l_{10}\}, \{l_{11}, l_{12}\}, \begin{array}{c|cc} & l_{11} & l_{12} \\ \hline l_{10} & W_{10,11} & W_{10,12} \end{array}, \vee(l_{10}) \rangle,$$

where

$W_{10,11}$ = "classification is performed";

$W_{10,12}$ = "classification is refused".

Token α obtains characteristic

"classification (normal, pathology) of the current patient"

in place l_{11} and it does not obtain any characteristic on entering place l_{12}.

$$Z_6 = \langle \{l_{12}\}, \{l_{13}, l_{14}\}, \begin{array}{c|cc} & l_{13} & l_{14} \\ \hline l_{12} & W_{12,13} & W_{12,14} \end{array}, \vee(l_{12}) \rangle,$$

where

$W_{12,13}$ = "a new classification procedure is necassary";

$W_{12,14} = \neg W_{12,13}$.

Token α does not obtain any characteristic when it enters any of the places l_{13} and l_{14}.

$$Z_7 = \langle \{l_{13}\}, \{l_{15}, l_{16}\}, \begin{array}{c|cc} & l_{15} & l_{16} \\ \hline l_{13} & W_{13,15} & W_{13,16} \end{array}, \vee(l_{13}) \rangle,$$

where

$W_{13,15}$ = "it is necessary to repeat for a second time the initial procedure";

$W_{13,16} = \neg W_{13,15}$.

Token α obtains characteristic

“list of patient’s speech parameters”

in place l_{15} and then do not obtain any characteristic in place l_{16}.

$$Z_8 = < \{l_{15}\}, \{l_{17}\}, \begin{array}{c|c} & l_{17} \\ \hline l_{15} & true \end{array}, \vee(l_{15}) >.$$

Token α obtains characteristic

“vector with the results of the digital acoustical signal analysis of the patient’s speech”

in place l_{17}.

$$Z_9 = < \{l_{16}\}, \{l_5, l_{18}\}, \begin{array}{c|cc} & l_5 & l_{18} \\ \hline l_{16} & true & true \end{array}, \vee(l_{16}) >.$$

Token α splits to two tokens α with no new characteristic in place l_{18}, and γ' with characteristic

“a new classifier”

in place l_5.

3. Conclusion

If a laryngeal pathology detection system is included as a module in a hospital data analysis and diagnostic system, the communication processes and the interactions between the different modules is effectively describable by the GN modelling.

Ten years ago the idea that all areas of the Artificial Intelligence could be described by unified mathematical tools, was introduced. In [11] it has been discussed why the GNs are a suitable apparatus for the above goal. The present communication is the second step of the authors towards the description of the processes of speaker identification and laryngeal pathology detection by means of voice analysis.

It is an extension of the GN-model from [12].

References

1. B a n c i, G., S. M o n i n i, A. F a l a s c h i, N. S a r i o. Vocal fold disorder evaluation by digital speech analysis. – J. Phonetics, **14**, 1986, 495-499.
2. B o y a n o v, B., D. D o s k o v, T. I v a n o v, S. H a d j i t o d o r o v. PC based system for analysis the voice of patients with laryngeal diseases. – ENT Journal, **70**, 1991, No 11, 767-772.
3. G u b r y n o w i c z, R. A fuzzy approach of pitch period analysis for evaluation of functional status of the larynx source. – J. Phonetics, **14**, 1986, 525-530.
4. H a d j i t o d o r o v, S., T. I v a n o v, B. B o y a n o v. Analysis of dysphony using objective voice parameter. – In: Proc. of 2nd Balkan Conf. on Operational Research, Thessaloniki, Greece, Oct. 1993, 911-917.
5. L a v e r, J., S. H i l l e r, J. M a c k e n z i e, E. R o o n e y. An acoustic screening system for the detection of laryngeal pathology. – J. Phonetics, **14**, 1986, 517-524.
6. L e i n o n e n, L., J. K a n g a s, K. T o r k k o l a, A. J u v a s. Dysphonia detected by pattern recognition of the spectral composition. – J. Speech \& Hear. Res., **35**, 1992, 287-295.
7. L u d l o w, C., C. B a s s i c h, N. C o n n o r, D. C o u l t e r, Y. L e e. The validity of using phonatory jitter and shimmer to detect laryngeal pathology. – In: Laryngeal Function in Phonation and Respiration. Boston, Brown & Co, 1987, 492-508.

8. Y u n i k, M., B. B o y a n o v. Method for evaluation the noise-to-harmonics ratios in pathologic and normal voices. – Acustica, **70**, 1990, 89-91.
9. A t a n a s s o v, K. Generalized Nets. Singapore, New Jersey, London, World Scientific, 1991.
10. A t a n a s s o v, K. Generalized Nets in Artificial Intelligence. Vol. 1. Generalized Nets and Expert Systems. Sofia, “Prof. M. Drinov” Academic Publishing House, 1998.
11. S h a n n o n, A., G. G l u h c h e v, K. A t a n a s s o v, S. H a d j i t o d o r o v. Generalized net representing process of handwriting identification. – Cybernetics and Information Technologies, **1**, 2001, No 1, Sofia, 71-80.
12. H a d j i t o d o r o v, S., P. M i t e v, K. A t a n a s s o v, B. K o l e v, G. G l u h c h e v, V. V a s s i l e v. A generalized net description for laryngeal pathology detection with refusal option. – In: Proceedings of the 3rd Int. Workshop on GNs, Sofia, 1 Oct. 2002, 14-17.

Обобщено мрежово описание на откриване на ларингеални патологии без възможност за отказ от класификация

Стефан Хаджитодоров[1], *Петър Митев*[1], *Красимир Атанасов*[1], *Георги Глухчев*[2], *Васил Василев*[2], *Антони Шенън*[3]

[1]*Център по биомедицинско инженерство, 1113 София, E-mails: sthadj@argo.bas.bg, pmmi@clbme.bas.bg, krat@argo.bas.bg*
[2]*Институт по информационни технологии, 1113 София, E-mails: gluhchev@iinf.bas.bg, vasko@iinf.bas.bg*
[3]*KvB Institute of Technology, North Sydney, 2060, University of Technology, Sydney, 2007, AUSTRALIA, E-mail: tony@kub.edu.au*

(Р е з ю м е)

Обобщените мрежи са разширение на мрежите на Петри. Чрез тях е построен модел на процеса на откриване на ларингеални патологии. Моделът е втори от серия статии върху приложение на обобщените мрежи в областта на разпознаване на гласови сигнали.