Third Int. Workshop on GNs, Sofia, 1 Oct. 2002, 14-17

A generalized net description for laryngeal pathology detection without refusal option[1]

S. Hadjitodorov[1], P. Mitev[1], K. Atanassov[1]
B. Kolev[1], G. Gluhchev[2] and V. Vasilev[2]

[1] Centre for Biomedical Engineering – Bulgarian Academy of Sciences,
Acad. G. Bonchev Str., Bl. 105, Sofia-1113, BULGARIA
e-mails: *sthadj@argo.bas.bg*, *pmmi@clbme.bas.bg*, *krat@argo.bas.bg*
[2] Institute of Information Technologies – Bulgarian Academy of Sciences
Acad. G. Bonchev Str., Bl. 29A, Sofia-1113, BULGARIA
e-mails: *gluhchev@iinf.bas.bg*, *vasko@iinf.bas.bg*

Abstract: A Generalized Net (GN – an extension of the Petri net) model of the process of speaker identification (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 recognition problems.

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 Fo and Pitch period To; various statistics of To and Fo; 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.

One of the main drawbacks of both 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 (class) "refusal to classify". In more sophisticated schemes a final definite decision (excluding the class of "refusals") is

[1]**Aknowledgements:** This work was supported partly by the National Foundation for Scientific Research in the frameworks of projects I-804/98 and TK-L-3/1998

sought. Here, the process of acoustical analysis and a simple two level classification scheme have been modelled by generalized nets.

2. The model

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

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 that we shall note by α enter sequentially place l_1 with initial characteristic

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

We shall use notation α for the later type of tokens, for brevity. The correct form is α_i, where i is the current number of the respective α-token.

GN-transitions have the following forms.

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

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.

$$Z_2 =< \{l_2\}, \{l_3, l_4\}, \begin{array}{c|cc} & l_3 & l_4 \\ \hline l_2 & W_{2,3} & W_{2,4} \end{array}, \vee(l_2) >,$$

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.

$$Z_3 =< \{l_3, l_5, l_7, l_8, l_{10}\}, \{l_5, l_6, l_7, l_8\}, \begin{array}{c|cccc} & l_5 & l_6 & l_7 & l_8 \\ \hline l_3 & true & false & false & false \\ l_5 & true & false & false & false \\ l_7 & false & false & false & true \\ l_8 & false & W_{8,6} & W_{8,7} & true \\ l_{10} & true & false & false & false \end{array},$$

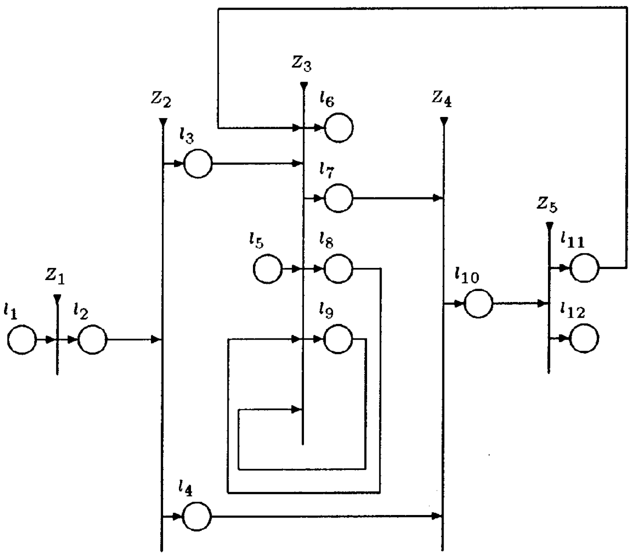
$$\wedge(l_5, \vee(l_7, l_8), \vee(l_3, l_{10})) >,$$

where
$W_{8,6}$ = "the patient (a token in place l_4) must be classified";
$W_{8,7}$ = "the DB of patients with known classification is sufficient for classifiers' training".

Token α arriving from place l_3 or l_{10} enters place l_5 and then unites with token β that stays there and that obtains the characteritic

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

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



If $W_{8,6} = true$, token γ splits to two tokens – the same token γ that stays in place l_8 and token δ with characteristic

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

in place l_6. If $W_{8,7} = true$, token γ from place l_8 enters place l_7 with characteristic

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

Token γ arriving from place l_7 enters place l_8 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.

$$Z_4 = < \{l_4, l_6\}, \{l_9\}, \begin{array}{c|c} & l_9 \\ \hline l_4 & true \\ l_6 & true \end{array}, \wedge(l_4, l_6) > .$$

Tokens α and δ unite in place l_9, generating a token α that obtains characteristic

"result of the data classification".

$$Z_5 =< \{l_9\}, \{l_{10}, l_{11}\}, \begin{array}{c|cc} & l_{10} & l_{11} \\ \hline l_9 & W_{9,10} & W_{9,11} \end{array}, \vee(l_9) >,$$

where
$W_{9,10}$ = "classification is performed";
$W_{9,11}$ = "classification is refused".

Token α obtains characteristic

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

in place l_{10} and it does not obtain any characteristic on entering place l_{11}, from where it leaves the GN.

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 [10] 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.

References:

[1] G. Banci, S. Monini, A. Falaschi and N.Sario, "Vocal fold disorder evaluation by digital speech analysis", J. Phonetics, vol. 14, pp. 495-499, 1986.

[2] B. Boyanov,D. Doskov,T. Ivanov and S.Hadjitodorov, "PC based system for analysis the voice of patients with laryngeal diseases", ENT Journal , vol.70, No 11 , pp. 767-772, 1991.

[3] R. Gubrynowicz, "A fuzzy approach of pitch period analysis for evaluation of functional status of the larynx source", J. Phonetics, vol.14, pp. 525-530, 1986.

[4] S. Hadjitodorov, T. Ivanov and B. Boyanov, "Analysis of Dysphony Using Objective Voice Parameter", In Proc.of 2-nd Balkan Conf. on Operational Research, Thessaloniki, Greece, pp. 911-917, Oct. 1993.

[5] J. Laver, S. Hiller, J. Mackenzie and E. Rooney, "An acoustic screening system for the detection of laryngeal pathology", J. Phonetics vol. 14, pp. 517-524, 1986.

[6] L. Leinonen, J.Kangas, K. Torkkola and A. Juvas , "Dysphonia Detected by Pattern Recognition of the Spectral Composition", J. Speech & Hear. Res. , vol.35, pp. 287-295, 1992.

[7] C. Ludlow, C. Bassich, N. Connor, D. Coulter and Y. Lee, "The validity of using phonatory jitter and shimmer to detect laryngeal pathology", In: Laryngeal function in phonation and respiration, Boston, Brown & co, 1987, pp. 492-508.

[8] M. Yunik and B. Boyanov, " Method for evaluation the noise-to-harmonics ratios in pathologic and normal voices", Acustica vol.70,pp. 89-91, 1990.

[9] Atanassov, K. Generalized Nets, World Scientific. Singapore, New Jersey, London, 1991.

[10] Atanassov, K. Generalized Nets in Artificial Intelligence. Vol. 1: Generalized nets and Expert Systems, "Prof. M. Drinov" Academic Publishing House, Sofia, 1998.