Fuzzy RST and RST Rules Can Predict Effects of Different Therapies in Parkinson's Disease Patients

Andrzej W. Przybyszewski[1,2(✉)]

[1] Polish-Japanese Institute of Information Technology, 02-008 Warsaw, Poland
przy@pjwstk.edu.pl
[2] Department of Neurology, UMass Medical School, Worcester, USA

Abstract. Neurodegenerative disorders (ND) such as Parkinson's disease (PD) are increasing in frequency with ageing, but we still do not have cure for ND.

In the present study, we have analyzed results of: neurological, psychological and eye movement (saccadic) tests in order to discover patterns (KDD) and to predict disease progression with fuzzy rough set (FRST) and rough set (RST) theories. It is a longitudinal study in which we have repeated our measurements every six months and estimated disease progression in three different groups of patients: BMT-group: medication only; DBS-group medication and deep brain stimulation (DBS); and POP–group same as DBS but with several years longer period of DBS. With help of above KDD methods, we have predicted UPDRS (Unified Parkinson's Disease Rating Scale) values in the following two visits on the basis of the first visit with the accuracy of 0.7 for both BMT visits; 0.56 for DBS, and 0.7-0.8 for POP visits. We could also predict UPDRS of DBS patients by rules obtained from BMT-group with accuracy of 0.6, 0.8, and 0.7 for three following DBS visits. Using FRTS we have predicted UPDRS of DBSW3 from DBSW2 with accuracy of 0.5. We could not predict by RST disease progression of POP patients from other groups but with FRST we could predict POPW1 on the basis of DBSW1 results (with accuracy of 0.33). In summary: long-term DBS (POP-group) in contrast to other-groups has changed brain mechanisms and only FRST found similarities between POP and other-groups in disease progressions.

Keywords: Neurodegenerative disease · Rough set · Decision rules
Granularity

1 Introduction

Parkinson (PD) is the second after Alzheimer most popular neurodegenerative diseases (ND) is caused by death of cells primary in the substantia nigra (SN) and lack of the dopamine. Therefore, the first help is to increase the level of the dopamine by its precursor L-Dopa. With disease development, there are adaptive mechanisms that inactivate high level of the dopamine and patients have to increase dosages of their medication till getting so-called ON-OFF effect. This effect can be cured by DBS. About 30 years ago, Alain Benabid in Grenoble (France) introduced the deep brain

M. Ceci et al. (Eds.): ISMIS 2018, LNAI 11177, pp. 409–416, 2018.
https://doi.org/10.1007/978-3-030-01851-1_39

stimulation (DBS) by targeting the subthalamic nucleus (STN). It was claimed as "the most important discovery since L-Dopa". But PD progression is also related with neurodegeneration in other structures like prefrontal cortex and related cognitive problems and limbic system leading to depressions and also related to lack of the dopamine as reward transmitter. However, the main symptoms in PD are movement disorders measured mainly by the UPDRS.

As disease starts about 20 years before first symptoms each patient has different plastic brain mechanism, and rate of the disease progression, which gives a problem to find an optimal therapy that depends on the doctor's experiences and precision of his/her tests. As results of different doctor's tests are partly subjective, we propose to use objective, doctor independent test of reflexive eye movement measurements that can be performed automatic without doctor's time.

In this study, we have developed methods of rules discovery (KDD) and symptoms prediction on the basis of the object classifications found in the visual system of primates [1]. There are many similarities of the unknown complex object classification to recognition of the disease complexity. Visual brain plasticity (learning mechanisms) mechanisms are universal in the whole brain. In general, classification processes are related to two different logic systems as we have demonstrated in the visual brain [1]. The brain solves problems of imprecisions and contradictions by using approach similar to rough set theory [2]. It means that descending and ascending pathways, by using different logics, interact in order to minimize the boundary region between known and actual object's properties. It is similar to testing hypothesis by neurologist: what are the most important symptoms of the actual patient.

This study is expansion of our previous work [3] by using additional to RST, fuzzy RST that are more universal in finding similarities. In the perspective, our methods should lead to the remote diagnosis and treatments (telemedicine).

2 Methods

We have analyzed tests from PD patients from three groups: (a) BMT – group: 23 patients that have therapy limited to medication; (b) DBS – group: 24 patients on medications and with the Deep Brain Stimulation (DBS) of subthalamic nucleus. For this group surgery was performed during the study; (c) POP-group: 15 patients with surgery performed earlier before our study. Patients from all groups were tested in the following sessions: MedOn/Off sessions: ses.1/3: without/after medication. In addition, patients from DBS and POP groups were tested in StimOn/Off sessions, StimOn: ses.2/4 without/with med. All tests: neurological, neuropsychological, and eye movement tests were performed in dept. of Neurology, Brodno Hospital, Warsaw Medical University. In this study, we have measured parameters of the fast, reflexive saccadic eye movements (EM). Detailed methodology was described earlier [3, 4]. In short, each patient has to follow horizontally moving randomly 10° to the right or 10° to the left dot. We have estimated the following saccadic parameters: delay – it is a time difference between the beginning of the stimulus and eye movements; relative amplitude: amplitude of the saccade related to the amplitude of the light spot movements; max saccade velocity; duration of the saccade as a time between the end and the beginning

of each saccade. Eye movements were recorded by the head-mounted saccadometer (Ober Consulting, Poland).

2.1 Theoretical Basis

Our rough set (RS) data mining analysis is based on the Pawlak's concept of RS theory (RST) (Zdzislaw Pawlak [5]). An information system [5] a pair $S = (U, A)$, where U, A are nonempty finite sets called the universe of objects U and the set of attributes A. If $a \in A$ *and* $u \in U$, the value $a(u)$ is a unique element of V (where V is a value set).

We define as in [5] the *indiscernibility relation* of any subset B of A or $IND(B)$ as: $(x, y) \in IND(B)$ or $xI(B)y$ iff $a(x) = a(y)$ for every $a \in B$ where the value of $a(x) \in V$. It is an equivalence relation $[u]_B$ that we understand as a *B-elementary granule*. The family of $[u]_B$ gives the partition U/B containing u will be denoted by $B(u)$. The set $B \subset A$ of information system S is a reduct $IND(B) = IND(A)$ and no proper subset of B has this property [6]. In most cases, we are only interested in such reducts that are leading to expected rules (classifications). On the basis of the reduct we have generated rules using four different ML methods (RSES 2.2): exhaustive algorithm, genetic algorithm [7], covering algorithm, or LEM2 algorithm [8].

A *lower approximation* of set $X \subseteq U$ in relation to an attribute B is defined as $\underline{B}X = \{u \in U : [u]_B \subseteq X\}$. The *upper approximation* of X *is* defined as $\overline{B}X = \{u \in U : [u]_B \cap X \neq \phi\}$. The difference of $\overline{B}X$ *and* $\underline{B}X$ is the boundary region of X that we denote as $BN_B(X)$. If $BN_B(X)$ is empty then set than X is *exact* with respect to B; otherwise if $BN_B(X)$ is not empty and X is not *rough* with respect to B.

A decision table (training sample in ML) for S is the triplet: $S = (U, C, D)$ where: C, D are condition and decision attributes [9]. Each row of the information table gives a particular rule that connects condition and decision attributes for a single measurements of a particular patient.

However, FRTS (fuzzy rough set theory) replaced defined above 'crisp' dependences by a tolerance or similarity relations $R_a(x, y)$ as a value between two observations x and y. As summarized in [10] there are several tolerance relationships as normalized difference (so-called *'Eq. 1'*) or Gaussian or exponential differences [10]. There are formulas related to normalized differences between pairs of attributes. The most common are *Lukasiewicz* and *t.cos t-norms* [10]. As decision attributes are nominative we used crisp relations between them.

We define *B-lower* and *B-upper* approximations for each observation x in FRST as following: *B-lower* approximation as: $(R_B \downarrow X)(x) = \inf_{y \in U} I(R_B(x, y), X(y))$, where I is an *implicator* [10]. The B-lower approximation for the observation x is then the set of observations which are the most similar to observation x and it can predict the decision attribute with the highest confidence, based on conditional attributes B.

The B-upper approximation is defined by $(R_B \uparrow X)(x) = \sup_{y \in U} \tau(R_B(x, y), X(y))$, where τ is the t-norm. The B-upper approximation is a set of observations for which the prediction of decision attribute has the smallest confidence [10].

Also rules in FRST have different construction than in RST. They are based on the tolerance classes and appropriate decision concepts. The *fuzzy rule* is a triple *(B, C, D)*,

where B is a set of conditional attributes that appear in the rule, C stands for fuzzy tolerance class of object and D stands for decision class of object. There are important differences that will be demonstrated in our results.

We have used the RSES 2.2 (Rough System Exploration Program) [9] with implementation of RST rules to process our data and "RoughSets" version 1.3-0 implemented in R [11]. As we have demonstrated earlier that the RST method is superior to other classical methods [3].

3 Results

As described in the Methods section, total 62 PD patients were divided into three groups: BMT-group (only medication), DBS-group (medication and STN stimulation, surgery during the study) and POP-group (medication and STN stimulation, long after the surgery). In BMT-group: 23 patients, the mean age was 57.8 ± 13 (SD) years; disease duration was 7.1 ± 3.5 years, UPDRS was 48.3 ± 17.9.

In POP-group: 15 patients, the mean age was 63.1 ± 18.2 (SD) years and disease duration was 13.5 ± 3.6 years (stat. diff. from BMT $p < 0.025$, and from DBS-group: $p < 0.015$), UPDRS was 59.2 ± 24.5 (stat. diff. than BMT-group: $p < 0.0001$).

In DBS-group: 24 patients, the mean age of 53.7 ± 9.3 years, disease duration was 10.25 ± 3.9 years; UPDRS was 62.1 ± 16.1 (stat. diff. than BMT-group: $p < 0.0001$).

These statistical data are related to the data obtained during the first session for each group: BMT W1 (visit one), DBS W1 (visit one) and POP W1 (visit one). It is clear that in visit one W1 UPDRS in different groups are different.

In DBSW1 visit before the surgery, there were only two sessions (MedOFF, MedON). In ses.1 mean UPDRS was 62.2 ± 16.1, in ses.3 was 29.9 ± 13.3 strongly ($p < 0.0001$) different from ses.1 (effect of medication). UPRDS of DBSW2 after the surgery in ses.1 is larger than before the surgery 65.3 ± 17.6 but there are not stat. sig. diff. UPDRS ses.1 of DBSW3 is 68.7 ± 17.7 and stat. diff. ($p < 0.03$) than in W2.

In POP-group UPDRS values are similar. There is an increase of the UPDRS ses.1 from W1: 63.1 ± 18.2 to W2: 68.9 ± 20.3 to W3: 74,2 ± 18.4 but there were smaller differences for ses.4 (both med. and stim. on) W1: 21 ± 11.3 to W2: 23.3 ± 9.5 to W3: 23,8 ± 10.7. It seems that groups DBS and POP are similar.

In BMT group UPDRS ses.1 W1: 48.3 ± 17.9; W2: 57.3 ± 16.8 ($p < 0.0005$ diff than W1); W3: 62.2 ± 18.2 ($p < 0.05$ diff. than W2). In ses.3 UPDRS was W1: 23.6 ± 10.3; W2: 27.8 ± 10.8; W3: 25 ± 11.6 (no stat. diff. between visits for ses.3).

3.1 FRST and RST Rules and ML Results for BMT Group

There were patient only on medication with two different sessions (MedOFF and MedON) measured every six months.

Using RST, after the discretization, we have divided UPDRS into 4 ranges: "(-Inf, 24.0)", "(24.0, 36.0)", "(36.0, 45.0)", "(45.0, Inf)", with help of the feature selection RSES software and by using machine learning and RST algorithms [7] we have obtained rules from BMTW1 and use them to predict UPDRS for BMTW2 and BMTW3. As in [3] we have used: parameters related to psychological testing (PDQ39

– quality of life, Epworth – quality of sleep), and parameters of saccades where in this only latency was significant (SccLat), decision attribute: UPDRS [3].

As each row gives a particular rule and by using RSES we have obtained more general rules like:

$$(S\# = 3)\&\left(PDQ39 = ''(-Inf, 50.5)''\right) = > \left(UPDRS = ''(-Inf, 24.0)''\right) \quad (1)$$

It states that if the session for ses.3 and PDQ39 is smaller than 50.5 then UPDRS will be smaller than 24.0.

In FRST also decision attribute was UPDRS and "in agreement with discretization in RST we got four values: (-Inf, 24.0) - > "1", "(24.0, 36.0)" - > "2", "(36.0, 45.0)" - > "3", (45.0, Inf)) - > "4". An example of analog FSRT rule:

$$(S\# = 3)\&\left(PDQ39 = ''(40)''\right)\&\left(RSAmp = ''(9.5)''\right)\&\left(RSLat = ''(281)''\right)\&(RSDur = ''(45)'' = > (UPDRS = 1) \quad (2)$$

It states that if the session for ses.3 and PDQ39 is near 40 and RSAmp near 9.5 and RSLat is near 281 and RSDur is near 45 then UPDRS is 1. Where RS are parameters of the reflexive saccades: RSAmp is amplitude, RSLat is saccade latency, and RSDur is saccade duration.

We have used these rules to predict UPDRS in BMTW2 and W3 using 6-fold cross validation we have obtained for both visits global accuracy 0.7, and the global coverage 1.0. For FRST we have obtained accuracy 0.63.

3.2 FRST and RST Rules and ML Approach for DBS Group

We have predicted UPDRS of DBSW3 by rules from DBSW2 (both groups have 4 sessions), and we have obtained the global accuracy 0.56 and global coverage 1.

In the next step, we have applied the same BMTW1 rules to the DBS group. It was successful for DBSW1 pre-operative patients as they were also in two sessions with a high dosage of medication. We have obtained the global accuracy 0.64 with the global coverage 0.5.

As we have noticed that there are large difference between UPDRS in ses.1 W1 therefore prediction of UPDRS from BMT group for DBSW2 and W3 groups were not possible as there are different numbers of sessions. Therefore, we have divided DBSW2, W3 patients into two subgroups: one without stimulation (StimOff) and another one with StimOn. We could not predict UPDRS in DBS groups without stimulation (StimOff) only with stimulation StimOn, and all our predictions for DBS and POP groups are only for StimOn.

UPDRS of DBSW2 were predicted from BMTW1 rules with global accuracy 0.85, but with coverage was 0.3 and some classes were not at all predicted (for UPDRS larger than 63). We have obtained similar results for UPDRS of DBSW3 from BMTW1 rules; the global accuracy was 0.74 but the global coverage 0.56.

3.3 RST and FRST Rules and ML Approach for POP Group

On the basis of rules obtained from POPW1 we have predicted UPDRS in POPW2 with the accuracy 0.67; UPDRS in POPW3 from POPW1 with accuracy 0.8 and global coverage 0.97 (for details see [3]).

In contrast to the DBS group we were not successful in using rules from the BMT patients to predict UPDRS of the POP group patients. In order to find possible reasons we have compared rules from the BMTW1 patients with rules of POPW3 group as UPDRS values for both groups were similar (see Methods).

BMTW1 rules:

$$\left(PDQ39 = ''(Inf, 50.5)''\right)\&\left(RSLat = ''(264.0, Inf)''\right)\&\left(dur = ''(Inf, 5.65)''\right) = > \left(UPDRS = \{''(-Inf, 33.5)''[4], ''(33.5, 43.0)''[1]\}\right)5 \tag{3}$$

POPW3 rules:

$$\left(PDQ39 = ''(-Inf, 48.0)''\right)\&\left(RSLat = ''(301.0, Inf)''\right)\&\left(RSPeak = ''(403.5, 522.0)''\right) = > \left(UPDRS = ''(66.0, Inf)''[1]\right)1 \tag{4}$$

Above rules (3) and (4) are contradictory as the same values of conditional attributes, limited by other attributes, give opposite results of the UPDRS.

We have performed computations for comparison of DBSW1 and POPW1 groups using FRTS. As decision attributes are nominal, we have changes range of UPDRS to small, medium, large and very large that was implemented as 1, 2, 3, 4 (Table 1).

We have obtained our predictions using the hybrid-fuzzy rules with aggregation by the *t.norm Lukasiewicz*, *tolerance Eq. 3* (modified Gaussian from [10]), and as *relation Gaussian kernel* (with 0.2 as a parameter). As example of fuzzy rules are:

Table 1. Confusion matrix for UPDRS of POPW1-subgroup (StimOn) by rules obtained from DBSW1-subgroup (StimOn) with FRST.

Actual	Predicted				
	"1"	"2"	"3"	"4"	ACC
"1"	5.0	0.0	0.0	4.0	0.556
"2"	3.0	0.0	0.0	6.0	0.0
"3"	3.0	0.0	0.0	4.0	0.0
"4"	1.0	0.0	0.0	4.0	0.8
TPR	0.42	0.0	0.0	0.29	

TPR: True positive rates for decision classes; ACC: Accuracy for decision classes: the global coverage was 1 and the global "accuracy was 0.33, the coverage for all decision classes was 1. Where: (-Inf, 17.0) - > "1", "(17.0, 27.0)" - > "2", "(27.0, 36.0)" - > "3", (36.0, Inf)) - > "4".

$$(Epworth\ 'is\ around'\ 11)\&(RSDur\ 'is\ around'\ 66) = > (UPDRS = 3) \quad (5)$$

$$(RSLat\ 'is\ around'\ 424)\&(RSDur\ 'is\ around'\ 48) = > (UPDRS = 2) \quad (6)$$

$$(RSLat\ 'is\ around'\ 286)\ v\ (RSLat\ 'is\ around'\ 292)\ v\ (RSLat\ 'is\ around'\ 251) = > (UPDRS = 1) \quad (7)$$

Above fuzzy rules (5–7) are different that RST crisp rules (3), (4) but they have 'fuzzy' properties that seems to be more general and this way might cover some contradictions between rough set rules.

4 Discussion

There is always actual discussion how can we achieve that all medical procedures would be optimal for an individual patient? Thanks to technology is significant progress in medical science and neurology with new procedures improving PD patient's treatments. But generally, the long lasting neurodegeneration processes with the compensatory; specific for each person plastic changes it is extremely difficult question. We propose to solve this problem by using similar mechanisms that were found in the visual system to categorize complex objects. We have used rough and fuzzy rough set theories to fit our granules in optimal way to the complexity of the disease symptoms. We were successful in prediction of the symptoms development in time (longitudinal study) for different therapies such as medication and DBS. However, we had problems to predict symptoms of patients with the long lasting brain stimulation (POP-group). We were partly successful by using more general FRST rules, but might try to increase number of attributes in order to find what are mechanisms of the long-lasting brain stimulation. We expect that in the near future more systems will be replacing, at least in part, medical doctors as they are more objective, automatic, more precise and they do not take doctor's time. Such system, when intelligent, could also help in doctor's decision and help in treatment optimization.

In summary, we have tested three groups of patients: BMT-group (medication only); DBS-group (medication and short brain stimulation), and POP-group (medication with long lasting DBS) in the longitudinal study. Mean UPDRS without medication and stimulation were:

BMT: W1-48; W2-57; W3–62; with the disease duration of 7 years

DBS: W1-62; W2-65; W3–69; 10 years

POP: W1-63; W2-69; W3–74; 13.5 years

Mean UPDRS increase for half of the year for BMT: 7, for DBS: 3.5, for POP: 5.5. From above, it is evident that deep brain stimulation is slowing down the progression of the disease as the disease duration for POP group is almost twice as for BMT group. These is also our major problem with prediction of UPDRS from BMT group to DBS and POP more advanced groups with significant difference in UPDRS. As electric brain stimulation significantly lowered UPDRS, we may think that DBS resets brain and symptoms in time into the beginning of the disease. It is why it was possible to

predict symptoms development in more advanced disease stages from earlier one. It might be the direction in finding the optimal treatment.

References

1. Przybyszewski, A.W.: Logical rules of visual brain: from anatomy through neurophysiology to cognition. Cogn. Syst. Res. **11**, 53–66 (2010)
2. Przybyszewski, A.W.: The Neurophysiological Bases of Cognitive Computation Using Rough Set Theory. In: Peters, James F., Skowron, A., Rybiński, H. (eds.) Transactions on Rough Sets IX. LNCS, vol. 5390, pp. 287–317. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-89876-4_16
3. Przybyszewski, A.W., Szlufik, S., Habela, P., Koziorowski, D.M.: Rules determine therapy-dependent relationship in symptoms development of Parkinson's Disease Patients. In: Nguyen, N.T., Hoang, D.H., Hong, T.-P., Pham, H., Trawiński, B. (eds.) ACIIDS 2018. LNCS (LNAI), vol. 10752, pp. 436–445. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-75420-8_42
4. Przybyszewski, A.W., Kon, M., Szlufik, S., Szymanski, A., Koziorowski, D.M.: Multimodal learning and intelligent prediction of symptom development in individual Parkinson's Patients. Sensors, **16**(9), 1498 (2016). https://doi.org/10.3390/s16091498
5. Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning About Data. Kluwer, Dordrecht (1991)
6. Bazan, J., Nguyen, H.S., Trung, T., Nguyen, Skowron A., Stepaniuk, J.: Desion rules synthesis for object classification. In: Orłowska, E. (ed.) Incomplete Information: Rough Set Analysis, pp. 23–57. Physica – Verlag, Heidelberg (1998). https://doi.org/10.1007/978-3-7908-1888-8_2
7. Bazan, J., Nguyen, H.S., Nguyen, S.H., Synak, P., Wróblewski, J.: Rough set algorithms in classification problem. In: Polkowski, L., Tsumoto, S., Lin, T. (eds.) Rough Set Methods and Applications, pp. 49–88. Physica-Verlag, Heidelberg New York (2000). https://doi.org/10.1007/978-3-7908-1840-6_3
8. Grzymała-Busse, J.: A new version of the rule induction system LERS. Fundamenta Informaticae **31**(1), 27–39 (1997)
9. Bazan, J.G., Szczuka, M.: The rough set exploration system. In: Peters, J.F., Skowron, A. (eds.) Transactions on Rough Sets III. LNCS, vol. 3400, pp. 37–56. Springer, Heidelberg (2005). https://doi.org/10.1007/11427834_2
10. Riza, L.S., et al.: Implementing algorithms of rough set theory and fuzzy rough set theory in the R package "RoughSets". Inf. Sci. **287,** 68–69 (2014)
11. Bazan, J.G., Szczuka, M.: RSES and RSESlib - a collection of tools for rough set computations. In: Ziarko, W., Yao, Y. (eds.) RSCTC 2000. LNCS (LNAI), vol. 2005, pp. 106–113. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-45554-X_12