

# Using the DTW method for estimation of deviation of care processes from a care plan

© Alexey Molodchenkov

© Mikhail Khachumov

Federal Research Center “Computer Science and Control” of Russian Academy of Sciences,  
Moscow

[aim@isa.ru](mailto:aim@isa.ru),

[khmike@inbox.ru](mailto:khmike@inbox.ru)

## Abstract

Hospitals increasingly use process models for structuring their care processes. Activities performed to patients are logged to a database or a log. These data can be used for managing and improving the efficiency of care processes and quality of care. In this article, we propose the method for estimation of deviation of care processes from a care plan. Care plan defines the steps of a patient treatment for a certain disease in a specific hospital. Care plan is built on the base of care process model. A care process model is built on the base of exemplars of care processes, stored in a database or a log. The Dynamic Time Warping (DTW) algorithm was used for estimation of deviation. The DTW algorithm measures a distance-like quantity between two given traces containing information about execution or not execution of actions defined by care plan.

## 1 Introduction

Explore of methods for management of care process (CP) as a flow of therapeutic and diagnostic activities allows us to analyze situations with patients and recommend to decision-makers (DM) appropriate action for the treatment of patients. In this connection, occurs increased interest in an automation of hospitals units (accounting, reception, offices, warehouse and so on) and an analysis and a formal representation of care processes in order to support doctors work.

The first area is deep enough elaborated - powerful medical information systems for supporting of therapeutic and diagnostic processes are developed. Flows of patients can be optimized by simulation in order to identify bottlenecks.

The second area requires additional researches to estimate stages of a care process and to develop recommendations for decision-makers. Currently tools for analysis and medical diagnostics, using different classifiers with high accuracy and precision are developed [3, 4, 8, 11, 17, 18, 24]. This area has a number of features and is of great interest for us. A care process,

despite the standards, always has an individual (personalized) character. There may be various deviation from the selected care plan, depending on the changing care conditions, concomitant deceases, etc. Therefore, there are not only problems replying of a care process, and operational management of a care process in terms of possible deviations. Management is a particular sequence of treatment actions (operations), which is based on states of a patient, a prescribed care plan and medical databases, i.e., precedents.

The aim of our research is to develop algorithms and tools that assist to a doctor by making proposals (recommendations) on the organization of care process in accordance with an actual patient’s state and care plan. One of the objectives is to assess the quality of the rendered medical services by comparing a patient’s care process and care plan. The care plan defines the actions of the doctor and is based on the care process model, discovered from precedents.

The multi-dimensional distance based on Dynamic Time Warping (DTW) algorithm is used to differences between care process and care plan. Experiments show that DTW algorithm is effective to compare sequences of actions in relation to care processes. We consider the use of DTW algorithm to detect deviations between real care process of a patient and care plan.

## 2 Formalization and optimization of care processes

Some methods for formalization and application of care processes are described in [1, 12, 17]. Methods and algorithms for discovery of models of care processes on the base of event logs and precedents are described in [3, 6, 12-16, 19]. The model includes all traces from event logs and precedents. Fig. 1 illustrates the Petri net workflow process definition for handling a medical complaint. In this figure we can identify the following routing constructs: transitions “Identification” and “Cardiologist” are AND-splits, “I diagnosis OK”, “I diagnosis NOK” and “decide surgery” are AND-joins, c4, c5, c8 and c10 are OR-splits and c6, c9 and c11 are OR-joins.

Automatic discovery of care process model is difficult and actual problem. Some methods and algorithms to solve this problem are described in articles [3, 6, 12-16, 19]. Every trace in event logs and in a care

process model are characterized, in general, the timing associated with the time of applying operations by physician, and is characterized by quality indicators (signs), defining the current state of the patient. In real life we have deviations of real care processes from care plans built on the base of care process model. These deviations can be associated with a change in the patient's state, lack or replacement of some drugs or other medical devices, influence of other disease or other causes.

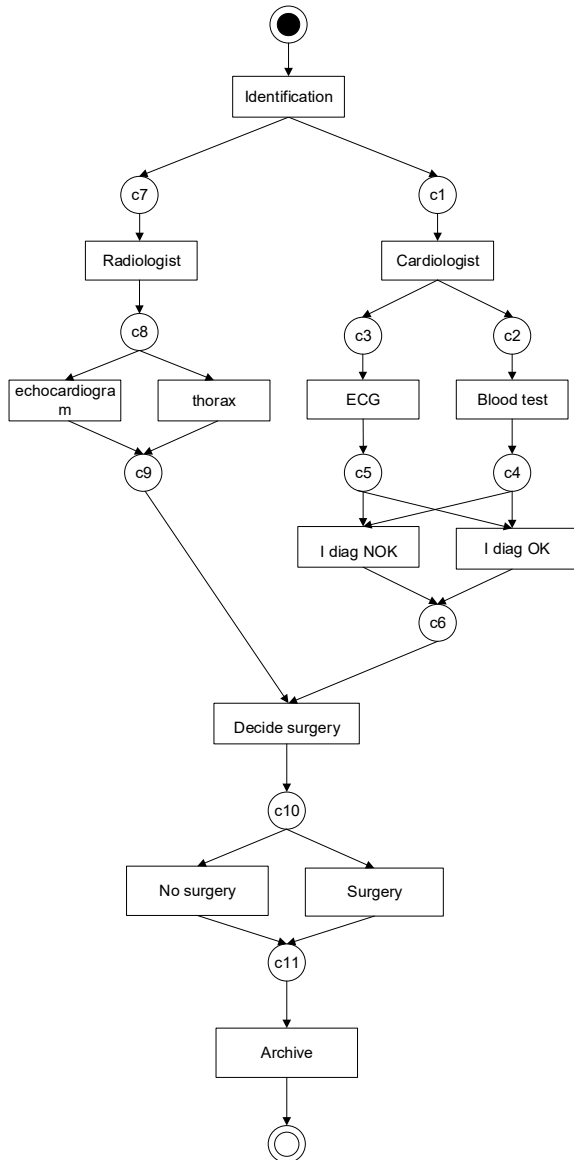


Figure 1 A Petri net of a care process.

### 3 Identification of possible deviations of a care process from the care plan

Qualitatively expert assessment of a trace as a way on the graph with temporary marks will allow to reveal and estimate various deviations from the course of medical and diagnostic process due to both objective and subjective reasons and to eliminate them in the subsequent realizations of care process. Non-performance of action which has to be surely executed in

care process, non-compliance of actions sequence, performance of actions not provided by care process, etc. [2, 9] are considered to be deviations. In care process it is necessary to emphasize the deviations associated with time limits imposed on actions. For example, some actions have to be performed on the first day of the patient arrival. The situation when the action has been performed after the specified time interval is the deviation in this case.

The method for detection and visualization of the deviations associated with performance of actions not included in the model of care process and non-performance of actions that need to be executed is described in the article [10]. In articles [1, 14, 21, 23] the fitness function is used to check conformance of a care process model and processes in event log. Petri nets are used as formal representation of the care process model. Let  $k$  is the number of different traces from the aggregated log. For each log trace  $i$ , ( $1 \leq i \leq k$ ),  $n_i$  is the number of process instances combined into the current trace,  $m_i$  the number of missing tokens,  $r_i$  the number of remaining tokens,  $c_i$  the number of consumed tokens, and  $p_i$  the number of produced tokens during log replay of the current trace. The token-based fitness metric  $F$  is defined as follows [1, 21, 23]:

$$F = \frac{1}{2} \left( 1 - \frac{\sum_{i=1}^k n_i m_i}{\sum_{i=1}^k n_i c_i} \right) + \frac{1}{2} \left( 1 - \frac{\sum_{i=1}^k n_i r_i}{\sum_{i=1}^k n_i p_i} \right)$$

In case of deviation detection in the course of treatment of a specific patient,  $n$  and  $k$  are equal to 1 since there is one trace and one instance. Function  $f$  will be as follows:

$$f = \frac{1}{2} \left( 1 - \frac{m}{c} \right) + \frac{1}{2} \left( 1 - \frac{r}{p} \right)$$

If  $f$  is equal to 1, then the trace is completely consistent with the care process model. Otherwise, there are deviations. In [23] it is shown that deviations from the care plan lead to increased cost of treatment. In [5] the method that could check for deviations of care process of a patient from a plan and locate specific points of these deviations is described.

Let's consider a method for detection of the deviations associated with the performance of actions not provided by the care plan and non-performance of actions provided by care process. The DTW method was applied calculate the deviation or distance. The method allows to find closeness between two measurement sequences for a certain period of time. Generally, the length of sequences can be different, and measurements can be made with different rates [7]. DTW method became widely spread in medicine. A theory of modified DTW algorithms and its applications are presented in [22]. In particular, recognition of human activity is considered by comparison of the gestures presented in

the form of two-dimensional time series. Recognition of such activity of the patients can be very useful in modern healthcare for monitoring of patients condition and automatic reporting creation for health workers. The results of experiments confirmed the sufficient accuracy of one modification.

DTW algorithm calculates the optimal sequence of transformation (deformation) between two time series [7, 20]. Let two numerical sequences  $(a_1, a_2, \dots, a_n)$ ,  $(b_1, b_2, \dots, b_m)$  are given. We obtain deviations matrix  $D$ , where  $d_{ij} = |a_i - b_j|$ ,  $i = 1, \dots, n, j = 1, \dots, m$ . At the second step we build deformations matrix. Each element  $r_{ij}$  is defined by means of dynamic programming algorithm and local optimization criteria:  $r_{ij} = d_{ij} + \min(r_{i-1, j-1}, r_{i-1, j}, r_{i, j-1})$ . The path in the deformation matrix defining a deviation begins in its left upper corner and ends in the right lower. The value  $R$  of deformation is defined by the sum of the minimum local deviations of each path element.  $R$  is divided by the number of path elements and is considered as distance estimation between sequences.

We consider the example of DTW algorithm application for comparison of processes of bronchial asthma treatment that is presented in Table 1 (data are provided by the Medical center of the Central bank of the Russian Federation). Table contains care plan that includes operations mandatory to perform and reports of real care processes of three patients.

To formalize care process reporting we will define the following variables: «executed»  $\rightarrow \alpha = 1$ , «not executed»  $\rightarrow \beta = -1$ , «not required»  $\rightarrow \sigma = 0$ . The distance between «executed» and «not executed» operations we will define as  $|\alpha - \beta| = 2$ , and the distance between «executed» («not executed») and «not required», respectively  $|\alpha - \sigma| = |\beta - \sigma| = 1$ .

Obtained care process parameters are included in Table 2 in the form of sequences

The sequences can now be compared. We will apply the DTW method to calculate care process deviations of care processes from the care plan. In Fig. 2 the process of comparison of the Patient 3 (P3) course of treatment with the plan (P) is presented as comparison of two sequences by the DTW method using the dynamic programming scheme. In the table the way determining the minimum value of deviation  $R$  is highlighted in color. In this case  $R = 2.257$ .

We will apply the DTW method to calculate deviations of Patients 1 and 3 (respectively P1 and P3) care processes. All necessary data for comparison of processes are shown in Fig. 3.

Table 3 demonstrates possibility of distances calculation between precedents. The more the distance, the more precedents are differed from each other. The table shows that Precedent 1 (care trace of the Patient 1) is close to the plan.

**Table 1** Bronchial asthma therapy process

Operations of the care plan		Progress report		
		Precedent 1	Precedent 2	Precedent 3
<b>Reception area / Intensive care unit (ICU)</b>				
1	Transfer from the reception area to the ward with beds /ICU through 2 hours or less	Executed	Executed	Executed
2	External respiration function or Peak expiratory flow rate	Executed	Not required	Not executed
3	Pulse oximetry	Executed	Not required	Executed
4	Chest radiography	Executed	Executed	Executed
<b>1 day in ICU / ward with beds</b>				
5	Pulse oximetry	Not required	Not required	Executed
6	Peak Flow Meter	Executed	Not required	Not executed
7	Inhaled short-acting $\beta_2$ -agonists or formoterol	Executed	Executed	Executed
<b>2-7 days in ward with beds</b>				
8	Consultation of an exercise therapy doctor	Executed	Executed	Executed
9	Consultation of a physiotherapist	Executed	Executed	Executed
10	Peak Flow Meter	Executed	Executed	Executed
11	External respiration function	Executed	Executed	Not executed
12	Inhaled glucocorticosteroids	Executed	Executed	Executed
13	Inhaled $\beta_2$ -agonists	Executed	Executed	Executed
<b>8-21 days in ward with beds</b>				
14	Peak Flow Meter	Executed	Not required	Executed
15	External respiration function	Executed	Not required	Executed
16	Systemic glucocorticosteroids	Not required	Not required	Not required
17	Inhaled glucocorticosteroids	Executed	Not required	Executed
18	Inhaled $\beta_2$ -agonists	Executed	Not required	Executed

**Table 2** The summary table of the formalized indicators of care process performance

Transaction number	Precedent 1	Precedent 2	Precedent 3	Care plan
1	1	1	1	1
2	1	0	-1	1
3	1	0	1	1
4	1	1	1	1
5	0	0	1	1
6	1	0	-1	1
7	1	1	1	1
8	1	1	1	1
9	1	1	1	1
10	1	1	1	1
11	1	1	-1	1
12	1	1	1	1
13	1	1	1	1
14	1	0	1	1
15	1	0	1	1
16	0	0	0	1
17	1	0	1	1
18	1	0	1	1

**Table 3** Deviations of care processes (pair distances)

	Care plan	Precedent 1	Precedent 2	Precedent 3
Care plan	0,000	0,486	2,143	2.257
Precedent 1	0,486	0,000	0,333	1.935
Precedent 2	2,143	0,333	0,000	4,421
Precedent 3	2.257	1.935	4,421	0,000

### 3 Conclusion

Medical processes describing care of patients are useful in daily work of physicians, especially in difficult situations. Certain step towards the creation of tools to support physician's work in the course of care process is taken in the article. At the current stage of research, methods for the analysis and evaluation of deviations associated with performance and non-performance of actions for elementary care processes are chosen and studied. The problem is solved on the existing generalized care process model and reduced to evaluation of deviations of care precedent from available traces. The proposed method allows verifying compliance of treatment with the care plan and procedures established by care standard. In addition it helps the decision-maker with the choice of a rational way of the treatment carried out with the use of strategies and rules. In reality, to make a choice of a rational way of treatment it is necessary to sort and estimate rather large number of admissible trajectories taking into account strict binding of operations at the time-point that certainly complicates process of comparison and determination of distances. Authors intend further to research similar processes.

The reported study was funded by RFBR according to the re-search project No 16-37-00034 «Research and development of methods for analysis of deviations of executed medical processes from their model»

### References

- [1] Wil M. P. van der Aalst, Process Mining: Discovery, Conformance and Enhancement of Business Processes, Springer Publishing Company, Incorporated, 2011, 352 p.
- [2] Adriansyah, A.: Aligning Observed and Modeled Behavior. Ph.D. thesis, Eindhoven University of Technology, 2014.
- [3] Breiman L., Friedman J. H., Olshen R. A., Stone C. T. Classification and Regression Trees.—Wadsworth, Belmont, California, 1984.
- [4] Carroll, Robert J., et al. "Portability of an algorithm to identify rheumatoid arthritis in electronic health records." Journal of the American Medical Informatics Association 19.e1 (2012): e162-e169.

P3	1	-1	1	1	1	-1	1	1	1	1	-1	1	1	1	1	0	1	1				
P	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18				
1	1	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	7	7	7			
2	1	2	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	7	7	7		
3	1	3	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	7	7	7	
4	1	4	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7
5	1	5	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7
6	1	6	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7
7	1	7	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7
8	1	8	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7
9	1	9	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7
10	1	10	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7
11	1	11	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7
12	1	12	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7
13	1	13	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7
14	1	14	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7
15	1	15	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7
16	1	16	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7
17	1	17	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7
18	1	18	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7

**Figure 2** Calculation of care process deviation from the care plan (on the example of Patient 3)

P3	1	-1	1	1	1	-1	1	1	1	1	-1	1	1	1	1	0	1	1					
P1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18					
1	1	1	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	7	7	7	7		
2	1	2	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	7	7	7	7	
3	1	3	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7	7
4	1	4	0	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7	7	7	7
5	0	5	1	1	2	3	3	3	4	5	5	5	5	6	7	7	7	7	6	7	8		
6	1	6	1	3	1	1	1	3	3	3	3	3	5	5	5	5	5	5	6	6	6		
7	1	7	1	3	1	1	1	3	3	3	3	3	5	5	5	5	5	5	6	6	6		
8	1	8	1	3	1	1	1	3	3	3	3	3	5	5	5	5	5	5	6	6	6		
9	1	9	1	3	1	1	1	3	3	3	3	3	5	5	5	5	5	5	6	6	6		
10	1	10	1	3	1	1	1	3	3	3	3	3	5	5	5	5	5	5	6	6	6		
11	1	11	1	3	1	1	1	3	3	3	3	3	5	5	5	5	5	5	6	6	6		
12	1	12	1	3	1	1	1	3	3	3	3	3	5	5	5	5	5	5	6	6	6		
13	1	13	1	3	1	1	1	3	3	3	3	3	5	5	5	5	5	5	6	6	6		
14	1	14	1	3	1	1	1	3	3	3	3	3	5	5	5	5	5	5	6	6	6		
15	1	15	1	3	1	1	1	3	3	3	3	3	5	5	5	5	5	5	6	6	6		
16	0	16	2	2	2	2	2	4	4	4	4	4	6	6	6	6	6	6	6	7			
17	1	17	2	4	2	2	2	4	2	2	2	2	4	4	4	4	4	4	4	5	5	5	
18	1	18	2	4	2	2	2	4	2	2	2	2	4	4	4	4	4	4	4	5	5	5	

**Figure 3** Calculation of care processes deviation from the care plan (on the example of Patients 1 and 3)

The distance according to the established scheme is  $R = 1.935$ . Results of pair comparison of all patients care processes are given in Table 3.

- [5] De Leoni, Massimiliano, Fabrizio Maria Maggi, and Wil MP van der Aalst. "Aligning event logs and declarative process models for conformance checking." *Business Process Management*. Springer Berlin Heidelberg, 2012. 82-97.
- [6] Gupta, Shaifali. "Workflow and process mining in healthcare." Master's Thesis, Technische Universiteit Eindhoven (2007).
- [7] Ing-Jr Ding, Chih-Ta Yen, Yen-Ming Hsu. *Developments of Machine Learning Schemes for Dynamic Time-Wrapping-Based Speech Recognition // Mathematical Problems in Engineering*. 2013.
- [8] Karthik, R., Menaka, R., Kulkarni, S., & Deshpande, R. (2014). Virtual doctor: an artificial medical diagnostic system based on hard and soft inputs. *International Journal of Biomedical Engineering and Technology*, 16(4), 329-342.
- [9] S. Leemans, D. Fahland, and W.M.P. van der Aalst. Exploring Processes and Deviations. In F. Fournier and J. Mendling, editors, *Business Process Management Workshops, International Workshop on Business Process Intelligence (BPI 2014)*, Lecture Notes in Business Information Processing, Springer-Verlag, Berlin, 2015.
- [10] S.J.J. Leemans, D. Fahland, and W.M.P. van der Aalst. Process and Deviation Exploration with Inductive Visual Miner. In L. Limonad and B. Weber, editors, *Business Process Management Demo Sessions (BPMD 2014)*, volume 1295 of CEUR Workshop Proceedings, pages 46-50. CEUR-WS.org, 2014.
- [11] Magoulas, George D., Andriana Prentza. "Machine learning in medical applications." *Machine Learning and its applications*. Springer Berlin Heidelberg, 2001, pp. 300-307.
- [12] Mans, Ronny S., Wil MP van der Aalst, and Rob JB Vanwersch. *Process Mining in Healthcare: Evaluating and Exploiting Operational Healthcare Processes*. Springer International Publishing, 2015.
- [13] Mans, R. S., Schonenberg, M. H., Song, M., van der Aalst, W. M., Bakker, P. J. "Application of process mining in healthcare—a case study in a dutch hospital." *Biomedical Engineering Systems and Technologies*. Springer Berlin Heidelberg, 2008, pp. 425-438.
- [14] Mans, R. S., van der Aalst, W. M., Vanwersch, R. J., & Moleman, A. J. "Process mining in healthcare: Data challenges when answering frequently posed questions." *Process Support and Knowledge Representation in Health Care*. Springer Berlin Heidelberg, 2013, pp. 140-153.
- [15] Mans, R. S., Schonenberg, M. H., Song, M., van der Aalst, W. M. P., Bakker, P. J. M. "Process Mining in Healthcare." Case study. Eindhoven University of Technology (2015).
- [16] Maruster, L., van der Aalst, W., Weijters, T., van den Bosch, A., Daelemans, W. Automated discovery of workflow models from hospital data. *Proceedings BNAIC-01, Amsterdam, October 25-26, 2001*, p.183-194.
- [17] Nazarenko, G. I., Kleimenova, E. B., Yashina, L. P., Molodchenkov, A. I., Payushchik, S. A., Konstantinova, M. V., Mokin, M.V., Otdelenov, V.A., Sychev, D. A. Development of the ontology of patient management technological records for modeling of clinical workflows in a general hospital. *Scientific and Technical Information Processing*, 2015, 42(6), pp. 455-462.
- [18] Isa, Nor Ashidi Mat. "Towards intelligent diagnostic system employing integration of mathematical and engineering model." *AIP Conference Proceedings*. Vol. 1660. No. 1. 2015.
- [19] Poelmans, Jonas, et al. "Combining business process and data discovery techniques for analyzing and improving integrated care pathways." *Advances in Data Mining. Applications and Theoretical Aspects*. Springer Berlin Heidelberg, 2010, pp. 505-517.
- [20] Romanenko A.A. Time series alignment: forecasting using DTW / Machine learning and data analysis, 2011. Vol. 1, No. 1, pp.77-85.
- [21] A. Rozinat and W.M.P. van der Aalst. Conformance Checking of Processes Based on Monitoring 9Real Behavior. *Information Systems*, 33(1):64–95, 2008.
- [22] Mohammad Shokoohi-Yekta, Bing Hu, Hongxia Jin, Jun Wang, Eamonn Keogh. Generalizing Dynamic Time Warping to the Multi-Dimensional Case Requires an Adaptive Approach (SDM 2015). - 2015 SIAM International Conference on Data Mining. - [http://www.cs.ucr.edu/~eamonn/Multi-Dimensional\\_DTW\\_Journal.pdf](http://www.cs.ucr.edu/~eamonn/Multi-Dimensional_DTW_Journal.pdf) (last access: 18.12.2015)
- [23] T.J.H. van de Steeg *Process Mining in Healthcare: Mining for cost and (near) incidents*, Eindhoven, 2015, 67 p.
- [24] Taha Samad-Soltani-Heris M. L., Mahmoodvand Z., Zolnoori M. *Intelligent Diagnosis of Asthma Using Machine Learning Algorithms*. – 2013.