JOHN ATANASOFF SOCIETY
OF AUTOMATICS AND INFORMATICS

JOHN ATANASOFF CELEBRATION DAYS

International Conference

AUTOMATICS AND INFORMATICS'2014

PROCEEDINGS

Edited by
Michail Petrov

Published by

JOHN ATANASOFF SOCIETY
OF AUTOMATICS AND INFORMATICS

Bulgaria, Sofia, October 1-3, 2014

International Conference
AUTOMATICS AND INFORMATICS'2014
1-3 October 2014, Sofia, Bulgaria

ANSI/ISA-95 PERSONNEL QUALIFICATION ADAPTIVE TESTING

P. Vasilev[1], E. Yanev[2], I. Gospodinov[2]

[1]*University of Chemical Technology and Metallurgy, Department of Industrial Automation*
8 Kl. Ohridski, 1756 Sofia, Bulgaria,
[2]*University of Chemical Technology and Metallurgy, Department of Programming and Computing Systems,*
8 Kl. Ohridski, 1756 Sofia, Bulgaria, E-mail: plamenvasilev2@gmail.com

Abstract: The abilities of personnel across a production plant have a significant impact over production and quality of products. Traceability of problems and their causalities is essential for every production and it is one of the major tasks of a MES/MOM system. The subject of this study is tracking and tracing the role of personnel abilities in quality products production. A Computerized Adaptive Testing algorithm using the Item Response Theory for adaptive testing of Personnel Qualification is presented in the article.

Keywords: ISA-95, MES/MOM, CAT, IRT, Qualification Test, Quality, Production

INTRODUCTION

Sustaining quality of products in Manufacturing Operations highly depends on the qualification of personnel and continuous education through qualification elevation programs and testing. Often these programs are underestimated because of the presumption that the operations in every-day ordinary work of personnel are routine and each operator should achieve equal results in the end of the day. But knowledge about the processes and procedures is not distributed equally throughout the personnel.

According to the standard ISO 9001:2008 the personnel that performs work, which affects the product conformity requirements, shall be competent on the basis of appropriate education, training, skills and experience [1]. Evaluating these abilities needs to be on a regular basis and should exclude external human factors, such that may introduce subjective attitude.

ANSI/ISA-95

ISA-95 is the international standard for integration of enterprise and control systems. ISA-95 consists of models and terminology. These can be used to determine which information has to be exchanged between systems for sales, finance, and logistics and systems for production, maintenance, and quality. This information is structured in UML models that are the basis for the development of standard interfaces between ERP and MES/MOM systems. The ISA-95 standard can be used for several purposes, for example as a guide for the definition of user requirements, for the selection of MES/MOM suppliers, and as a basis for the development of MES/MOM systems and databases [2].

Personnel Model in ISA-95

The Personnel Model in the ISA-95 standard defines the connections between the Person, his properties and Personnel Class (Roles) in the production, and the properties of the class.

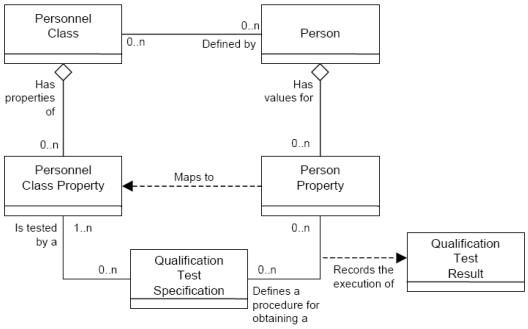

Fig. 1 UML Diagram of Personnel Model

In the model in Fig. 1:

Person is a specifically identified individual.

Person properties may include the current availability of a person and other current information, such as location and assigned activity, and the unit of measure of the current information.

Personnel class is a means to describe a grouping of persons with similar characteristics for purposes of scheduling and planning.

Each personnel class may have zero or more properties. Examples of personnel class properties for the personnel class "operators" may be "class 1 certified", "class 2 certified", "night shift", and "exposure hours". Production requests may specify required personnel class property requirements for a product segment.

Qualification Test Specification may be associated with a personnel class property or person property. This is typically used where a qualification test is required to ensure that a person has the correct training and/or experience for specific operations. A qualification test specification may test for one or more properties.

Qualification Test Result records the results from a qualification test for a specific person [3].

Based on the Personnel Model a *Qualification Test* is designed to correspond to specific Personnel Class properties that each Person, which is a member of this Class, must achieve. The Qualification Test Result may be further interpreted for providing specific certifications or classifications, needed for preparation of *Production Schedule* as *Personnel Requirement* and *Personnel Requirement Properties*.

For Example: there may be a requirement for one mechanic with a specified level of certification achieved, available five hours after start of production. There would be one personnel requirement for the requirement for the mechanic and two personnel requirement properties - one for the certification level, and one for the time requirement. [3]

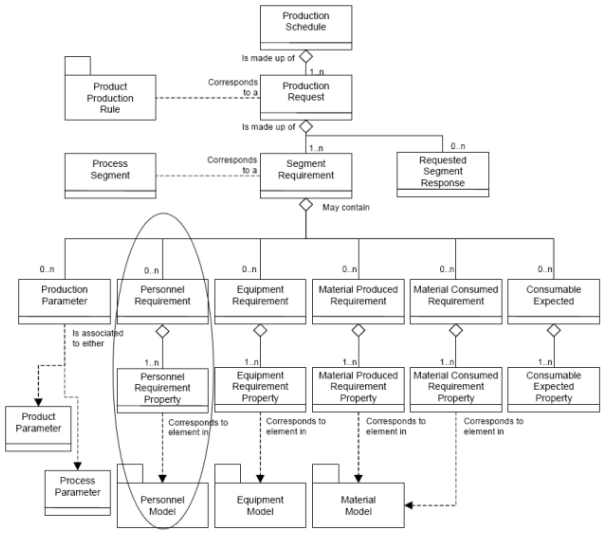

Fig. 2 UML diagram of Production Schedule Model [3]

Personnel Model Attributes

The second part of ANSI/ISA-95 standard defines the attributes of the object information model. The attributes represent a minimum set of industry-independent information. Attributes can be added depending of the needs of the specific production. Examples of additional attributes that can be added for Qualification Test Specifications can be: "Qualification Type" with values – "Operational", "Safety", "Foreign Language", etc. or "Validated by" with values – "Name of responsible person".

The attributes given for Qualification Test Specification are:

- Name: identifies the test for certifying one or more values of the person properties;
- Description: Additional information;
- Version: if there is more than one version of qualification test specification.

The attributes given for Qualification Test Results are:

- ID: A unique instance identification that records the results from the execution of a test identified in a qualification test specification for a specific person;
- Description: Additional information;
- Date: Date/Time of the test;
- Result: Pass/Fail, Alphabetical score, etc.;
- Result Unit of Measure
- Expiration: Date/Time of expiration, may be used for appointing next test (when personnel is being periodically tested) [4].

QUALIFICATION TESTS CLASSIFICATION

The purpose of Qualification Test Specifications in a MES/MOM system is to provide the necessary tool for tracking and tracing personnel knowledge and skills that correspond to specific personnel class (Cutting machine operator, Electrical engineer, etc.). For each personnel class there may be several test specifications depending of which personnel class property the test is related to. For Example: a Cutting machine operator may need operational certificate, occupational safety certificate, or computer skills certificate in his every-day work, so he must pass an exam every year (or every six months).

Each test must contain a bank of questions, covering knowledge range in the given area, with different levels of difficulty by which the skills of a given person to be evaluated. Example of questions with different difficulty level for operational test of cutting machine operator may be:

- What is the maximum width of a coil that can pass through the cutting machine MAS?
- What is the maximum number of strips that can be cut on Fröhling?
- What is the RI (Return Index) of GRS?
- What is the maximum speed of cutting material with Alloy – 850 (Semi-red brass), Temper – Extra Hard and Thickness – 0,5 mm?

A good practice is the bank of these questions to be updated and expanded continuously with examples taken from the documentation describing the processes (job descriptions, workflows, technical specifications, production rules, etc.) and real situations emerging in production process. The degrees of freedom of the test are proportional to the number of questions, while accuracy of the test results is proportional to the number of difficulty levels.

COMPUTERIZED ADAPTIVE TESTING

Computerized adaptive testing (CAT) utilizes computers to administer effective tests that are tailored to each examinee's ability. During every test taking step CAT uses an adaptive algorithm for finding a suitable test item (question) based on the information obtained from the answers on the previous items. Typically CAT employs the Item Response Theory (IRT) [5] for searching of optimal items and for evaluating the examinee's ability. IRT is a modern theory that assigns a probability P of an examinee with an ability θ to correctly answer a test item, given by:

$$P(\theta) = c + \frac{1-c}{1+e^{-a(\theta-b)}} \quad (1)$$

Every item has three parameters:

- a – discrimination
- b – difficulty in range [-3 ÷ 3]
- c – pseudo-guessing chance [0 ÷ 1]

The ability θ is measured in z-score and usually varies in the interval [-3 ÷ 3]. The value of $\theta = -3$ corresponds to very low ability whereas the value of $\theta = 3$ corresponds to very high ability. IRT is often used for adaptive testing because personal trait of the examinee θ is on the same scale as the item

difficulty *b*. This allows us to get an estimate for examinee's ability from the answer to each test item and to select the next test item with difficulty that is close to this estimated ability. The probability of a correct answer *P* to an item with discrimination $a = 2$, difficulty $b = 0$, and pseudo-guessing chance $c = 0.25$, the so called Item Response Function, is shown in Figure 3.

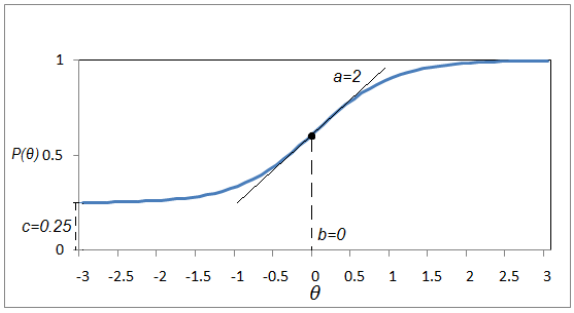

Fig. 3 Item Response Function

In order to create a software platform for adaptive testing five main components [6] that realize the algorithm must be present:

- A calibrated item bank.
- A starting point of the test.
- An item selection algorithm.
- A scoring procedure.
- Termination criteria.

The *item bank* contains many items, each with a particular set of IRT parameters – discrimination, difficulty, and pseudo-guessing chance. Usually CAT employs Multiple Choice items with one correct answer. Developing of item banks is an involved process. The bank needs to contain properly calibrated and validated items, but this is not enough. Also, the bank needs to be reliable, i.e. it needs to be able to produce precise CAT ability estimates (that are close to the true abilities) after administering reasonably short sequence of items to each examinee, regardless of his/her ability.

The *starting point* of the test is the initial estimate of the examinee's ability that determines the difficulty of the first test item. If no prior information about the examinee's ability is available it is simply guessed. Typically, the most sensible guess is to start with an item with difficulty somewhere in the middle of the ability range.

The *item selection algorithm* is usually based on IRT because it places examinees and item difficulty on a same scale [-3 ÷ 3]. The ability estimation is a real-time process. The ability estimate is updated after each test item using the likelihood function *L(θ)*. The item response *u* takes a value of 1 if the item is answered correctly and a value of 0 if otherwise. For a sequence of *N* administered items a sequence of responses u_i, $i=1\ldots N$, is obtained (*i* is the item index). For this sequence, the likelihood function *L(θ)* is found by:

$$L(\theta) = \prod_{i=1}^{N} P_i^{u_i}(\theta)(1 - P_i(\theta))^{(1-u_i)} \qquad (2)$$

There are several ways of extracting estimated ability $\hat{\theta}$ from *L(θ)* - the maximum likelihood estimation (MLE) method [7], the expected a posteriori (EAP) method, the maximum a posteriori (MAP) method, etc.

After obtaining the value of $\hat{\theta}$ for *N* question items a search algorithm searches the item bank for an item with difficulty close to it.

For *Scoring procedure* can be used alphabetical or percentage scale. For any of the pre-requested scoring scales proper z-score normalization of the estimated ability must be made.
The CAT test length varies from examinee to examinee. There are several ways to terminate the testing process:

- Reaching a plateau in $\hat{\theta}$
- Reaching below some prescribed Maximum value of Standard Error - SEmax.
- Exceeding some prescribed minimal number of administered items Nmin,
- Reaching some prescribed maximal number of administered items Nmax,
- Reaching a prescribed maximal test time Tmax.

The basic adaptive testing algorithm (Figure 4) includes the following iteration steps:

1. The Item Bank is searched for the optimal item, based on the current estimate of the examinee's ability.
2. The chosen item is administered to the examinee.
3. The ability estimate is updated based upon all prior answers.
4. Steps 1–3 are repeated until one or more of the termination criteria, set by the test administrator, are satisfied.

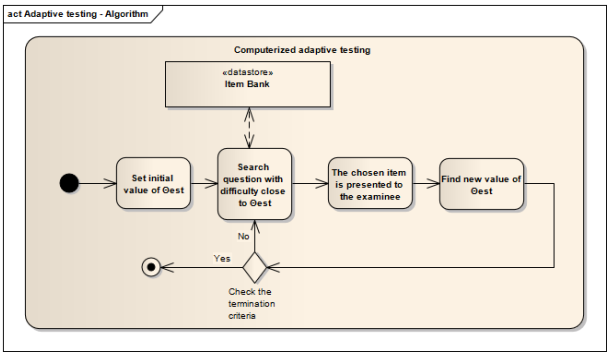

Fig. 4 Basic Adaptive Testing algorithm

QUALIFICATION ADAPTIVE TEST AND RESULTS

The first item (the starting point) is typically chosen to be with medium difficulty (i.e. $b = 0$). If the answer to this item is correct the adaptive algorithm finds a next item with a difficulty higher than 0. If otherwise, the difficulty of the next item is lower than 0.

Fig. 5 CAT Qualification Test on the fourth question

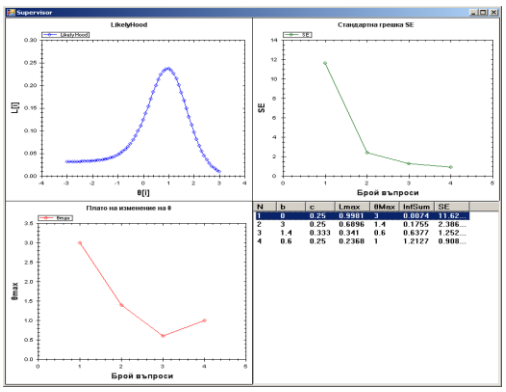

Fig. 6 CAT Qualification Test Results (Server side) in real-time

The results from the test may be seen by the test administrator on the server side. They include charts of the likelihood function, the Standard Error, the evolution of $\hat{\theta}$, and a grid with the values of the discrimination, the difficulty, and the pseudo-guessing chance of all the administered items, as well as the estimated ability, the Standard Error, etc.
The end of the Qualification Test (i.e. SE ≤ 0,29) brings up a pop-up with the final result (score), which is normalized by a predefined grading scale (i.e. alphabetical result C+).

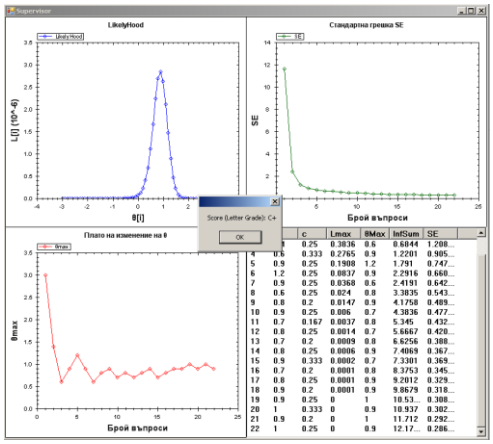

Fig. 7 CAT Qualification Test final result

CONCLUSIONS

The importance of Qualification Tests has been previously examined. This kind of adaptive testing in the production thou make sense only in composition of MES/MOM ISA-95-based system, where the results of adaptive testing of personnel can be further confronted with the rest of the production information, put together in a deep analysis for searching causalities in quality for every production step. On Fig. 8 and Fig. 9 with Pareto charts are represented the defects admitted by two different operators on a same machine for same period of time. The reason for these two significantly different indicated factors may be correlated with the operators' abilities estimated by a qualification test.

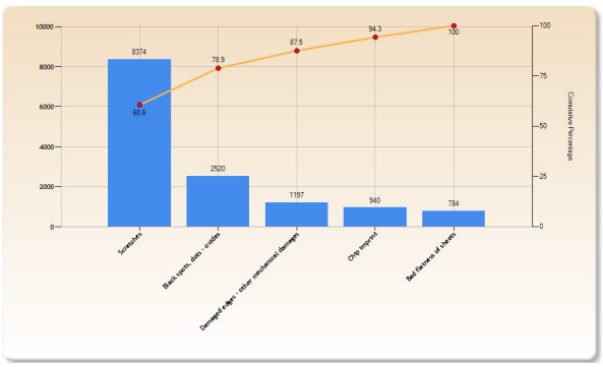

Fig. 8 Pareto chart of the defects in the production (Operator 1)

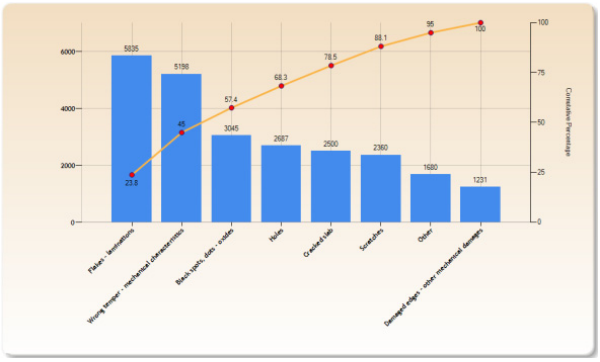

Fig. 9 Pareto chart of the defects in the production (Operator 2)[8][9]

REFERENCES

1. ISO 9001:2008 Quality management systems - Requirements
2. http://www.isa-95.com/
3. ISA–95.00.01–2000 Enterprise-Control System Integration Part 1: Models and Terminology
4. ISA–95.00.02–2005 Enterprise-Control System Integration Part 2: Object Model Attributes
5. Hambleton, R., Swaminathan, H., Rogers, H.: Fundamentals of Item Response Theory, CA: Sage Press, 1991.
6. Weiss, D. J., Kingsbury, G. G.: Application of computerized adaptive testing to educational problems. Journal of Educational Measurement, 1984.
7. Thompson, N.A.. "Ability estimation with IRT", 2009
8. NearSoft MOM4 – www.nearsoft.eu
9. V. Filipov, N. Christova, A Solution for Integrated Manufacturing Operation Management, Proceeding of International Conferences on Computational Intelligence for Modelling, Control and Automation (CIMCA), Austria, Vienna, 10-12 December 2008