A Model for E-voting Systems Evaluation

Aristides Dasso, Ana Funes

Universidad Nacional de San Luis, Ejército de Los Andes 950

D5700HWZ San Luis, Argentina

{arisdas, afunes}@unsl.edu.ar

Abstract. This work presents a model for the evaluation of electronic voting systems (e-voting). Although these systems are used more and more there have been reports of different kind of problems, underlining the need of having models for their evaluation. The model we present here is based primarily on the Logic Score of Preference (LSP) method, and the requirements are taken from the requirements developed by the Technical Guidelines Development Committee of the Election Assistance Commission of the United States. An overview of the LSP method is given as well as the reasons behind the decision to choose these requirements.

Keywords: Electronic voting, e-voting, System Evaluation, Continuous Logic, Electronic Government, E-government, LSP Method.

1 Introduction

Voting systems come in various forms and characteristics. The most common is the well-known ballot paper. However along the years others methods have been adopted seeking to improve the process by speeding it –both the actual vote emission and the vote counting operation– and also trying to make it less vulnerable both to unintentional errors as well as fraudulent attacks.

Introduction of different machinery have been the main form of improving the ballot. Lately the preferred machinery has been computers, either to allow the ballot emission remotely or directly.

At the same time, innumerable examples of problems with all kind of machinery, including computers, have prompted the need to evaluate not only the machinery, but also the method that the machinery institutes. Taking into account this, different government institutions have realized the need to established clear requirements for the new methods. The problems encountered have also forced the creation of governmental bodies charged with the task of producing sets of test requirements that can serve as a basis to construct a general model to evaluate systems that relies on computers.

In this work, we present a model for the evaluation of voting systems that rely on computers. To achieve this goal we have used specifications given by different organisms, but specially those produced by the Technical Guidelines Development Committee (TGDC) of the Election Assistance Commission (EAC). As the EAC says in its site: “The U.S. Election Assistance Commission was established by the Help America Vote Act of 2002 (HAVA). As an independent, bipartisan commission, EAC develops guidance to meet HAVA requirements, adopts voluntary voting system guidelines, and serves as a national clearinghouse of information on election administration. EAC also certifies voting systems, accredits test laboratories, and audits the use of HAVA funds” [13].

Note that this commission “certifies voting systems, accredits test laboratories” and for that very reason has a Technical Guidelines Development Committee [10] that develops comprehensive technical guidelines to which the voting systems must comply. Bear in mind that the USA is probably the country where most different machinery has been used and tried and where computers have been extensively employed.

In general, systems that employ computers to assist on the ballot are called Direct Record Equipment (DRE). They are also commonly called e-voting systems. We have taken from the report issue by TGDC (see [10]), the requirements that most apply to DRE systems.

The rest of this work is organized as follows: Section 2 presents an overview of the LSP method and at the same time, we introduce the different parts of our model. In Subsection 2.1, we show part of our requirement tree; some elementary criteria are given in Subsection 2.2 and part of the aggregation structure is presented in Subsection 2.3. Finally, in Section 3, we close the work with some conclusions and future work.

2 The LSP Method and its Application to e-voting Systems Evaluation

The method employed to construct the model presented here is the Logic Scoring of Preference (LSP) [2], [3], [4], [5]. This is a method for the realization of complex criterion functions and their application in the evaluation, optimization, comparison and selection of general complex systems.

The LSP method can be used to evaluate complex systems but since it is a general evaluation method, it can also be employed in evaluation processes involved in the e-voting area.

Since this method is not a simple additive scoring method but allows the use of complex and/or decisions, is especially useful where these conditions apply.

As a starting point in LSP, it must be clearly determined what are the user requirements, the main attributes of the system and their value preferences. These attributes are called *performance variables*. Each one of these variables is mapped into an *elementary preference* by defining and applying the corresponding *elementary criteria*.

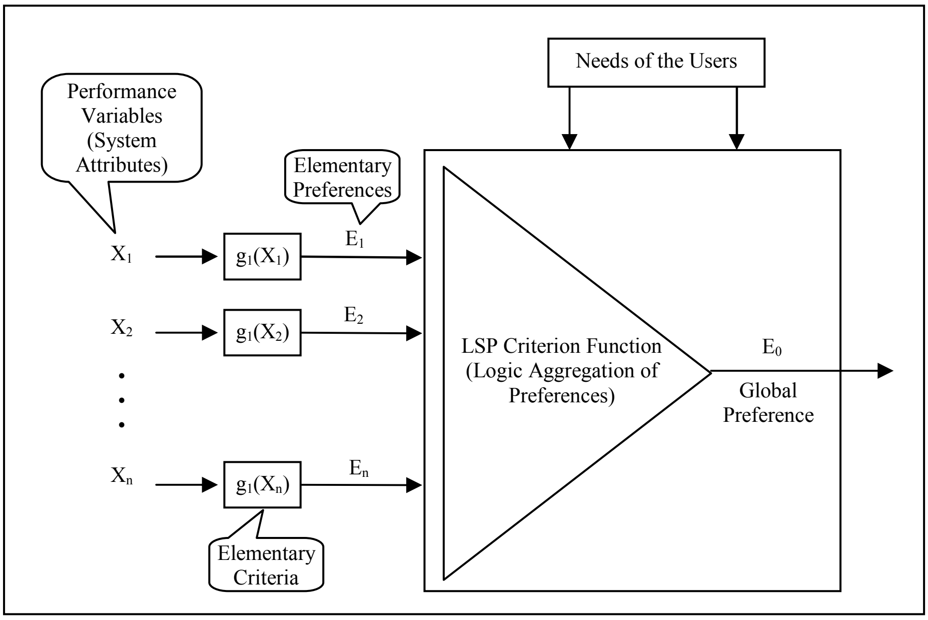

Figure 1. An overview of the LSP evaluation process.

An elementary criterion is a function that transforms a real value, coming from a performance variable, into a value belonging to the [0,100] interval called elementary preference. It represents the degree of fulfilment of a requirement. Therefore, to define an elementary criterion it is necessary to have some previous experience to determine what the acceptable range of values is for the corresponding performance variable.

All the elementary preferences are used as input to the *LSP criterion function*. This function yields a single global indicator of the degree of fulfilment of the system requirements. The LSP criterion function is built by aggregating the elementary preferences. To aggregate preferences means to replace a group of preferences (the input preferences) by a single preference (the output preference). It denotes the degree of satisfaction of the evaluator with respect to the whole group of input preferences.

To calibrate the LSP criterion function it is necessary to take into account the needs of the end users. The process of calibration obviously represents the most complex phase in the whole evaluation. The global preference –obtained as output of the LSP function– is the result of the combination of the elementary preferences taking into account both the relative importance of each preference and the necessary logic relationship between them.

Once the calibration of the LSP criterion function is finished, each competitive system can be evaluated. By giving as input to the model the set of values corresponding to its performance variables a global performance indicator is obtained as output for each competitive system. A condensed overview of the applied approach is presented in Figure 1.

We followed the steps below to create the final evaluation model (or LSP criterion function), which reflects the requirements a reasonable e-voting system must fulfil, and yields a quantitative indicator of the requirement satisfaction level (the global preference):

- Development of the system requirement tree.
- Definition of elementary criteria.
- Aggregation of preferences.

2.1 System Requirement Tree

In order to develop an exhaustive list of requirements, we applied a hierarchical decomposition process for requirement derivation. At the beginning, we defined all major groups of requirements, and then through successive decompositions each group was decomposed into subgroups. By repeating this process, the *system requirement tree* was obtained, where the tree leaves correspond to the performance variables.

In the case being considered here, the requirement tree comes out of the "Voluntary Voting System

Table 1. Item 3.1 of the Requirement Tree: General Usability (Performance, Functional Capabilities and Privacy)

3.1 General Usability
3.1.1. Performance
3.1.1.1. Overall performance metrics
3.1.1.1.1. General
3.1.1.1.1.1. Total completion performance
3.1.1.1.1.2. Perfect ballot performance
3.1.1.1.1.3. Voter inclusion performance
3.1.1.1.2. Usability metrics
3.1.1.1.2.1. Effectiveness metrics for usability
3.1.1.1.2.2. Voting session time
3.1.1.1.2.3. Average voter confidence
3.1.1.2. Usability testing by manufacturer for general population
3.1.2. Functional capabilities
3.1.2.1. General
3.1.2.1.1. Notification of effect of overvoting
3.1.2.1.2. Undervoting to be permitted
3.1.2.1.3. Correction of ballot
3.1.2.1.4. Notification of ballot casting
3.1.2.2. Editable interfaces
3.1.2.2.1. A Prevention of overvotes
3.1.2.2.2. Warning of undervotes
3.1.2.2.3. Independent correction of ballot
3.1.2.2.4. Ballot editing per contest
3.1.2.2.5. Contest navigation
3.1.2.2.6. Notification of ballot casting failure
3.1.2.3. Non-Editable interfaces
3.1.2.3.1. Notification of overvoting
3.1.2.3.2. Notification of undervoting
3.1.2.3.3. Notification of blank ballots
3.1.2.3.4. Ballot correction or submission following notification
3.1.2.3.5. Handling of marginal marks
3.1.2.3.6. Notification of ballot casting failure
3.1.3. Privacy
3.1.3.1. Privacy at the polls (System support of privacy)
3.1.3.1.1. Visual privacy
3.1.3.1.2. Auditory privacy
3.1.3.1.3. Privacy of warnings
3.1.3.1.4. No receipts
3.1.3.2. No recording of alternative format usage
3.1.3.2.1. No recording of alternative languages
3.1.3.2.2. No Recording of accessibility features

Guidelines Recommendations to the Election Assistance Commission" [10].

As it is said in the document's overview, "This document represents a recommendation from the Technical Guidelines Development Committee to the Election Assistance Commission for a voting system standard written to address the next generation of voting equipment. It is a complete re-write of the Voluntary Voting System Guidelines (VVSG) of 2005 and contains new and expanded material in many areas, including reliability and quality, usability and accessibility, security, and testing. The requirements are more precise, more detailed, and written to be clearer to voting system manufacturers and test laboratories. The language throughout is written to be readable and usable by other audiences as well, including election officials, legislators, voting system procurement officials, various voting interest organizations and researchers, and the public at large."

We have used this document as a basis to construct the preference tree of our Continuous Logic Model to evaluate e-voting systems. The document is an extensive and thorough list of requisites. It was created by a government body of the United States, that was charged by law with the task of producing testing requisites for e-voting systems, in a country that have an extensive experience in using different voting systems for a long time (remember that, in USA, the individuals states can decide locally how to implement the voting system) and that have had its share of failures with voting systems, particularly e-voting ones.

Be aware that the document in question applies not only to DRE, that is a denomination used for voting systems using electronic devices to record and/or print the vote, but also to other systems. However, DRE devices that are directly linked with what is called e-voting are in the most general class of devices and systems (see [10] INTRODUCTION, CH 2, Page 4), therefore most categories applied to DREs.

For reasons of space, we are presenting in this paper only a part of the model. We have chosen to illustrate our model with part of the requirement tree, that is shown in Table 1, and that corresponds only to the items "Performance Requirements", "Functional Capabilities" and "Privacy" of section "Usability, Accessibility, and Privacy Requirements" of the report. Bear in mind that the report has more than a thousand items, and although some of them are specific to the particular test to be performed and were deemed by us not suitable to be included in our model there are a significant number of requirements left. We refer the interested reader to the already mentioned document for a complete explanation of every item.

2.2 Definition of Elementary Criteria

Once we developed the requirement tree and determined the performance variables, we defined the elementary criteria.

Table 2. Some Elementary Criteria for Items 3.1.1. and 3.1.2.

3.1.1. Performance
- 3.1.1.1. Overall performance metrics
  - 3.1.1.1.1. General
    - 3.1.1.1.1.1. Total Completion Performance (TCP) $E(TCP) = (TNU_{cb} / TNU) \times 100$
      where TNU is the Total Number of Users,
      TNU_{cb} is the TNU that successfully cast a ballot.
    - 3.1.1.1.1.2. Perfect Ballot Performance (PBP)

$$E(PBP) \begin{cases} 100 \text{ si } TCB_{Err} \leq 1 \\ 100/TCB_{Err} \quad \text{si} \quad (TCB_{NErr}/TCB_{Err}) \\ 0 \quad \text{si} \quad (TCB_{NErr}/TCB_{Err}) \quad \leq \end{cases}$$

      where TCB is the Total number of Cast Ballots,
      TCB_{NErr} is the TCB without errors,
      TCB_{Err} is the TCB containing one or more errors.
    - 3.1.1.1.1.3. Voter Inclusion Performance: see [8], page 20.
  - 3.1.1.1.2. Usability metrics
    - 3.1.1.1.2.1. Effectiveness metrics for usability (EMU)
      E(EMU) = Obtained from a questionnaire
    - 3.1.1.1.2.2. Voting Session Time (VST)

$$E(VST) = \begin{cases} 0 \text{ si } ATB > MAT \\ ((MAT - ATB)/MAT) \times 100 \\ \qquad \text{si } ATB \leq MAT \end{cases}$$

      where TCB is the Total number of Cast Ballots,
      TTBC is the Total Time of Ballot Casting,
      ATB = TTBC / TCB,
      MAT is the Maximum Acceptable ballot casting Time.
    - 3.1.1.1.2.3. Average Voter Confidence (AVC)
      E(AVC) = Obtained from a questionnaire
- 3.1.1.2. Usability testing by manufacturer for general population (UTM).
  E(UTM) = Obtained from a questionnaire

3.1.2. Functional capabilities
- 3.1.2.1. General
  - 3.1.2.1.1. Notification of effect of overvoting (NOv)

$$E(NOv) = \begin{cases} 0 \text{ si Notification } \neg\exists \\ 100 \text{ si Notification } \exists \end{cases}$$

  - 3.1.2.1.2. Undervoting to be permitted (UP)

$$E(UP) = \begin{cases} 0 \text{ si permitted } \neg\exists \\ 100 \text{ si Permitted} \end{cases}$$

  - 3.1.2.1.3. Correction of ballot (CB)

$$E(CB) \begin{cases} 0 \text{ si Correction } \neg\exists \\ 100 \text{ si Correction } \exists \end{cases}$$

  - 3.1.2.1.4. Notification of ballot casting (NBC)

$$E(NBC) = \begin{cases} 0 \text{ si Notification } \neg\exists \\ 100 \text{ si Notification } \exists \end{cases}$$

An elementary criterion is defined as a mapping from a performance variable value into an elementary preference value. Since the interpretation of an elementary preference is the degree of fulfilment of a given requirement, this is a real number belonging to the unit interval [0, 100]. Therefore, the value 0 corresponds to a situation where the performance variable does not satisfy the requirements and the value 100 to the case that the requirement is completely fulfilled, while the values between 0 and 100 denote a partial satisfaction of requirements.

Given that only the leaves of the requirement tree correspond to performance variables, the elementary criteria must be applied only to the requirement tree's leaves.

Some of the elementary criteria used for obtaining the elementary preferences in our model are shown in Table 2. The table shows some of the elementary criteria for items 3.1.1. "Performance Requirements" and 3.1.2. "Functional capabilities". They are mainly based on the Voting Performance Protocols defined in [8] or [10]. When neither [8] nor [10] give a measure or a Voting Performance Protocol (VPP) is not suitable or clear enough we devised one of our own. For a more complete understanding of Table 2 we refer the reader to the explanation given in the referenced documents.

2.3 Aggregation of Preferences

Once we finished the requirement tree, we began with the aggregation of preferences. This process uses the structure of the system requirement tree to build a new tree structure, the final aggregation structure or LSP criterion function.

The process starts by aggregating groups of related elementary preferences and generating, in this way, subsystem preferences. Therefore, the elementary preferences, corresponding to the requirement tree leaves, are aggregated in new preferences. This bottom up process is repeated with the resulting groups of preferences until a single global preference can be computed. The logic aggregation structure created by applying the described process must reflect the user requirements that in this case are the requirements for an e-voting system.

If we want to aggregate n elementary preferences $E_1,...,E_n$ in a single preference E, the resulting preference E –interpreted as the degree of satisfaction of the n requirements– must be expressed as a function having the following properties:

1. The relative importance of each elementary preference E_i (i= 1...n) can be expressed by a weight W_i,

2. $min(E_1,...,E_n) \leq E \leq max(E_1,..., E_n)$.

These properties can be achieved using the weighted power means:

$$E(r) = (W_1 E_1^r + W_2 E_2^r + ... + W_n E_n^r)^{1/r}, \text{ where}$$

$$0 < W_i < 100, \; 0 \leq E_i \leq 100, \; i = 1, ..., n, \; W_1 + ... + W_n = 1, \; -\infty \leq r \leq +\infty$$

The choice of r determines the location of $E(r)$ between the minimum value $E_{min}=min(E_1,...,E_n)$ and the maximum value $E_{max}=max(E_1,...,E_n)$. For $r = -\infty$ the weighted power mean reduces to the pure conjunction (the minimum function) and for $r = +\infty$ to the pure disjunction (the maximum function), giving place to a Continuous Logic Preference (CPL). The range between pure conjunction and pure disjunction is usually covered by a sequence of equidistantly located CPL operators named: C, C++, C+, C+–, CA, C–+, C–, C– –, A, D– –, D–, D–+, DA, D+–, D+, D++, D. For a more detailed description of the technique for selection of r see [6] and [7].

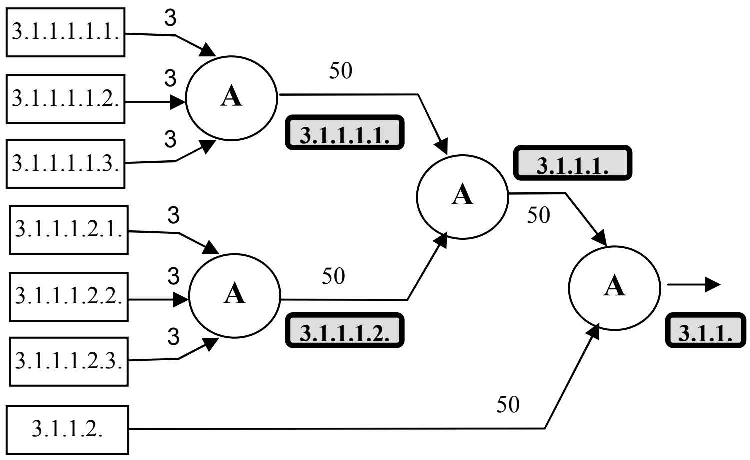

Figure 2. Aggregation Structure for item 3.1.1. "Performance".

In Figures 2 to 4 we show the aggregation structures corresponding to items 3.1.1. "Performance", 3.1.2. "Functional capabilities", 3.1.3. "Privacy" and finally in Figure 5 the item 3.1. "General Usability", of the requirement tree given in Table 1. In the figures, circles represent the values of r (CPL operators), rectangles correspond to the elementary preferences and the numbers on the edges to the corresponding weights. Rounded rectangles in light grey do not form part of the aggregation structure. We have introduced them to indicate the corresponding subsystem in the requirement tree.

In Figure 2, we have considered all items as non mandatory meaning that if any one of them is missing (zero) then the whole structure will not necessarily return zero. We have made this decision since they are items that refer to performance values in the requirement tree that are desirable but non mandatory.

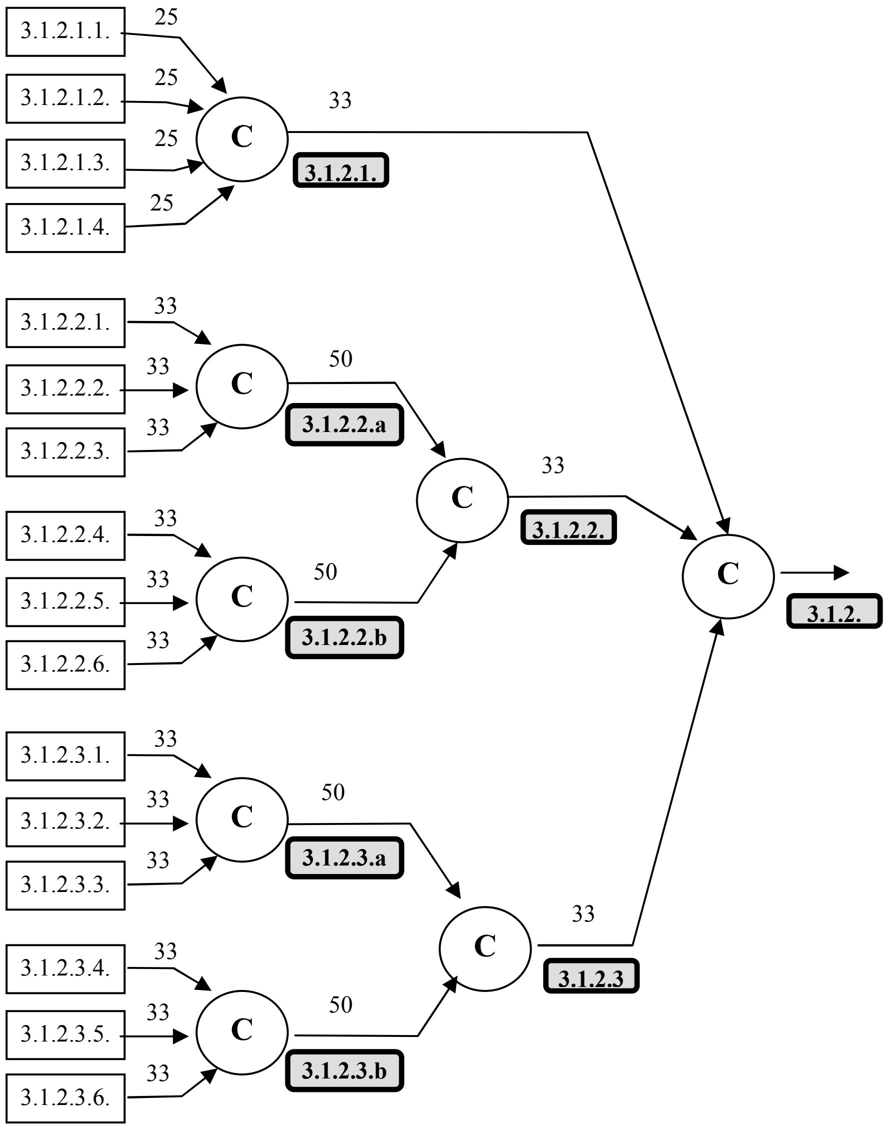

Figure 3. Aggregation Structure for item 3.1.2. "Functional capabilities"

This is the reason for the choice of the **A** operator that allows to compute an arithmetic media of the input preferences.

In Figure 3, we show the elementary preferences for item 3.1.2. “Functional capabilities”. Since there are six elementary preferences for Items 3.1.2.2. “Editable interfaces” and 3.1.2.3. “Non-Editable interfaces” and the CPL operators are calculated for five parameters, we have separated them in groups of three and applied the same operator three times (once for each group of three items and once for the result of both). Hence, the light grey rounded rectangles 3.1.2.2.a, 3.1.2.2.b, 3.1.2.3.a, 3.1.2.3.b are there to show this disposition.

To choose the CPL operators for the structure shown in Figure 3, we have followed a strict policy. Since we have considered that all the performance variables in item 3.1.2. “Functional capabilities” are essential, namely none of these items must be lacking, i.e. we have considered them mandatory, and that is to say, all of them must be present. This sub structure will be rejecting most systems except those that comply strictly with all the requirements in item 3.1.2., at least partially if not completely. All the used CPL operators are **C** (strict conjunction) so if any of the items is zero (not present) then the whole structure will evaluate to zero regardless of the other item’s values.

The aggregation structure for item 3.1.3. “Privacy” appears in Figure 4. We considered this item very important for the actual ballot casting since it has to do with the privacy of such an act, therefore the item in general is mandatory as well as the items that integrate it, that is why we have chosen a **C** operator to aggregate its items.

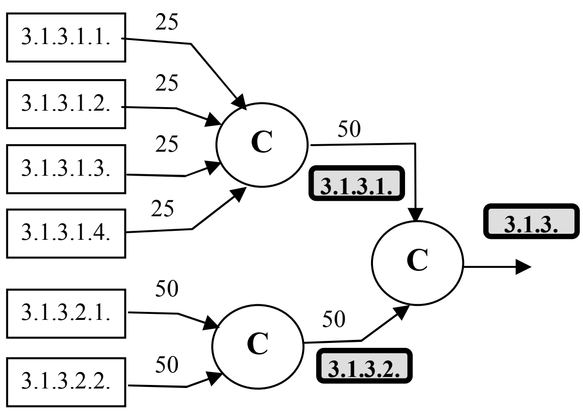

Figure 4. Aggregation Structure for item 3.1.3. “Privacy”

Figure 5 is the aggregation structure for item 3.1. “General Usability (Performance, Functional

Capabilities and Privacy)" that aggregarets the structures shown in Figures 2 to 4. We have aggregated two if its sub items (3.1.2, and 3.1.3) using a mandatory **C** operator. This means that if any of them is zero the result will be zero.

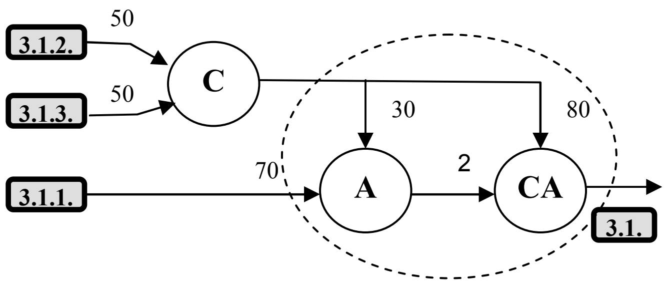

Figure 5. Aggregation Structure for item 3.1. "General Usability".

Finally, we have aggregated item 3.1.1. with the output of the **C** operator by applying a structure called *partial absorption* (circled with a dotted line in the figure). Using a *partial absorption* structure will balance the strict policy adopted for items 3.1.2 and 3.1.3. *Partial absorptions* are useful when trying to join a mandatory item *m* with optional items. If the mandatory item *m*=0 then the result of the partial absorption is zero whatever the value of the optional items are. Otherwise, the output is the mean of the range ($m-\delta^-$, $m+\delta^+$). δ^- y δ^+ determining the weights to be used. δ^- y δ^+ are obtained from a pre calculated table; see [16] for more on this.

Conclusions and Future Work

We have presented a model for the evaluation of e-voting systems based on the recommendations given in the report issued by the Technical Guidelines Development Committee of the Election Assistance Commission of the USA. Now that more and more these systems are starting to be used globally, it is important having tools to compare and evaluate their different capacities (security, accuracy, etc.).

We have built the evaluation model by following the steps proposed by the LSP Method, which is based on a Continuous Logic.

There is a number of other basic requirements produced by different institutions, some of them governmental institutions some not (see, for example, [1], [9], [12], [14] and [15]). Part of our present and future work is focused on revising our model considering some of the proposals made in those documents.

It is important to remark that the document used as a basis for our model does not include systems' cost as an item. Cost is an important aspect to be considered when implementing a system, however its evaluation is very complex and it warrants a complete and extensive model, since the different facets to be considered are numerous and not trivial. Items as being considered here are, not only equipment's cost, but also maintenance, amortization, storage, transportation, etc. This is another area that is being approached and where we expect to have results to show in future publications.

There is also more work to be done both in improving the model in the light of the experience gained in the application of e-voting systems –see for instance [11] and [17]– but also in the actual laboratory testing of the machinery involved. As such, governmental institutions in every country implementing these systems should get involved in the testing and eventually in the certification of such systems. These are important issues relating directly with the well-being of our institutions and our democracy.

References

[1] Council of Europe. Council of Europe activities in the field of e-democracy http://www.coe.int/t/dgap/democracy/Activities/GGIS/CAHDE/Default_en.asp

[2] Jozo J. Dujmovic, "Continuous Preference Logic for System Evaluation", IEEE Transactions on Fuzzy Systems, Vol. 15, Nº 6, December 2007.

[3] J. J. Dujmovic, "Preferential Neural Networks. Chapter 7 in Neural Networks – Concepts, Applications, and Implementations", Vol. II. Edited by P. Antognetti and V. Milutinovic. Prentice Hall Advanced Reference series, Prentice – Hall, pp 155-206, 1991.

[4] J. J. Dujmovic, "A Method for Evaluation and Selection of Complex Hardware and Software Systems", The 22nd International Conference for the Resource Management and Performance Evaluation of Enterprise Computing Systems. CMG96 Proceedings, vol. 1, pp.368-378, 1996.

[5] J. J. Dujmovic, "Quantitative Evaluation of Software", Proceedings of the IASTED International

Conference on Software Engineering, edited by M.H. Hamza, pp. 3-7, IASTED/Acta Press, 1997.

[6] J. J. Dujmovic and A. Bayucan, "Evaluation and Comparison of Windowed environments", Proceedings of the IASTED International Conference Software Engineering (SE'97), pp 102-105, 1997.

[7] J. J. Dujmovic and R. Elnicki, "A DMS Cost/Benefit Decision Model: Mathematical Models for Data management System Evaluation, Comparison, and Selection", National Bureau of Standards, Washington, D.C., No. NBS-GCR-82-374, NTIS No. PB82-170150 (155 pages), 1982.

[8] Human Factors and Privacy Subcommittee of the TGDC, "Usability Performance Benchmarks for the VVSG," August 2007. Available from http://vote.nist.gov/meeting-08172007/Usability-Benchmarks-081707.pdf.

[9] International Foundation for Electoral Systems. http://www.ifes.org

[10] Technical Guidelines Development Committee, "Voluntary Voting System Guidelines Recommendations to the Election Assistance Commission". August 31, 2007. http://www.eac.gov/testing_and_certification/voluntary_voting_system_guidelines.aspx

[11] María I. Tula (coordinadora), Voto Electrónico, Ariel, Buenos Aires 2005.

[12] ACM. http://usacm.acm.org/usacm/Issues/EVoting.htm

[13] US Election Assistance Commission. http://www.eac.gov/

[14] USENIX. http://www.usenix.org

[15] Verified Voting Foundation. http://www.verifiedvotingfoundation.org

[16] Stanley Y. W. Su, Jozo Dujmovic, D. S. Batory, S. B. Navathe, and Richard Elnicki. "A Cost-Benefit Decision Model: Analysis, Comparison, and Selection of Data Management Systems". ACM Transactions on Database Systems, Vol. 12, No. 3, September 1987.

[17] Council of Europe. E-voting documentation. http://www.coe.int/t/dgap/democracy/Activities/GGIS/E-voting/EVoting_Documentation/Default_en.asp#TopOfPage. (Retrieved May 10, 2011)