Analysis of the Behavior and Reliability of Voting Systems Comprising Tri-State Units

Sherif Yacoub
Information Infrastructure Laboratory
HP Laboratories Palo Alto
HPL-2002-117
April 25th , 2002*

E-mail: sherif_yacoub@hp.com

voting, reliability analysis, decision making

Voting is a commonly used technique in combining results from peer experts. In distributed decision making systems, voting mechanisms are used to obtain a decision by incorporating the opinion of multiple units. Voting systems has many applications in fault tolerant systems, mutual exclusion in distributed systems, and replicated databases. We are specifically interested in voting systems as used in decision-making applications. The voting system studied in this paper consists of N units, each has three states: correct (success), wrong (failed), and abstain (did not produce an output). The final output of the decision-making (voting) system is correct if a consensus is reached on a correct unit output, abstain if all units abstain from voting, and wrong otherwise.

In this paper, we describe a synthetic experimental procedure to study the behavior of voting systems using a simulator that we developed to: analyze the state of each expert, apply a voting mechanism, and analyze the voting results. For this analysis, we study the following behaviors of a voting system: 1) the reliability of the voting system, "R"; 2) the probability of reaching a consensus, "Pc"; 3) certainly index, "T"; and 4) the confidence index, "C". The configuration parameters controlling the analysis are: 1) the number of participating experts, "N", 2) the possible output states of an expert, and 3) the probability distribution of each expert states. Results of this study unleash several behaviors of a decision-making system with tri-state experts as function of various configuration parameters.

* Internal Accession Date Only

Analysis of the Behavior and Reliability of Voting Systems Comprising Tri-State Units

Sherif Yacoub
Hewlett-Packard Laboratories
1501 Page Mill Rd., MS 1126
Palo Alto, CA 94304
sherif_yacoub@hp.com

Abstract

Voting is a commonly used technique in combining results from peer experts. In distributed decision making systems, voting mechanisms are used to obtain a decision by incorporating the opinion of multiple units. Voting systems has many applications in fault tolerant systems, mutual exclusion in distributed systems, and replicated databases. We are specifically interested in voting systems as used in decision-making applications. The voting system studied in this paper consists of N units, each has three states: correct (success), wrong (failed), and abstain (did not produce an output). The final output of the decision-making (voting) system is correct if a consensus is reached on a correct unit output, abstain if all units abstain from voting, and wrong otherwise.

In this paper, we describe a synthetic experimental procedure to study the behavior of voting systems using a simulator that we developed to: analyze the state of each expert, apply a voting mechanism, and analyze the voting results. For this analysis, we study the following behaviors of a voting system: 1) the reliability of the voting system, "R"; 2) the probability of reaching a consensus, "Pc"; 3) certainly index, "T"; and 4) the confidence index, "C". The configuration parameters controlling the analysis are: 1) the number of participating experts, "N", 2) the possible output states of an expert, and 3) the probability distribution of each expert states. Results of this study unleash several behaviors of a decision-making system with tri-state experts as function of various configuration parameters.

Keywords: voting, expert combination, reliability analysis, decision-making, and fault-tolerance.

1 INTRODUCTION

Voting is a general technique that finds application and acceptance in many domains including software systems. Voting systems are also used in fault tolerant applications to achieve reliability [Nordmann et.al. 1999]. It is a widely used technique combining classifiers in the pattern recognition field [Ho et.al. 1994, Xu et.al. 1992, Yacoub et.al. 2002].

In distributed systems, majority voting has several applications. Voting systems are often used in distributed systems to control mutual exclusion among groups of nodes [e.g. Paris 1986] where each node is assigned a vote and weight and only a group of nodes with a majority vote can perform a restricted operation. They are used to control the update procedure when clusters of workstations are isolated due to network problems and to vote on data replication [Ahammad et.al. 1987]. They are used in replicated databases to prevent simultaneous read and write actions, or simultaneous write and write in disjoint subsets of data copies [Thomas 1979].

In a distributed decision making system, several experts are used to reach a decision. Each expert work independently on the problem and, reach a decision, and communicate the decision to the coordinator or the voting unit. An earlier study of voting in distributed applications with that nature is discussed in [Kumar et.al. 1991]. However, Kumar did not address the abstention state. The voting system studied in this paper consists of N s-independent experts (units) as illustrated in figure 1. For a given input proposition say "P", we want the system to take a decision. Each expert produces an output; the status of the unit could be correct "c", wrong "w", or abstain from producing an output "a". We call this model the "cwa" model. In practical applications, the abstention could be implemented by timing out each unit. Hence, the decision function for the j^{th} expert dj(P) is given by:

$$dj(P) = \begin{cases} 1, & \textit{iff the } j^{th} \textit{ expert output is correct} \\ 0, & \textit{iff the } j^{th} \textit{ expert output is wrong} \\ x, & \textit{iff the } j^{th} \textit{ expert abstain from producing an output.} \end{cases} \qquad \ldots \textit{Eq. 1}$$

As an application of such model consider the file replication in distributed systems. The client obtains a permission to write data to all replicated systems by sending a request to all units, each unit responds with one of three states: permission to write (correct), unable to accept request (wrong), or no reply (abstention).

Another example of the application of such system is in target detection where a target is to be identified and the identification task requires comparison of the input data to known features of the target. The decision is either target detected, cannot detect target, or cannot tell. A similar application in safety monitoring where a decision has to be made whether the situation is critical or not. Similarity in any organization, after a candidate is interviewed and the final decision is based on the number of favorable votes. In general, a system with tri-state units is very popular in decision-making applications.

The reliability of a unit "j" is characterized by three probabilities: Pj(c), Pj(w), and Pj(a), where Pj(c) + Pj(w) + Pj(a) = 1. To make a decision about the proposition acceptance, in a generalized weighted k-out-of-n system the decision D is governed by the following rule:

$$D(P) = \begin{cases} 1, & iff \quad \sum_{dj(P)\neq x} wjdj(P) \geq \lambda \sum_{dj(P)\neq x} wj \quad \&\& \quad \sum_{dj(P)\neq x} wj \neq 0 \\ 0, & iff \quad \sum_{dj(P)\neq x} wjdj(P) < \lambda \sum_{dj(P)\neq x} wj \quad \&\& \quad \sum_{dj(P)\neq x} wj \neq 0 \\ x, & iff \quad \sum_{dj(P)\neq x} wj = 0 \end{cases} \qquad \text{. . . . . } \textit{Eq. 2}$$

where w_j is a positive weight of an individual unit j and is used to express the importance of the results given by unit j; and λ is the threshold that should be exceeded to consider the output correct ($0 \leq \lambda \leq 1$).

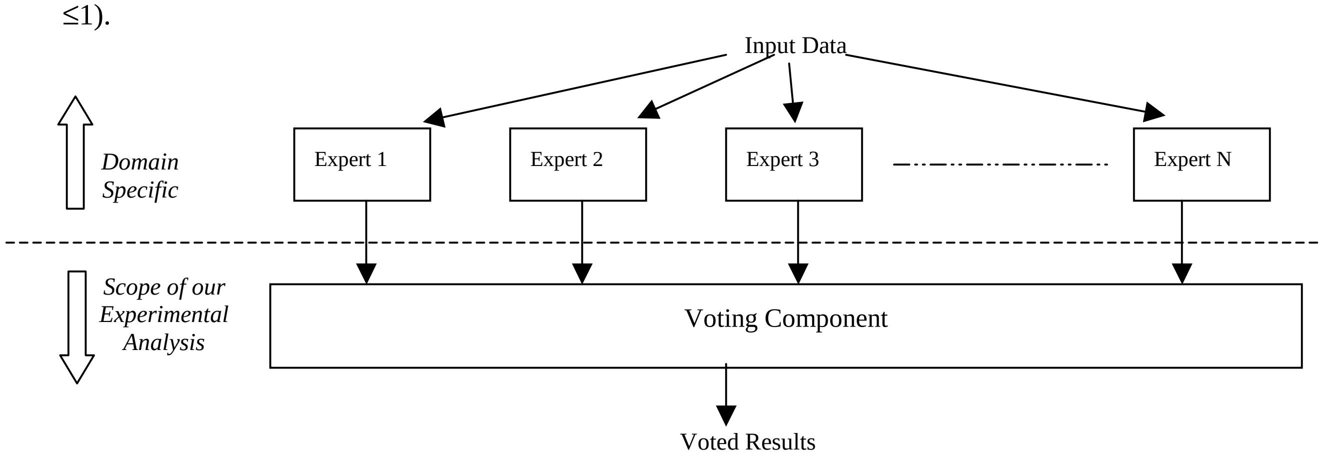


Figure 1 The scope of the analysis

Levitin et. al. [Levitin 2001 and Levitin et.al. 2001] studied techniques for selecting weights for individual experts (wj) and a consensus threshold (λ) that will maximize the reliability of the system given *predefined* number of experts and *predefined* probability distribution for the experts' output. In our study, we are more interested in the effect of the number of experts and the probability distribution functions of each expert's state on the voting system reliability. However, we run our analysis for a predefined set of weights and a threshold. All the analysis conducted in this paper is for *equally* weighted experts and for *majority* voting between non-abstained experts.

We take a synthetic experimental approach to study the behavior the voting system under such setup. We study the effect of the number of experts and probability distribution of the experts' states on: the probability of reaching a consensus; the reliability of the voting system; and confidence in the correct results. We illustrate a general synthetic approach that can be used to analyze any weighted k-out-of-n voting systems, however, we will focus in this paper on the "cwa" model discussed above.

Section 2 defines the terms that we use throughout the paper. Section 3 describes a mathematical model for the problem we are analyzing. Section 4 describes the synthetic analysis setup and procedure. Section 5 describes the results of the experiments. We organize the results under subsection where each subsection is dedicated to study the effect of one parameter on the voting reliability and behavior. Finally we conclude the paper and summarize findings in section 6.

2 DEFINITIONS AND ASSUMPTIONS

Definition 1: An Expert or A Unit

Voting is used to combine data that is produced from multiple experts. The words "experts", "engines", "classifiers", and "units" are used interchangeably. The word "classifier" is often used in the pattern recognition domain, "expert" and "unit" are used in the reliability analysis domain, and "engine" and "algorithm" is used in the algorithm combination domain. Usually several experts are used and their results are combined to produce data that is more accurate (or with higher confidence) than data produced from one individual expert.

Definition 2: Number of Units, N

We define "N" as the number of experts (or units) used in the voting system. We will study the behavior of the voting component as a function of N.

Definition 3: An Expert's Output

The following table illustrates a simple explanation and probability of the output of each expert. We refer to the decision of an expert j as dj(P) where P is the input proposition.

dj(P)	1	0	x
Expert State	Correct	Wrong	Abstained
Probability	Pj(c)	Pj(w)	Pj(x)

Table 1 Output from an expert

Hence, Pj(c) reflects an absolute measure for the reliability of unit j assuming that an abstained condition is considered an unreliable condition.

Definition 4: Majority Voting

Majority voting is a voting schema that has the following consensus rule:

$$\begin{cases} \textit{Consensus} & \textit{iff } \lfloor Y/2 \rfloor + 1 \text{ experts agree} \\ \textit{No Consensus} & \text{otherwise.} \end{cases} \quad \ldots\ldots\ldots\ldots \textit{Eq. 3}$$

where Y is the number of experts that did not abstain from producing an output. We use "N" for the total number of experts, "A" for the number of abstained experts, and hence Y=N-A.

Definition 5: Reliability of the Voting System "R"

We define "R" as the probability that the output from the voting system is correct. According to equation 2, *R=P(D(P) = 1)*. R is calculated using the probability of reaching consensus and the consensus is the correct output; i.e. equals 1.

Definition 6: Probability that a Consensus is reached "Pc"

We define "P_C" as the probability of reaching a consensus. We note here that the probability of reaching a consensus is not the same as the probability that the output is correct. For example, assume that we are using a majority voting of three experts A, B, C. Results from experts A and B are considered "w" (for wrong) and result from expert C is considered "c" (for correct). The output from the majority voting in this case is the "w". Hence, the voting system reached a consensus though this consensus is not correct. Hence, Pc and R are different random variables. Therefore, we study the behavior of the voting system in terms of both "Pc" and "R"

Definition 7: Certainty Index, T

We define "T" as the certainty index, which is certainty in the consensus result of the experts. T is a function of the following factors:

- Probability that a specific consensus case is reached. A consensus case is defined as a case where we have the set of votes sufficient to make a consensus. For instance, with three experts A, B, C and a majority voting, the case where A and B agree and differ from C is a consensus case.
- Number of experts participating in making a specific consensus. An expert is said to participate in making a consensus if the output of the expert is the same as the consensus results.

For example, assume that we have three experts, A, B, and C. Two experts produce the same result and the third produced a different result. The result of the majority voting for this case will be the result agreed upon by the two experts. Then the number of experts participating in reaching the consensus is 2 out of 3. Probability of that case can be calculated by expanding the decision tree as shown in the following section using the probability distributions of all experts, assume 6/9. Now assume that the three experts produce the same result then we have three out of three experts producing the consensus. The probability that the case where the three experts agree can be calculated from the decision tree, assume 1/9. The certainty index T3 for these considering two cases is : (6/9 * 2/3 + 1/9 * 3/3)/(1/9+6/9). The general formula for T is:

$$T = \sum_{s \in \{\text{Consensus Cases}\}} \text{P(s/c) x} \frac{\text{NE}_s}{\text{N}} = \sum_{s \in \{\text{Consensus Cases}\}} \frac{\text{P(s, c)}}{\text{Pc}} \text{ x } \frac{\text{NE}_s}{\text{N}} \quad \ldots\ldots\ldots\ldots \; Eq.4$$

where P(s/c) is the probability of the case “s” given that consensus is reached, “NEs” is the number of experts making a consensus in the case “s”, and Pc is the consensus probability.

Note that the certainty index that we calculate means how certain we are about the consensus reached but it does not say whether we are certain it is “correct” or “wrong”. If T is calculated for those cases only where a consensus is reached on the correct state (and Pc is replaced with R in the previous formula), then T measures the certainty in the correct result as opposed to certainty in the consensus. We will use the two variables T(Pc) and T(R) to distinguish the two cases. As illustrated in equation 4, T is considered a measure for weighted consensus.

Definition 8: Confidence index, C

The confidence index “C” represents how confidence we are in the results produced by the voting system. It takes into consideration an important factor, which is how many experts are participating in reaching a consensus. In this case it is similar to the certainty index but it is NOT normalized to the total number of experts. The formula for calculating C is:

$$C = \sum_{s \in \{\text{Consensus Cases}\}} \text{P(s/c) x NE}_s = N \times T \quad \ldots\ldots\ldots\ldots \; Eq.\ 5$$

Since this factor is not normalized to the number of experts, then it will be bigger than unity. The number is significant in comparing confidence produced using several N-experts systems. Note that C is calculated for all consensus case, which does not distinguish between the correct, wrong, or abstention consensus. To calculate the confidence in the correct results we will consider the consensus cases in which the consensus results are correct. We will use the two variables C(Pc) and C(R) to distinguish confidence in the voting system consensus and confidence in the voting system reliability respectively.

3 MATHEMATICAL MODEL

To mathematically model the analysis process, we will use the universal moment generating function technique (a universal Z-transform) that is defined in [Ushakov 1986] and used in [Levitin 2001] to model system reliability. The universal generating function is defined as :

$$U(z) = \sum_{k=1}^{K} S_k z^{x_k} \qquad \text{. . . . . } Eq.6$$

for a random variable X that can take K-possible values. S_k is used to model the probability that X is equal to some value x_k. To calculate the probability that some condition on X is met, such that $X \in \theta$ all coefficients of the polynomial U(z) is summed for every term that satisfies the condition $x_k \in \theta$. Hence,

$$\Pr\{X \in \theta\} = \delta(U(z), \theta) = \sum_{x_k \in \theta} S_k \qquad \text{. . . . . } Eq.7$$

We will use the polynomial $U_j(z)$ to describe the output state of the j^{th} expert. Each possible output state from the expert unit is characterized by two values: S_{jk} which define the probability of the output state whose index is k and the total score for the result of that expert G_{jk}. Recall that each unit has three possible output states: "c", "w", or "a". Hence S_{jk} and G_{jk} are given as follows:

Index	1	2	3
dj(P)	1	0	x
State	Correct	Wrong	Abstain
Probability (S_{jk})	$S_{j1} = P_j(c)$	$S_{j2} = P_j(w)$	$S_{j3} = P_j(x)$
Score (G_{jk})	$G_{j1} = 1$	$G_{j1} = 0$	$G_{j1} = ½$

Table 2 Parameters for the Uj(z)

To understand why the abstained value is considered by ½, let us assume that we have N experts. If a number of A experts abstain from voting, then we have a number of (N-A) experts that are considered for voting. Assume we have H experts producing the correct result, then the condition of producing a correct result is H > ½(N-A), which can be refined to H + ½ A > N/2. As a result if an expert is correct it contributes to H by a single share or weight (equals 1 for our studies) while if the expert abstains then it contributes with ½ of its weight. As a result we can calculate the quantity H + ½ A and compare it to N/2.

Now assume we have two experts. Each expert will have its universal function $U_1(z)$ and $U_2(z)$. To obtain the universal function of the system we use the composition operator:

$$\begin{aligned} U_{\{1,2\}}(z) &= \Omega(U_1(z), U_2(z)) \\ &= \Omega(\sum_{j=1}^{3} S_{1j} z^{G_{1j}}, \sum_{r=1}^{3} S_{2r} z^{G_{2r}}) \\ &= \sum_{j=1}^{3}\sum_{r=1}^{3} S_{1j} S_{2r} z^{G_{1j}+G_{2r}} \end{aligned} \qquad \text{. . . . . } Eq.8$$

The general rule for composing a number of experts {j} to an exiting universal function $U_\lambda(z)$ is given by the rule :

$$U_{\lambda \cup \{j\}}(z) = \Omega(U_\lambda(z), U_{\{j\}}(z)) \qquad \text{. . . . . } Eq.9$$

Hence for a system composed of N experts, the universal function would be:

$$U_N(z) = \prod_{r=1}^{N} \sum_{k=1}^{3} S_{rk} z^{G_{rk}} = \sum_{k_1=1}^{3}\sum_{k_2=1}^{3} \ldots\ldots \sum_{k_N=1}^{3} S_{1k_1} S_{2k_2} \ldots\ldots S_{Nk_N} z^{G_{1k_1}+G_{2k_2}\ldots\ldots G_{Nk_N}}$$

This polynomial represents all possible combination of the outcome from each expert. Now if we want to obtain the reliability of the system R, we use the equation:

$$R = \delta(U_N(z), G > N/2) = \sum_{G_k > N/2} S_k \qquad \text{. . . . . } Eq.10$$

which is the commonly used simple majority vote rule.

Although the mathematical model given in this section could be very effective for numerical calculations, we find that it does not provide enough flexibility to study variations in the number of experts used (since this requires change in the polynomial representation). Therefore, we use the mathematical model for modeling the problem and the consensus rule and we use a simulator that we discuss in the next section to facilitate the calculation procedures for voting reliability, consensus, and confidence.

4 SYNTHETIC EXPERIMENT SETUP

We have built a simulation system for the purpose of conducting various experiments to study the behavior and reliability of various voting systems as a function of several parameters. In this section we describe the inputs, outputs, and the operation of the simulator.

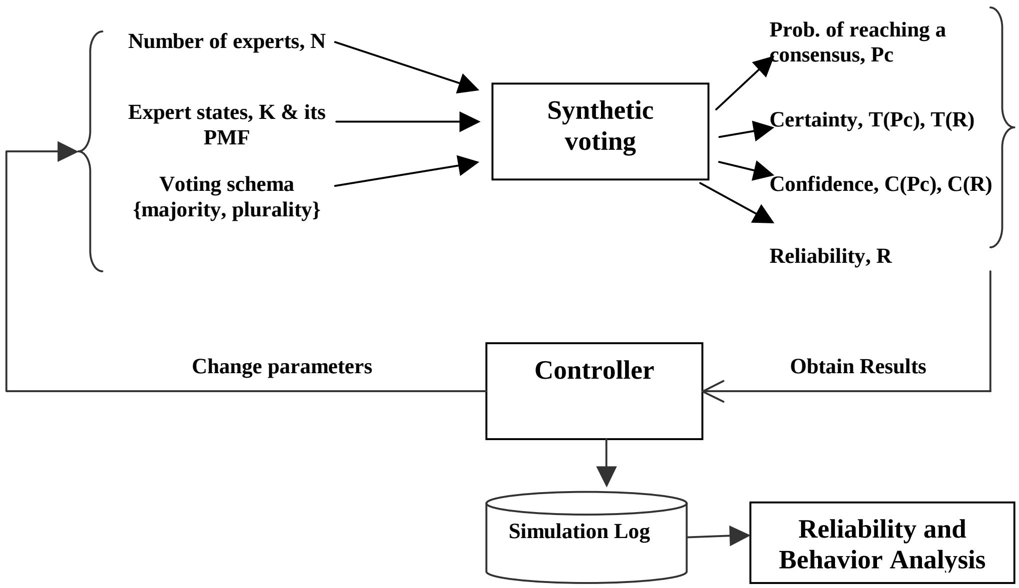


Figure 2 The simulation system

The simulator accepts the following inputs:

a) the number of experts participating in the experiment (previously defined as N),
b) the number of possible states of an expert K, and the probability distribution for each state.
c) the type of voting system (currently, the simulator supports only two types: majority and plurality).

For the type of voting system we discuss here, K=3 since each expert can be correct, wrong, or abstain. The probability distribution is hence reduced to Pj(c), Pj(w), and Pj(a). To calculate the system reliability, we exclude those units that abstained and hence we compare the correct and wrong unit results, hence majority and plurality voting are reduced to the same thing for equal weight units.

The output of a simulation is:

a) the reliability of the voting system, R
b) the probability of reaching a consensus, Pc
c) the certainty index, T(Pc), T(R)
d) the confidence index, C(Pc), C(R)

The following is the procedure followed by the simulator to calculate Pc, R, T, and C.

a) Build a Decision Tree

First, the simulator uses the number of experts N, the number of each expert states K, the set of PMFi for each expert states, to build a decision tree. The tree explores all possible combination of the experts' output.

The number of experts defines the depth of the tree. Each node at one level will be expanded to K nodes in the lower level. Going from one level to another is the same as exploring all possible output states from an expert. At the lowest level of the tree, we have all possible combinations that can be produced by N experts for K possible states. The nodes at the lowest level are called *terminal nodes*. The arcs in the decision tree are labeled with the probability that the expert is in specific state. This is usually specified in the PMF distribution for each expert.

As an example, consider the case where we have three experts (N=3) and two possible output states (K=2). We can use two symbols to model this case; assume they are "q" and "y". The PMF distribution is uniform for the first and third experts and one state "q" is favorable for the second expert than "y", PMF2(q)=2/3, and PMF2(y)=1/3. We have three experts; hence the depth of the tree is three. The following figure illustrates the expansion of the tree.

Note: Expert Order is Irrelevant

The order of the experts in the tree is not relevant. This is because: a) there is no explicit dependency between experts, b) the decision made by one expert does not affect the decision made by the other (experts are not collaborative), and c) the order of states (a unit output) at the terminal nodes is not the one affecting the decision, instead it is the count of the states that matters.

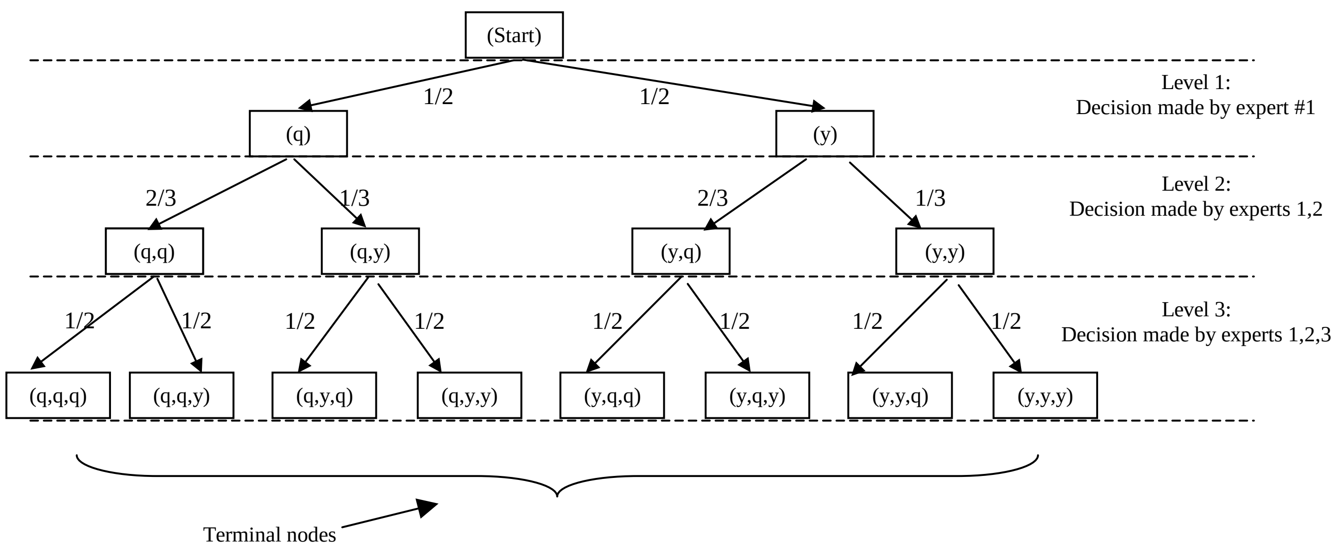


Figure 3 Decision tree for three experts with two possible decisions (q or y)

Note: Number of Terminal Nodes

The number of terminal nodes obtained from the decision tree is function of the number of experts N and the number of possible states of an expert, K. The total number of terminal nodes will be given by:

$$\text{NumberOfTerminalNodes} = M = K^N \quad \text{. . . . . } \textit{Eq. 11}$$

b) Assigning probability to various combinations.

For each terminal node, we calculate the probability of reaching that node by back propagating the tree and accumulating the probability of each branch. Since independence between experts is assumed, the probability of reaching a terminal node is the *product* of the probability of all branches traversed to reach that node. Hence:

$$PT(m) = \prod_{n=1}^{N} PMF_n(S_m(n)) \quad \text{. . . . . } \textit{Eq. 12}$$

where $S_m(n)$ is the n[th] state in the m[th] terminal node result. For instance if result for the first terminal node is (q,q,y), then $S_1(1)=S_1(2)=$"q" and $S_1(3)=$"y".

For the above decision tree, the probability of the terminal node (q,q,y) is: ½* 2/3 * ½

Special case: Uniform Distribution

In a simple situation, all states might be equally probable for all experts. Hence, the probability that the i[th] expert is at state kj is given by:

$$PMF_i(K_j / UniformDistribution) = \frac{1}{K} \quad \text{. . . . . } \textit{Eq. 13}$$

This means that all the terminal nodes are equally probable and hence each decision has a probability of 1/8 for the case shown in figure 3. Note that the number of terminal nodes according to equation 11 is : K^N , hence the probability of the m[th] terminal node PT(m) will be 1/ K^N:

$$PT(m) = \frac{1}{K^N} \quad \text{. . . . . } \textit{Eq. 14}$$

c) Evaluating Probability of reaching a consensus (Pc).

Each one of the terminal nodes is checked against the consensus rule for the required voting schema. Consensus rules for majority voting are given in equation 3. The probability of reaching a consensus is then calculated as follows:

$$Pc = \sum_{m=1}^{M} \mathrm{PT(m)} \times \delta\mathrm{(m)} \qquad \text{. . . . . } Eq.15$$

where PT(m) is calculated using equation 6 and λ(m) is given by:

$$\delta\mathrm{(m)} = \begin{cases} 1 \ \ \textit{iff} \text{ m is a consensus terminal node} \\ 0 \ \ \text{otherwise} \end{cases}$$

Special case: Uniform Distribution

In a simple situation when all states might be equally probable, we evaluate the number of nodes reaching a consensus and the probability of reaching a consensus is simplified to:

$$Pc = \frac{\text{Number of terminal nodes reaching a consensus}}{\text{Total number of terminal nodes}} = \frac{N_C}{K^N} \qquad \text{. . .. . . . } Eq.\ 16$$

d) Evaluating the certainty index, T

To calculate the certainty index, we only consider the terminal nodes labeled in procedure (c) above as reached a consensus. For each terminal node we calculate the percentage of experts participating in the consensus relative to the total number of experts. As an example, consider the terminal node (q,q,y). The consensus result in this case is “q”. Hence the certainty in this particular consensus decision is 2/3. Certainty index is then calculated by summing all consensus cases and weighing it with the probability that the consensus case is reached.

$$T = \sum_{m=1}^{M} \frac{NEm}{N} \times \frac{PT(m)}{Pc} \times \delta(m) \qquad \text{. . .. . } Eq.\ 17$$

where: NEm is the number of experts making a consensus for the mth terminal node, PT(m) is calculated from equation 12 and δ(m) is given from equation 15.

Special case: Uniform Distribution

In a simple situation when all states might be equally probable, we evaluate the number of nodes reaching a consensus and the certainty index is simplified to:

$$T = \sum_{m=1}^{M} \frac{NE_m}{N} \times \frac{1}{K^N} \times \delta(m) \times \frac{1}{Pc} \quad = \quad \frac{1}{N \times Nc} \times \sum_{m=1}^{M} NE_m \times \delta(m) \qquad \text{. . . . . } Eq.18$$

Note equations 17 and 18 are calculated for the all consensus cases, hence the values obtained there are T(Pc). To calculate T(R) we consider a subset of the consensus cases in which the consensus is reached on a correct state in this case “1”.

e) Evaluating the confidence index C.

To calculate the confidence index, we only consider the terminal nodes labeled in procedure (c) above as reached a consensus. For each consensus case, we count the number of experts participating in the consensus. As an example, consider the terminal node (q,q,y). The consensus result in this case is "q". Hence the confidence in this particular consensus decision is the absolute number 2. The confidence index is then calculated by summing all consensus cases and weighing it with the probability that the consensus case is reached.

$$C = \sum_{m=1}^{M} NEm \times \frac{PT(m)}{Pc} \times \delta(m) \qquad \text{. . . . . Eq.19}$$

Special case: Uniform Distribution

In a simple situation when all states are equally probable, we evaluate the number of nodes reaching a consensus and the confidence index is simplified to:

$$C = \sum_{m=1}^{M} NE_m \times \frac{1}{K^N} \times \frac{K^N}{Nc} \times \delta(m) \quad = \quad \frac{1}{Nc} \times \sum_{m=1}^{M} NE_m \times \delta(m) \qquad \text{. . . . . Eq. 20}$$

From equations 12 and 14, then

$$C = N \times T \qquad \text{. . .. . Eq. 21}$$

Hence C and T are tightly related. The certainty index T measures the confidence we have in the consensus given that we know the best result will be produced by all experts agreeing on the output. The confidence index C is an absolute measure for confidence used to compare confidence in the results produced by multiple N-expert systems.

Note equations 19 and 21 are calculated for the all consensus cases, hence the values obtained there are C(Pc). To calculate C(R) we consider a subset of the consensus cases in which the consensus is reached on a correct state in this case "1".

f) Evaluating Reliability, R

Each one of the terminal nodes is checked against the consensus rule for the required voting schema. Consensus rule for majority voting is given in equation 3. For those nodes that reached a consensus, we determine the consensus state and compare with the correct state. R is then calculated using equation 15 for those cases where the consensus state is the correct one.

5 ANALYSIS RESULTS

5.1 Probability of reaching a Consensus (Pc) and Reliability (R) as function of N

5.1.1 Purpose

The purpose of this analysis is to study the effect of changing the number of experts “N” on the probability of reaching a consensus “Pc” and the reliability of the voting system “R”. For a given number of experts and their PMF (P(c), P(w), and P(a)), the simulator expands the decision tree and assesses the consensus rule at the terminal nodes. The probability of reaching consensus and reliability is then calculated as discussed earlier. Results are shown in the following figure.

5.1.2 Results

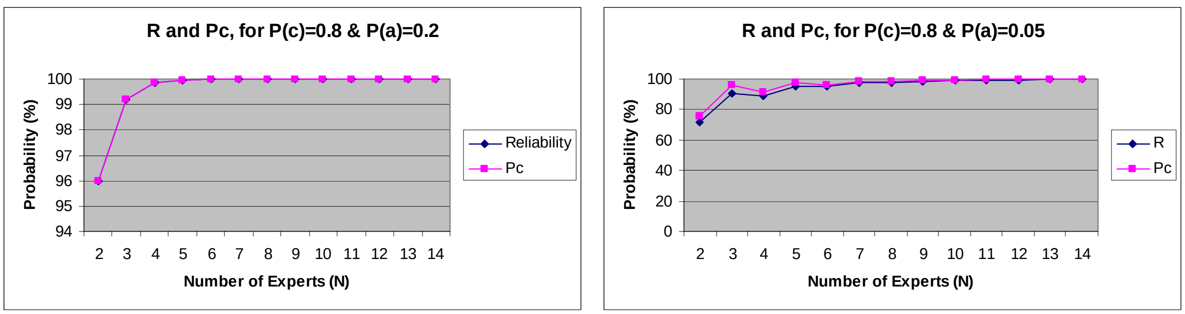


Figure 4 Reliability and Pc, for P(c)=0.8 and P(a)=0.2,0.05 as function of number of experts

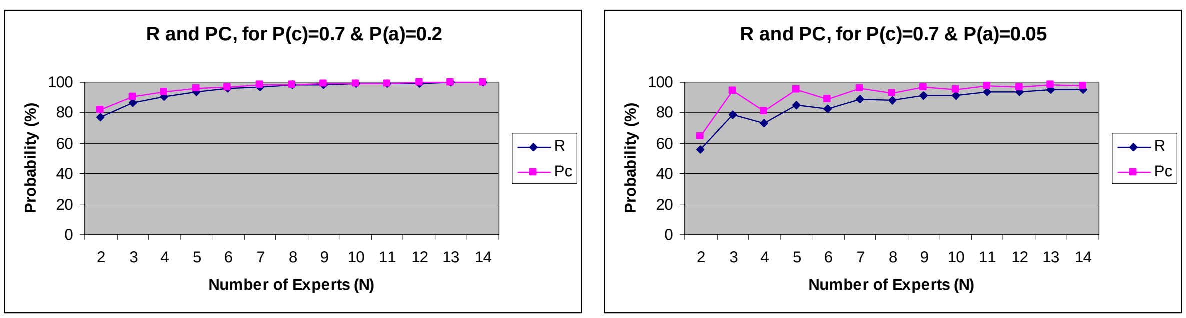


Figure 5 Reliability and Pc, for P(c)=0.7 and P(a)=0.2,0.05 as function of number of experts

From figures 4 and 5, we find that:

- As the number of experts increases, the reliability of the system improves for a given probability of an expert being correct, e.g. for the first graph in figure 4 where P(c)=0.8. It is also notable that the reliability approaches unity for a large number of experts. However, by using few experts (for instance N = 5) the reliability is considerably high and there is no need to develop or acquire more experts since the improvements in the reliability is not notable.

- For a given P(c), as the probability of abstention decrease, the probability of an experts being wrong increases (a direct result from the probability rule: P(c)+P(w)+P(a)=1). This produces the gap between the Pc and R in the previous figures. The gap between the two curves represents the situation where a consensus is reached but the consensus is wrong. For instance, in the first graph of figure 4, P(c)=0.8 and P(a)=0.2 and hence P(w)=0, there is no gap; i.e. when a consensus is reached it is correct. For the first graph in figure 5, P(c)=0.7 and P(a)=0.2 and hence P(w)=0.1, we should expect a gap between Pc and R. Note that the gap is not the unreliability of a specific unit; it is the unreliability of the voting system.

The above results are obtained for cases where the experts are reliable; i.e. P(c) is near unity, 0.9, 0.8, etc. For the case where the experts are unreliable, the following figures illustrate some results.

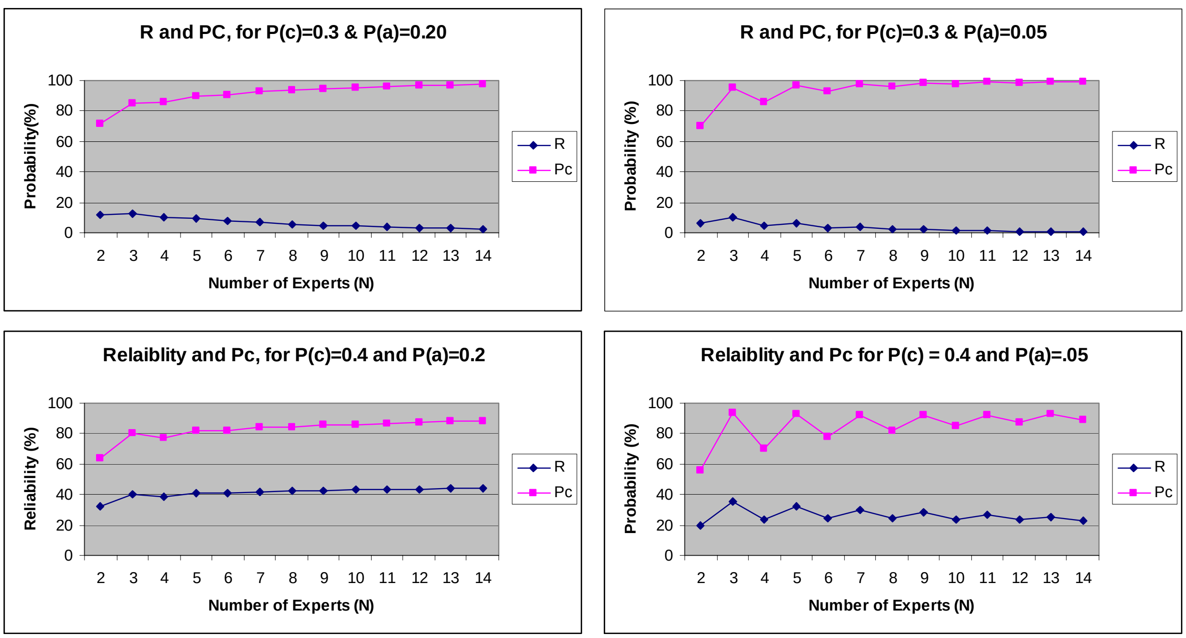


Figure 6 Reliability and Pc, for P(c)=0.3,0.4 and P(a)=0.2,0.05 as function of number of experts (N)

From figure 6, we find that:

- The gap between Pc and R is larger than it was for reliable units. This is obvious since Pc converges to 100% consensus but in this case it is consensus on the wrong result.

- For the cases where the individual experts are not reliable there is no benefit of adding new experts since the increase in the number of expert decreases the reliability of the system. For example, in the last graph of figure 6, the maximum reliability is achieved with a system of three experts only.

5.2 *Reliability of a Voting System as function of P(c) and N*

5.2.1 Purpose

The purpose of this analysis is to study the reliability of the voting system as function of the reliability of each unit and the number of experts used. Results are shown in the following figures.

5.2.2 Results

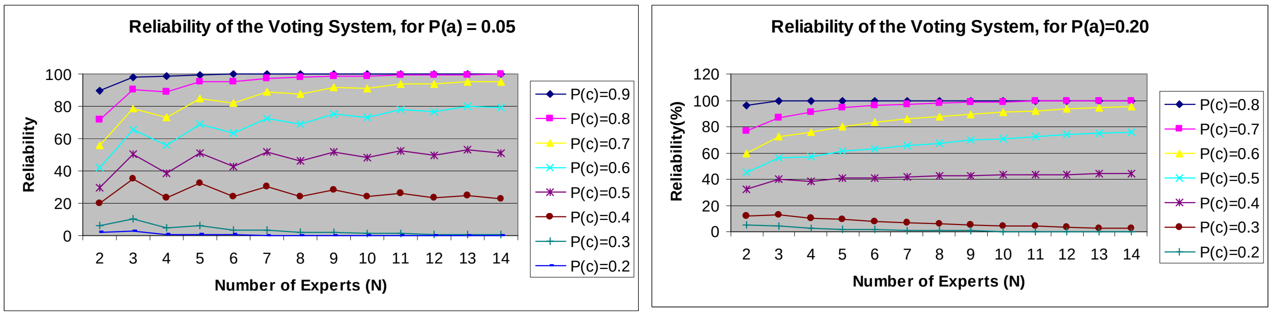


Figure 7 Reliability for P(a)=0.05,0.20 for various P(c) and number of experts N

Figure 7 illustrates the reliability of the voting system as a function of the reliability of its individual experts (P(c)). As the reliability of individual experts increases, additional experts would increase the reliability of the overall system. However, as the reliability of individual experts decreases, additional experts would degrade the reliability of the system. The condition [P(c) + P(a) > 0.5] can be used to determine whether increments in number of experts would be useful.

Another view of the reliability of the system as function of reliability of individual experts and for various numbers of experts is illustrated in the following figure.

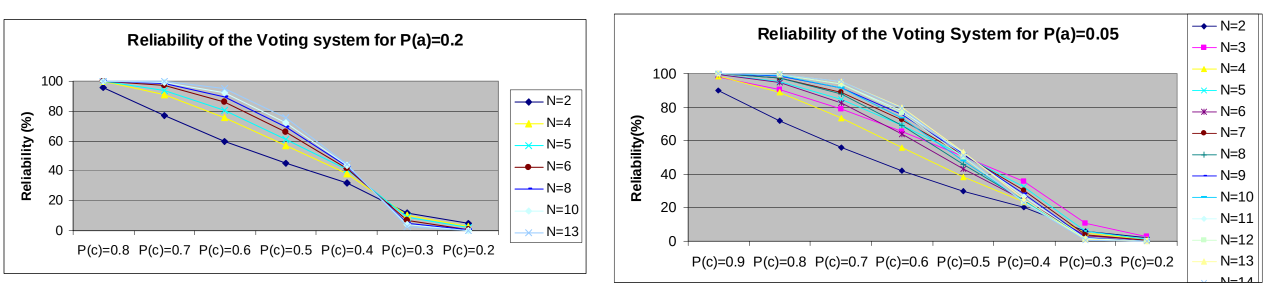


Figure 8 Reliability for P(a)=0.05,0.20 for various P(c) and number of experts N

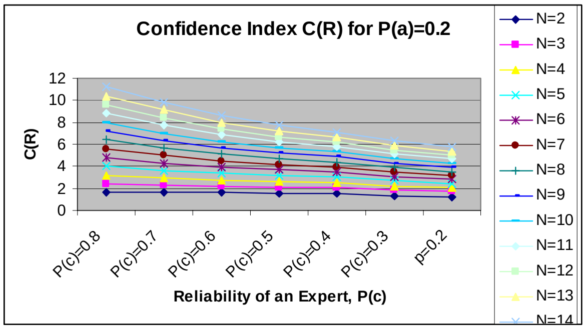
The same results regarding the condition P(c) + P(a) > 0.5 can be deduced from figure 8. For instance, in the first graph of figure 8, P(a) = 0.2, then for P(c)>0.3 the increase in the number of experts improves the reliability of the system and for P(c)<0.3 the more experts we use, the more the system becomes unreliable.

5.3 *Confidence Index, C(R) as function of Reliability of Individual Experts*

5.3.1 Purpose

The purpose of this analysis is to study the confidence index C of the voting system as a function of the number of experts as well as the reliability of individual experts. We study the confidence index of the correct output C(R).

5.3.2 Results



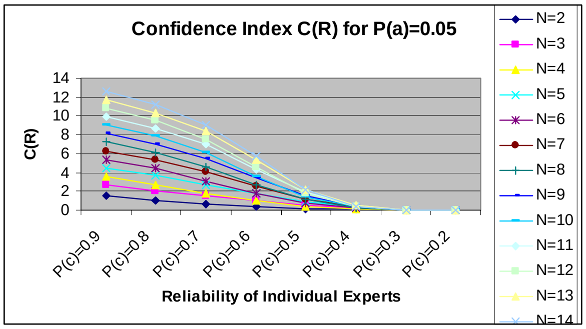


Figure 9 Confidence of the Voting system as function of P(c) and number of experts N

From figure 9, we note that:

- for a given expert reliability P(c), the confidence index increases as we increase the number of experts used in the system.
- the confidence in the correct output decreases as the reliability of individual experts decreases.

5.4 *Comparing Confidence (and Certainty) for Reliability and Consensus results.*

5.4.1 Purpose

The purpose of this analysis is to study the confidence index C and the certainty index T of the voting system as a function of the number of experts as well as the reliability of individual experts. We study the confidence/certainty index of the correct output C(R)/T(R) and confidence/certainty index of the consensus results C(Pc)/T(Pc).

5.4.2 Results

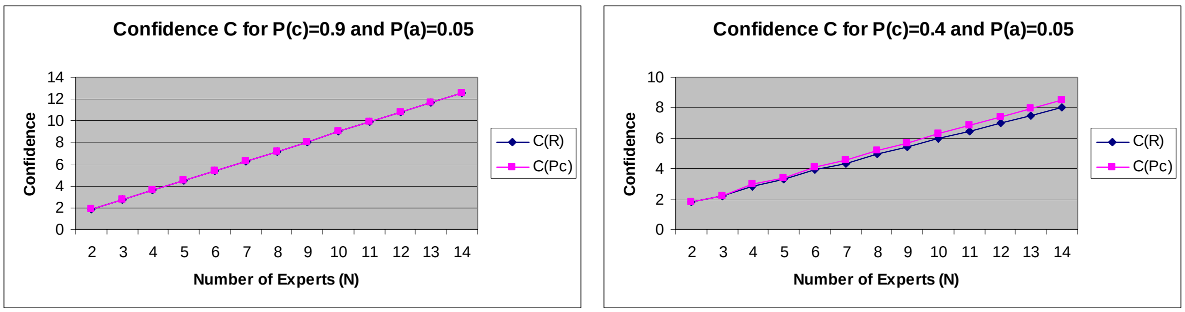


Figure 10 Confidence index C(R) and C(Pc) as function of the number of experts N

Figure 10 illustrates results from calculating the confidence in the reliability results C(R) as well as the consensus results C(Pc) for P(c)=0.9 and P(c)=0.4. We find that:

- The confidence increases by increasing the number of experts.
- When the experts used are reliable enough the C(R) and C(Pc) would coincide as illustrated in the first graph of figure 10. However, for the case where the experts are unreliable there is a gap between confidence in the consensus and confidence in correct output. This has been explained earlier in section 5.1.

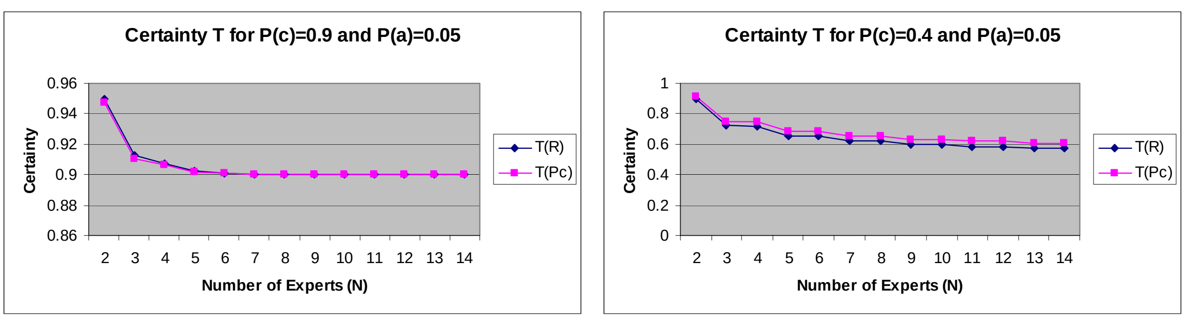


Figure 11 Certainty index T(R) and T(Pc) as function of the number of experts N

Figure 11 illustrates results from calculating the certainty in the reliability results T(R) as well as the certainty in the consensus results T(Pc) for P(c)=0.9 and P(c)=0.4. We find that:

- As the number of experts increases, the certainty index stabilizes to a constant value.
- We get the same "gap" effect in certainty similar to the gap effect in the confidence index discussed in figure 10.

5.5 *Comparing Certainty versus Confidence*

5.5.1 Purpose

The purpose of this experiment is to study and compare confidence and certainty. We compare the confidence index of the correct output C(R) and confidence index of the consensus results C(Pc) as well as the certainty index in the correct output T(R) and the certainty index in the consensus T(Pc).

5.5.2 Results

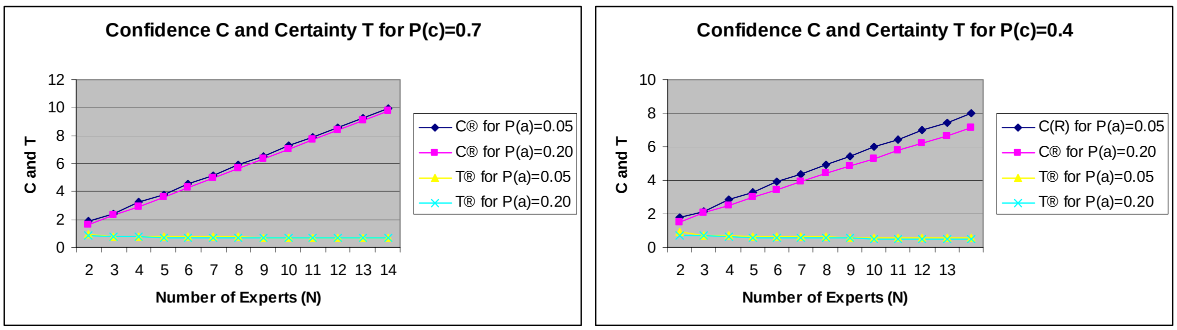


Figure 12 Comparing Confidences and Certainty

From figure 12, we find that:

- the confidence index C is monotonically increasing as we use more experts
- the certainty index converges to a constant as we increase the number of experts, which is also illustrated in figure 11.
- the confidence index gets the same "gap" effect between confidence in consensus C(Pc) and confidence in reliability C(R).

As a result of this analysis, we realize why we need to indices T and C. Certainty index is referred to as a weighted consensus; it is always less than unity. Confidence index is used to compare several n-experts by relatively measuring confidence index of one g-experts system with another h-experts system. As an example, consider the case where we have 3-experts system which produce the output (c,c,w); T in this case is 2/3=67% and C is 2. Now assume another 5-experts system with the result (c,c,c,w,w); T in this case is 3/5=60% and C=5. Without knowing the internal structure of the two systems, the second system produces a result confidence higher than the first (5 compared to 2); hence the confidence index is useful. On the other hand, if we are interested in taking into consideration the ratio between correct and wrong experts, the first system is higher than the second (67% compared to 60%); hence the certainty index is useful.

5.6 Reliability of a Voting System as function of Abstention Probabilities

5.6.1 Purpose

The purpose of this analysis is to study the reliability of the voting system as a function of the abstention probabilities of the experts.

5.6.2 Results

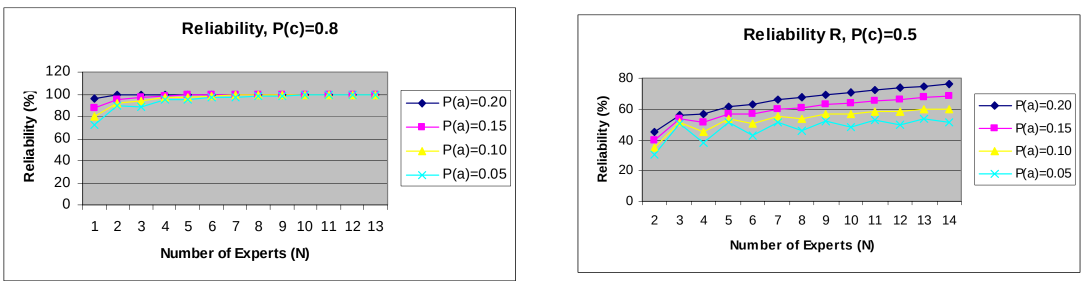


Figure 13 Reliability of the voting system as function of the number of experts for various P(a).

From figure 13, we find that:

- For a given P(c), as the P(a) decreases, the reliability of the system decreases. This is logically acceptable since P(w) increases with P(a) decreases; i.e. the units are more unreliable. This leads to the well-known conclusion that we would better have an expert abstaining from producing an output rather than producing the wrong output.

- For a given P(c), as the units become more unreliable (on the expense of being absent), the reliability of the system becomes more sensitive to the even/odd number of experts. For instance, in the second graph of figure 10, for P(a)=0.05, the oscillation in the reliability is more obvious. Therefore, in cases where the units are unreliable, we better use an odd number of experts to obtain higher system reliability than even number of experts.

6 CONCLUSION

In this paper, we analyzed the reliability and behavior of a majority voting system. We have taken a synthetic experimental approach in which a simulator is used to expand all possible decisions using a decision tree. The decision tree is function of the number of experts, the states of the output of each expert, and the probability distribution for these states. We have studied a voting system comprising tri-state experts where the expert state could be "correct", "wrong", or "abstain". We conducted a series of studies to analyze the behavior of the voting system in terms of reliability, probability of reaching a consensus, confidence and certainty in its consensus, and confidence and certainty in its correctness. The results of these studies are discussed in details together with their logical interpretation. We study the effect of the number of experts and the reliability and abstention of each expert on the behavior and reliability of the voting system.

The results that we obtain in this study reveal the behavior and reliability of a majority voting schemas and enable us to make better choices in terms of the number of experts given some domain requirements such as "accuracy" or "consensus".

There are still several problems that can be analyzed using the same approach that we used in this study. For example, in the next phase we will study the behavior and reliability of the more generalized weighted *k-out-of-n* voting system.

A long-term objective of this research is to create reliability lookup tables for voting systems. Ultimately by studying different probability distributions of the output of the experts and running synthetic experiments like the ones we developed in this paper, we can construct lookup tables that the analyst uses to make decisions about the number of experts to use in implementing a voting system. Alternatively, a tool could be developed to study what-if scenarios by submitting parameters about the experts and obtaining results in terms of the reliability, confidence, certainty, and probability of reaching a consensus.

REFERENCES

[Ahammad et.al. 1987] M. Ahammad and M. Ammar, "Performance Characterization of Quorum-Consensus Algorithms for Replicated Data," *Proceedings of the 7th Symposium on Reliability in Distributed Software and Database Systems*, 1987, pp161-167.

[Ho et.al. 1994] T.K. Ho, J. J. Hull, and S. N. Srihari, "Decision Combination in Multiple Classifier Systems", IEEE Trans. on PAMI, 16(1), pp.66-75, 1994.

[Kumar et.al. 1991] Akhil Kumar and Kavindra Malik, "Voting Mechanisms if Distributed Systems", *IEEE Transactions on Reliability*, Vol 40, No. 5, 1991, pp593-600

[Lam et.al. 1994] L. Lam and C.Y. Suen, “A Theoretical Analysis of the Application of Majority Voting to Pattern Recognition”, *Proceedings of IEEE*, pp. 418-420, 1994.

[Lam et.al. 1997] L. Lam and C. Suen, “Application of Majority Voting to Pattern Recognition: An Analysis of Its Behavior and Performance,” IEEE Transaction on Systems, Man., and Cybernetics, Vol 27, No 5, September 1997.

[Levitin 2001] G. Levitin “Analysis and Optimization of Weighted Voting Systems consisting of Voting Units with Limited Availability”, Reliability Engineering and System Safety, 73(2001), pp 91-100.

[Levitin et.al. 2001] G. Levitin and A. Lisnianski, “Reliability Optimization for Weighted Voting System”, Reliability Engineering and System Safety, 71(2001), pp 131-138.

[Lin et.al. 2002]. X. Lin, S. Yacoub, and J. Burns, “Performance Analysis of Classifier Combination by Plurality Voting and their application in Pattern Recognition” Hewlett-Packard Laboratories, Technical Report, March 2002.

[Lorczak et.al. 1989] P. Lorczak, A. Caglayan, and D. Eckhardt, “A Theoretical Investigation of Generalized Voters for Redundant Systems”, IEEE Computer Soc. Press, 21-23 June 1989, pp 444-451

[Nordmann et.al. 1999] L. Nordmann and H. Pham, “Weighted Voting Systems,” *IEEE Transactions on Reliability*, Vol 48, No 1, 1999.

[Parhami 1994] B. Parhami, “Threshold voting is fundamentally simpler than plurality voting” *International Journal on Reliability, Quality, and Safety Engineering*, Vol 1(1), 1994, pp95-102.

[Parhami 1994b] B. Parhami, “Voting Algorithms” *IEEE Transaction on Relaibility* , Vol 43, 1994, pp617-629.

[Paris 1986] J. Paris, “Voting with WiTesses: A Consistency Scheme for Replicated Files,” *Proceedings of the 6th International Conference on Distributed Computing Systems*, May 1986, pp606-616.

[Thomas 1979] R. Thomas, “A Majority Consensus Approach to Concurrency Control” *ACM Transactions on Database System*, June 1979, pp 180-209.

[Ushakov 1986] Ushakov IA, “Universal Generating Function” Soviet Journal of Computing System Science, 24(5), 1986, pp118-129.

[Xu et.al. 1992] L. Xu, A. Krzyzak, and C. Suen, “Method of Combining Multiple Classifiers and Their Application to Handwriting Recognition”, *IEEE Transactions on Syst. Man Cybernet*, Vol. 22(5), pp. 418-435, 1992.

[Yacoub et.al. 2002] S. Yacoub and S. Simske, “Meta-Algorithmic Patterns”, Hewlett-Packard Laboratories, Technical Report, April 2002.

[Yacoub et.al. 2002b] S. Yacoub, X. Lin, and J. Burns, “Analysis of the Reliability and Behavior of Majority and Plurality Voting Systems”, submitted to the 13th *International Symposium on Software Reliability Engineering*, ISSRE2002, Annapolis MD, Nov. 2002.