Project Description

Part-I: An Overview

1 Introduction

1.1 Problem statement: *Quantification of Expert Opinion*

Although disaster control is executed by human agents, their expert opinions are seldom integrated as quantitative variables in industrial application of infrastructure management, ref. [5, 23, 34]. The primary obstacle originates in the text-based presentation of subjective opinion. A fuzzy logic engine, ref. [47, 60], circumvents this difficulty by converting qualitative policy statements into probability-like measures[1] with subjectivity ranges. Subsequently, a Bayesian updating procedure, ref. [8], ensures that the complex system, ref. [41], will 'learn from mistakes,' ref. [33].

1.2 The end product — A Computer Program: development/validation

Program segments,[2] to be delivered as computer algebra and *object-oriented*-procedure modules, are:
(a) Equation and eigenvalue solvers with `Interval Arithmetic`[3] indicating subjectivity bounds;
(b) Fuzzy logic quantifiers for qualitative information and fuzzy logic arithmetic routines;
(c) Bayesian integration modules (for a large number of variables the algebraic form is too slow);
(d) Extreme value statistical computation package.
These components will yield the final prototype: the intelligent engine, *vide* §2, page 2.

The computer program architecture, Figure 2.1, page 2, will be validated for:
(i) managing the progress of infrastructure projects and aiding international donors (*e.g.* §9, page 11);
(ii) planning rescue and evacuation after a manmade or natural disaster (*e.g.* §7, page 9).

1.3 Novel and distinctive features of this proposal

In safety engineering, conventional reliability analyses are *mean driven*. Within the context of *security engineering* the probabilistic central moments[4] are of little value since threats always originate from extremes. Here, Bayesian Integral form is employed to handle a large number of *extreme value* variables.

1.4 PI's prior attempt

After September 11, 2001, the principal investigator submitted three proposals to the NSF ITR (Information Technology Research). After the Tsunami of December 2004 and Hurricane Katrina (of August 2005) two NSF proposals addressed the particular issues of emergency management. Following the reviewers comments, a general methodology is addressed here as an intelligent infrastructure framework.

1.5 Items at no cost to the project

1. the principal investigator will devote his entire sabbatical semester to this research;
2. foreign participants from Finland and Japan have secured their own funding;
3. a non-profit international real estate development agency will advise this research *pro bono.*

1.6 Future plans to extend this research

The team will explore the possibility of semi-automated transformation from existing resources. An anticipated automatic translator will facilitate easy deposit of a wide variety of disaster-related numerical responses and policy materials into a common repository by engineers, social scientists, civic authorities, international agencies and the public-at-large. Close cooperation will be sought with various Computer Science Departments, *vide* §10.2, page 12, to clearly separate the semantics and the display aspects. Without the extensive use of `content markup`, the archive will not attain the intended levels of functionality in synthesizing engineering data and policy texts. A flexible and versatile standard language-based representation is necessary to support the kind of indexing and referencing needed via an XML-based document format, ref. [6].

[1] known as *membership functions* in the literature (relevant bibliographic references are included in respective sections)
[2] for positive definite systems, to treat correlation matrices, and carry out perturbation calculations
[3] computer programing terms are written in the `tt` (teletype) font
[4] mean, standard deviation, skewness, kurtosis, etc.

2 Flow Diagram of the Computer Program to be Delivered

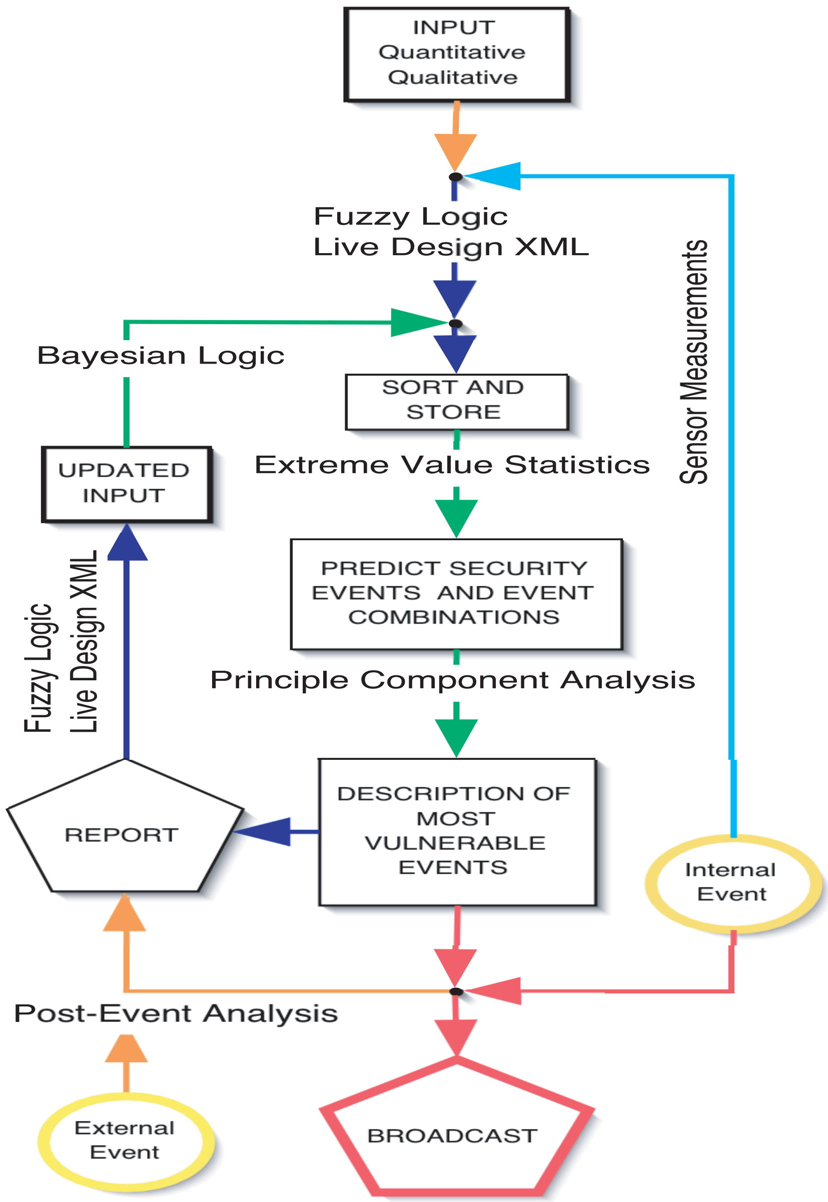


Figure 2.1: Main Flow Diagram of the (Core Technology) Prototype

Figure 2.1 shows the conceptual layout of the proposed methodology. The computer program to be delivered will accept, organize, document and broadcast warning messages, which are shown in the OUTPUT (depicted by pentagons), based on the (*most up-to-date*) data bank described in the body of the figure. For a large and extremely complex civil engineering system the results of the extreme value statistical analysis will be translated into plain a natural language that is familiar to the users.

Symbolic code (a conceptual *Mathematica*, [52], package) for computational statistical aspects, *e.g.* Bayesian updating (in the algebraic form), extreme value distributions and principal component analysis, all with interval arithmetic constructs (ref. [11]) have already been programmed separately by the principal investigator. This project will provide an opportunity to transform the package into a robust (C^{++}) application.

The foreign participants of the project (using their own university funding resources) will develop additional modules to validate the model and investigate emergency management related cases.

Part-II: Mathematical Details for Scientific Tasks

3 Interval Arithmetic

As far as the numerical values are concerned, it is more realistic to assign a *range*, ref. [35], (with `interval arithmetic` constructs[5]) to a decision variable rather than a single number. For example, it is more pertinent to describe the probability of failure as lying between 30% and 40% rather than an unambiguous statement that it is strictly 34%, *vide* ref. [16].

In this research a named variable X will be assigned a range (x_{min}, x_{max}) — called an *interval* in terms of symbolic computation — and within the interval a frequency function $f(x)$ will be assumed. For example, the following $\beta-$ distribution with two independent parameters m and n will be postulated:

$$f(x) = \frac{(\frac{x-x_{min}}{x_{max}-x_{min}})^{m-1}(1-\frac{x-x_{min}}{x_{max}-x_{min}})^{n-1}}{\beta(m,n)}; \quad a < x < b : \text{x=Interval[\{a,b\}] in } \textit{Mathematica} \tag{3.1}$$

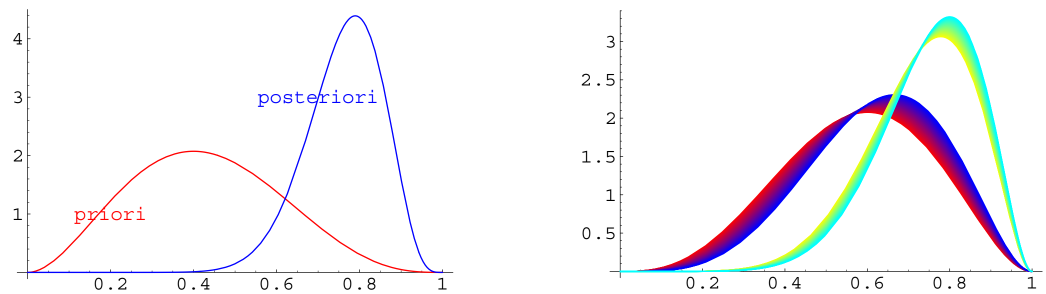
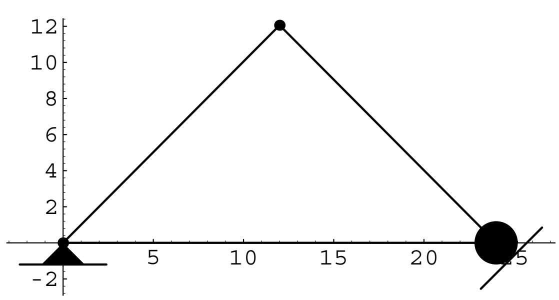


Figure 3.1: (a) Conventional distributions and (b) Distributions with Intervals

3.1 Problem completed



In a truss analysis program, written in *Mathematica*, the height of the crown (in inches) is simply entered as: `12*Interval[{1, 1.1}]`.
Without modifying the *Mathematica* program, (also see §5.3, page 6) for some cases (different orientations of the inclined members and roller sliding angles) the horizontal force in the horizontal tie came out with a negative lower bound.
The following numerical values were obtained when explicit rules related to additive and multiplicative neutrals ($x - x \to 0$; $x/x \to 1$) were enforced.

Figure 3.2: Truss with 10% range for the location of the crown node, *vide* ref. [10, 13]

For a unit downward vertical force at the crown the member force ranges are calculated as:
Interval[{3.90526, 5.4709}], Interval[{-214.249, -142.143}], Interval[{-201.258, -154.359}].

3.2 Tasks for this project

A *robust* numerical method requires to guarantee:

$$\int_{x_{min}}^{x_{max}} \frac{f(x)\, dx}{x_{max} - x_{min}} = 1, \quad \text{when } m \text{ and } n \text{ are intervals: } (m_{min}, m_{max}) \text{ and } (n_{min}, n_{max}) \tag{3.2}$$

A computer program module (a numerical integration routine) will be developed using discretized intervals for x, m and n (during the first six months of the project). The convergence criteria will be examined both numerically and analytically via benchmark examples.

[5] with $z = (z_{min}, z_{max}) : x = (1,2); y = (3,4)$ then $x + y = (4,6), y - x = (1,3); x - y = (-3,-1); \frac{x}{y} = (\frac{1}{4}, \frac{2}{3})$

4 Extreme Value Statistics

4.1 Summary of Completed Review

The theoretical development of this model dates back to the early 1920s with Frèchet 1927, Fischer and Tippett 1928, Gnedenko 1943, and the first comprehensive work by Gumbel, ref. [17,18,21]. Conventional applications range from natural phenomena (*e.g.* floods, droughts, earthquakes) to strength of materials and reliability studies. Within the context of this proposal we have extended the domain to encompass: (i) man-made disasters; (ii) failures in emergency management; and (iii) early warning systems. The theory and application of *extremes* fall in two broad categories: (i) the exact theory; and (ii) the asymptotic theory.

4.1.1 Mathematical definitions and notations to be used in the research

The distribution of extremes differs from its parent distribution. Consider N samples each of size n. Let the j^{th} sample be: $x_1^j, x_2^j, \ldots, x_n^j$ (superscripts identify the sample) the smallest and the largest extremals are $x_1^j, j = 1, \ldots, N$ and $x_n^j, j = 1, \ldots, N$, respectively. If the parent distribution $F_X(x)$ of the random variable X is known, then one can express the distribution of the extremes analytically. If $Y = \max\{x_1, x_2, , \ldots, , x_n\}$, then the cumulative distribution function of Y is $F_y(t)^n$. Conversely, if $Z = \min\{x_1, x_2, \ldots, x_n\}$, then $F_z(t) = 1 - [1 - F_x(t)]^n$. The difficulties for a large n are: (i) whether asymptotic forms exist; (ii) what are they; and (iii) how quickly they are achieved. The following table identifies: Type I – an exponential Gumbel Distribution with no bounds, Type II – Cauchy Distribution, Type III – a limited class of distributions (asymptotically stable according to the Frèchet, and Fisher and Tippet stability postulate[6]).

Initial Type	Largest Value	Smallest Value	Symmetry
Exponential Type I	$\Phi(x) = \exp\left(1 - e^{-\alpha_n(x-u_n)}\right)$ $\alpha_n > 0$	$\Phi(x) = 1 - \exp(1 - e^{\alpha_1(x-u_1)})$ $\alpha_1 > 0$	$\alpha_1 = -\alpha_n$ $u_1 = -u_n$
Cauchy Type II	$\Phi(x) = \exp\left(-\left(\frac{v_n-\epsilon}{x-\epsilon}\right)^{k_n}\right)$ $k_n > 0,\ x \geq \epsilon,\ \nu_n > \epsilon \geq 0$	$\Phi(x) = 1 - \exp\left(-\left(\frac{\omega-\nu_1}{\omega-x}\right)^{k_1}\right)$ $k_1 > 0,\ x \leq \omega,\ \nu_1 \leq \omega$	$\nu_1 = \nu_n$ $k_1 = k_n$ $\omega = -\epsilon$
Bounded Type III	$\Phi(x) = \exp\left(-\left(\frac{\omega-x}{\omega-\nu_n}\right)^{k_n}\right)$ $k_n > 0,\ x \leq \omega,\ \nu_n < \omega$	$\Phi(x) = 1 - \exp\left(-\left(\frac{x-\epsilon}{\nu_1-\epsilon}\right)^{k_1}\right)$ $k_1 > 0,\ x \geq \epsilon,\ \nu_1 > \epsilon \geq 0$	$\nu_1 = \nu_n$ $k_1 = k_n$ $\omega = -\epsilon$

4.2 Tasks in this proposal

The purpose here is to analyze observed extremes and to forecast further extremes. Following the research in [54–59], the team will undertake an analysis of *man-made* security breaches (different from natural disasters). Experience shows no evidence of the Type II model as a descriptor of *commonly known natural disasters* [9,19]. This will be explored in this proposal as preliminary results show the man-made disasters are best modeled as Type-II cases[7]. Distributions of extremes displayed an important property of interest here: the maxima distribution possessed a lower bound but no upper one.

4.2.1 Development of a Seminar Course

There are very few courses currently offered anywhere in the world that focus on the use of extreme value statistics in predicting non-natural disasters. The researchers from Finland plan to develop a short course for health care rescuers. Associates from Japan have demonstrated simulated scenarios for a complex combination of catastrophic security breaches. The numerical results from evacuation and rescue models will be analyzed in the light of the all three previously mentioned types of extreme value statistics.

[6] distribution of the largest value in a sample of size Nn, should be the same as the distribution of the large value in a sample of size n, except for a linear transformation, i.e., $F^n(x) = F(a_n x + b_n)$.

[7] The study will not entirely focus on Type-II, all three types will be examined (for the Bayesian analysis of §6, page 7).

5 Fuzzy Logic Background

5.1 The fundamental notation

Conventional reliability based design, which hinges on mean and standard deviation, is adequate for *safe design*, but furnishes merely the baseline necessary condition for *secured design* — the main theme of this proposal. Extreme value statistics with `interval arithmetic` constructs encompass ranges of uncertainty and provide a complete digital representation of security variables. However, subjective preferences and human intuition cannot be overlooked when mean value and standard deviations mislead security related calculations. This difficulty is circumvented here by the *fuzzy logic paradigm* summarized below. While establishing a robust design algorithms, Arakawa [Arakawa, 1999] proposed and validated the effectiveness of both direct[8] and inverse operators for fuzzy numbers.

5.1.1 A case study — for clarification and a simplified introduction

The example of deriving positional data from digital images associated with the *stereo imaging method* was examined for data reliability. A visual implementation of the proposed *fuzzy logic computer program module* incorporates qualitative (subjective) heuristics within a Bayesian computational environment[9] and is illustrated here. The major challenge of removing optical data contamination[10] is addressed here.

After carefully aligning two cameras, the position data was extracted from digital stereo images of a cluster of buildings. The essential steps of the proposed image sharpening algorithm from 'low reliability' sources is describes here.

The blue dots indicate the fixed reference points.
The capability of stereo view generator is shown by its ability to display the photographs from different viewpoints.
The refinement algorithm reproduced details (right picture) in the second floor from blurry input (left picture).

Figure 5.1: Before (left) and after (right) data decontamination

5.2 Quantification of subjective judgments in intelligent civil infrastructure

After 9/11, 2001, the principal investigator initiated research into *security engineering*. The core technology, which converts *relative priorities* from expert opinions into *probabilistic weights* via the fuzzy logic *membership functions*, found applications in "Early Warning System of the Tsunami," (December 2004) and "Emergency Management after Katrina," (August 2005).

After the Kobe earthquake (January 17, 1995 — Hyogo-Ken Nanbu Earthquake) in Japan, the performance of conventional safety-factor-based-designs was concluded to be inadequate. Research on *fuzzy logic*-based design emerged as a viable alternative. A robust design environment was developed to deploy *fuzzy logic* to relate numerical values of design variables with known dynamic behavior of systems when the computational model of ref. [2, 3] was demonstrated to be numerically more efficient than Monte Carlo simulations. Associates of this project, who are faculty members in Kagawa University, Japan, have developed general purpose *fuzzy logic* software tools. In addition to employing the concept in machine design they have successfully implemented *fuzzy logic*-based management computer programs for communication through visualization and monitoring on-line sensor instrumentation for help and secure facilities from deadly public threats. It has now been conclusively established that subjective design expertise of specialists can be captured within a *fuzzy logic* environment, which can seamlessly interface with probabilistic-based computational framework, *vide* Figure 2.1, page 2, that employs the concept.

[8] The direct method often amplified fuzziness at the function level.
[9] to be elaborated in §6, page 7
[10] Security breaches, lack of coordinations and unpreparedness display similar inconsistencies.

5.3 Mathematical Definitions and Notations

Numerical problems arise in the way available computer algebra programs treat an `Interval Arithmetic` object, cf. §3.1, page 3 for a 'remedy.' For example, if $x =$`Interval[{3,4}]` then $\frac{2\,x}{x}$ is calculated as the *general result* of `Interval[{3/2,8/3}]` and not the *specific result* that is 2. Fuzzy logic can circumvent this anomaly with a high weight $(= 1)$ for all $\frac{x}{x}$ expressions.

Interrelations between fuzzy logic, probabilistic notions and `Interval Arithmetic` are described below, within the context of this proposal.

5.3.1 LR- numbers

In the interest of brevity only the *LR-fuzzy number* representation is described here. Three real numbers $(x, x^L, x^R)_{LR}$, and two non-increasing functions $L(*), R(*)$ identify a variable $\tilde{x}$ that has membership value of unity. Therein x^L and x^R represent the range of the fuzzy number (in *Mathematica*: `Interval[{xL, xR}]`) that determines the spread of $\tilde{x}$. A simple LR (hat function) within the spread is: $L(\xi) = R(\xi) = 1 - |\xi|, |\xi| < 1,\ 0$ otherwise. The following transformation functions is from ref. [2]:

$$f^L(\tilde{x}) = \sum_{\frac{\partial f}{\partial x_i}>0}^{N} \frac{\partial f}{\partial x_i} x_i^L - \sum_{\frac{\partial f}{\partial x_i}<0}^{N} \frac{\partial f}{\partial x_i} x_i^R; \quad f^R(\tilde{x}) = \sum_{\frac{\partial f}{\partial x_i}>0}^{N} \frac{\partial f}{\partial x_i} x_i^R - \sum_{\frac{\partial f}{\partial x_i}<0}^{N} \frac{\partial f}{\partial x_i} x_i^L \tag{5.1}$$

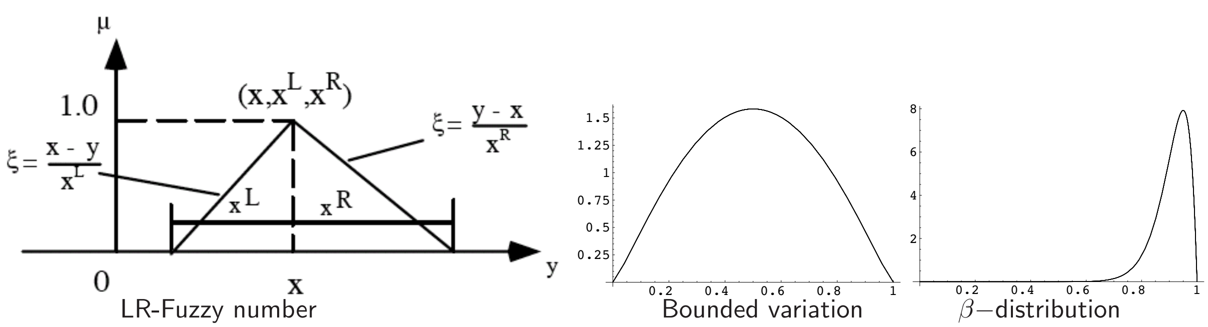


Figure 5.2: Possible distribution of a *fuzzy variable* within bounds

5.3.2 How a photograph can be sharpened — Selection of Highly Reliable Data

The inherent error in any practical measurement (*e.g.* the GPS pixel position digital values pertaining to a building) contributes fuzziness. For a point (x, y), in the spirit of §5.3.1, the LR- fuzzy representation is $((x, x_{min}, x_{max}), (y, y_{min}, y_{max}))$. From moving locations, the GPS data is tracked for the same positions (in the polar (r, θ) coordinate) every at ten minute intervals. The difference in θ indicates fuzziness of the positional data. A Gaussian distribution is assumed for the fuzziness in r. Probabilistic measures are then assigned from expert's (text-based) subjective guidelines. Each piece of data is transformed by a fuzzy function. For a uniform fuzziness in the data, the end results will display different fuzziness. Initially, one uses all combinations of points and calculates fuzzy set sum of membership functions leading to the maximum value of integrated membership function as "*maximum level.*" With a cutoff level, say, $\frac{1}{3}$ — the a value expounded in equation (5.2) — one can reliably identify (those areas above that limit) and determines the data bounds.

5.4 Tasks for this project

Since a narrow interval indicates good data, the weight for designating a reliability index is based on the interval width w of a fuzzy number. Hence, the prototype to be delivered will address the following three issues: (1) assigning the fuzzy number to data, considering the ambiguity of uncertain elements; (2) extracting the data using the fuzzy set sum; (3) implementation of fuzzy weighted averages, such as:

$$w = \frac{x}{x^L} \text{ or } \frac{x}{x^R} \text{ with } x_{estimated} = \frac{\sum_{i=1}^{N} w_i x_i}{\sum_{i=1}^{N} w_i}; \text{ and with a cut-off } a \;\; x_{estimated} = \frac{\sum_{i=1, w_i \geq a}^{N} w_i x_i}{\sum_{i=1, w_i \geq a}^{N} w_i} \tag{5.2}$$

6 Bayesian Network: A Security Engineering Example

The cost of security depends (nonlinearly) on the associated payoff that can be assessed exclusively in the probabilistic sense. The complexity in algebraic formulation grows exponentially with the number of variables. The common and familiar algebraic form (to be described in equation (6.3)) becomes intractable. A Bayesian network, ref. [24, 38, 42], furnishes an elegant representation of joint probabilities in a graphical communication language based on the *semantics of causality*, ref. [39]. This computer program *Modelica* ref. [20, 48] can create a complex system using the formal specification of the links in the database, *vide* examples from ref. [39], pp 355. This motivated the proposed integral representation when the number of variables are very large that are observed in any realistic situation.

6.1 A summary review of the basic Bayesian concept – algebraic form

A probabilistic setting built on the experiences from a *somewhat remote* past is utilized to evaluate the 'current security state' of the civil infrastructure. It is then prudent to effect a mechanism to accommodate observations from the immediate past. Thomas Bayes' posthumously published work of 1763 laid the ground work, which Tribus, ref. [49], commented to be the initiation of the era of mathematical inductive reasoning, as follows: "modification of 'prior opinion' by recent outcomes." Any standard textbook (*e.g.* Harr, ref. [22], Silva ref. [45]) summarizes the following forms as incorporated in PI's seminar classnotes in Finland, ref. [43]. The conditional probability of an outcome $A_i, i = 1 \dots N$, which is mutually exclusive and collectively exhaustive, given that an observed outcome B, is:

$$\textit{posterior}\text{ probability} = \frac{\textit{prior}\ \text{ probability} * \text{likelihood}}{\sum_i \textit{prior}\ \text{ probability}_i * \text{likelihood}_i} \tag{6.1}$$

$$P[A_i|B] = \frac{P[A_iB]}{P[B]}; \quad \sum_i^N P[A_i] = 1; \qquad \sum_i^N P[A_iB] = P[B] = \sum_i^N P[A_i]P[B|A_i] \tag{6.2}$$

therein $P[A_iB]$ is the joint probability that both A and B will occur and $P[B]$ is that of B, thus:

$$\begin{aligned} &P[A_iB] = P[A_i|B]P[B] = P[BA_i] = P[B|A_i]P[A_i] \\ &P[A_i|B] = P[B|A_i]\frac{P[A_i]}{P[B]} = \frac{P[B|A_i]P[A_i]}{\sum_i^N P[A_i]P[B|A_i]}; \qquad \text{Bayes' result} P[A_i|B] \text{ in terms of } P[B|A_j] \end{aligned} \tag{6.3}$$

Since $P[B|A_i]$ have already occurred it is *a priori* thus $P[A_j|B]$ is *a posteriori.* A numerical example follows. A security system predicts fire A_f or toxic chemical A_c to be the cause of a disaster after the fact. Today it predicts a 25% probability of fire. Previously it was correct 80% for fire and 70% for toxicity. Let B be previous prediction for fire. The revised estimate is:

$$P[A_f|B] = \frac{P[B|A_f]P[A_f]}{P[B|A_f]P[A_f] + P[B|A_c]P[A_c]} = .47, \text{ an improved estimation from } .25$$

6.2 Large number of variables

6.2.1 Tasks completed (Bayesian integral form)

This is a crucial step in the *principal component analysis*[11], ref. [25], to yield the *threat index* from a composite scenario of many (in the order of thousand) *security variables* all that are only conceivable in the statistical sense. The generic form (programmed symbolically in *Mathematica*) is:

$$f(x|y) = \int f(x|y,\eta)f(\eta|y)d\eta \text{ and } f(x|y) = \int f(x|\eta)f(\eta|y)d\eta \rightarrow \text{integral equations:}$$

$$f(x) = \int f(x|y)f(y)dy; \;\; f(y) = \int f(y|x)f(x)dx \quad \text{that yields the fixed point : } f(x) = \int h(x,x')f(x')dx';$$

$$\text{where the kernel:} h(x,x') = \int f(x|y)f(y|x')dy \tag{6.4}$$

[11] Eigenvalue computation with `Interval Arithmetic` has been published by the principal investigator in ref. [12, 14].

6.2.2 Tasks completed (Bayesian network — a need to integrate graphic and numerics)

The interesting extension to many Bayesian variables, ref. [44], invokes Gibb's sampling, ref. [1]. The following adjustment problem (ref. [39], pp. 355), which highlights the decision process, ref. [51], has been reproduced using computer algebra with graphics (in *Mathematica*) to depict the network, ref. [36]:

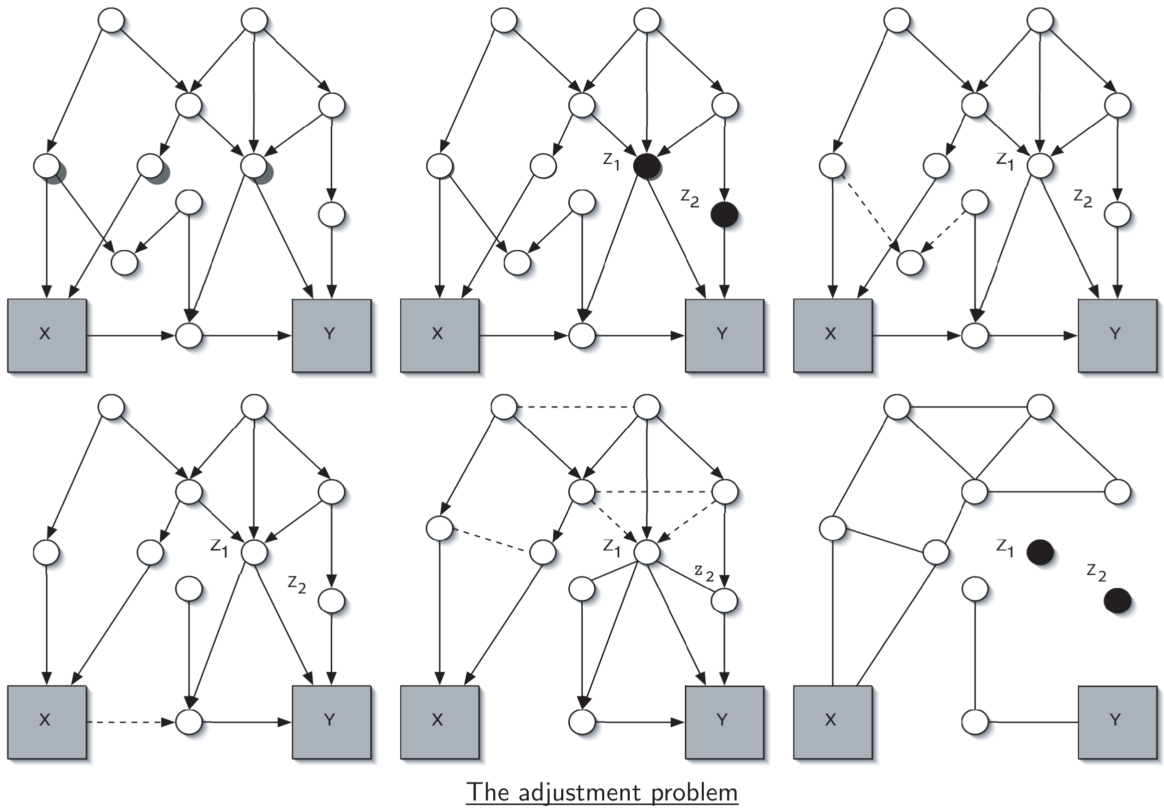


Figure 6.1: Network model for Bayesian analysis

The sequence of six figures shows: (i) the adjustment problem; (ii) testing the sufficiency of Z_1 and Z_2 as appropriate measures; (iii) deletion of all non-ancestors of X, Y, Z; (iv) deletion of all arcs emanating from X; (v) connecting all two parents that share a common child; (vi) verification — X disconnected from Y implies sufficiency of Z_1 and Z_2.

To build the prototype to be delivered the existing *Mathematica* routine will be modified to accept fuzzy links. Probabilistic estimations will then be carried out according to the Bayesian integrals presented in equation (6.4).

6.2.3 Tasks for this project

In Bayesian integration of equation (6.4) the most challenging task is how to implement `Interval Arithmetic` operations without sacrificing tightest bounds[12].

A Bayesian network modeling procedure has already been preliminarily written in *Mathematica* by the PI. The PhD student will implement graph theory constructs with *Modelica*, ref. [20], to address the qualitative features that emphasize (i) 'extreme value statistics' and (ii) generation of quantitative probabilistic measures using 'fuzzy logic'engines on qualitative (policy) information[13].

A technical development will require the creation of a graphics environment where an `Interval Arithmetic` object is depicted with a continuous color distribution, shown in Figure 3.1-b, page 3.

Object-oriented ($\mathtt{C}^{++}$) routines to carry out extreme value statistical computation with a large number of security variables is the main yield of this proposal. These modules will be integrated in the flow diagram described in Figure 2.1, page 2.

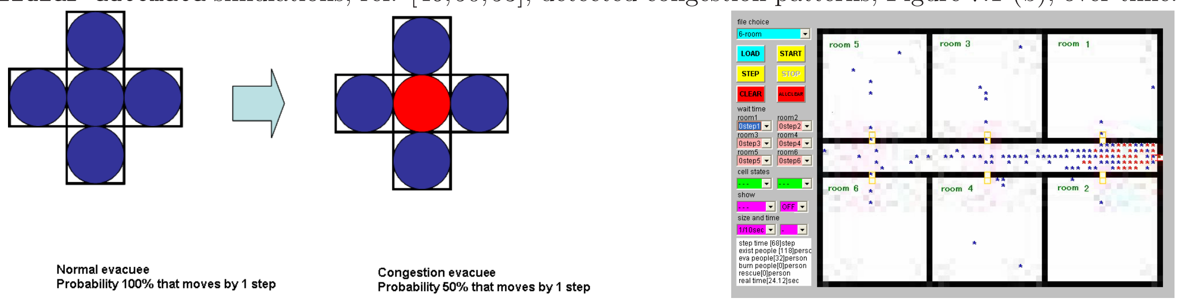
[12] Characteristic difficulties are mentioned in §5.3, page 6. An *ad-hoc* 'remedy' is suggested in §3.1, page 3.

[13] As an alternative, the team will also explore the 'machine learning'paradigm, ref. [26].

Part-III: Application, Proposal Merit and Impact

7 Evacuation based on Quantification of Experts' Guidelines

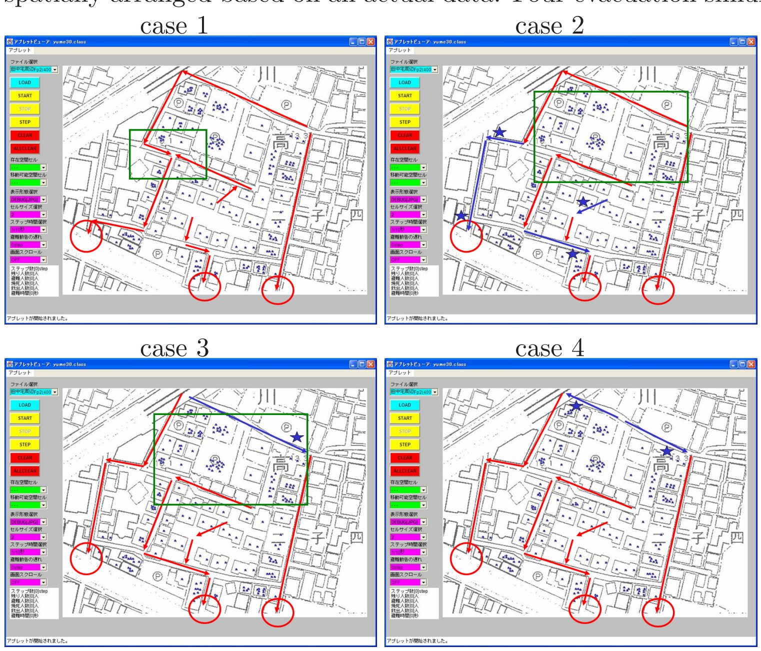
Japanese researchers with this project[14] have published a set of evacuation examples in ref. [4] by quantifying the text-base evacuation guidelines. By discretizing the floor space into cells (*e.g.* occupied, empty, obstacle and engulfed by fire or smoke) and then by updating each cell state at each clock tick, all possible dangerous events are detected from the exhaustive search. Using the rule described in Figure 7.1-a, `cellular automata` simulations, ref. [46, 50, 53], detected congestion patterns, Figure 7.1-(b), over time.

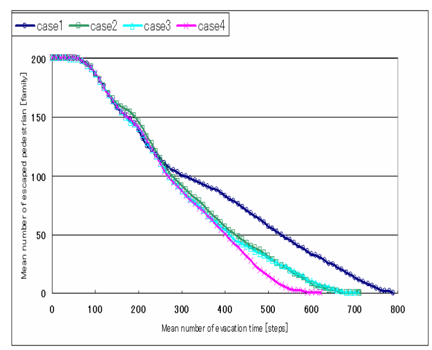


(a) Observed probabilities of evacuations (b) Corridor congestion in a six room floor

Figure 7.1: Text-based rules transformed by *fuzzy logic* in simulating evacuation

A more ambitious example[15] in Figure 7.2 depicting a populous residential area is modeled by 400 by 300 cells where a single family (of 2 to 5) occupies one cell, Figure 7.2. Two hundred and one families are spatially arranged based on an actual data. Four evacuation simulations through three exits (red circles).





Time steps vs. number yet to evacuate

Case 1 depict non-emergency routes, case 2 learns from the mistakes of case 1, case 3 learns from case 2 and finally case 4 learns from 2 and 3.
This Bayesian concept dramatically enhances evacuation:
from 788 steps to 711, 707 and 620,
and reduces estimated times:
from 284, 256, 254 to 233 seconds.
An enlarged PDF view can show more clearly the red (original escape routes) and blue (improved routes), and green (overcrowding) colors.

Figure 7.2: Enhancing community evacuation by successive refinement using discrete simulations

7.1 Tasks for this project

Pedestrian abilities will be considered in greater detail along with panic and fatigue that are inevitable. Individual speed will enter[16] as a fuzzy logic variable (*e.g.* individual health condition and age). For a large number of evacuees, the integral form of Bayesian updating, equation (6.4), will be be employed.

[14] financially supported by the Fire and Disaster Management Agency of Japan

[15] For brevity, another example of evacuation of a multistoried building is not included in the proposal.

[16] aforementioned examples, from ref. [4], apply a uniform rule to all evacuees at all time.

8 Discovering *extreme events* — Simulations during Idle Time

When no new critical information is coming, the computer program *connects the dots* by detecting 'not at all obvious' patterns, ref. [7], that can point out future threats, disasters or catastrophic failures in emergency management, ref. [40]. This background simulation "generates and examines" rarely occurring scenarios using the exhaustive search technique on combinatorics on words, ref. [28, 29]. Here in the interest of clarity and brevity a four letter alphabet $X = \{a, b, c, d\}$, is selected to illustrate the basic task.

8.1 Hypothesizing *extreme events* as abelian square-free words

Let us take an example of a set S of abelian square-free (`a2f`)[17] words, ref. [30], of fixed length over X of length 30. Reducing the redundant structures[18], (renaming of letters[19] and mirror images), the number of words in S would still be in millions. Let us extend (within preset boundaries for lengths) all words in S, in an `a2f` fashion, alternately to the right and to the left (with one letter — word of length 1 — or with longer words at a time). Usually, for a given word in S, experience shows that one finds relatively quickly that the extensions are either impossible to achieve, or that they are easily within reach up to given boundaries for lengths. Interestingly, at rare times it will be extremely hard to decide whether an extension exists or not. Even more rarely, some words like "abcdacbabdabacdacbcdad" of length 22, partly explained below, may require billionfold - or more - computation time in order to decide their true character. It should be emphasized that in spite of over 45 years of intense study, people have not been successful in understanding these processes and structures in a mathematical way.

8.1.1 A numerical example to depict an extreme event — need for exhaustive searches

Consider an abelian square-free (`a2f`) word "abcdacbabdabacdacbcdad" of length 22. Let us try to extend it in an `a2f` fashion with all possible ways alternately to the right and to the left. Here, at a time, the extension length is one letter only. The numbers of extensions (of all words in the bi-directional tree at a time), when the length ranges from 22 to 118 is as follows:

```
1, 2, 2, 5, 14, 23, 14, 26, 10, 16, 8, 9, 9, 16, 16, 27, 27, 54, 54, 68, 136, 194, 291,
444, 296, 450, 225, 331, 331, 474, 948, 1242, 2484, 3072, 1536, 1586, 1586, 1920, 4602,
5597, 4342, 5790, 8838, 11356, 26476, 36306, 32976, 47199, 28989, 39392, 25914, 33008,
30481, 42651, 40586, 53293, 66368, 92574, 88038, 120104, 81647, 108862, 168542, 236432,
254523, 348894, 258148, 349467, 248844, 354942, 247461, 342131, 334879, 463388, 457469,
632051, 166316, 233239, 233187, 324829, 649602, 907191, 445441, 611721, 611063, 840479,
1679287, 2301836, 2302465, 3157227, 3154210, 4306159, 8466798, 11575001, 5779271, 7866918,
0
```

The death of all (nearly 8 million) branches of this bi-directional tree at length 117 is quite dramatic and comes as an enormous surprise. Without an exhaustive search: (i) such a conclusion is not possible; and (ii) some important extreme and critical events may go undetected; thus its indispensability is recognized.

8.1.2 An Algebraic Form to detect *extremely rare events*

Let us find `a2f` *endomorphisms*[20] $g : X^* -> X^*$ restricted to *uniform length* (the same word size, *i.e.* $|g(x)| = |g(y)|$ for all x, y in X). Here, the image words for letters can be obtained by cyclic permutation of letters (thus they would have the same structure as g(a)). Let us start the search from a certain length, say $|g(x)| = 101$[21], or, from $|g(a)| = 10$, under general endomorphisms of *uniform length*. At each iteration, let us increase the length by one, and go through all the alternatives.

The aforementioned may not or may produce an `a2f` endomorphism, in some far away future. During the last 16 years of intensive worldwide research and computation only one example of an `a2f` endomorphism has been found[22] that can be readily identified by the 'security engineers' as an extremely rare event. Such successful simulations in the background, perhaps the only means to detect *a yet unknown mishap*, will discover a new mathematical structure to characterize a pattern akin to a generic disaster.

[17] form $\alpha\beta\gamma\delta$ where the positions of letters in the word β cannot be permuted to yield the word γ as an anagram, α, δ arbitrary (including empty) words

[18] *e.g.* `abacaba` and `bcbabcb`have the same structure

[19] in *Mathematica*: `bcbabcb == StringReplace["abacaba","a"->"b","b"->"c","c"->"a"] -> True`

[20] Informally, `a2f` morphisms are mappings which preserve the `a2f` property of words

[21] if only cyclic permutations are considered (the range 2,...,100 has been checked already)

[22] by the Finnish colleagues of this project, *vide* Figure 9.1, page 11

9 An application for international development agencies

The method will be tested[23] in a multi-million dollar project to build West African university campuses.

9.1 Prelude

International sponsors and donors, without exception, demand a multi-million dollar proposal to furnish evidence that development will proceed as expected to meet the goals. However, only subjective discussions and mainly text-based non-quantitative information are cited to justify management decisions. This model, which will be implemented via a platform independent program to run in personal computers, will yield quantitative measures to depict progress over time. Focus on the three main points, viz., transparency, objectivity and accountability, summarizes the unique feature of this research, ref. [15].

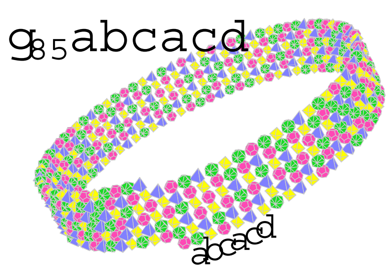
9.2 Methodology

In the interest of brevity, four independent characteristic features are selected to measure and track the progress of a project. For example, in the case of monitoring socio-economic advances in a developing country, quantitative measures related to: (i) nutrition, (ii) sanitation, (iii) education and (iv) recreation, could be adequate; equivalent notions from the “United Nations Millennium Declaration” of (i) illiteracy, (ii) poverty, (iii) hunger and (iv) disease are quite prevalent in contemporary discourses (cf. in weather monitoring (i) temperature, (ii) atmospheric pressure, (iii) relative humidity and (iv) wind speed are the four criteria that have been successfully used). Mathematically speaking, such a formalism constitutes a branch of theoretical computer science known as ‘Combinatorics on Words over a Four Letter Alphabet,’ ref. [30], Figure 9.1 (a) and (b) shows a corresponding sequence of processes.

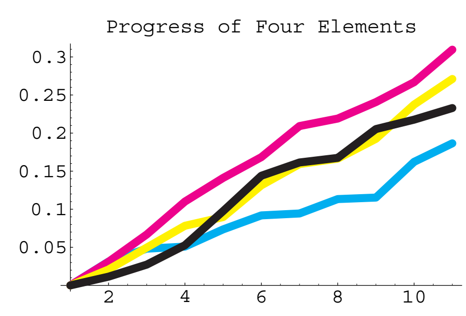
9.2.1 Quantification of progress during the execution of a project

The performance index is a single score, which indicates the overall ‘health of the project,’ starts from zero and evolves as follows:

1. Formalized and well documented text based policy decisions are transformed into probability type numerical weights of positive fractions (mathematically known as ‘membership functions’) that reflect all consensus as well as any difference of opinion according to their relative importance.
   All parties agree in deciding the numerical value that will be attached to each policy element. If all sides agree, modification of such numerical weights (mathematically known as the Bayesian updating) are possible during the execution of the project.
2. The current action is resolved into the four preselected items (*e.g.* (i) through (iv) as above) yielding four positive fractions.
3. Depending upon the correlation with the preceding actions these four fractions contribute a non-negative change to the performance index.
4. Any allowable observation error will provide a range of acceptable values (as per the Interval Arithmetic construct of Computer Algebra). All correlation coefficients will thus be calculated with upper and lower bounds.



(a) The g_{85} Chain with Four Color Beads



(b) Temporal Progress with Four Criteria

Figure 9.1: Modeling with Combinatorics on Words over a Four Letter Alphabet

[23] by this project’s ‘Business Associates,’ The Christie Wareck Company, New Haven, CT (this participation is *pro bono*)

9.3 Transparency

Any discrepancy between the set objective and what really happened can be traced back to track how it evolved. The decision taken and discretion exercised by the local authority can be precisely identified. This will answer in very definite terms 'who did what and when.' In all stages of the development and at all times no action is hidden from the scrutiny by the donor agencies.

9.4 Objectivity

Subjective decisions during the execution of conventional projects are based on the experiences and feelings of management personnel. If those salient features are included during the planning stage, when the Fuzzy Logic numerics are assigned, no major controversial step will be suggested by the computer program. At each step 'what is important' is replaced by 'why is it important.' Consequently any catastrophe arising out of 'human error' can be completely eliminated.

9.5 Accountability

At each step of the development process the theoretical progress trajectory of the 'performance index,' is compared with the field situation. Such a reality check can readily detect any financial corruption in the implementation process. The self-documenting nature of the computer execution precisely points to the responsible human mentor and seeks explanations in accordance with set policies. Not just merely 'what happened' but 'who caused it and why' can be ascertained in an unbiased fashion.

10 *Intellectual Merit* of the proposed research

10.1 Engineering Analysis: numerical methods and computer mathematics

This proposal combines numerical technics pertaining to extreme value statistical computation within a Bayesian updating framework where a range of subjective variability in opinion is represented by `Interval Arithmetic` and fine-tuned with *fuzzy logic.* A typical four by four correlation matrix:

$$\begin{bmatrix} (1.0, 1.0) & (0.498674, 0.851677) & (0.0138573, 0.175681) & (0.122179, 0.14596) \\ (0.498674, 0.851677) & (1.0, 1.0) & (0.120763, 1.08911) & (0.0889141, 0.190372) \\ (0.0138573, 0.175681) & (0.120763, 1.08911) & (1.0, 1.0) & (0.403755, 0.48195) \\ (0.122179, 0.14596) & (0.0889141, 0.190372) & (0.403755, 0.48195) & (1.0, 1.0) \end{bmatrix} \tag{10.1}$$

was obtained for a case of Figure 9.1. An `object oriented` 'eigenvalue' module that can account for fuzzy distribution shown in Figure 5.2, page 6, will significantly enhance the state-of-the-art numerical modeling. The challenge is to obtain all eigenvalues of any correlation matrix as positive intervals.

10.2 Collaborating fields for practical infrastructure management software

Multidisciplinary cooperation between civil (project management), mechanical (sensor development) and computer engineering (software architecture) will create a new discipline of *security engineering* in the spirit of this proposal. The team's on-going collaboration, which intends to build practical infrastructure management software, is summarized below.

1. Computer science (Finland and Japan): to develop an efficient XML-application that provides an extensible data format for interchange, deep domain modeling, and content-based searches; optimized object oriented codes written for data interchange, efficient execution, and easy extension and maintenance; an index to represent subjective threat data, driven by fast numerics, from intelligence reports as opposed to a single deterministic value for an object in conventional CAD;
2. Graphic design (Columbia Architecture School, Prof. Sanchez): automatic forecasting and subsequent generation of threat alarms with suggested rescue and evacuation plans as animation graphics; demonstration of smart signs; and generating warning messages as 3-D graphics with sound;
3. Industrial physics (Université de Franche-Comté, France, Prof. Gharbi): fabrication of lab-on-chip and advanced MEMS (micro electrical mechanical systems) to detect chemical and biological agents;
4. Industrial mathematics (Oslo University College, Norway: Profs. Jonassen and Weierholt): construction of a parallel computing environment and a network of fiber-optic sensors; and modules to estimate the computation time between receiving signals from planted instrumentation.

11 *Broader Impact*

A management tool for donor agencies has been adequately described in §9, page 11.

11.1 Emergency Management

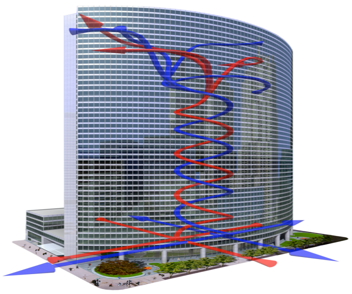


Figure 11.1 shows different paths, red for rescuers and blue for evacuees, depending on the 'current' spatial distribution of the life threatening object. The on-line sensor technology makes it possible to keep all digital information up to date. The most effective measures can be broadcast according to the text-based safety policies.

The proposed paradigm of computing probabilities of successful evacuations and rescue strategies separately, in order to compute the overall effectiveness of an emergency service mission, was instrumental in establishing working relationship with Finnish authorities for providing emergency medical care in the Arctic region.

Figure 11.1: Unlike conventional **EXIT** signs, evacuations and rescues paths are *separately* optimized

11.2 First Responders

11.2.1 Domestic

The team discussed the possibilities with *local/federal* authorities to help in conducting safety drills in buildings and fields on Maine's Ayers Island, in cooperation with the University of Maine, Orono, ME. After this *proof-of-concept* endeavor, *the fractional cost* caused by a single disaster could cover subsequent NSF research, with closer cooperation with civic authorities when special purpose computing modules can be integrated with the product of this proposal to yield life saving security strategies.

11.2.2 Foreign

The team's Finnish researchers will help design and implement a special purpose program, ref. [37], for medical emergencies in the Arctic region.

11.3 Training Evacuation

Evacuation specialists have noted that, unfortunately, people behave quite differently during emergency evacuation drills and real situations. There have been no means to quantify such 'lag factor' (inverse of the safety factor). For example, to really evacuate in 1 minute the exercise should set the target to be 40 seconds, if the 'lag factor' is $\frac{2}{3}$. The proposal provides one such important application technology via quantification of qualitative experiences and opinions.

11.4 System-of-Systems Research

Under complex systems the conventional studies are restricted to digital information. The current state of large-scale concurrent and distributed systems whose components are complex systems themselves cannot handle qualitative information. The current proposal addresses the core technology that makes it possible to widen the domain of *System-of-Systems* applications, ref. [27, 31, 32], to eradicate diseases in poor and developing countries. The team members are currently collaborating with international agencies such as the UN and OXFAM.

11.5 Consumer Protection — Reading the Fine Print

A public advocacy agency requested assistance in the following area. Consumers are many time baffled by the different restrictions and options posted on products. Especially for senior citizens, understanding the medical benefits they are entitled to and the restrictions in different plans have recently posed an enormous hardship. Since the medicine and care costs are stated in terms of numbers accompanied by policy texts it is hard for individuals to determine their best course of action. The exhaustive search technique of this proposal can attend to personal needs that depend upon an existing medical condition.

The proposed computer program can be accessed via the internet to serve a wide group of citizens.

12 PIs NSF project record (2000 —)

This proposal is in a new area compared to PI's NSF funded research on engineering mechanics. All publications generated from NSF grants are classified and listed separately after the section "References."

12.1 Civil Engineering and Engineering Mechanics

12.1.1 Convex Polygonal Finite Macroelements

July 1, 2000 to May 31, 2002 — CMS-9820353: $198,000
PhD Graduate: Andre S. Publico, Current Affiliation:(Departments of Civil, Mechanical and Bio engineering)
Associate Professor, University of the Philippines Diliman, Quezon City 1101, Philippines

12.1.2 Concave Finite Element Shape Functions

June 1 , 2002 and expires August 31, 2007 [24] , CMS-0202232 : $275,104
PhD Graduate: Elisabeth A. Malsch Current Affiliation:
Thornton-Tomasetti Group, Inc., Corporate Headquarters: 51 Madison Ave. New York, New York 10010
Distinguished awards for Thesis work: AvHumboldt-Stipendiatin, Institute of Applied Mechanics, Technische Universität, Braunschweig Postfach 3329, 38023 Braunschweig, Germany.

12.1.3 Research Experiences for Undergraduates

June to August 2004, Grant No. CMS-0202232, Amendment No. 001 Proposal No. CMS-0434465
Two students studied structures and foundations of historic structures in France.

12.1.4 Workshop: Nano, Continuum, Material and Computational Mechanics

October 15, 2003 to September 30, 2004, CMS-0350433: $14,788
In conjunction with the Asia Institute of Technology, Thailand, the workshop initiated long term projects in continuum mechanics applied to:

(i) nano-mechanics: charge density calculation on curved surfaces
(ii) carbon fiber reinforcement in vibration isolation in buildings and rail roads
(iii) NSF sponsored analysis of structures and foundations at World Heritage sites

12.2 International Cooperation

12.2.1 US-France Cooperative Research: Engineering Shape Calculation for Surgery, Biology and Anthropology

March 1 , 2003 to February 28, 2007 – US-France – INT-0233570: $16,920 (to Columbia U)
and CNRS ca. $20,000 to U. Bordeaux-I

The study of orthomorphic sampling has revealed an architectural balance for the anatomical components in Treil's maxillo-facial model that has now become a standard reference in Anatomy, Surgery, Orthodontics, Anthropology and Forensic medicine. A special volume of the *International Journal of Evolutionary Optimization* will publish papers [25] of two seminars in New York, 2003 and Avignon, 2006.

12.2.2 Workshop for Scientists and Engineers on Structural Deformations at the Historic Site of Angkor, in Cambodia

April 15, 2005 to March 31, 2007 – OISE-0456406: $33,000

Workshops are scheduled for 2007. Faculty members, graduate and undergraduate students are now being contacted. Build Bright University, Phnom Penh, Cambodia will host seminars on the structural engineering aspects. Institute of Environment and Global Technology Solution, Phnom Penh, Cambodia will jointly host seminars with OXFAM and UN-NGOs on local economy and impact of tourism.

[24] Due PI's medical leave from June 2005 to December 2005, all NSF projects were granted one year no-cost extension.

[25] peer reviewed archive journal publication from the lectures by: José Braga, University of Bordeaux I; Jacques Faure, University Paul Sabatier, Toulouse III; James T. Goodrich Albert Einstein College of Medicine, NY; Brian Morgan, UCLA; Héctor M. Pucciarelli, La Plata National University, Argentina; Joan T. Richtsmeier, The Pennsylvania State University; Francis Thackeray, Transvaal Museum, Pretoria, South Africa; Jacques Treil, Clinique Pasteur Toulouse; A. J. M. Spencer, U. Nottingham; A. P. S. Selvadurai, McGill; H. Irschik and B. Buchberger, U. Linz; E. Tachibana, Osaka U.; W. Shiraki, M. Arakawa and H. Inomo, Kagawa U., Takamatsu. (undergraduate participants did not present any paper)

Appendix: Task Table

	Year-1		Year-2		Year-3	
	July to December 2007	January to June 2008	July to December 2008	January to June 2009	July to December 2009	January to June 2010
Doctoral Student	a) Review Image focusing papers	a) Review Interval Arithmetic	a) Develop Bayesian Integrals, network	a) Four letter model to simulate disasters	a) Numerical simulations	a) Dissertation
	b) Review Extreme value statistics	b) Code large number of Bayesian variables	b) Implement Interval Arithmetic for Fuzzy Logic variables	b) Examples of Fuzzy Logic in Bayesian settings	b) Dissertation	b) Journal and Conference Papers
					c) Conference Papers	
PI	Develop Interval Arithmetic codes	Develop Extreme value statistics codes	Fuzzy Bayesian variables	Consolidating symbolic program modules	Journal and Conference Papers	Journal and Conference Papers
Associates (Japan)	Develop Evacuation simulations	Updating Evacuation strategy	Organizing a Conference	Evacuation with four features	Test constraints with Intervals	Write Journal Papers
Associates (Finland)	Work with Emergency medical rescuers	Avoidance of disaster patterns in four letter model	Four letter fuzzy logic model	Four letter Bayesian model	Tutorials of all mathematical aspects	Write Conference Papers
Business Associates	Researching case histories	Presenting case studies	Presenting emergency model	Project management model	Presentations to International donors	Presentations to International Planners