Proceedings of Ninth World Conference on Earthquake Engineering
August 2-9, 1988, Tokyo-Kyoto, JAPAN (Vol.VIII)

SK-12

A KNOWLEDGE-BASED SEISMIC RISK EVALUATION SYSTEM FOR THE INSURANCE AND INVESTMENT INDUSTRIES (IRAS)

Weimin DONG[1], Jong Eup KIM[1], Felix S. WONG[2] and Haresh C. SHAH[1]

1 Department of Civil Engineering, Stanford University, Stanford, California, USA

2 Weidlinger & Associates, 4410 El Camino Real, Suite 110, Los Altos, California, USA

SUMMARY

This paper summarizes the development of the Insurance and Investment Risk Analysis Systems (IRAS) which provide consultation on earthquake risk for the insurance and investment banking industries. Major features of IRAS will be briefly described, including: interactive input/output facilities, graphic data retrieval, hierarchical knowledge-based management, integration of independent program modules, combinations of backward-chaining and forward-chaining inference mechanisms, and approximate reasoning schemes based on fuzzy set theory to deal with linguistic and/or incomplete information.

INTRODUCTION

Recent advances in earthquake engineering have permitted experts to rationally estimate earthquake hazard, vulnerability, and risk. However, this expertise is not readily available to managers in insurance and banking industries because it is difficult for non-engineers to define the relevant data for seismic risk evaluation, obtain these data, and apply these data for decision making. Thus, a project is initiated at Stanford to develop a knowledge-based expert system called IRAS (Insurance and Investment Risk Analysis Systems), which provides consultation on seismic vulnerability of existing buildings for use by real-estate investment analysis, insurance underwriters and reinsurers, decision makers and portfolio managers, and structural engineers and appraisers in evaluating and revising investment strategies and insurance policies or premiums.

IRAS has several features. First, it adopts the commercial software I/O PRO as the main Input/Output facilitator. The screen development system facilitates creation of text and graphic screens used as the input and output media for interactive programs. The slides displayed on screen are used to communicate with the user for input and output and explanations. Input data formats can be numerical, linguistic, or graphical based upon the context. Second, the control strategy for the inference mechanism in IRAS is a combination of backward chaining (goal-driven) and forward chaining (data-driven). The system uses backward chaining to satisfy diverse goals (inquiries). However, if the goal is specified, the system uses forward chaining to collect the relevant data. Since goal specification significantly reduces the search space, only minimum search effort is required. Third, unlike the conventional rule-based systems, IRAS recognizes the fact that seismic risk evaluation needs both judgemental expertise and well established mathematical procedures. Hence, IRAS incorporates both rule-based systems and algorithmic programs which saves a great deal of computation. Fourth, to increase the ease of upgrading of IRAS and to ease the restriction of internal memory, the submodules of IRAS are written as independent programs. The programs are compiled independently and are then called into memory when needed by the driver. Fifth, IRAS adopts the current probabilistic approach for hazard analysis to handle uncertainties in the prediction of ground shaking for the site. However, in evaluating the vulnerability of a building, design detail, construction quality and other factors will affect the performance of a building during earthquakes and must be identified in order to get a reasonable evaluation. To reflect the judgemental knowledge of the effect of different factors on building damage, IRAS uses an uncertainty model based on fuzzy set theory.

IRAS is designed in such a way that the knowledge bases can be updated as new information is made available. Furthermore, the systems are user friendly, and hence their repeated usage can make them a standard for evaluating seismic hazard and risk for the banking and insurance industries.

DOMAIN KNOWLEDGE

Seismic risk is defined as the likelihood of loss due to earthquakes and involves four basic components: hazards, exposure, vulnerability, and location. These factors are further defined below (Miyasato et al, 1986):

* The hazards or dangerous situations may be classified as follows:

  - Primary hazards (fault break, ground vibrations):
  - Secondary hazards which are potentially dangerous situations triggered by the primary hazards. For example, a fault break can cause a tsunami or ground shaking can result in foundation settlement, foundation failure, liquefaction, landslides, etc.:
  - Tertiary hazards produced by flooding by dam break, fire following an earthquake and the like.

  All these hazards lead to damage and losses. They may be expressed in terms of severity, frequency, and location.

* The exposure is defined as the value of the structures and contents, business interruption, lives, etc.

* The vulnerability is defined as the sensitivity of the exposure to the hazard(s) and the location relative to the hazards(s).

* The location is defined as the position of the exposure relative to the hazard.

For the purpose of insurance and the real estate industry, property losses are the major concern. Property loss is usually measured by the damage ratio which is defined as the repair cost of the damaged facility divided by the initial cost of the facility. Due to uncertainties in predicting structural behavior during future earthquakes, the current practice of the insurance industry in California is to use PML (probable maximum loss). (Steinbrugge, 1982) as the basis for premium calculation. PML is defined as the damage ratio so that during the "maximum probable" earthquake, 9 out of 10 buildings will experience damage less than the value given by PML.

PML does not consider the randomness of earthquake occurrence with respect to time, location, or earthquake size. In order to reflect the uncertain nature of earthquake occurrence, a second index, called the damage threshold (DT) is used which combines the uncertain response of the building with the random occurrence of future earthquakes (Chiang et al, 1984, ATC, 1986). Both indices are used in IRAS.

IRAS is divided into three subsystems which corresponds to the major components of seismic risk: SHES, SRES1 and SRES2. The seismic hazard evaluation system (SHES) combines hazard and location components to obtain the seismic hazard estimation. The main flow chart for SHES is shown in Figure 1. SRES1, the seismic risk evaluation system, is used to screen the property loss from exposure and vulnerability of the building. In this level, only building type (classification) is required. The flow chart for SRES1 is shown in Figure 2. Figure 3 shows the flow chart of SRES2, which performs the second level of seismic risk evaluation taking into consideration specific information of the buildings. Data management and the inference mechanism of these subsystems will be described in the sequential sections.

INFERENCE MECHANISM (INFERENCE ENGINE)

An inference engine incorporates reasoning methods which act upon input data and knowledge from the knowledge base to solve the desired problem and produce an explanation when requested. Control strategy for the inference engine could be forward-chaining, backward-chaining or a mixture of both.

In the IRAS application, the system should be able to satisfy diverse goals (inquiries) such as "what is the real estate investment portfolio risk for a given region due to a catastrophic earthquake ?" or "what is the probable maximum loss of a particular building due to all contributing fault seismicity?". The goal specifies the reasoning path that should be pursued. Hence, it is natural that backward chaining (goal-driven) should be used. However, when the goal is specified and the reasoning path to achieve this goal is identified, the systems will use forward chaining (data-driven) to collect the relevant data either by querying the user or searching and retrieving it from the knowledge base. Thus, the control mechanism is a combination of backward chaining and forward-chaining. Since goal specification significantly reduces the search space, only a minimum search effort is required.

KNOWLEDGE (DATA) BASE

The knowledge (data) base for the IRAS systems consists of raw data, production rules, engineering and analysis programs and approximate reasoning schemes. Unlike conventional rule-based systems which use If-Then rules only, IRAS recognizes the fact that seismic risk evaluation needs both judgemental expertise and well-established mathematical procedures. Hence, IRAS incorporates both If-Then production rules and algorithimic programs. For instance, model selection depends heavily on the expert's subjective judgement, and If-Then rules are suitable to guide the user to select the appropriate model. After the model is selected, the relevant procedures are executed using algorithmic programs.

Combining inference rules with algorithmic programs is also necessary for the following reason. In most cases during inference, when the facts match the antecedents of a particular rule, the rule is triggered and the consequent can be retrieved directly from the knowledge base without further computing. When the conditions do not match the antecedents of any rule in the knowledge base, the systems will refer to the relevent programs to calculate the consequents (results). This approach saves a great deal of computation, a consideration especially important for micro-computer implementation. Obviously, it is applicable only for problems where the inference mechanism is well-defined (as regular computational programs). For loosely structured inference mechanism, the partial matching problem is resolved through default (applying prior information). In this case, the reliability of the consequent is reduced. The process of uncertainty propagation will be described presently.

INTEGRATING INDEPENDENT PROGRAMS

A common practice in programming is to have a main driver and many subroutines. The driver and subroutines are compiled into a global executable program. However, when the problem to be solved is complex, many subroutines and submodules are needed. The size of the program increases rapidly and soon the capacity of the internal memory of a microcomputer is exceeded. It is then necessary to rely on fancy input/output manipulation and peripheral storage to fit the program into the computer memory. Furthermore, when any submodule of the program needs to be changed due to technological or engineering advances, the relevant routines must be changed and recompiled. The fitting must be reconstructed.

SEISMIC HAZARD EVALUATION SYSTEM
(SHES)

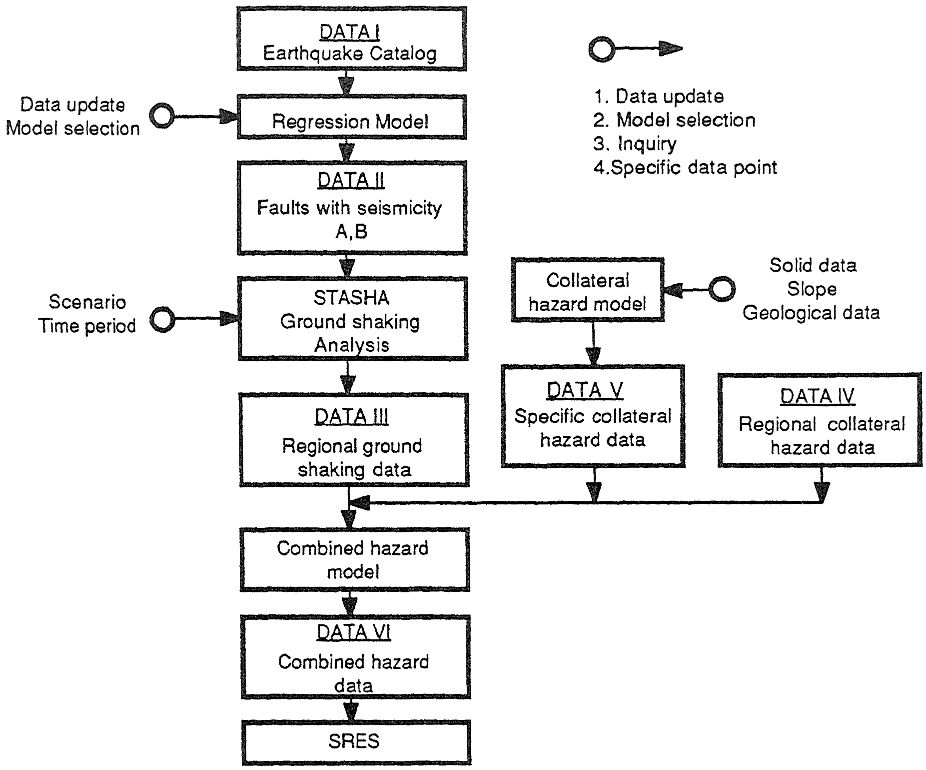


Fig.1

SEISMIC RISK EVALUATION SYSTEM
(SRES-1)

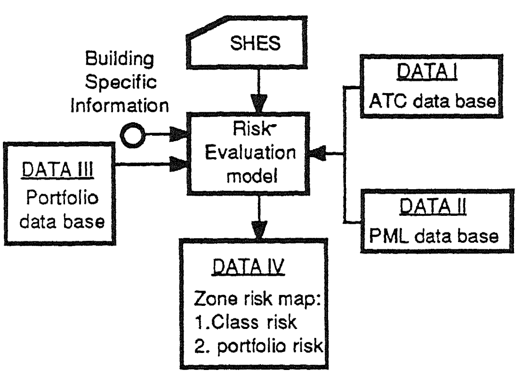


Fig. 2

SEISMIC RISK EVALUATION SYSTEM
(SRES-2)

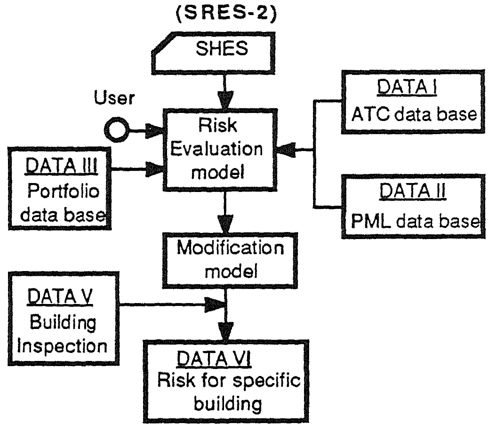


Fig.3

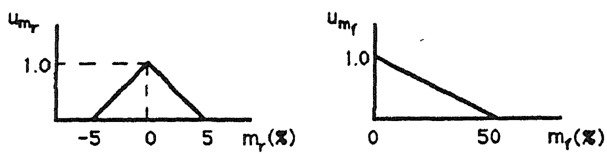
To facilitate upgrading of IRAS and to ease the restriction of internal memory on the IBM AT, the submodules of the systems are written as independent programs. Most of these programs have already been developed during the past ten years by staff and students at the Blume Earthquake Engineering Center, and they are simply ported to the microcomputer. The programs are compiled independently and individually, and are then called into memory when needed by the driver, much like a subroutine is used. The retrieval, execution and then return of the external programs is easily achieved on the IBM AT using the interrupt feature of DOS. Each external program can be as large as the total internal memory of the AT (currently at 640K).

UNCERTAINTY TREATMENT

As mentioned in the previous sections, there are uncertainties involved at each stage of the evaluation process. Earthquake occurrence is random in nature; so is its size. For this type of uncertainty, the probabilistic approach has been well established and the data in California is reasonably good to support estimation using this approach. Hence, IRAS adopts the current probabilistic approach for hazard analysis to handle uncertainties in prediction of ground shaking for the site. The program STASHA, developed at Stanford University for hazard analysis, was incorporated into IRAS using the approach described in the previous section.

There is yet another type of uncertainty in the evaluation which cannot be handled using probabilistic methods. In evaluating the vulnerability of a building, design detail and construction quality will affect the performance of the building. The damage degree will vary in a wide range from bad engineering design to good engineering design. All these factors will significantly influence the building performance during earthquakes and must be identified in order to get a reasonable evaluation. When the user fails to answer the inquiry on these factors from the systems, it is then expected that the system will give an answer with a wider spread due to the larger uncertainty. Because data regarding damage from diverse building types is scarce and is not sufficient to support a probability distribution, IRAS uses an uncertainty model based on fuzzy set theory to reflect the judgmental knowledge of the effect of different factors on building damage.

Fuzzy sets with different membership functions are used to represent the prior information on these effects. Some examples are given in Figure 4.



Modification for roof system   Modification for fault rupture

Fig. 4 Fuzzy representations for modification factors

Whenever the response to a query for data is unknown, the system will use the fuzzy set instead of a crisp number to count its effect. The Vertex method (Dong and Wong, 1987, Dong and Shah, 1987) is used to combine all these effects and to calculate the total effect, resulting in a certainty factor which reflects the degree of uncertainty. When the system gives the evaluation result, it also indicates the reliability of the result (certainty factor) and how the reliability can be improved.

INPUT/OUTPUT FACILITIES

IRAS adopts as the main I/O facility the commercial software I/O PRO, developed by MEF Enviromental (MEF, 1985). I/O PRO is a modular set of software development tools and utilities which together create a high productivity environment for FORTRAN, C and Pascal programmers. The screen development system facilitates creation of text and graphic screens used as the input and output media for interactive programs. The slides displayed on screen are used to communicate with the user for input and output and explanations if requested. Input data formats can be numerical, linguistic or graphical, based upon the context.

Besides the interactive mode, the user can also choose the batch mode in which all data are read in together using a format such as the Lotus 1-2-3 spreadsheet or Dbase III. This mode facilitates the data transfer from insurance and investing banking company data bases.

In order to display the regional risk, IRAS also incorporates another commercial software, ATLAS (Strategic Locations Planning, 1985), to show the thematic map of regional risk. All I/O options are built into the master program and can be exercised according to the user's goal and decision needs.

CONCLUSION

This paper summarizes the development of the seismic risk analysis systems, IRAS, for insurance and real estate investment industry. The systems has been completed and demonstrated to the clients and users. The response from potential users has been excellent. The systems now are in operation and will be further improved in practice

IRAS is developed for buildings located in California, due to its sponsors' main interest. However, in view of the modularity and flexibility of its design, IRAS can be readily adapted to other regions of the world when the appropriate data/knowledge bases are available.

ACKNOWLEDGEMENTS

The authors wish to express their sincere gratitude to Aetna and Salomon Brothers for their support. Partial support was also provided by the National Science Foundation under grant ECE-8515252 to Stanford University.

REFERENCES

1. Applied Technology Council, Earthquake Damage Evaluation Data (ATC-13), (Redwood City, CA, 1986).
2. Chiang, W.L. et al, Computer Programs for Seismic Hazard Analysis, Report No. 62, The John A. Blume Earthquake Engineering Center, Stanford University (1984).
3. Dong, W. and Shah, H., Vertex Method for Computing the Function of Fuzzy Variables, Fuzzy Sets and Systems (1987), Vol 24, No. 1, pp. 65-78.
4. Dong, W. and Wong, S., Fuzzy Weighted Averages and Implementation of the Extension Principle, Fuzzy Sets and Systems (1987), Vol. 21, No. 2, pp. 183-199.
5. MEF Enviromental Inc., I/O PRO -Screen Development System and Utilities, Version 1.1 (1985), Austin, TX.
6. Miyasato, G., Dong, W., Levitt, R., Boissonnade, A., Implementation of a Knowledge Based Seismic Risk Evaluation System on Microcomputers, Artificial Intellegence in Engineering, (1986) Vol. 1, No. 1, pp.29-35.
7. Steinbrugge, Karl V., Earthquakes, Volcanoes, and Tsunamis - An Anatomy of Hazards (Skandia America Group, New York, 1982).
8. Strategic Locations Planning, ATLAS-AMP-EGA (1985), San Jose, CA.