Visual book review[1]

"Safe and Sound, AI in hazardous applications" by John Fox and Subrata Das

Boris Kovalerchuk
Dept. of Computer Science
Central Washington University, Ellensburg, WA 98926-7520
borisk@cwu.edu

An abbreviated version will appear in IEEE Engineering in Medicine and Biology Magazine, 2001

01.19.2001

A recent book "Safe and Sound, AI in hazardous applications" by John Fox and Subrata Das (AAAI Press/ The MIT Press, 2000, 293p., ISBN 0-262-06211-9) attracts attention of the research community and practitioners to the problem of safety of traditional and computer-aided medical diagnosis and treatment. The authors use medical applications as a focal point for the general safety problem through variety of hazardous applications. At first glance, the problem is already well known. However, they show that discovery and analysis of the sources of danger in hazardous applications are far from having rigorous solutions. The book uses an artificial intelligence (AI) approach, which allows one to express different types of statements in consistent logical fashion. Specifically the authors consider two main types of statements for making medical decisions: claims and their grounds. In addition, a confidence value is assigned to a claim using ground statements. For instance, the claim can be that Mr. P has a gastric ulcer and the grounds can be that Mr. P has pain after meals & ulcers are painful because of an increase in acidity. The word "support" can express the level of confidence. In the book, the safety issue for defining diagnosis and treatment is viewed as adding restrictions on inference. The goal of restrictions is to avoid dangerous actions for patients. This review should help a reader to see conditions for successful applications of **logic of argument** (LA), which is the central theme of the book.

The book contains three parts (see Figure 1):

Part 1: Method for building software agents, which produce Rigorously Engineered Decisions.

Part 2: Technique for deploying agents in general and hazardous environment with medical examples.

Part 3: Formal and logical aspect of the method.

The spread of the central theme over the book is presented in Figure 1. Chapter 4 provides an informal description of LA, Chapters 13 and 15 contain formalisms and Chapter 15 implementation of LA in Prolog language. The critical component of LA is exposed in Axiom 8 (Chapter 13). This axiom is the central assumption of LA. It actually formalizes the requirement of independence of arguments used in LA in inferring diagnosis or treatment recommendation. The property is also known as truth-functionality in Artificial Intelligence literature [1] and was a subject of intensive discussions in AI community for years [1,2,3].

Figure 2 presents main concepts introduced in the book such as

- Domino autonomous agent model. The model includes several components: Goals, Situation (patient data), Actions (clinical orders), Candidate solutions, Decisions (diagnosis, treatment), and Plans (treatment and care plans) (Chapter 6).
- PROForma -- the development environment for the Domino model. PROForma includes techniques for reasoning under uncertainty, formulating goals, identifying and arguing solutions. (Chapter 5).
- Guardian agent and safety bag. These components are watching out for hazards and safety (Chapter 7).
- Safety protocol. This protocol allows one to communicate safety conditions in terms of modalities such as safe, authorized, preferred, permitted, and obligatory (Chapter 10).
- Reasoning procedure. This procedure permits explicit reasoning about safety in terms of actions such as anticipate, alert, avoid, augment, diminish, monitor, schedule, react, authorize (Chapter 8).

Map of the book with the central theme

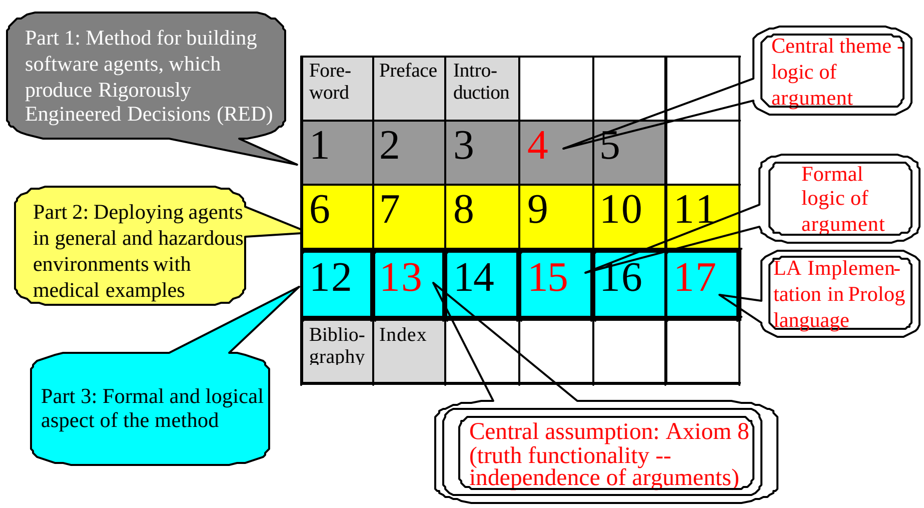

Figure 1.

The authors coined the term RED -- Rigorously Engineered Decisions in order to distinguish diagnostic and treatment decisions made in disciplined way from less controlled decisions. The RED approach is based on making decisions from mentioned claims, their grounds, and confidence values. Specifically, the authors suggest using the logic of argument. A modern AI terminology of autonomous software agents is used in the book. The Figures below analyze LA an instantly mathematical subject by using a conceptual visualization approach.

Technical results presented in the book have two components: the general logic of argument is covered in Chapters 4, 12-15, and 17, and a formal logic of safety is covered in Chapter 16 (see Figure 3). Two languages are introduced for LA: language R^2L (Chapter 12) for rigorously engineered decisions, and its formal equivalent Temporal

Propositional Language R2L (Chapter 13). The transformation procedure from R2L to R2L is given in Chapter 14. The first language is convenient for describing tasks and the second one for implementation in Prolog (Chapter 17).

Main concepts introduced

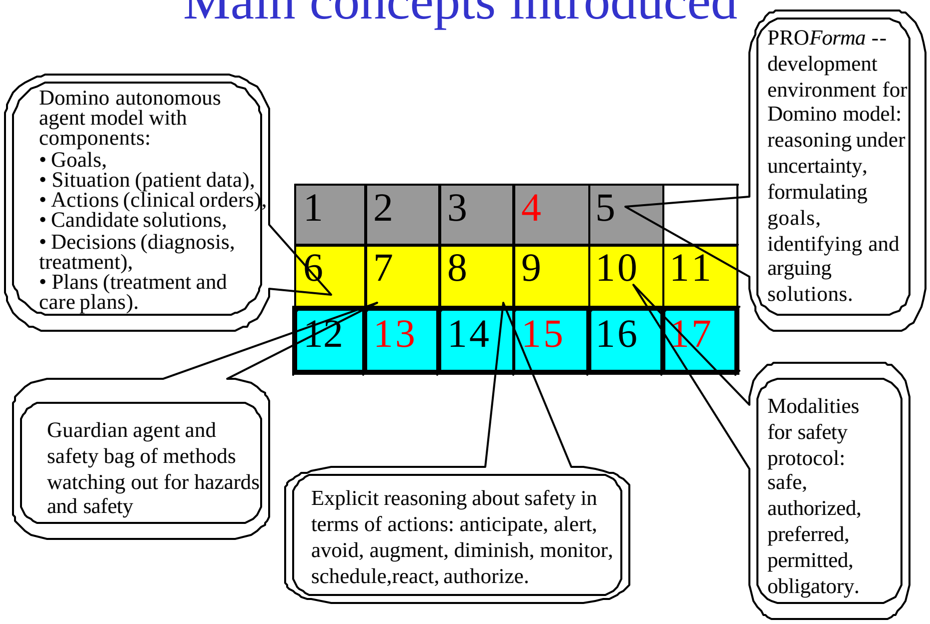

Figure 2.

Main technical results

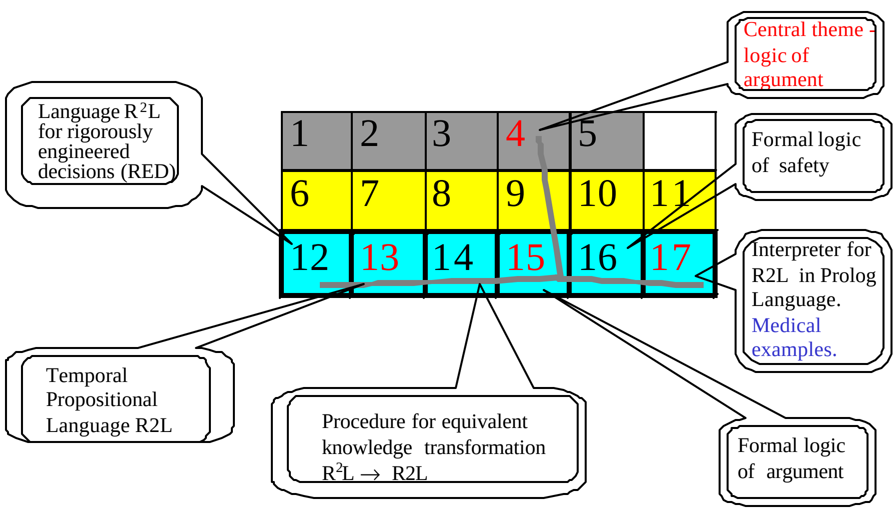

Figure 3.

The authors see AI and cognitive science professionals as their primary audience. Note that the people working in medical applications can gain a greater appreciation of the potential to be found in the rigorous analysis of safety issues as well. Unfortunately, medical funding agencies in the US have underestimated this potential. The book provides enough successful medical examples to convince a reader that Rigorously Engineered Decisions make sense and should have a strong future potential in medical applications.

However, as with any technique, the RED approach is not a panacea. It has its own restrictions. The rest of this review is devoted to description of the central theme of the book -- logic of argument (Figures 4 and 5) and analysis of restrictions of the approach (Figures 6 and 7). The term argument is defined as a triple: <Claim: Grounds: Qualifier>, where a Qualifier is the confidence value assigned to the Claim based on the Grounds. The authors emphasize that an argument has a form of a logical rule,

$$\text{Grounds} \Rightarrow_{\text{supports}} \text{Claim}$$

but does not have not its force, that is, an argument does not sanction a definite conclusion (see Figure 4).

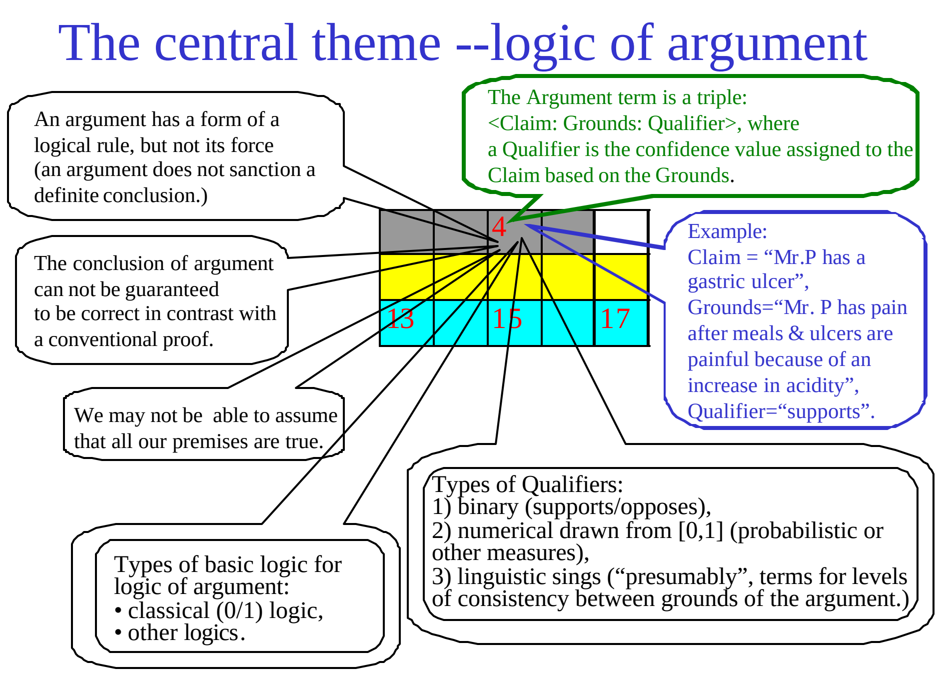

Figure 4.

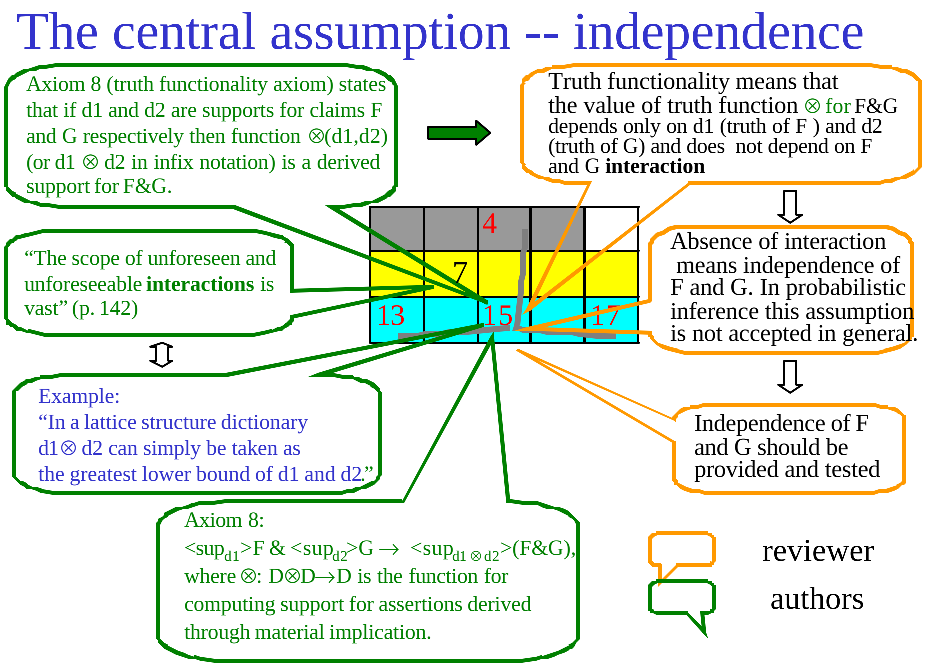

Figure 5.

The authors state that the conclusion of argument process can not be guaranteed to be correct in contrast with a conventional proof. This, of course, is the typical situation in medical arguments.

Several types of Qualifiers are considered to express the level of support for the claim: binary (supports/opposes); numerical drawn from [0,1] (probabilistic or other measures); linguistic signs such as “presumably” or other terms for expressing levels of consistency between the grounds of the argument.

The authors base their assumption of independence (Axiom 8) on the statement that in many applications there is not enough data for estimating dependencies in the form of conditional probabilities P(F|Q) and P(Q|F). This is obviously true. Therefore, the authors select tasks were arguments are most likely independent. For instance, argument A= “ Mr. P has pain after meals” and argument B= ”ulcers are painful because of an increase in acidity” are relatively independent. Ulcers can be painful independent of the pain suffered by the Mr. P after meal. However, this is not something that can be guaranteed for all possible medical statements.

Figure 5 explains the logic of axiom 8 and its manner for expressing independence of arguments through truth-functionality.

Figure 6 analyzes future of the central assumption of the book. Authors mention that "The scope of unforeseen and unforeseeable **interactions** is vast" (p. 142). For some tasks, designers can provide independent (non-interacting) arguments, but it is not a universally rigorous design approach.

Fuzzy logic control success justifies this statement [3]. Control engineers were able to find thousands applications with minor interactions between arguments and/or make adjustments in logical connectors/operators using neural networks and other learning techniques to accommodate interactions. Medical field needs similar techniques. Study of breast cancer diagnosis shows many interactions between arguments, which can not be ignored. For instance in [4] these interactions were restored by the interactive monotone Boolean function approach and used for developing diagnostic rules.

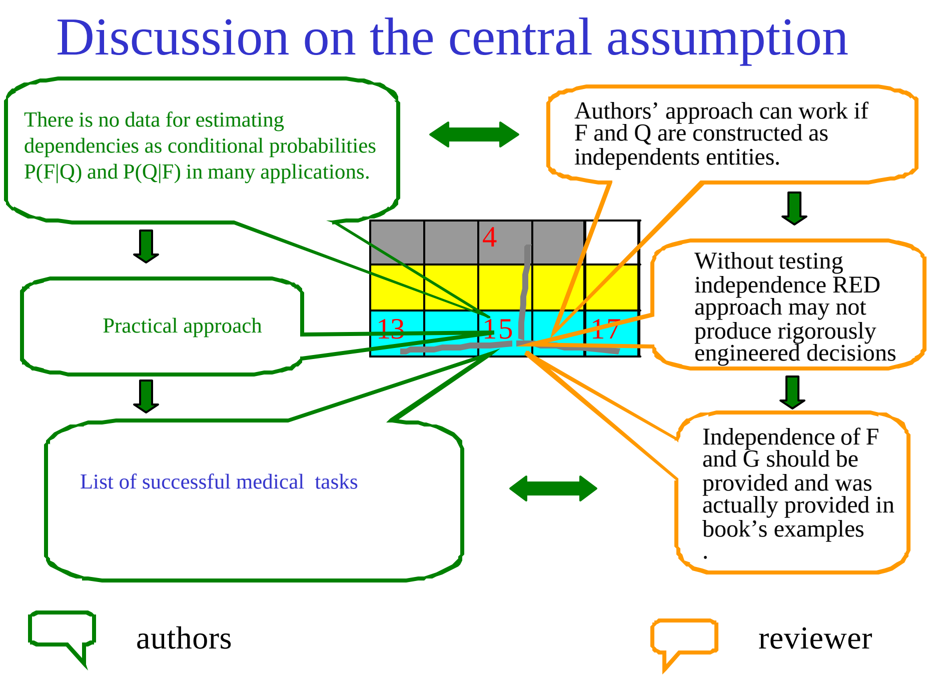

Figure 6.

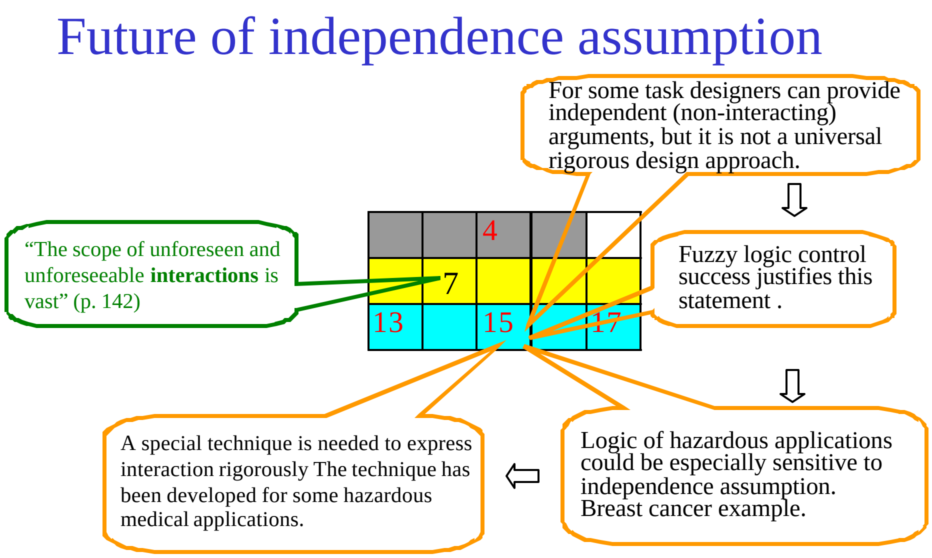

Figure 7.

Concluding remarks. The problem of safety in hazardous applications is a growing area of research and applications. The authors' approach is fruitful and the book can be a useful reference for medical informatics professionals.

References

[1] Russell, S., Norvig, P., *Artificial Intelligence: Modern Approach*, Prentice Hall, 1995,

[2] Bellman, R., Giertz, M. On the analytical formalism of the theory of fuzzy sets, *Information Sciences*, N. 5, 1973, pp. 149-156.

[2] Kovalerchuk, B., Context Spaces as Necessary Frames for Correct Approximate Reasoning, *International Journal of General Systems*, vol. 25(1), pp. 61-80, 1996.

[3]. Kovalerchuk, B., Vityaev E., Ruiz J.F., Consistent Knowledge Discovery in Medical Diagnosis, *IEEE Engineering in Medicine and Biology Magazine*, vol. 19, N, 4, 2000, pp. 26-37.